IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 5 January 2024, accepted 20 January 2024, date of publication 25 January 2024, date of current version 6 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3358412

RESEARCH ARTICLE

Deep Reinforcement Learning Based Control Strategy for Voltage Regulation of DC-DC Buck Converter Feeding CPLs in DC Microgrid

ANUGULA RAJAMALLAIAH[1], (Graduate Student Member, IEEE), SRI PHANI KRISHNA KARRI[1], (Member, IEEE), AND YANNAM RAVI SHANKAR[2]

[1]National Institute of Technology Andhra Pradesh, Tadepalligudem, Andhra Pradesh 534101, India
[2]School of Electrical and Computer Engineering, Dire Dawa University, Dire Dawa 1362, Ethiopia

Corresponding author: Yannam Ravi Shankar (yannam.ravi@ddu.edu.et)

ABSTRACT A DC microgrid's tightly regulated DC/DC converter encounters significant challenges in voltage stability, primarily due to the negative incremental resistance of constant power loads (CPLs). Conventional controllers often struggle with load variations and changes in system parameters. Therefore, there has been growing interest in adaptive machine learning algorithms, such as Deep Reinforcement Learning (DRL), to improve voltage regulation. This paper presents an end-to-end DRL framework based on a modified Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm. The framework is designed to directly control power switches for regulating the voltage of a DC/DC buck converter that supplies power to CPLs. Real-time experiments were conducted using OPAL-RT to validate the approach under diverse load cycles and converter parameter changes.Comparative analysis against other DRL-based control strategies, including Deep Q-learning (DQN) and Deep Deterministic Policy Gradient (DDPG), demonstrated the superior static and dynamic voltage response of the proposed modified TD3 DRL controller, particularly in scenarios involving load and parametric variations.

INDEX TERMS Voltage stability, micro grid, constant power loads, dc-dc buck converter, deep Q-learning, deep deterministic policy gradient, twin delayed deep deterministic policy gradient.

I. INTRODUCTION

Dc Microgrids have gained significant attention in the past decade, driven by factors such as the integration of renewable energy sources and the rise in DC loads. DC microgrids offer numerous advantages, including enhanced efficiency, improved power quality, and increased flexibility. By leveraging the benefits of DC microgrids, developing countries can effectively address energy crises, enhance energy access, promote sustainability, and stimulate economic growth [1]. A typical DC microgrid, as depicted in Figure 1, represents a standard configuration for a DC microgrid involves the interconnection of distributed renewable energy sources and various electrical loads. These energy sources encompass renewable systems like solar panels or wind turbines, as well as conventional sources such as batteries and diesel generators. The loads in DC microgrids span residential, commercial, and industrial facilities, encompassing lighting, machinery, and other electrical devices. However, DC microgrids supplying CPLs face voltage-related challenges [2]. When CPLs are connected to a DC microgrid, they exhibit negative impedance characteristics, leading to a voltage droop phenomenon [3]. This phenomenon introduces rapid and significant voltage fluctuations, causing issues with voltage instability [4]. To tackle these challenges, various passive and active damping approaches have been proposed [5].

Passive damping techniques offer inherent stability without the need for active control or external energy sources.

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

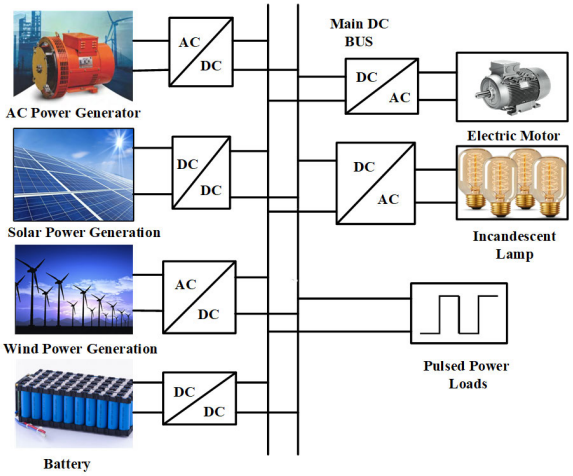

FIGURE 1. A standard configuration for a DC microgrid involves the interconnection of distributed renewable energy sources and various electrical loads.

