IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 31 July 2024, accepted 19 August 2024, date of publication 27 August 2024, date of current version 5 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3450793

RESEARCH ARTICLE

Direct Current Control of Grid Connected Two Level Inverter With LCL-Filter Using Deep Reinforcement Learning Algorithm

ANUGULA RAJAMALLAIAH[1], (Graduate Student Member, IEEE), SRI PHANI KRISHNA KARRI[1], (Member, IEEE), MAMDOUH L. ALGHAYTHI[2], (Member, IEEE), AND MESHARI S. ALSHAMMARI[2], (Member, IEEE)

[1]National Institute of Technology Andhra Pradesh, Tadepalligudem, Andhra Pradesh 534101, India
[2]Department of Electrical Engineering, College of Engineering, Jouf University, Sakaka, Aljouf 72388, Saudi Arabia

Corresponding author: Mamdouh L. Alghaythi (mlalghaythi@ju.edu.sa)

This work was supported by the Deanship of Graduate Studies and Scientific Research at Jouf University under Grant DGSSR-2024-02-01126.

ABSTRACT This work presents a novel control paradigm to improve the Direct Current Regulation (DCR) of two-level inverters that are connected to the grid with LCL filters. The Deep Reinforcement Learning (DRL) based Deep Deterministic Policy Gradient (DDPG) algorithm is utilized to address the constraints of traditional control methods, such as Proportional Integral (PI) controllers and Model Predictive Control (MPC). The suggested method tackles challenges like as non-linearity, model dependency, and parameter fluctuations, which have a substantial impact on the performance of the DCR of grid connected two-level inverters of electric power system. The DDPG algorithm offers a flexible control technique that enables adaptive learning and optimization of policies. The results are validated in Real time mode Hardware In Loop (HIL) using Opal-RT & Texas instrument launchpad. Simulations performed on the MATLAB platform provide a reliable testing environment to assess the effectiveness of the proposed controller compared to traditional alternatives. The simulation results clearly show that the DDPG-based controller performs better than any other controller. This strategy surpasses traditional methods, demonstrating an increased resistance to reliance on specific models and uncertainties in parameters.

INDEX TERMS Deep reinforcement learning, direct current regulation, grid-connected inverters, LCL filters, deep deterministic policy gradient, PI controller, model predictive control.

I. INTRODUCTION

Over the past five years, Deep Reinforcement Learning (DRL) has significantly accelerated progress in the areas of power electronics and electrical engineering [1], [2]. DRL has made substantial contributions to enhancing the performance of electrical applications by optimizing the switching of converters and inverters within the field of power electronics [3], [4], [5]. Its applications span various domains, including motor drives, renewable energy systems, and electric vehicles [6], [7], [8], [9], [10], [11]. Continuous research and advancements in DRL are anticipated to profoundly reshape the power electronics industry and the broader discipline of electrical engineering. Notably, DRL has emerged as a potent and influential force within the realm of electrical engineering, particularly concerning renewable energy sources [12], [13], [14]. Figure 1 illustrates the evolving applications of DRL in various electrical engineering domains such as electric drives, DC microgrids, Maximum Power Point Tracking (MPPT) in solar photovoltaics, and electric vehicle charging stations [2], [15].

In the context of electric drives, Deep Reinforcement Learning (DRL) has proven effective in augmenting speed control for Permanent Magnet Synchronous Machines (PMSMs) [16], [17], [18], [19], [20], [21], [22]. The adaptability of DRL algorithms becomes evident in scenarios

The associate editor coordinating the review of this manuscript and approving it for publication was Inam Nutkani.

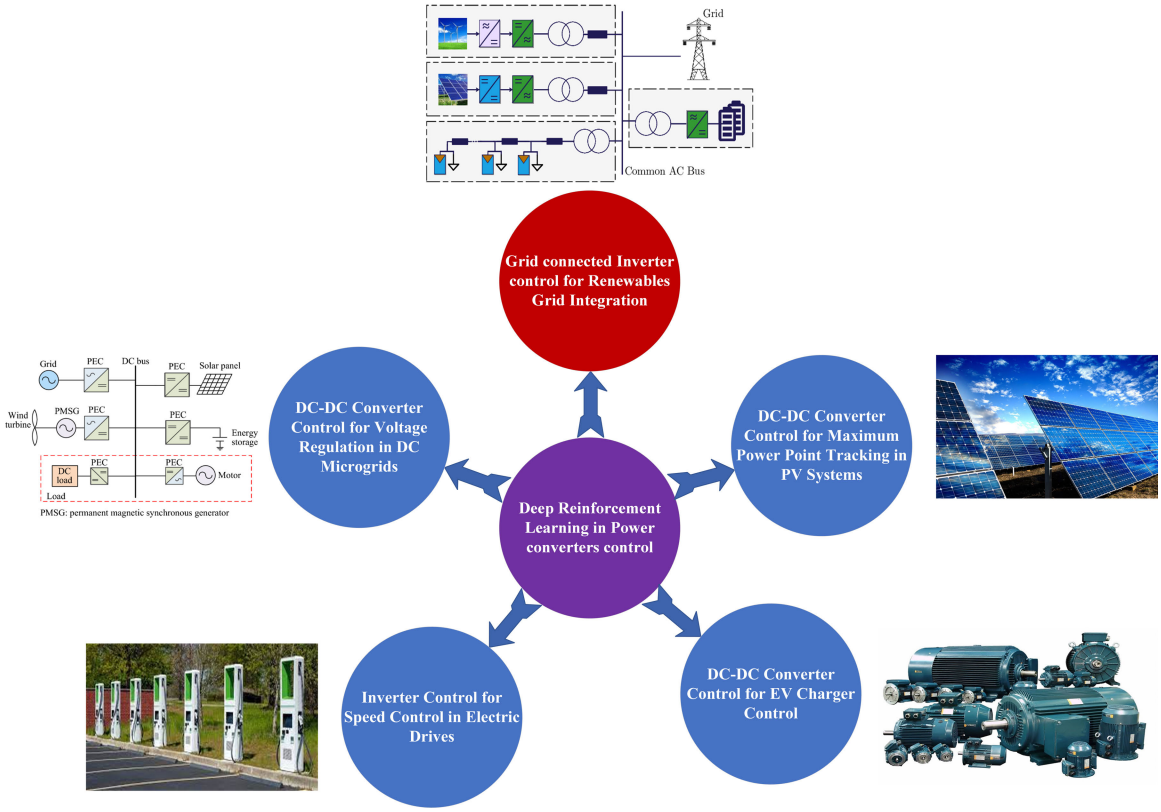

FIGURE 1. Implementing DRL in various electrical engineering applications showcases its effectiveness in renewable grid integration, MPPT in solar photovoltaics, electric drives, EV charging locations, and DC microgrids.

such as voltage regulation within DC microgrids, where they exhibit robust performance, contributing to the stable and efficient operation of microgrids [23], [24], [25], [26], [27], [28], [29], [30], [31], [32]. DRL algorithms have demonstrated remarkable efficacy in tackling the complex obstacles related to Maximum Power Point Tracking (MPPT) for photovoltaic modules, showcasing their ability to optimize power extraction in various environmental situations [33], [34], [35]. Additionally, DRL has demonstrated excellence in optimizing grid-connected inverters, playing a crucial role in achieving accurate current regulation and ensuring the seamless integration of these inverters into larger power systems [36], [37]. In particular, the application of DRL in the control of battery charging for electric vehicles stands out, showing its ability to handle dynamic changes, uncertainties, and variations in the charging process, ultimately contributing to the efficient and reliable operation of electric vehicles in real-world scenarios [38], [39].

This research aims to advance our understanding of the intricate performance of Deep Reinforcement Learning (DRL) in various electrical engineering applications, thereby contributing to the ongoing development in this field. The future of renewable energy grid integration systems is anticipated to be significantly influenced by DRL, leading to the development of more efficient, adaptive, and sustainable energy solutions [40], [41]. DRL, with its dynamic and evolving nature, has the potential to significantly impact the electrical engineering field. It can provide innovative methods and solutions to address the intricate issues associated with power systems and the integration of renewable energy sources [42], [43], [44], [45].

Grid-connected two-level inverters with LCL filters play a crucial role in modern power systems, serving as integral components in renewable energy applications and distributed generation systems [46], [47]. The motivation behind employing these inverters lies in their ability to efficiently convert DC power from renewable sources, such as solar panels or wind turbines, into AC power compatible with the electrical grid. The inclusion of LCL filters (inductor-capacitor-inductor) in these inverters is motivated by the need to address grid requirements for low harmonic distortion and improved performance in terms of current and voltage regulation. LCL filters help mitigate high-frequency harmonics, reduce total harmonic distortion (THD), and enhance the overall stability and reliability of the grid-connected system [48], [49].

One of the significant advantages of grid-connected two-level inverters with LCL filters is their capacity to provide higher quality power output, meeting stringent grid codes and standards. The LCL filter not only improves the system's ability to filter out harmonic content but also aids in damping high-frequency resonances, contributing to enhanced grid

integration. These inverters find applications in a wide range of scenarios, including photovoltaic systems, wind energy converters, and energy storage systems [50]. By ensuring a cleaner and more stable connection to the grid, these inverters facilitate the seamless integration of renewable energy sources, promoting sustainability and reducing the environmental impact of power generation. The ongoing development and optimization of grid-connected two-level inverters with LCL filters underscore their importance in the evolving landscape of power electronics and renewable energy integration [51].

The necessity of DCR in a GCI stems from the imperative to adhere to grid standards and regulations. The LCL filter, while essential for attenuating high-frequency harmonics, introduces resonances that can jeopardize stability if not properly controlled. Current control strategies, such as proportional-integral (PI) controllers, are implemented to address these challenges, ensuring that the inverter's output adheres to specified grid codes. This control is crucial to maintain high-quality power output, minimize harmonic distortions, and stabilize the network connection. Furthermore, optimal power transfer and grid stability depend on effective current control, making it an indispensable component for the reliable integration of renewable energy sources into the electrical grid [52], [53].

Various current control techniques are utilized in the research domain to enhance the performance of these inverters [54], [55], [56], [57]. The merits of these techniques include precise current regulation, reduced harmonic distortion, and improved overall system efficiency. Model Predictive Control (MPC), sliding mode control (SMC), and proportional-integral control (PI) are among the commonly explored approaches. MPC offers advantages in handling non-linearity and uncertainties, providing a dynamic control strategy [58], [59], [60], [61], [62]. Sliding mode control techniques excel at maintaining precise current tracking and responding effectively to changes in the grid or load conditions. One of the significant merits is the robustness of sliding mode control to parameter variations and external disturbances, crucial for ensuring the stable and reliable operation of inverters connected to the grid [63], [64], [65]. On the other hand, PI control, a well-established technique, offers simplicity and ease of implementation [66]. These techniques contribute to achieving grid compliance and power quality standards, essential for the seamless integration of inverters into the electrical grid.