However, these techniques may have limitations when dealing with rapid and significant load changes or variations in CPL characteristics [6], [7]. Furthermore, careful design and tuning are essential to ensure the effectiveness of passive damping techniques in addressing voltage challenges unique to DC microgrids supplying CPLs [8].

Active damping approaches encompass small signal-based and eigenvalue methods, commonly employed in the analysis and design of control systems for voltage regulation in DC-DC Buck converters supplying CPLs [9], [10], [11]. Although these methods offer valuable insights and initial design guidelines, it is crucial to validate their results through simulations and, if possible, experimental testing to ensure accuracy and applicability in practical systems. While these methods serve as valuable tools for analyzing and designing control systems, they do have certain limitations and potential drawbacks, such as the linearity assumption and design iterations [12], [13], [14]. It is important to be mindful of these constraints and drawbacks in order to make informed decisions during the analysis and design phases.

On the other hand, large signal-based methods, also known as time-domain or nonlinear analysis methods, offer alternative approaches to analyzing and designing control systems, addressing certain limitations of small signal-based and eigenvalue methods [15], [16]. These methods specifically focus on the dynamic behavior of the system in response to significant variations and nonlinearities.

Similar to small signal-based methods, large signal-based methods may involve iterative design processes to optimize control parameters and achieve the desired system performance. Fine-tuning the control strategy often requires multiple simulations and iterations. Large signal-based analysis heavily relies on simulation tools and models, necessitating accurate modeling of nonlinear components and control strategies using appropriate techniques and detailed system models. The accuracy of the analysis results depends on the reliability of the models used, leading to the emergence of model-free techniques [17]. It is crucial to employ precise modeling techniques to enhance the reliability of large signal-based analysis results.

Feedback control techniques utilize information about the system output and compare it to a desired setpoint to generate a control signal that adjusts the system inputs. For example, PID control continuously monitors the error between the setpoint and the actual output, adjusting the control signal based on proportional, integral, and derivative terms to regulate system behavior. Although PID control requires careful tuning to achieve optimal performance, it may not suffice for addressing complex control requirements or compensating for significant parameter variations.

Prediction methods play a crucial role in the voltage regulation of DC-DC Buck converters supplying constant power loads (CPLs), providing insights into load transients, system dynamics, control horizons, constraint handling, and disturbance rejection. These methods empower the control system to anticipate future behavior, optimize control actions, and maintain stable and accurate voltage regulation in the presence of varying load conditions and system dynamics.

Model Predictive Control (MPC) finds common use in various applications, particularly in the voltage regulation of DC-DC Buck converters supplying CPLs. MPC offers several advantages, including predictive capabilities, constraint handling, adaptability, and multivariable control [18]. Among different variants, Finite Control Set MPC stands out due to its computational efficiency and implementation simplicity. Nonlinear MPC enables more precise voltage regulation and control over a broader operating range. Distributed MPC is applicable when the control system involves multiple interconnected DC-DC Buck converters or when CPLs are distributed across different locations. Stochastic MPC enhances control robustness by considering the probabilistic nature of disturbances and incorporating risk measures into the optimization problem [19].

However, the successful implementation of MPC requires careful consideration of computational requirements, modeling accuracy, and tuning complexity. Adequate implementation, modeling, and parameter tuning are crucial elements for achieving effective and reliable voltage regulation in DC-DC Buck converters supplying CPLs using MPC.

Unlike linear control techniques, such as PID control, sliding mode control (SMC) is specifically designed to handle nonlinearities in system dynamics and uncertainties in system parameters [20], [21]. SMC serves as a potent control technique for voltage regulation in DC-DC Buck converters supplying CPLs, offering robustness against uncertainties, disturbances, and nonlinearities, while ensuring a fast and accurate system response [22]. Various variations of sliding mode control, including higher-order sliding mode control, adaptive sliding mode control, and fuzzy sliding mode control, can be employed to address challenges associated with CPLs and nonlinear load characteristics. For example, adaptive sliding mode control is effective in handling parameter variations and uncertainties common in

DC-DC converters operating with CPLs. Higher-order sliding mode control provides improved accuracy and robustness, whereas fuzzy sliding mode control offers a flexible and adaptive control approach in the presence of imprecise or uncertain information. Integral sliding mode control can be employed to eliminate steady-state errors and enhance voltage regulation accuracy, which is particularly crucial in systems supplying CPLs with varying power demands [23]. However, it is crucial to carefully address the chattering phenomenon and design complexity associated with sliding mode control to ensure effective implementation and satisfactory control performance [24]. Proper attention to these aspects is essential for achieving successful and reliable control outcomes with sliding mode control in DC-DC Buck converters supplying CPLs.

Fuzzy Logic Control (FLC) finds wide application in various control systems, including the voltage regulation of DC-DC Buck converters supplying CPLs. FLC proves to be a versatile control technique capable of handling imprecise information, adapting to changing conditions, and providing interpretable control strategies. Adaptive FLC is particularly useful in scenarios where CPL characteristics or operating conditions vary over time [25]. Fuzzy Sliding Mode Control combines the concepts of FLC and sliding mode control, wherein sliding mode control ensures robustness against parameter variations and disturbances, while FLC provides linguistic rules and fuzzy logic-based decision-making. By combining these approaches, Fuzzy Sliding Mode Control achieves enhanced robustness and improved tracking performance for voltage regulation in the presence of CPLs. Fuzzy Gain Scheduling involves utilizing FLC to adjust the control gains or parameters of the DC-DC Buck converter based on the operating conditions. A hierarchical fuzzy control structure can be implemented in DC-DC Buck converters supplying CPLs, where the higher-level controller sets the desired output voltage, and the lower-level fuzzy controller adjusts the converter's duty cycle to regulate the output voltage [26]. However, the successful implementation of FLC requires the appropriate selection of linguistic variables, membership functions, rule bases, and defuzzification methods to ensure accurate and effective control performance [27]. Careful consideration of these elements is crucial for achieving successful and reliable control outcomes with FLC in DC-DC Buck converters supplying CPLs.

Artificial Neural Network (ANN) techniques, specifically within the neural networks or deep learning subfield, are integral components of the machine learning and AI toolbox. ANNs are trained using data, enabling them to learn patterns, relationships, and mappings between input and output variables. In the context of voltage regulation for DC-DC Buck converters supplying constant power loads (CPLs), ANN techniques find applications in modeling, control system design, adaptive control, and fault detection. ANNs have the potential to enhance voltage regulation performance by leveraging their ability to learn from data and make predictions or control decisions based on learned patterns and relationships. However, the successful application of ANN techniques necessitates careful data collection, pre-processing, network architecture design, and training [28]. Adequate training data, representing various operating conditions and possible scenarios, is crucial for achieving accurate and reliable results. Additionally, the complexity and interpretability of ANNs should be considered, as they may require substantial computational resources and may not offer intuitive insight into the control decision-making process [29]. It is essential to balance the potential advantages of ANN techniques with the practical considerations of data complexity and computational resources during their application in voltage regulation systems.

DRL is a machine learning technique employed to develop control strategies by interacting with an environment and receiving feedback in the form of rewards or punishments. In the context of voltage regulation for DC-DC Buck converters supplying CPLs, DRL can be applied to develop control strategies that optimize voltage regulation performance. The DRL agent interacts with the converter system, taking actions such as adjusting control parameters or duty cycle to regulate the output voltage. It receives rewards or penalties based on how effectively it achieves the desired voltage regulation objective [30]. DRL serves as a powerful tool for enhancing control strategies in voltage regulation systems, allowing for adaptive and optimized performance based on real-time interactions with the converter system. The DRL agent interacts with the converter system, taking actions such as adjusting control parameters or duty cycle to regulate the output voltage. It receives rewards or penalties based on how effectively it achieves the desired voltage regulation objective [30]. DRL serves as a powerful tool for enhancing control strategies in voltage regulation systems, allowing for adaptive and optimized performance based on real-time interactions with the converter system.