However, these current control techniques are not without their drawbacks and challenges. One notable challenge is the trade-off between achieving fast dynamic response and minimizing overshoot, particularly in transient conditions. The computational complexity associated with advanced control techniques such as MPC can be a drawback in real-time applications [67], [68], [69], [70], [71]. Additionally, the need for accurate modeling of the system dynamics poses a challenge, as inaccuracies can lead to suboptimal performance [72], [73], [74], [75], [76]. Furthermore, in the context of LCL filters, resonance issues may arise, and mitigating these resonances without compromising other aspects of performance remains a research challenge. The notable drawback of sliding mode control in grid-connected inverters is the occurrence of chattering, involving high-frequency oscillations that may introduce harmonic content and degrade current waveform quality. In the context of grid-connected systems prioritizing power quality and adhering to stringent grid code requirements, mitigating chattering poses a critical challenge. Addressing chattering is essential to prevent increased stress on inverter components, ensuring sustained system efficiency and reliability [77], [78]. As researchers explore novel control strategies for LCL-powered grid-connected inverters, addressing these challenges and finding a balance between complexity and performance are crucial for the continued advancement of this technology in power systems.

This article contributes to the existing literature in the following ways:

- The paper presents a control architecture based on the DDPG algorithm for the dq-axis current control of a two-level inverter coupled to a power grid.
- The effectiveness of the proposed control technique has been confirmed by deviating from training conditions.
- The proposed controller has been successfully deployed in real time utilizing OP4512 simulator and a Texas Instrument Launchpad in HIL mode.
- The proposed controller has been compared analytically in steady state and dynamic performance characteristics against well-established and advanced control techniques.

The article is structured in the following manner: Section I explores the implementation of DRL in different applications and emphasizes the significance of GCIs, investigates LCL filter configurations, analyzes the need for current controller strategies, and evaluates existing control mechanisms documented in the available literature. Section II introduces the mathematical model for the GCI, the PI control structure, and provides an overview of the MPC controller structure. Section III provides an explanation of the DRL structure, including its basic framework, as well as the training methods used. Section IV explores the assessment of DRL's performance in both training and testing scenarios and conducts a comparison analysis of its performance. Section V provides a summary of the main discoveries in GCI current controller approaches utilizing DRL and examines possible future research scope avenues in this domain.

II. MATHEMATICAL FORMULATION OF THE CONTROLLER CONFIGURATION FOR THE TWO-LEVEL GCI

The structure of the two-level GCI with a LCL filter configuration can be noticed in Figure 2, displaying the standard configuration for a grid-connected inverter and its associated closed-loop control structure, highlighting the dq-axis. This arrangement consists of three legs, with two MOSFET switches per leg, leading to 8 possible switching

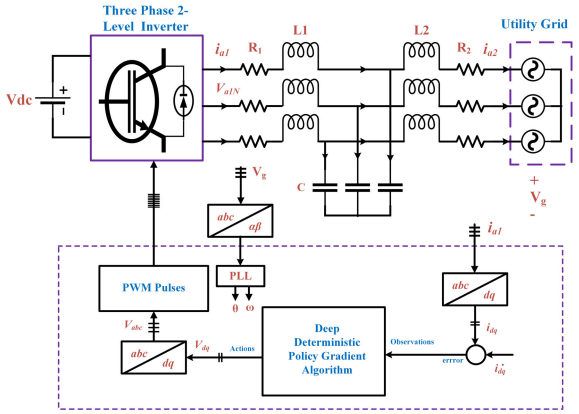

FIGURE 2. An illustrative presentation of the standard configuration for a grid-connected inverter and its associated closed-loop control structure, emphasizing the dq-axis.

states. This configuration entails the participation of multiple components.

The continuous time dynamics of the mentioned configuration are represented by inverter terminal voltage, $V_{inv} = \left[V_{a1N}\ V_{b1N}\ V_{c1N}\right]^T$, the switching state, $S = [S_a\ S_b\ S_c]$, inverter currents, $i_1 = \left[i_{a1}\ i_{b1}\ i_{c1}\right]^T$, grid currents, $i_2 = \left[i_{a2}\ i_{b2}\ i_{c2}\right]^T$ and grid voltages $V_{gabc} = \left[V_{ga}\ V_{gb}\ V_{gc}\right]^T$. The inverter-side filter inductor is denoted by L_1, lumped resistance is denoted by R_1, the filter capacitor by C, and the grid-side inductor is denoted by L_2, lumped resistance is denoted by R_2. The output voltage of the inverter and the switching state are related by $V_m = S_m V_{dc}$ where $m \in \{a, b, c\}$.

The mathematical modeling of a grid-connected two-level inverter with an LCL filter involves capturing the dynamic behavior of the system, which is crucial for designing current control tasks using PI and MPC controllers. Let's denote state variables as i_1, i_2, v_c, input variables as v_{abc}, output variables i_2 and provide a simplified representation of the system. Consider the state-space representation of the grid-connected two-level inverter with LCL filter:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{aligned} \tag{1}$$

where:

$\mathbf{x}(t)$ represents the state vector,

$\mathbf{u}(t)$ represents the input vector,

$\mathbf{y}(t)$ represents the output vector.

where

$$A = \begin{bmatrix} -\frac{R_1}{L_1} & 0 & -\frac{1}{L_1} \\ 0 & -\frac{R_2}{L_2} & \frac{1}{L_2} \\ \frac{1}{C} & -\frac{1}{C} & 0 \end{bmatrix}, B = \begin{bmatrix} \frac{1}{L_1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, D = 0.$$

For ease of controller design, the abc-frame components are projected into the dq-frame, as defined by the following relationship. $V_{dq} = \frac{2}{3} P V_{abc}$, where

$$P = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

A. PI CONTROLLER DESIGN

The current control task using a Proportional-Integral (PI) controller involves regulating the inverter current based on the error signal. Let $i_{\text{ref}}(t)$ be the reference current, and $i(t)$ be the actual inverter current. The control input $u_{\text{dq}}(t)$ for PI is given by:

$$u_{\text{dq}}(t) = K_p \left(e_{\text{dq}}(t)\right) + K_i \int \left(e_{\text{dq}}(t)\right) dt \tag{2}$$

where K_p and K_i are the proportional and integral gains. The error signal $e_{\text{dq}}(t)$ represents the discrepancy between the desired and actual values along the d and q axes. The PI controller requires frequent retuning of the controller gains to handle uncertainties and parameter variations in the system, which can be resource-intensive and impractical for dynamic environments.

B. MPC CONTROLLER DESIGN

Finite Control Set Model Predictive Control (FCS-MPC) is a sophisticated control strategy employed in grid-connected inverters to regulate the grid currents and ensure the stability of the inverter grid integration. The design process for FCS-MPC involves several intricate steps that collectively contribute to its effective implementation in real-world applications. To commence the design process, it is imperative to construct accurate mathematical models of both the grid-connected inverter and the electrical grid itself. These models encapsulate the system dynamics, power electronic components, and the intricate interactions between the inverter and the grid as shown in equation 1. Following this, the continuous-time system model undergoes discretization, converting it into a discrete-time framework suitable for the discrete nature of MPC given by equation 3.

$$V_{dq}(t) = i_{dq}(t)R + L\frac{di_{dq}(t)}{dt} + V_{cdq}(t) \tag{3}$$

A critical aspect of FCS-MPC design involves the creation of a prediction model, which forecasts the system's evolution over a predetermined prediction horizon as shown in equation 4-5. This predictive capability enables the optimization of control inputs based on anticipated future states. Simultaneously, a finite control set is defined, representing the permissible set of control inputs, typically voltage vectors, that the inverter can apply during each control interval.

$$\frac{i_{dq}(k+1) - i_{dq}(k)}{\tau_s} = \frac{V_{dq}(k) - V_{cdq}(k)}{L} - \frac{i_{dq}(k)R}{L} \tag{4}$$

where R, lumped resistance, τ_s, sampling time.

Central to the design of FCS-MPC is the formulation of a comprehensive cost function shown in equation 5, reflecting the desired control objectives. This cost function encompasses terms related to current reference tracking, minimization of grid current harmonics.

$$i_{dq}^{P}(k+1) = i_{dq}(k) + \frac{\tau_s}{L}[V_{dq}(k) - V_{cdq}(k)] - \frac{\tau_s}{L}[i_{dq}(k)R] \tag{5}$$

With the cost function, prediction model, and finite control set in place, the next step involves formulating an optimization problem. This problem aims to minimize the cost function over the prediction horizon, subject to the system dynamics, constraints on control inputs, and compliance with grid requirements. An appropriate optimization solver is then selected to efficiently solve this constrained optimization problem, considering real-time constraints.

The actual control input is calculated by the optimization solver, reflecting the optimal voltage vectors that minimize the defined cost function. This optimal switching control input is applied to the inverter, and the control process is executed in a receding horizon manner. This recalculation of the finite control set, considering the updated system state, ensures adaptability to dynamic conditions.

The drawback of MPC lies in its model dependency, computational complexity, and the need for solving optimization problems in real-time, which may lead to increased computation time. DRL can address this challenge by leveraging neural networks to approximate complex control policies learned through reinforcement learning. DRL methods, such as DDPG, offer a more efficient approach, enabling the system to learn and adapt control strategies in dynamic environments, thus mitigating the model dependency with traditional FCS MPC.

III. CONTROLLER DESIGN

The DRL framework for the two-level inverter connected to the grid with an LCL filter involves an agent that interacts with an environment. The agent makes decisions based on observed states, takes actions to control the inverter switches, and receives rewards as feedback, aiming to optimize the inverter's performance in terms of grid stability, current regulation, and power quality. Figure 3 presents a standard configuration of the DRL framework, highlighting the interactions among the environment (grid-connected inverter), agent (DDPG), and their associated components. The components of a DRL framework in the context of a grid-connected two-level inverter with an LCL filter are defined as:

A. ENVIRONMENT

In this context, the environment refers to the grid-connected two-level inverter with an LCL filter. It encapsulates the dynamics and behavior of the power electronic system, taking into account factors such as the grid conditions, LCL filter parameters, and the overall electrical network.

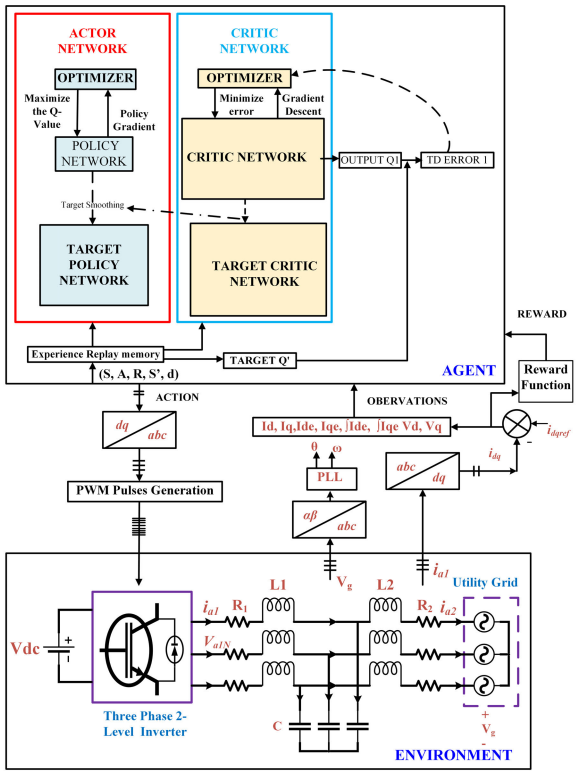

FIGURE 3. Presenting a standard configuration of the DRL framework, highlighting the interactions among the environment (grid-connected inverter), agent (DDPG), and their associated components.

B. AGENT

The agent is the DRL-based controller designed to regulate the dq-axis currents of the inverter. It interacts with the environment, making decisions (actions) to optimize the system's current tracking performance based on the observed states. Here, the agent is a DDPG neural network as shown in Table 1. The table provides detailed configurations for the DDPG neural networks, comprising actor and critic components. The actor network features three layers with neuron counts of 32, 16, and 1, respectively, and utilizes the ReLU activation function, with a learning rate of 0.001. Similarly, the critic network has the same layer configuration but employs the Tanh activation function, with a lower learning rate of 0.0001. General parameters for the DDPG setup include a discount factor of 0.9, a mini-batch size of 512, and an experience buffer length of 2 million, facilitating efficient learning and experience replay. The action space is continuous, indicating the network's capacity to handle a range of values rather than discrete actions.

C. STATES

States represent the relevant variables and parameters that describe the current condition of the system. In the case of the grid-connected inverter, states include the current DC voltage, inverter output current, grid voltage, LCL filter currents, and other pertinent variables that influence the system's behavior. The observation set encompasses feature

TABLE 1. DDPG Neural network configurations.

Type	Variable	Value
Actor	Learning rate	0 0.001 s
	Activation function	ReLU
	Layer wise neurons	$L_1 = 32, L_2 = 16, L_3 = 1$
Critic	Layer wise neurons	$L_1 = 32, L_2 = 16, L_3 = 1$
	Activation function	Tanh
	Learning rate	0.0001 s
General	Discount factor	0.9
	Action space	Continuous
	Experience buffer length	$2X10^6$
	Mini batch size	512

parameters, including the grid voltages in the dq-frame ($V_{dg}(t), V_{qg}(t)$), and the inverter's output currents in the d-q frame ($I_d(t), I_q(t)$). The set additionally incorporates the error signals along the d-q axis, represented as $e_d(t)$ and $e_q(t)$, together with their corresponding integral values $\left(\int e_d(t), \int e_q(t)\right)$. Furthermore, the observational set is precisely articulated through the formal expression given in Equation (11):

$$\mathrm{O} = \left[V_{dg}(t), V_{qg}(t), I_d(t), I_q(t), e_d(t), e_q(t), \int e_d(t), \int e_q(t) \right] \tag{6}$$

Herein, $V_{dg}(t)$ and $V_{qg}(t)$ represent the voltages of the grid in the dq-frame.

D. ACTIONS

Actions refer to the control inputs that the agent is capable of impacting the system. In the DC control of the inverter, actions may involve adjusting the modulation index, changing the switching frequency, or manipulating other control parameters to regulate the output and respond to variations in the grid or load conditions. In the context of inverter current control, the tasks may involve estimating the dq-voltage values (V_d, V_q), these then be transformed into benchmark values in the abc frame for generating PWM pulses, as seen in Figure 3.

E. REWARD

The reward signal serves as feedback to the agent, indicating the desirability of its actions in the given state. In the context of the inverter, the reward could be designed to encourage objectives such as achieving precise current regulation by minimizing the error signal and responding effectively to dynamic changes in the grid. The reward signal is given by the equation 7.

$$R(t) = \begin{cases} 1 - |e_{dq}(t)|, \text{ if } |e_{dq}(t)| \leq 0.001 \\ 0.01 - |e_{dq}(t)|, \text{ if } 0.001 < |e_{dq}(t)| \leq 0.01 \\ 0.001 - |e_{dq}(t)|, \text{ if } 0.01 < |e_{dq}(t)| \leq 0.1 \\ -5|e_{dq}(t)|, \text{ if } 0.1 < |e_{dq}(t)| \end{cases} \tag{7}$$

The given equation defines the reward function $R(t)$ based on the error $|e_{dq}(t)|$ in a control system. The reward function assigns different values depending on the magnitude of the error:

- If the error is very small ($|e_{dq}(t)| \leq 0.001$), the reward is high ($1 - |e_{dq}(t)|$), incentivizing precise control.
- For errors slightly larger ($0.001 < |e_{dq}(t)| \leq 0.01$), the reward decreases to $0.01 - |e_{dq}(t)|$.
- For moderate errors ($0.01 < |e_{dq}(t)| \leq 0.1$), the reward further reduces to $0.001 - |e_{dq}(t)|$.
- If the error exceeds 0.1, the reward becomes significantly negative ($-5|e_{dq}(t)|$), penalizing large deviations.

This structure encourages the control system to minimize errors, rewarding precision and penalizing significant inaccuracies.

Algorithm 1 Deep Deterministic Policy Gradient (DDPG)

Input: Initial policy parameters θ, Q-function parameters ϕ, empty reply buffer D
Set target parameters equal to main parameters $\theta_{\text{targ}} \leftarrow \theta$, $\phi_{\text{targ}} \leftarrow \phi$
repeat
  Observe state s and select action $a = \text{clip}\left(\mu_\theta(s) + \epsilon, a_{\text{Low}}, a_{\text{High}}\right)$, where $\epsilon \sim \mathcal{N}$
  Execute action a in the environment
  Observe next state s', reward r, and done signal d to indicate whether s' is terminal
  Store (s, a, r, s', d) in reply buffer D
  If s' is terminal, reset environment state
  it's time to update **then**
  for however many updates **do**
    Randomly sample a batch of transitions $B = \{(s, a, r, s', d)\}$ from D
    Compute targets: $y\left(r, s', d\right) = r + \gamma(1 - \mathrm{d})\, Q_{\phi_{\text{targ}}}\left(s', \mu_{\theta_{\text{targ}}}\left(s'\right)\right)$
    Update Q-function by one step of gradient descent using:
$$\nabla_\phi \frac{1}{|B|} \sum_{(s,a,r,s',d) \in B} \left(Q_\phi(s, a) - y\left(r, s', d\right)\right)^2$$
    Update policy by one step of gradient ascent using:
$$\nabla_\theta \frac{1}{|B|} \sum_{s \in B} Q_\phi\left(s, \mu_\theta(s)\right)$$
    Update target networks with:
    $\phi_{\text{targ}} \leftarrow \rho\phi_{\text{targ}} + (1 - \rho)\phi$,
    $\theta_{\text{targ}} \leftarrow \rho\theta_{\text{targ}} + (1 - \rho)\theta$
  end for
  end if
until convergence

F. TRAINING PROCEDURE

DDPG is a robust deep reinforcement learning algorithm specifically developed to address reinforcement learning issues involving continuous action spaces. In contrast to discrete action spaces, which allow for the explicit enumeration and selection of actions, continuous action spaces necessitate the use of more advanced methodologies. DDPG tackles this difficulty by integrating components from policy gradient techniques and Q-learning.

At its core, DDPG maintains two important components: an actor and a critic. The actor is responsible for selecting actions, and the critic evaluates these actions by estimating their expected return. The algorithm operates iteratively, continuously improving both the actor and the critic to learn an optimal policy.

The training procedure of the DDPG algorithm consists of several sequential steps, as outlined in Algorithm 1 (citation). To start, as shown in Step 1, the algorithm initializes its policy parameters (θ) and Q-function parameters (ϕ). Additionally, it creates an empty replay buffer (D) to store experiences, which is crucial for breaking the temporal correlations in the data and improving the stability of learning.

The concept of target networks is a key feature of DDPG: target actor and target critic. These target networks, which include both a target policy network θ_{targ} and a target Q-function network ϕ_{targ}, are assigned the same values as the primary parameters θ and ϕ in Step 2. However, they are only updated periodically and with a soft update using a parameter ρ, which helps stabilize training.

The main loop commences in Step 3 and iterates through the following phases until a specified stopping criterion is satisfied. In Step 4, the current state s is observed, and an action a is selected by introducing clipped noise to the output of the policy function. Step 5 executes the selected action in the environment, followed by observing the next state s', the reward r, and the done signal d in Step 6. In Step 7, these observations are subsequently saved in the replay buffer D. If the next state is a terminal state, the environment state is reset in Step 8. During the model update process, as specified in Step 9, a set of transitions is randomly selected from the replay buffer for updating, as explained in Step 10.

The core of DDPG's training loop involves selecting actions, executing them in the environment, and storing the resulting experiences in the replay buffer. These experiences consist of the current state (s), the selected action (a), the received reward (r), the next state (s'), and a done signal (d) that indicates if the next state is terminal. Importantly, DDPG encourages exploration by adding noise sampled from a normal distribution to the actions. This noise helps the agent explore the state space effectively.

The update phase is where DDPG shines. When it's time to update, the algorithm randomly samples a batch of transitions from the replay buffer (D) as shown in Step 11. For each transition, it computes the target Q-values (y) based on the observed reward, next state, and the target networks as shown in Step 12. This step is essential for training the critic network (ϕ) effectively.

DDPG then performs two critical updates. First, it updates the critic network (ϕ) using gradient descent to minimize the mean squared error between the current Q-values and the target Q-values as shown in Step 13. This step ensures that the critic provides accurate value estimates.

Second, it updates the actor network (θ) using gradient ascent, aiming to maximize the expected Q-value for the current state as shown in Step 14. This update encourages the actor to select actions that lead to higher returns.

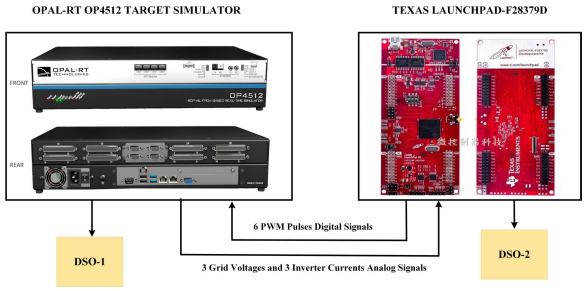

FIGURE 4. An arrangement demonstrating the step-by-step process of implementing the proposed DDPG algorithm online.

Finally, DDPG performs the target network updates by softly blending the main networks and the target networks as shown in Steps 15 to 17. This gradual update helps stabilize learning by reducing target volatility.

The algorithm repeats this process until convergence as shown in Step 20, typically monitored by tracking the performance improvement or a predefined stopping criterion. After completing the training successfully, the agent is saved as an optimal policy. It is then deployed onto the controller board for real-time experimentation in the following section.

IV. KEY FINDINGS AND ANALYSIS

This section specifically examines the results of DCR in a 3-phase two-level GCI. The utilized control techniques include the application of the DDPG algorithm, along with traditional PI and MPC approaches. The PI gains are determined using the Ziegler-Nichols strategy for initial estimation, followed by a trial-and-error method to obtain the best gains. The gains are represented as follows:

$$K_{P_1} = 5, \quad K_{I_1} = 5000, \quad K_{P_2} = 5, \quad K_{I_2} = 5000 \tag{8}$$

The execution and verification of the DRL control framework require an extensive range of simulations and real-time experiments. During the training phases of DRL, simulations are performed using MATLAB 2021 software on a powerful Dell system 5820 (32GB RAM, 16GB GPU). The training approach yields an optimized Deep DRL agent utilizing the Math Kernel module for Deep Neural Networks (MKL-DNN) module.

Real-time validation is a critical aspect, and for this purpose, the OP4512 OPAL-RT simulator and Texas Launchpad F28379D are employed. The hardware setup, illustrated in Figure 4, ensures seamless coordination between components. Experimental verification involves meticulous evaluation using the OP4512 Opal-RT Target simulator, F28379D Texas Launchpad, and host computer, as depicted in Figure 4.