The literature explores several methods regarding the utilization of the DRL approach in DC/DC converters. A recent study introduced a self-adaptive system [31] employing a proximal policy optimization algorithm (PPO) for parameter tuning, aiming to mitigate the destructive effects of Constant Power Loads (CPLs). In contrast, another study [32] adapts a DDPG algorithm to generate a duty ratio compensation signal capable of controlling the output voltage of a DC/DC buck-boost converter feeding CPLs. However, these DRL management strategies are partially dependent on the mathematical model, the accuracy of which is influenced by efficiency.

A novel approach involves the design of an auxiliary DDPG controller fused with an intelligent Proportional-Integral (PI) controller and sliding mode observer [32]. This design aims to reduce observer estimation errors and further enhance the dynamic characteristics of the DC/DC buck converter feeding CPLs. While the DDPG algorithm has found widespread implementation in robotics

and autonomous driving, particularly for continuous control problems, its remarkable performance is accompanied by challenges. The overestimation of Q-values in the critic network and the dependence on hyperparameter tuning can lead to unstable operation. Addressing these issues is crucial for ensuring the stability and reliability of the implemented DDPG-based controller.

To enhance the bus voltage regulation performance of DC/DC buck converters feeding CPLs, an implementation of a model-free DRL-based DQN algorithm has been introduced [33]. This algorithm is primarily designed to address stabilization issues and ensure the required system performance. Additionally, it significantly reduces settling time in the event of a disturbance. The method employs a discrete action space in the continuous-time domain to match the switching speed of the switch element. However, the practical implementation of digital controllers demands a discrete-time model [34], [35]. Consideration of these aspects is essential for the successful and effective implementation of digital controllers in real-world scenarios.

This paper introduces a modified Twin Delayed Deep Deterministic Policy Gradient (TD3) control algorithm designed for DC microgrids supplying CPLs. The goal is to address instability issues arising from negative impedance characteristics. The TD3 algorithm not only mitigates instability but also enhances the dynamic characteristics of the system. In contrast to the DQN algorithm, which is suited for discrete control tasks, the TD3 algorithm is effective for continuous control tasks. Moreover, TD3 differs from the DDPG algorithm in three key ways: i) it incorporates double Q-learning without clipping the action signal, ii) it introduces delays in the updates of the actor for stable training, and iii) it provides better action noise regularization. These distinctions contribute to the improved performance and stability of the TD3 algorithm in regulating DC microgrids supplying CPLs.

The main contributions of this work are outlined below:

- Introduction of an end-to-end continuous action space-based TD3 DRL algorithm modified for practical deployment. This algorithm aims to enhance voltage regulation in DC microgrids with constant power loads.
- Conducted real-time validations of the proposed TD3 controller using OPAL-RT in the context of microgrid control applications. This step ensures the practical applicability and effectiveness of the algorithm.
- Conducted static, dynamic, and robust performance analyses of the modified TD3 DRL controller. The analysis focused on voltage response in the presence of load power variations and parameter changes. Comparative assessments were made against other state-of-the-art DRL-based control algorithms, including DQN and DDPG.

The remainder of the paper is organized as follows: the mathematical modeling of the system and control statement are discussed in Section-II. The details of the DRL elements and the workflow of the controller methodology are described in Section-III. The real-time validation, simulation results and comparative analysis of different DRL control algorithms for voltage regulation are presented in Section-IV. Finally, Section-V summarises of the work.

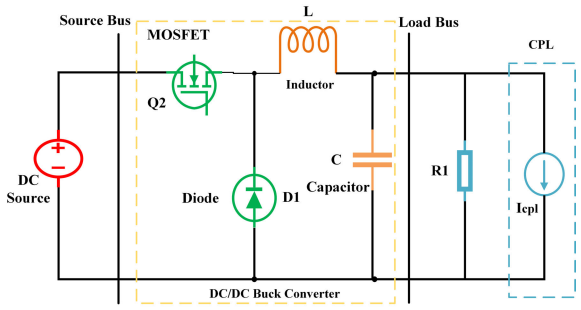

FIGURE 2. In a typical DC microgrid configuration, there is a DC/DC buck converter integrated to supply power to CPLs.

II. MATHEMATICAL DISSEMINATION OF DC/DC BUCK CONVERTER FEEDING CPL

The DC microgrid serves as a common interface for various distributed energy resources (DERs), energy storage elements, and diverse electrical loads. Within DC microgrids, DC-DC converters play a crucial role, facilitating voltage conversion and regulation to align with the requirements of different sources and loads. Figure 2 illustrates a typical onboard microgrid configuration, where a DC/DC buck converter feeds CPLs. In this representation, the lumped CPL is depicted as a controlled current source, while the lumped resistive load is denoted as R. DC microgrids are capable of accommodating a range of loads that operate on DC power. CPLs, a specific load type commonly encountered in DC microgrids, find applications in various fields. Examples include motor drives, data centers, telecommunications equipment, industrial processes, and medical devices. CPLs maintain a constant power demand irrespective of changes in voltage or current. They are designed to consume a fixed amount of power, ensuring stability in power consumption even when input voltage or current fluctuates.

The V-I characteristics of the constant power loads is given by equation (1)

$$i_{\text{cpl}} = \frac{p_{\text{cpl}}}{v_{\text{cpl}}} \tag{1}$$

where, p_{cpl} is CPL load power, v_{cpl} is CPL voltage, i_{cpl} is CPL current.

The state-space representation of the dynamic model based on the average switching scheme is given by equation (2)-(4) [32]

$$\frac{\text{di}_L}{\text{dt}} = \frac{1}{L}\left[\, v_{\text{in}}.d - v_c\right] \tag{2}$$

$$\frac{\text{dv}_c}{\text{dt}} = \frac{1}{C}\left[\, i_L - \frac{v_c}{R_1} - \frac{P}{v_c}\right] \tag{3}$$

$$y = v_{\text{out}} \tag{4}$$

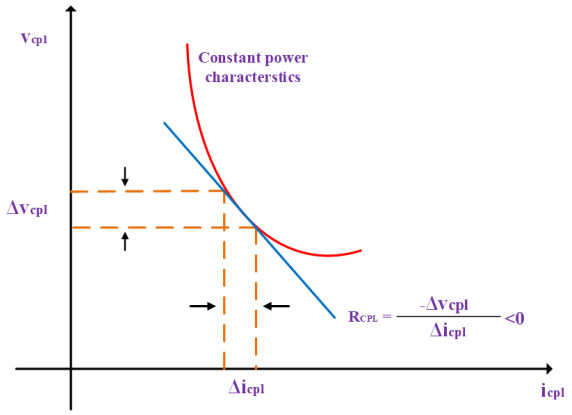

FIGURE 3. The V-I characteristics of CPLs demonstrate negative incremental impedance behavior, causing voltage fluctuations in DC microgrids.

Here, v_c – capacitor voltage, i_L – inductor current, L – inductance, C – capacitance, P – constant power, d – duty ratio, v_{in}– input voltage, v_{out} – output voltage.