The OP4512 simulator is equipped with 16GB of RAM, an Intel Xeon E3 processor, and 256GB of SSD memory, facilitating real-time simulations. It boasts a range of input and output channels, including Analog In, Analog Out,

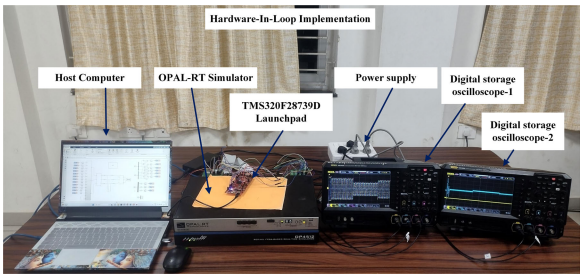

FIGURE 5. The Real-Time Hardware Setup is implemented utilising OP4512 simulator, F28379D microprocessor, Dell computer.

Digital In, and Digital Out, ensuring flexible data transfer. The TMS320F28379D Launchpad, the Host computer, the OPAL-RT simulator, and it's software platform work together to coordinate the verification of the Real-Time simulation.

The TMS320F28379D Launchpad, with a processing speed of 200MHz includes two C28xCPUs, two CLAs, 1MB Flash memory, and features like ADCs, DACs, and PWMs. The Exploration advances through a structured process, starting with training of the DDPG agent in offline using simulink and transitioning to real-time implementation on both the Op4512 simulator and the TMS320F28379D microprocessor.

The hardware connections, as illustrated in Figure 5, facilitate that a OP4512 system delivering grid voltages and inverter currents, and the launchpad receiving and interpreting these signals. Real-Time execution, detailed in the flowchart in Figure 4, involves configuring the processor, integrating the model onto the OPAL target, setting I/O parameters, and designating the target, confirming the control technique's efficacy in a real-time setting. In the HIL setup, Opal-RT simulates the real-time behavior of a 2-level GCI, while the Texas Instruments F28379D LaunchPad executes the DDPG algorithm for control. The process begins with Opal-RT initializing the simulation model and establishing communication with the F28379D LaunchPad. During simulation, Opal-RT generates real-time signals representing voltage and current outputs of the inverter, which are then transmitted to the LaunchPad.

On the LaunchPad, the DDPG algorithm processes these signals to generate control commands, which are sent back to Opal-RT. The algorithm, comprising actor and critic networks, continuously evaluates and refines the control strategy based on feedback received from Opal-RT. This feedback loop enables the algorithm to dynamically adjust the inverter's operation to meet performance objectives such as voltage stability and harmonic reduction. By leveraging the computational capabilities of the F28379D LaunchPad, the DDPG algorithm ensures efficient and adaptive control of the inverter in real-time, as depicted in the provided flowchart.

The outcomes derived from the application of the DDPG technique in both the training and testing scenarios are deliberated in Case-A. The effectiveness of the proposed DDPG method is evaluated in Case-B by comparative assessment. The results are Verified utilizing the real-time Opal-RT experimental setup. The control approach described in Case-C is evaluated for its ability to track fluctuating references, simulating real-world settings. Case-D involves conducting a thorough study to compare and evaluate the controllers' collective dynamic responses. This analysis aims to discover the relative degrees of excellence of each controller under typical working circumstances and when there are fluctuations in parameters. Lastly, in Case-E, the dq-currents' tracking stability is evaluated during grid dynamic conditions. The characteristics of the GCI are obtained using well-established real-time research techniques, as detailed in [61] and [70]. Table 2 provides a summary of the parameters to ensure reproducibility. The table outlines key configuration parameters for a two-level inverter with an LCL filter, indicating a power rating of 100 KVA, an output voltage of 415 volts Line-to-Line, a DC link voltage of 700 volts, and operating at a frequency of 50 Hz. Notably, the LCL filter resonates at 1 KHz, while the inverter switches at 10 KHz. Capacitance and inductance values of 100 microfarads and 1.5 millihenries, respectively, underscore the filter's design importance. The sampling time of 1 microsecond highlights control system time resolution. These parameters collectively inform system capabilities and design considerations for efficient and effective operation.

TABLE 2. The configuration parameters for 2 level GCI.

S. No	Variable	Value
1	Power capacity of an inverter	100 KVA
2	Output voltage	415 V_{L-L}
3	DC link voltage	700 V
4	Frequency (f)	50 Hz
5	Resonating frequency (f_{res})	1 KHz
6	Switching frequency (f_{sw})	10 KHz
7	Capacitance (C)	100 μ Farad
8	Inductance (L)	1.5 mH
9	Sampling time (T_s)	1 μ sec

A. CASE-A: PERFORMANCE OF DDPG CONTROLLER IN TRACKING REFERENCE CURRENT

The DRL agent is trained using the DDPG algorithm for a total of 910 episodes on a 3-phase two-level inverter equipped with an LCL filter. The training incorporates fluctuations in dq-reference current as described by equations (9) - (10).

$$i_d(t) = \begin{cases} -180\,A, & \text{if } 0 < t \leq 0.4 \\ -80\,A, & \text{if } 0.4 < t \leq 0.7 \\ -160\,A, & \text{if } 0.7 < t \leq 1.0 \end{cases} \tag{9}$$

$$i_q(t) = \begin{cases} 0\,A, & \text{if } 0 < t \leq 1.0 \end{cases} \tag{10}$$

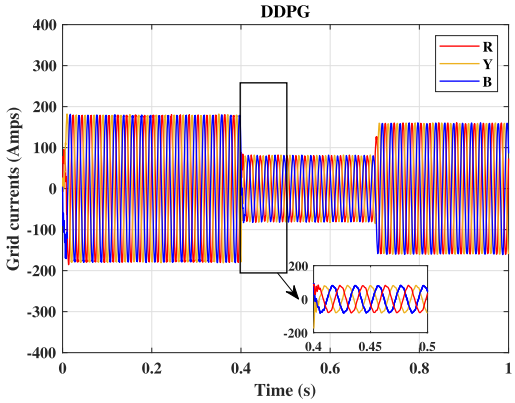

FIGURE 6. The 3-phase grid currents utilizing the DDPG method under training conditions.

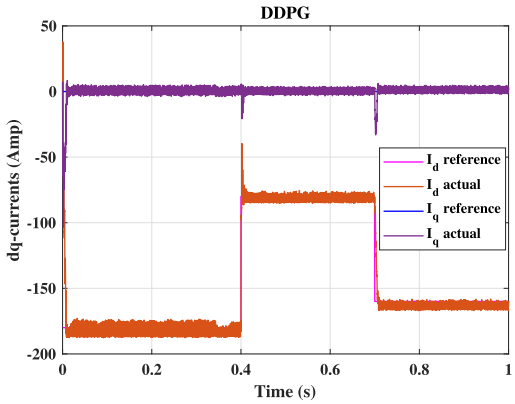

FIGURE 7. Using the DDPG approach, performance tracking of dq-axis currents under training conditions.

After successfully completing the training, testing is conducted on the agent using distinct reference current waveforms, which deviate from the waveform values used during training, as described by equations (11) - (12).

$$i_d(t) = \begin{cases} -160\,A, & \text{if } 0 < t \leq 0.4 \\ -110\,A, & \text{if } 0.4 < t \leq 0.7 \\ -60\,A, & \text{if } 0.7 < t \leq 1.0 \end{cases} \tag{11}$$

$$i_q(t) = \begin{cases} 0\,A, & \text{if } 0 < t \leq 1.0 \end{cases} \tag{12}$$

Figure 6-8 display waveforms depicting variations in dq-axis currents, grid currents, and THD of the two-level GCI. These figures represent the training condition. Similarly, Figure 9-11 show the same waveforms during the testing phase. These graphs clearly demonstrate that the grid currents effectively align with the reference waveforms and that the grid current's THD complies with the conditions specified in the IEEE 519 grid code. Moreover, it's clear that the trained agent has developed a skilled comprehension of the fundamental patterns and connections inherent in the training data, enabling its application to to unfamiliar data.

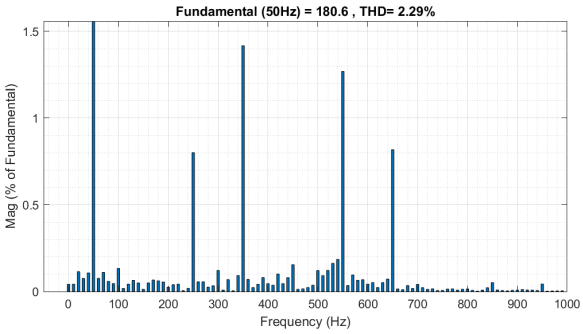

FIGURE 8. Using the DDPG approach in a training case, the THD of 3-phase grid currents is 2.29%.

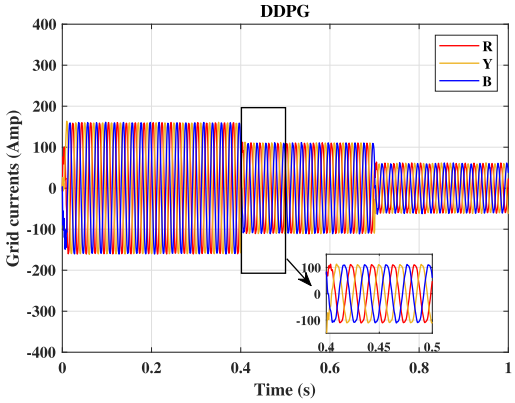

FIGURE 9. Grid currents in three phases under test conditions with the DDPG approach.

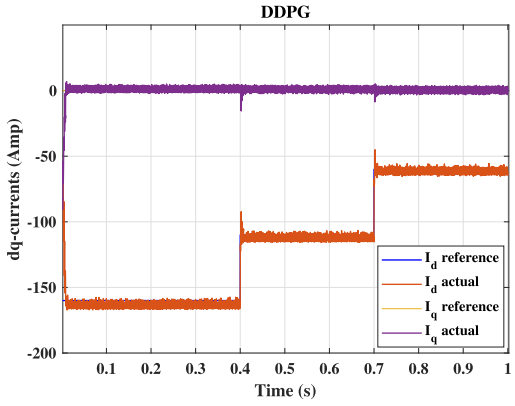

FIGURE 10. Performance tracking with the DDPG approach for dq-axis currents under test conditions.

B. CASE-B: HIL VERIFICATION OF CONTROLLERS EMPLOYING DDPG, PI CONTROL, MPC STRATEGIES

The PI, MPC and DDPG algorithms are deployed in F28379D and it's real-time results are showed in following manner. Figure 12 illustrates the real-time validation of 3-phase grid-side current waveforms using the DDPG method, showcasing its effectiveness in accurately capturing and reproducing grid currents. In Figure 13, the DDPG method's real-time

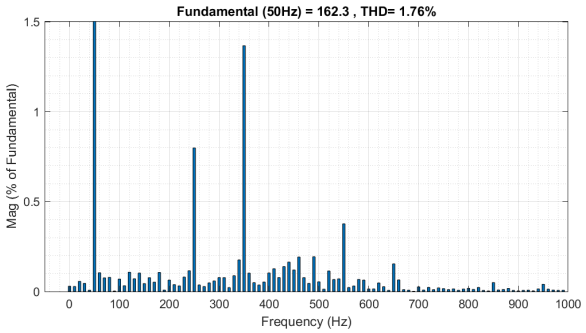

FIGURE 11. Using the DDPG approach in a testing example, grid currents THD is 1.76%.

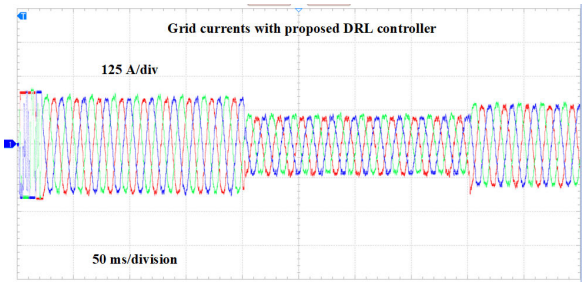

FIGURE 12. DDPG method-based real time validation of 3-phase grid side current waveforms.

validation of DQ-axis current tracking performance is depicted, demonstrating its ability to precisely track desired current trajectories. Similarly, Figure 14 showcases the utilization of the PI approach to validate three-phase grid-side current waveforms in real-time, providing insights into the method's performance in current regulation. Figure 15 complements this by exhibiting the real-time validation of DQ-axis current tracking performance using the PI method, highlighting its capability in accurately tracking desired current trajectories. Moreover, Figure 16 demonstrates the real-time validation of three-phase grid-side current waveforms employing the MPC technique, underscoring its effectiveness in regulating grid currents. Lastly, Figure 17 illustrates the real-time validation of DQ-axis current tracking performance using the MPC approach, showcasing its ability to accurately track desired current trajectories over time.

Figures 18 through 20 show the THD graphs for each of the three ways. In normal operation, all controllers show respectable dynamic and steady-state performance. Nonetheless, the following scenario will assess their resilience, static, and dynamic performance features.

C. CASE-C: ADAPTABILITY TO VARIATIONS IN REFERENCE CURRENT

In the preceding two assessments, step waveforms have been employed as reference waveforms. However, in practical applications such as grid-connected inverters used in solar PV or wind grid integration, the reference waveforms fluctuate due to uncertainties in irradiance, temperature, and

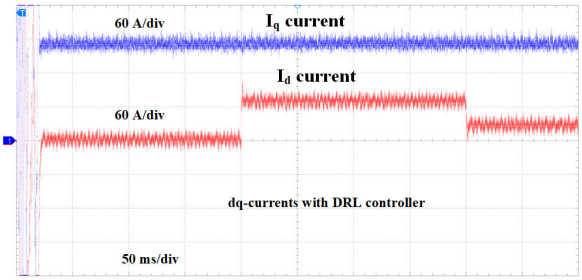

FIGURE 13. DDPG method-based real-time validation of DQ-axis current tracking performance.

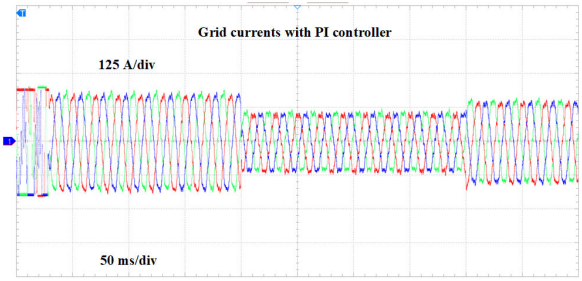

FIGURE 14. Using the PI approach to validate three-phase grid-side current waveforms in real-time.

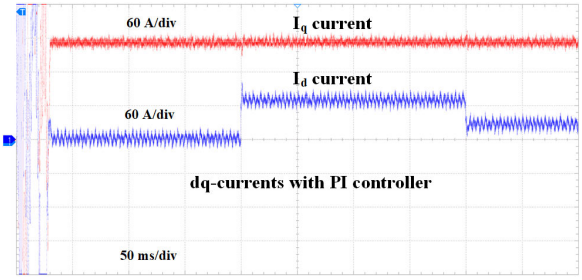

FIGURE 15. PI method-based real-time validation of DQ-axis current tracking performance.

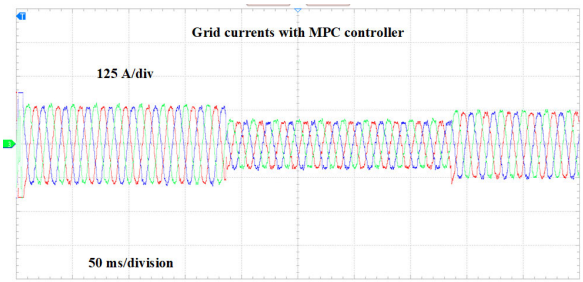

FIGURE 16. Three-phase grid-side current waveforms are validated in real time using the MPC technique.

wind speed. Equation (13) defines the randomly fluctuating reference waveform in such scenarios.

Equation (13) defines a fluctuating reference waveform $S(t)$ commonly seen in practical applications like grid-connected inverters for solar PV or wind grid integration. It consists of a constant DC component S_{dc} and a sum of sinusoidal components, each characterized by its amplitude

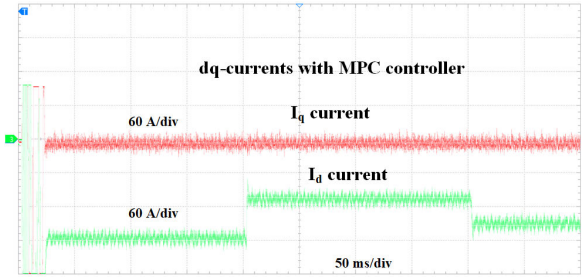

FIGURE 17. Utilising the MPC approach to validate DQ-axis current tracking performance in real time.

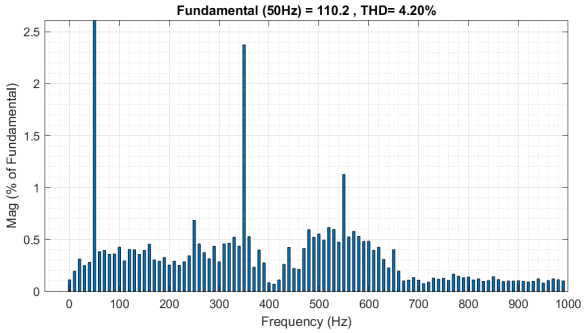

FIGURE 18. Using the DDPG approach, the THD for grid-side three phase currents is 4.20%.

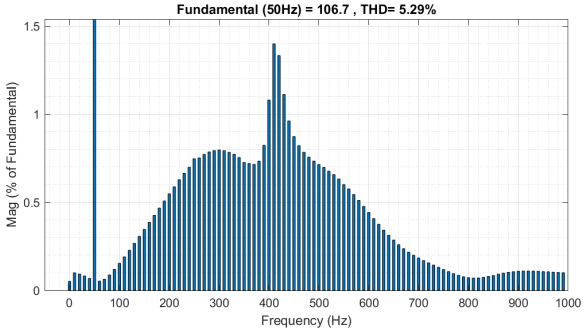

FIGURE 19. Using the PI approach, the THD for grid-side three phase currents is 5.29%.

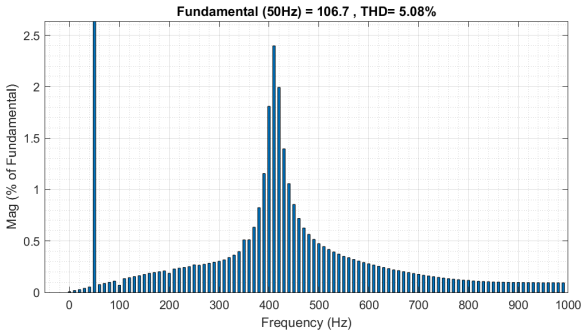

FIGURE 20. Using the MPC approach, the THD for grid-side three phase currents is 5.08%.

A_p, angular frequency ω_p, and phase angle ϕ_p. This equation captures the variability of reference waveforms due to factors

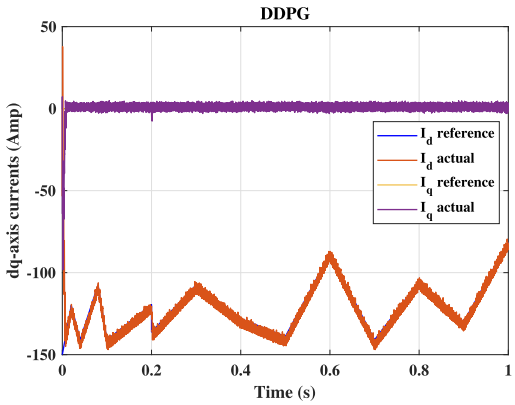

FIGURE 21. The tracking the dq axis currents performance under varying reference currents using the DDPG approach.

such as irradiance, temperature, and wind speed fluctuations.

$$S(t) = S_{dc} + \sum_{p=1}^{M} A_p \sin(\omega_p t + \phi_p) \tag{13}$$

In this unstable and chaotic wind environment, a solar PV generates electricity by utilizing a negative d-axis value and a zero q-axis value, indicating the absence of reactive power.

In addition, Figure 21 demonstrates the remarkable efficiency of the controller based on DRL in accurately following the fluctuating reference current waveform under optimal switching conditions in the power converter.

D. CASE-D: COMPARATIVE ASSESSMENT OF THE PROPOSED METHOD PERFORMANCE IN CASE OF UNCERTAINTY IN PARAMETERS

The comparative assessment evaluates the performance of the proposed method under conditions of parameter uncertainty, specifically focusing on variations in filter inductance. This analysis demonstrates how the proposed method adapts to changes in inductance values, which are critical in applications like GCIs. The method's robustness is highlighted by its ability to maintain stability and performance despite these variations, showcasing superior adaptability compared to traditional control strategies [79]. This robustness ensures reliable operation in real-world scenarios where filter inductance can fluctuate due to manufacturing tolerances, aging components, or environmental factors, underscoring the proposed method's effectiveness in handling practical uncertainties [80].

In this scenario, a comprehensive assessment is conducted to compare the proposed DDPG technique with traditional PI and MPC methods in terms of their performance in handling dynamic parameter changes. The study specifically investigated two parameter variations: modifications in the resistance and inductance of the grid filter. The results revealed a notable effect on the controller, with observable effects arising from alterations in the grid filter inductance rather than changes in resistance. The controller exhibited

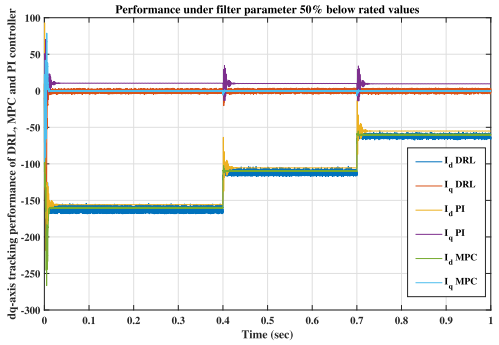

FIGURE 22. Examining the static & dynamic performance of dq current tracking using PI, MPC & DRL control methods with nominal parameters.

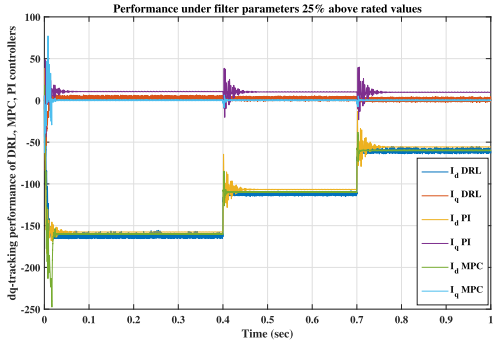

FIGURE 23. PI, MPC, and DRL control algorithms' static and dynamic responses to dq-current tracking under parameter mismatch situations are compared, with values 25% above normal levels.