DC-DC Buck converters are widely utilized in power electronics to step down the voltage from a higher input voltage to a lower output voltage. These converters play a crucial role in providing stable and regulated voltage output to power loads, including constant power loads. Voltage regulation in a buck converter entails adjusting the duty cycle of the switching mosfet to maintain the desired output voltage. CPLs typically adapt their impedance to sustain a constant power draw. However, abrupt changes or fluctuations in the load can lead to voltage fluctuations. If the load demands more power, the converter may struggle to meet the increased power requirement, resulting in a temporary drop in voltage. Conversely, if the load abruptly reduces its power demand, the converter may overshoot the voltage output. In cases of sudden voltage drops or rises, constant power loads exhibit a negative incremental impedance characteristic, as illustrated in Figure. 3. This characteristic introduces instability issues into the microgrid supplying CPLs [9].

The primary objective is to design a controller for the tightly regulated output voltage of the converter, ensuring system stability even under significant variations in load power and system parameters. The employed control mechanisms can impact crucial factors such as response time, accuracy, and the load's capability to handle transient conditions. Consequently, the entire system is characterized by nonlinearity and switching dynamics, with its components, especially the converter load, exhibiting notable uncertainties. Therefore, the application of modern and intelligent controllers becomes essential to effectively manage and regulate the system.

III. DRL CONTROLLER METHODOLOGY

DRL comprises an environment and an agent. The interaction between the environment (a DC/DC converter feeding CPLs) and the agent (TD3) is illustrated in Figure. 4. The agent makes decisions (actions) by retrieving observations from

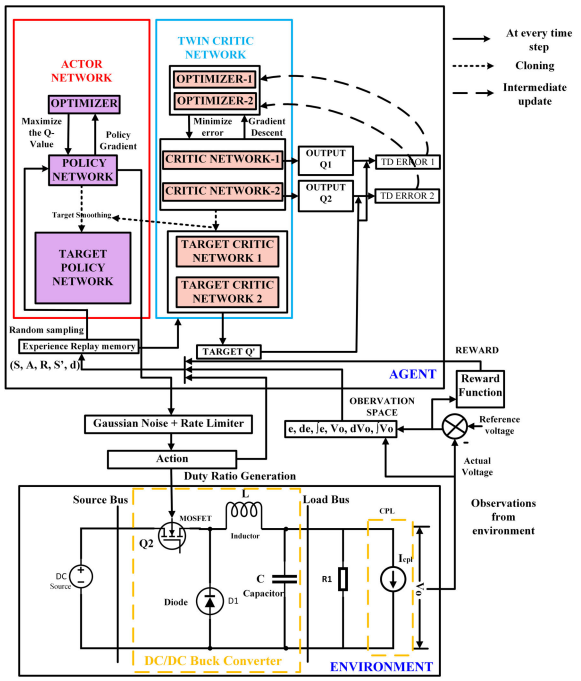

FIGURE 4. The DRL methodology for agent training illustrates the dynamic interaction between the agent and the environment as the agent learns to optimize microgrid performance.

the environment and receives a reward for the action taken. The agent's objective is to maximize the cumulative reward by providing the most suitable action signals to satisfy the control objective. The agent's decisions (control actions) progressively improve through a rigorous learning process, as explained in Section III-E. The various elements of the DRL methodology are discussed as follows:

A. OBSERVATION SPACE

The measurable parameters and state variables within the environment can serve as observations for training the agent. In this context, observations for the agent include the voltage error signal (represented as the difference between desired and actual output voltage, denoted as $e(t)$), the derivative of the error signal, the integration signals of the error signal, the output voltage (v_{out}), the derivative of the output voltage, and the integral signals of the output voltage. The observation set is defined by Equation (5).

$$O_t = \left\{ \mathrm{e}\,(t)\,, \frac{de(t)}{\mathrm{dt}}, \int e\,(t)\,, v_{\text{out}}\,(t)\,, \frac{dv_{\text{out}}(t)}{\mathrm{dt}}, \int v_{\text{out}}(t) \right\} \tag{5}$$

B. ACTION SPACE

The controller regulates the duty ratio of gate pulses to the MOSFET switch. A continuous control signal is preferred, as it offers a suitable range of duty ratios for various operating conditions, thereby enhancing dynamic performance. In contrast to conventional end-to-end DRL algorithms used for controlling DC/DC buck converters, which have a discrete action space with a limited number

of action sets to reduce training complexity, the proposed method is an end-to-end DRL-based controller. This method operates without a conventional controller and necessitates a continuous action space, accommodating different values of duty ratio.