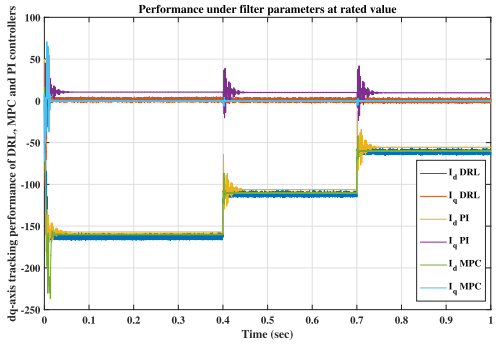

FIGURE 24. PI, MPC, and DRL control approaches are compared with the static and dynamic responses of dq-current tracking in a parameter mismatch scenario where values exceed nominal values by 25%.

remarkable durability in the face of fluctuations in the resistance of the grid filter. An important discovery showed that the controller had less influence when the grid filter inductance surpassed the specified value; however, adverse outcomes occurred when the inductance fell below the set threshold. The examination focuses on evaluating the effectiveness of performance in both fixed and moving conditions, as well as compliance with grid code criteria. Figure 22 exhibits the current management of the dq-axis accomplished by utilizing the DDPG technique, the PI, and MPC controllers, and their comparison under typical

TABLE 3. Evaluating steady & dynamic performance measures for DRL, MPC, and PI control methods across different operating situations.

Control method	Variation in parameters	Steady state error (%)	Settling time (ms)	% Overshoot	THD (%)
PI	50 % below	8.7	49.5	41.7	5.08
	Nominal	4.4	36	40.9	4.39
	25 % above	10.5	53	42	5.01
MPC	50 % below	11.6	22	22.7	4.90
	Nominal	7.3	10	20.9	4.80
	25 % above	14.5	15	22.9	6.7
DRL	50 % below	4.5	7	12	4.45
	Nominal	1.8	4	10	4.20
	25 % above	6.0	8	20	4.32

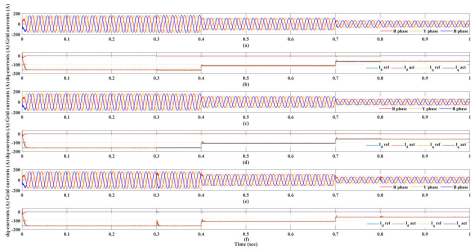

FIGURE 25. Demonstrating the dq-current's stability in the face of phase angle change and dynamic grid voltage situations.Variation in parameters.

operating conditions. Figures 23 and 24 depict the current tracking performance under two conditions: when the filter inductance parameter is less than 50% of the nominal value and when it is greater than 25% of the nominal value, respectively. The analysis of both stable and changing performance is displayed in Table 3 for all three situations.

The table evaluates the static and transient performance of PI, MPC, and DRL control methods across various operating conditions (50% below nominal, nominal, and 25% above nominal). Key metrics include percentage overshoot, settling time, steady-state error, and THD. The DRL control method demonstrates superior performance with consistently lower percentage overshoot, faster settling times, reduced steady-state error, and competitive THD values compared to PI and MPC across all evaluated scenarios, highlighting its effectiveness in handling diverse and dynamic operational conditions.

The performance evaluation reveals that the DRL control method outperforms PI and MPC methods across various operating conditions. DRL consistently achieves lower percentage overshoot, faster settling times, and reduced steady-state errors, indicating its superior stability and efficiency. Additionally, DRL maintains competitive THD values, further underscoring its effectiveness in managing diverse and dynamic operational scenarios. These findings suggest that DRL is a more robust and reliable control method for optimizing performance in variable grid conditions.

- The ranking of robustness in terms of overshoot is as follows: DDPG is the most robust, followed by MPC, and then PI. The DDPG controller consistently exhibits the lowest percentage of overshoot, maintaining a value of less than or equal to 10% even when subjected

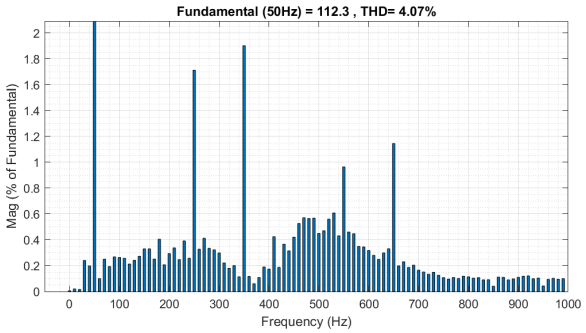

FIGURE 26. Grid current THD study in the context of dynamic voltage rise.

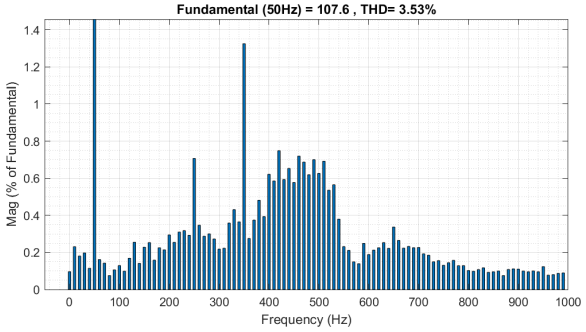

FIGURE 27. Grid currents' THD investigation with dynamic voltage dip circumstances.

to higher step changes across all parameter variables. It demonstrates a high level of resilience regarding overshooting.

- The settling time is ordered as follows: DDPG < MPC < PI. The DDPG-based controller has a more rapid response.
- The controllers can be ordered in terms of steady-state error as follows: DDPG < MPC < PI. Out of all the controllers, the DDPG controller demonstrates the most efficient and stable performance in the long term. It efficiently reduces steady-state error, showcasing outstanding steady-state performance.
- The controllers, ranked in ascending order based on their THD values, are DDPG, MPC, and PI. The DDPG controller regularly delivers decreased THD values.

This study emphasizes the enhanced stability and efficiency of the DDPG controller compared to PI and MPC.

E. CASE-E: EVALUATION OF THE PROPOSED METHOD IN VARIOUS CHANGING GRID SCENARIOS

In current scenario, the stability of the suggested controller has thoroughly evaluated while considering evolving grid voltage settings. These settings are defined by voltage fluctuations spanning from 0.9 to 1.1 V P.U. The evolving condition is deliberately initiated, starting at time t=0.05 s and concluding at t = 0.35 seconds, in order to mimic scenarios including both voltage escalation and reduction, as well as phase angle alterations. Figure 25 demonstrates

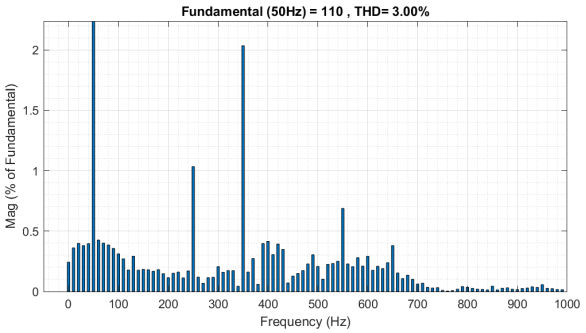

FIGURE 28. Grid current THD in the event of phase angle distortion.

that the controller effectively maintains stability by precisely tracking the reference current, even in the presence of voltage fluctuations caused by swell, sag, and phase angle variation dynamics. Additionally, the THD checked under these specific circumstances adheres to the grid code standards specified by the IEEE 519. The information is depicted in Figures 26 to 28. This thorough assessment highlights the controller's skill in maintaining stability and enhancing performance across a variety of dynamic and ever-changing operational conditions.

V. CONCLUSION

In conclusion, this article significantly contributes to the existing body of literature by introducing a novel control architecture based on DRL for the DCR of a inverter integrated into a power grid. The paper presents the technical details of introduction of DRL into the grid connected inverter with training and testing case studies. The efficacy of the proposed control technique has been validated through experimentation, revealing its superior performance compared to several traditional control strategies. The successful real-time deployment of the DRL-based controller, utilizing Opal-rt and a F28379D microprocessor in HIL mode, underscores its practical applicability. The performance of the proposed controller has been validated for real time scenarios of renewable applications. Furthermore, The analytical comparison of the DDPG controller's performance for GCI DCR, evaluating both steady-state and dynamic characteristics has been carried out. This comparison included conventional and advanced control methodologies, such as PI and MPC controllers, to assess robustness against parameter variations. This research advances the understanding of advanced control methodologies in power electronics, provides a tangible and validated solution for enhancing the DCR of inverters in practical grid integration applications. Future work focuses on implementing the DRL on physical hardware setups.

REFERENCES

[1] O. Zandi and J. Poshtan, "Voltage control of a quasi Z-source converter under constant power load condition using reinforcement learning," *Control Eng. Pract.*, vol. 135, Jun. 2023, Art. no. 105499. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0967066123000680

[2] O. Zandi and J. Poshtan, "Voltage control of DC–DC converters through direct control of power switches using reinforcement learning," *Eng. Appl. Artif. Intell.*, vol. 120, Apr. 2023, Art. no. 105833. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0952197623000179

[3] P. I. N. Barbalho, V. A. Lacerda, R. A. S. Fernandes, and D. V. Coury, "Deep reinforcement learning-based secondary control for microgrids in islanded mode," *Electr. Power Syst. Res.*, vol. 212, Nov. 2022, Art. no. 108315. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779622004977

[4] C. Li, Y.-A. Chen, C. Jin, R. Sharma, and J. Kleissl, "Online PV smart inverter coordination using deep deterministic policy gradient," *Electr. Power Syst. Res.*, vol. 209, Aug. 2022, Art. no. 107988. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779622002164

[5] S. Saadatmand, P. Shamsi, and M. Ferdowsi, "Adaptive critic design-based reinforcement learning approach in controlling virtual inertia-based grid-connected inverters," *Int. J. Electr. Power Energy Syst.*, vol. 127, May 2021, Art. no. 106657. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061520342022

[6] X. Zhou, J. Wang, X. Wang, and S. Chen, "Deep reinforcement learning for microgrid operation optimization: A review," in *Proc. 8th Asia Conf. Power Electr. Eng. (ACPEE)*, Apr. 2023, pp. 2059–2065.

[7] J. Gan, S. Li, C. Wei, L. Deng, and X. Tang, "Intelligent learning algorithm and intelligent transportation-based energy management strategies for hybrid electric vehicles: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 10, pp. 10345–10361, May 2023.

[8] O. A. Alimi, K. Ouahada, and A. M. Abu-Mahfouz, "A review of machine learning approaches to power system security and stability," *IEEE Access*, vol. 8, pp. 113512–113531, 2020.

[9] Y. Gao and N. Yu, "Deep reinforcement learning in power distribution systems: Overview, challenges, and opportunities," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT)*, Feb. 2021, pp. 1–5.

[10] D. Hai, T. Zhu, S. Duan, W. Huang, and W. Li, "Deep reinforcement learning for volt/VAR control in distribution systems: A review," in *Proc. 5th Int. Conf. Energy, Electr. Power Eng. (CEEPE)*, Apr. 2022, pp. 596–601.

[11] T. Berghout, M. Benbouzid, X. Ma, S. Djurovic, and L.-H. Mouss, "Machine learning for photovoltaic systems condition monitoring: A review," in *Proc. 47th Annu. Conf. IEEE Ind. Electron. Soc.*, Oct. 2021, pp. 1–5.

[12] P. S. Tan, T. B. Tang, and E. T. W. Ho, "Explainable artificial intelligence applied to deep reinforcement learning controllers for photovoltaic maximum power point tracking," in *Proc. Int. Conf. Future Trends Smart Communities (ICFTSC)*, Dec. 2022, pp. 1–5.

[13] L. Avila, M. De Paula, I. Carlucho, and C. S. Reinoso, "MPPT for PV systems using deep reinforcement learning algorithms," *IEEE Latin Amer. Trans.*, vol. 17, no. 12, pp. 2020–2027, Dec. 2019.

[14] K. Wang, J. Ma, K. L. Man, K. Huang, and X. Huang, "Sim-to-real deep reinforcement learning for maximum power point tracking of photovoltaic systems," in *Proc. IEEE Int. Conf. Environ. Electr. Eng. IEEE Ind. Commercial Power Syst. Eur.*, Sep. 2021, pp. 1–4.

[15] S. Abedi, S. W. Yoon, and S. Kwon, "Battery energy storage control using a reinforcement learning approach with cyclic time-dependent Markov process," *Int. J. Electr. Power Energy Syst.*, vol. 134, Jan. 2022, Art. no. 107368. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061521006074

[16] R. Anugula and S. P. Krishna Karri, "Deep reinforcement learning based adaptive controller of DC electric drive for reduced torque and current ripples," in *Proc. IEEE Int. Conf. Technol., Res., Innov. Betterment Soc. (TRIBES)*, Dec. 2021, pp. 1–6.

[17] A. Traue, G. Book, W. Kirchgässner, and O. Wallscheid, "Toward a reinforcement learning environment toolbox for intelligent electric motor control," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 3, pp. 919–928, Mar. 2022.

[18] Y. Wang, S. Fang, J. Hu, and D. Huang, "A novel active disturbance rejection control of PMSM based on deep reinforcement learning for more electric aircraft," *IEEE Trans. Energy Convers.*, vol. 38, no. 2, pp. 1461–1470, Aug. 2023.

[19] Y. Wang, S. Fang, J. Hu, and D. Huang, "Multiscenarios parameter optimization method for active disturbance rejection control of PMSM based on deep reinforcement learning," *IEEE Trans. Ind. Electron.*, vol. 70, no. 11, pp. 10957–10968, Nov. 2023.

[20] S. Bhattacharjee, S. Halder, Y. Yan, A. Balamurali, L. V. Iyer, and N. C. Kar, "Real-time SIL validation of a novel PMSM control based on deep deterministic policy gradient scheme for electrified vehicles," *IEEE Trans. Power Electron.*, vol. 37, no. 8, pp. 9000–9011, Aug. 2022.

[21] Y. Wang, S. Fang, and J. Hu, "Active disturbance rejection control based on deep reinforcement learning of PMSM for more electric aircraft," *IEEE Trans. Power Electron.*, vol. 38, no. 1, pp. 406–416, Jan. 2023.

[22] Y. Wang, S. Fang, and D. Huang, "An improved model-free active disturbance rejection deadbeat predictive current control method of PMSM based on data-driven," *IEEE Trans. Power Electron.*, vol. 38, no. 8, pp. 9606–9616, Aug. 2023.

[23] C. Cui, N. Yan, B. Huangfu, T. Yang, and C. Zhang, "Voltage regulation of DC–DC buck converters feeding CPLs via deep reinforcement learning," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 3, pp. 1777–1781, Mar. 2022.

[24] M. Hajihosseini, M. Andalibi, M. Gheisarnejad, H. Farsizadeh, and M.-H. Khooban, "DC/DC power converter control-based deep machine learning techniques: Real-time implementation," *IEEE Trans. Power Electron.*, vol. 35, no. 10, pp. 9971–9977, Oct. 2020.

[25] S. Jiang, Y. Zeng, Y. Zhu, J. Pou, and G. Konstantinou, "Stability-oriented multiobjective control design for power converters assisted by deep reinforcement learning," *IEEE Trans. Power Electron.*, vol. 38, no. 10, pp. 12394–12400, Oct. 2023.

[26] M. Gheisarnejad, H. Farsizadeh, and M. H. Khooban, "A novel nonlinear deep reinforcement learning controller for DC–DC power buck converters," *IEEE Trans. Ind. Electron.*, vol. 68, no. 8, pp. 6849–6858, Aug. 2021.

[27] X. Meng, Y. Jia, Q. Xu, C. Ren, X. Han, and P. Wang, "A novel intelligent nonlinear controller for dual active bridge converter with constant power loads," *IEEE Trans. Ind. Electron.*, vol. 70, no. 3, pp. 2887–2896, Mar. 2023.

[28] A. Fathollahi, M. Gheisarnejad, B. Andresen, H. Farsizadeh, and M.-H. Khooban, "Robust artificial intelligence controller for stabilization of full-bridge converters feeding constant power loads," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 70, no. 9, pp. 3504–3508, Jun. 2023.

[29] C. Cui, T. Yang, Y. Dai, C. Zhang, and Q. Xu, "Implementation of transferring reinforcement learning for DC–DC buck converter control via duty ratio mapping," *IEEE Trans. Ind. Electron.*, vol. 70, no. 6, pp. 6141–6150, Jun. 2023.

[30] Y. Tang, W. Hu, D. Cao, N. Hou, Z. Li, Y. W. Li, Z. Chen, and F. Blaabjerg, "Deep reinforcement learning aided variable-frequency ontrol for dual-active-bridge converter," *IEEE Trans. Ind. Electron.*, vol. 70, no. 10, pp. 10506–10515, Oct. 2023.

[31] Y. Tang, W. Hu, B. Zhang, D. Cao, N. Hou, Y. Li, Z. Chen, and F. Blaabjerg, "Deep reinforcement learning-aided efficiency optimized dual active bridge converter for the distributed generation system," *IEEE Trans. Energy Convers.*, vol. 37, no. 2, pp. 1251–1262, Jun. 2022.

[32] O. E. Egbomwan, S. Liu, and H. Chaoui, "Twin delayed deep deterministic policy gradient (TD3) based virtual inertia control for inverter-interfacing DGs in microgrids," *IEEE Syst. J.*, vol. 17, no. 2, pp. 2122–2132, Sep. 2023.

[33] K. Wang, D. Hong, J. Ma, K. L. Man, K. Huang, and X. Huang, "Maximum power point tracking of photovoltaic systems using deep Q-networks," in *Proc. IEEE 18th Int. Conf. Ind. Informat. (INDIN)*, vol. 1, Jul. 2020, pp. 100–103.

[34] K.-Y. Chou, C.-S. Yang, and Y.-P. Chen, "Deep Q-network based global maximum power point tracking for partially shaded PV system," in *Proc. IEEE Int. Conf. Consum. Electron.-Taiwan (ICCE-Taiwan)*, Sep. 2020, pp. 1–2.

[35] M. Arianborna, J. Faiz, and A. Erfani-Nik, "MPPT control of a PMSG connected to the wind turbine based on deep Q-network," in *Proc. 10th Iranian Conf. Renew. Energy Distrib. Gener. (ICREDG)*, Mar. 2023, pp. 1–5.

[36] O. E. Egbomwan, H. Chaoui, and S. Liu, "Power delivery capability improvement of voltage source converters in weak power grid using deep reinforcement learning with continuous action," in *Proc. IEEE 31st Int. Symp. Ind. Electron. (ISIE)*, Jun. 2022, pp. 412–417.

[37] J. H. Woo, L. Wu, S. M. Lee, J.-B. Park, and J. H. Roh, "D-STATCOM d-q axis current reference control applying DDPG algorithm in the distribution system," *IEEE Access*, vol. 9, pp. 145840–145851, 2021.

[38] M. Mahazabeen, A. J. Abianeh, S. Ebrahimi, and F. Ferdowsi, "Performance evaluation of an RL-aided DAB-based EV charger control," in *Proc. North Amer. Power Symp. (NAPS)*, Oct. 2022, pp. 1–5.

[39] M. Nicola and C.-I. Nicola, "Improved control of DC–DC converter for PV array and battery hybrid system based on sine cosine algorithm and RL-TD3 agent," in *Proc. 3rd Int. Conf. Electr., Comput., Commun. Mechatronics Eng. (ICECCME)*, Jul. 2023, pp. 1–6.

[40] Z. Zhang, D. Zhang, and R. C. Qiu, "Deep reinforcement learning for power system applications: An overview," *CSEE J. Power Energy Syst.*, vol. 6, no. 1, pp. 213–225, Mar. 2020.

[41] D. Cao, W. Hu, J. Zhao, G. Zhang, B. Zhang, Z. Liu, Z. Chen, and F. Blaabjerg, "Reinforcement learning and its applications in modern power and energy systems: A review," *J. Mod. Power Syst. Clean Energy*, vol. 8, no. 6, pp. 1029–1042, Nov. 2020.

[42] P. Qashqai, M. Babaie, R. Zgheib, and K. Al-Haddad, "A model-free switching and control method for three-level neutral point clamped converter using deep reinforcement learning," *IEEE Access*, vol. 11, pp. 105394–105409, 2023.

[43] A. Bag, B. Subudhi, and P. K. Ray, "A combined reinforcement learning and sliding mode control scheme for grid integration of a PV system," *CSEE J. Power Energy Syst.*, vol. 5, no. 4, pp. 498–506, Dec. 2019.

[44] C. Xiang, X. Zhang, T. Qie, T. K. Chau, J. Ye, Y. Yu, H. H. C. Iu, and T. Fernando, "A novel deep deterministic policy gradient assisted learning based control algorithm for three-phase DC/AC inverter with an RL load," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 11, no. 6, pp. 5529–5539, Dec. 2023.

[45] X. Fu, S. Li, and I. Jaithwa, "Implement optimal vector control for LCL-filter-based grid-connected converters by using recurrent neural networks," *IEEE Trans. Ind. Electron.*, vol. 62, no. 7, pp. 4443–4454, Jul. 2015.

[46] N. Babu, "Adaptive grid-connected inverter control schemes for power quality enrichment in microgrid systems: Past, present, and future perspectives," *Electr. Power Syst. Res.*, vol. 230, May 2024, Art. no. 110288. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779624001767

[47] W. Wu, K. Hu, M. Zhang, and G. Han, "Fully discrete-time domain model and damping characteristics accurate analysis for grid-connected inverter with LCL filter," *Electr. Power Syst. Res.*, vol. 221, Aug. 2023, Art. no. 109388. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779623002778

[48] Y. Zhao, W. Hu, G. Zhang, Q. Huang, Z. Chen, and F. Blaabjerg, "Novel adaptive stability enhancement strategy for power systems based on deep reinforcement learning," *Int. J. Electr. Power Energy Syst.*, vol. 152, Oct. 2023, Art. no. 109215. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061523002727

[49] P. Mei, H. R. Karimi, H. Xie, F. Chen, C. Huang, and S. Yang, "A deep reinforcement learning approach to energy management control with connected information for hybrid electric vehicles," *Eng. Appl. Artif. Intell.*, vol. 123, Aug. 2023, Art. no. 106239. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0952197623004232

[50] Y. Zhu, J. Zhao, Z. Zeng, L. Mao, and K. Qu, "SISO impedance modeling and stability comparison of grid-connected inverter control system in different time domains," *Electr. Power Syst. Res.*, vol. 228, Mar. 2024, Art. no. 110069. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779623009562

[51] S. Jahan, S. P. Biswas, S. Haq, Md. R. Islam, M. A. P. Mahmud, and A. Z. Kouzani, "An advanced control scheme for voltage source inverter based grid-tied PV systems," *IEEE Trans. Appl. Supercond.*, vol. 31, no. 8, pp. 1–5, Nov. 2021.