C. AGENT

The TD3 algorithm represents an advanced reinforcement learning technique utilized for training agents in continuous action spaces. Serving as an extension of the original DDPG algorithm, TD3 incorporates several improvements aimed at enhancing stability and learning efficiency. Specifically designed for scenarios where an agent must learn to execute continuous actions in an environment to maximize a reward signal, TD3 has demonstrated successful applications across various tasks. These tasks include robotic control, autonomous driving, and simulated physics environments. The structure of the TD3 algorithm can be segmented into several components, as depicted in Figure. 4, each fulfilling a specific role in the learning process.

- Actor Network: The actor network is responsible for mapping states to actions. It takes the current state as input and generates a continuous action based on a learned policy. The goal of the actor network is to maximize the expected cumulative reward over time. The actor network parameters are represented with θ.
- Critic Networks: TD3 utilizes two critic networks, often referred to as twin critics. The critic networks estimate the value or quality of a given state-action pair. They take both the current state and the action as inputs and output a value estimate. By having two separate critics, TD3 reduces the overestimation bias that can occur in value-based methods. The twin critic networks parameters are represented with ϕ_1, ϕ_2.
- Target Networks: TD3 employs target networks for both the critics and the actor. These target networks are periodically updated to slowly track the learned values. The critic target networks are used to estimate the target Q-values for the actor's policy updates, while the actor target network is used to estimate the gradient of the value function with respect to the actions. The target actor, twin critic network parameters are represented with θ_{target}, $\phi_{target1}$, $\phi_{target2}$ respectively.
- Replay Buffer: The replay buffer is a memory structure that stores past experiences of state-action-reward-next state transitions. It allows the algorithm to break the sequential correlation in the data, mitigating issues associated with temporally correlated samples. During training, samples are randomly sampled from the replay buffer to create training batches.
- Exploration: To address the exploration-exploitation trade-off, TD3 typically employs noise perturbations during action selection. This noise can be added directly to the actions or incorporated into the policy function to encourage exploration of different actions in the environment.
- Policy and Value Updates: TD3 utilizes a variant of the deterministic policy gradient algorithm to update the actor network. The updates are performed less frequently than the critic updates, as a delayed policy update helps stabilize learning and prevent policy oscillations.

TABLE 1. DQN, DDPG and TD3 neural network configurations.

Category	Parameter	DQN	DDPG	TD3
Actor	Layer wise neurons	$L_1 = 16, L_2 = 8$	$L_1 = 16, L_2 = 8$	$L_1 = 16, L_2 = 8$
	Activation function	ReLU	ReLU	ReLU
	Learning rate	0.001 s	0.001 s	0.001 s
Critic	Layer wise neurons	-	$L_1 = 16, L_2 = 8$	$L_1 = 16, L_2 = 8$
	Activation function	Tanh	Tanh	Tanh
	Learning rate	0.0001 s	0.0001 s	0.0001 s
General	Discount factor	0.9	0.95	0.95
	Mini batch size	512	512	512
	Experience buffer length	$2X10^6$	$2X10^6$	$2X10^6$
	Action space	Discrete	Continuous	Continuous

The overall training structure of TD3 involves an iterative process that includes collecting experiences, updating the critic networks, updating the actor network, and periodically updating the target networks. This process continues until the agent learns an optimal policy that maximizes the expected cumulative reward in the continuous action space. The parameters of the actor and critic network for the DQN, DDPG, and TD3 methods are presented in Table 1.