[52] T. E. K. Zidane, A. S. Aziz, Y. Zahraoui, H. Kotb, K. M. AboRas, Kitmo, and Y. B. Jember, "Grid-connected solar PV power plants optimization: A review," *IEEE Access*, vol. 11, pp. 79588–79608, 2023.

[53] X. Wang, F. Blaabjerg, and P. C. Loh, "Grid-current-feedback active damping for LCL resonance in grid-connected voltage-source converters," *IEEE Trans. Power Electron.*, vol. 31, no. 1, pp. 213–223, Jan. 2016.

[54] S. He, D. Zhou, X. Wang, and F. Blaabjerg, "Passivity-based multisampled converter-side current control of LCL-filtered VSCs," *IEEE Trans. Power Electron.*, vol. 37, no. 11, pp. 13848–13860, Nov. 2022.

[55] V. N. Kumar, N. Babu P., R. Kiranmayi, P. Siano, and G. Panda, "Improved power quality in a solar PV plant integrated utility grid by employing a novel adaptive current regulator," *IEEE Syst. J.*, vol. 14, no. 3, pp. 4308–4319, Sep. 2020.

[56] A. Akhavan, J. C. Vasquez, and J. M. Guerrero, "A robust method for controlling grid-connected inverters in weak grids," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 68, no. 4, pp. 1333–1337, Apr. 2021.

[57] R. Zhao, Q. Li, H. Xu, Y. Wang, and J. M. Guerrero, "Harmonic current suppression strategy for grid-connected PWM converters with LCL filters," *IEEE Access*, vol. 7, pp. 16264–16273, 2019.

[58] P. Li, X. Tong, Y. He, and T. Chen, "Adaptive model predictive control for PUC grid-connected inverter system with inaccurate parameters," *Control Eng. Pract.*, vol. 139, May 2023, Art. no. 105652. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0967066123002216

[59] N. Panten, N. Hoffmann, and F. W. Fuchs, "Finite control set model predictive current control for grid-connected voltage-source converters with LCL filters: A study based on different state feedbacks," *IEEE Trans. Power Electron.*, vol. 31, no. 7, pp. 5189–5200, Jul. 2016.

[60] S. Mariethoz and M. Morari, "Explicit model-predictive control of a PWM inverter with an LCL filter," *IEEE Trans. Ind. Electron.*, vol. 56, no. 2, pp. 389–399, Feb. 2009.

[61] P. Falkowski and A. Sikorski, "Finite control set model predictive control for grid-connected AC–DC converters with LCL filter," *IEEE Trans. Ind. Electron.*, vol. 65, no. 4, pp. 2844–2852, Apr. 2018.

[62] F. Piotr, "Finite control set model predictive control for grid-connected NPC converter with LCL filter and novel resonance damping method," in *Proc. 19th Eur. Conf. Power Electron. Appl. (EPE ECCE Europe)*, Sep. 2017, pp. P.1–P.8.

[63] G.-X. Zhong, Z. Wang, J. Li, and Q. Su, "Robust fault diagnosis for closed-loop grid-connected inverter based on sliding mode observer and identifier," *Electr. Power Syst. Res.*, vol. 217, Apr. 2023, Art. no. 109097. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779622011464

[64] O. Lopez-Santos, G. Garcia, L. Martinez-Salamero, and L. Cortes-Torres, "Suppressing the effect of the DC-link voltage ripple on the current control of a sliding-mode controlled microinverter," in *Proc. CHILEAN Conf. Electr., Electron. Eng., Inf. Commun. Technol. (CHILECON)*, Oct. 2015, pp. 447–452.

[65] D. Foito, V. F. Pires, A. Cordeiro, and J. F. Silva, "Sliding mode vector control of grid-connected PV multilevel systems based on triple three-phase two-level inverters," in *Proc. 9th Int. Conf. Renew. Energy Res. Appl. (ICRERA)*, Sep. 2020, pp. 399–404.

[66] C. Jin and W. Zhenhua, "Pan Boolean PI control of grid-tied two-level inverter for wind power generation," in *Proc. IEEE Workshop Adv. Res. Technol. Ind. Appl. (WARTIA)*, Sep. 2014, pp. 427–430.

[67] P. Falkowski and A. Sikorski, "Comparative analysis of finite control set model predictive control methods for grid-connected AC–DC converters with LCL filter," in *Proc. IEEE 27th Int. Symp. Ind. Electron. (ISIE)*, Jun. 2018, pp. 193–200.

[68] T. Dragicevic, C. Zheng, J. Rodriguez, and F. Blaabjerg, "Robust quasi-predictive control of *LCL*-filtered grid converters," *IEEE Trans. Power Electron.*, vol. 35, no. 2, pp. 1934–1946, Feb. 2020.

[69] I. Poonahela, S. Bayhan, H. Abu-Rub, M. M. Begovic, and M. B. Shadmand, "An effective finite control set-model predictive control method for grid integrated solar PV," *IEEE Access*, vol. 9, pp. 144481–144492, 2021.

[70] J. Scoltock, T. Geyer, and U. K. Madawala, "A model predictive direct current control strategy with predictive references for MV grid-connected converters with *LCL*-filters," *IEEE Trans. Power Electron.*, vol. 30, no. 10, pp. 5926–5937, Oct. 2015.

[71] M. E. Zarei, D. Ramirez, M. Prodanovic, and G. Venkataramanan, "Multivector model predictive power control for grid connected converters in renewable power plants," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 10, no. 2, pp. 1466–1478, Apr. 2022.

[72] C. R. D. Osório, D. A. Schuetz, G. G. Koch, F. Carnielutti, D. M. Lima, V. F. Montagner, and H. Pinheiro, "Modulated model predictive control applied to LCL-filtered grid-tied inverters: A convex optimization approach," *IEEE Open J. Ind. Appl.*, vol. 2, pp. 366–377, 2021.

[73] N. N. Nam, N. D. Nguyen, C. Yoon, M. Choi, and Y. I. Lee, "Voltage sensorless model predictive control for a grid-connected inverter with LCL filter," *IEEE Trans. Ind. Electron.*, vol. 69, no. 1, pp. 740–751, Jan. 2022.

[74] M. Tebaa and M. Ouassaid, "Finite set model predictive control of three phase 2L-VSC for grid connected photovoltaic system," in *Proc. IEEE 21st Medit. Electrotech. Conf. (MELECON)*, Jun. 2022, pp. 1229–1234.

[75] B. Zhang, W. Wu, Y. Yang, N. Gao, J. Chen, E. G. Koutroulis, H. S. Chung, M. Liserre, and F. Blaabjerg, "A novel simplified finite control set repeat model predictive control for grid-connected inverters," *IEEE Trans. Ind. Electron.*, vol. 70, no. 11, pp. 11324–11333, Nov. 2023.

[76] G. V. Hollweg, R. V. Tambara, J. R. Massing, L. C. Borin, E. Mattos, G. G. Koch, C. R. D. Osório, and V. F. Montagner, "Model reference adaptive controllers with improved performance for applications in LCL-filtered grid-connected converters," *Control Eng. Pract.*, vol. 138, Sep. 2023, Art. no. 105591. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0967066123001600
[77] A. Bag, B. Subudhi, and P. K. Ray, "An adaptive sliding mode control scheme for grid integration of a PV system," *CPSS Trans. Power Electron. Appl.*, vol. 3, no. 4, pp. 362–371, Dec. 2018.
[78] C. Dang, X. Tong, and W. Song, "Sliding-mode control in dq-frame for a three-phase grid-connected inverter with LCL-filter," *J. Franklin Inst.*, vol. 357, no. 15, pp. 10159–10174, Oct. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0016003219309093
[79] S. Alepuz, S. Busquets-Monge, J. Bordonau, J. A. Martinez-Velasco, C. A. Silva, J. Pontt, and J. Rodriguez, "Control strategies based on symmetrical components for grid-connected converters under voltage dips," *IEEE Trans. Ind. Electron.*, vol. 56, no. 6, pp. 2162–2173, Jun. 2009.
[80] E. Figueres, G. Garcera, J. Sandia, F. Gonzalez-Espin, and J. C. Rubio, "Sensitivity study of the dynamics of three-phase photovoltaic inverters with an LCL grid filter," *IEEE Trans. Ind. Electron.*, vol. 56, no. 3, pp. 706–717, Mar. 2009.

ANUGULA RAJAMALLAIAH (Graduate Student Member, IEEE) received the bachelor's degree in electrical and electronics engineering from JNTU Hyderabad and the master's degree in instrumentation and control from NIT Calicut. He is currently pursuing the Ph.D. degree with NIT Andhra Pradesh, specializing in deep reinforcement learning implementation for power converter control applications.

He is a dedicated Ph.D. Scholar with a passion for machine learning, power electronics, and control systems. He resides in Tadepalligudem, continuing his research in machine learning implementation in dc and ac microgrids, poised for further contributions to artificial intelligence in power electronics applications. He remain dedicated to intellectual exploration and the pursuit of academic excellence.

SRI PHANI KRISHNA KARRI (Member, IEEE) received the M.Tech. degree in signal processing and the Ph.D. degree in multidimensional signal processing and machine learning from IIT Kharagpur, in 2011 and 2017, respectively. Following his Ph.D., he expanded his expertise during a Postdoctoral Fellowship with IIT Kharagpur, from 2017 to 2018. He is currently an Assistant Professor with the Electrical Engineering Department, NIT Andhra Pradesh. During this phase, he delved into developing deep reinforcement learning (RL) based image analysis frameworks, pushing the boundaries of image analysis by integrating advanced deep RL techniques. His research interests include signal processing, machine learning, and image analysis.

MAMDOUH L. ALGHAYTHI (Member, IEEE) received the B.S. degree in electrical engineering from Jouf University, Sakakah, Saudi Arabia, in 2012, the M.S. degree in electrical engineering from Southern Illinois University, Carbondale, IL, USA, in 2015, and the Ph.D. degree in electrical engineering from the University of Missouri, Columbia, MO, USA, in 2020. He is currently an Assistant Professor with the Department of Electrical Engineering, Jouf University. He is the Chair of the Electrical Engineering Department, Jouf University, and the Coordinator of the Master's Program in Renewable Energy. He has authored and coauthored many journal articles and conference papers. His current research interests include design and modeling of power electronics converters, renewable energy systems, reliability aspects of dc-dc converters, smart grids, energy management, and microgrids. He is a member of the IEEE Power Electronics Society.

MESHARI S. ALSHAMMARI (Member, IEEE) received the B.Sc. degree in electrical engineering from the University of Hail, Saudi Arabia, in 2014, the M.E. degree in energy systems engineering and the Ph.D. degree in electrical and electronics engineering from the University of Galway, in 2017 and 2022, respectively. He is currently an Assistant Professor with the Electrical Engineering Department, Jouf University, Sakaka, Saudi Arabia. His research interests include the analysis of ac and dc distribution systems, dc/ac power converters for utility grid and renewable energy applications, and renewable energy integration systems. In addition, he was awarded the Best Poster at the 7th NUI Galway Research Day, in 2017, and the Best Paper at ICRERA 2021.