D. REWARD MODEL

The reward signal is designed based on the error signal because the control objective depends on the line tracking performance. A discrete reward signal, suitable for the given task, is provided by Equation (6)

$$r(t) = \begin{cases} L_1 - |e(t)|, & e(t) < 0.1 \\ L_2 - |e(t)|, & 0.1 < e(t) < 1 \\ L_3\ |e(t)|, & 1 < e(t) \end{cases} \tag{6}$$

where, L_1, L_2, and L_3, are constant. In this work L_1, L_2, and L_3 have been taken by trial and error method and chosen as 5,1,-5 respectively, guided by insights derived from relevant literature [36].

E. LEARNING PROCESS OF AGENT

The TD3 algorithm represents a recent breakthrough in artificial intelligence, particularly excelling in solving problems characterized by a continuous action space within the environment. The flow chart depicting the operation of the TD3 algorithm, involving various sequential steps, is presented in Figure 5.

The agent is initially initialized with random policy parameters and Q-function parameters, and a replay memory with a capacity of 10^6 transitions is established. Target network parameters are set to the same values as the main network parameters. The agent then observes the current

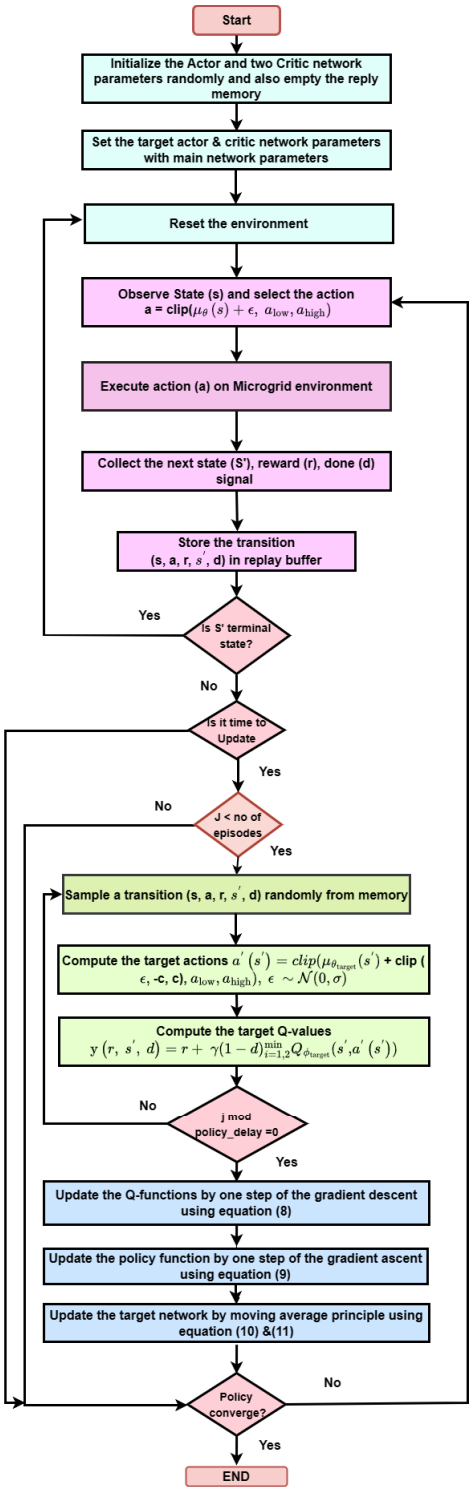

FIGURE 5. The flowchart of DRL training methodology involves sequential steps of initialisation of actor & critic network, action selection, policy optimization and actor & critic network updates for agent learning for performance improvement.

environment state (s) and selects an action (a) within predefined limits (a_{low}, a_{high}) by following the policy μ and introducing a normal distribution noise. This action is executed in the environment, yielding the next state (s') and a corresponding reward (r).

The observation, action, reward, next state, and done status (d) are stored as a tuple (transition) in the experience replay memory. Subsequently, the networks undergo a learning process. The agent samples a batch of transitions randomly from the memory and computes the target action (a'), incorporating truncated ($(-c,\ c\ ),\ c\ >\ 0$) normally distributed noise ($N(0,\sigma)$, where c is a constant) for each action to prevent overfitting.

For the critic network, target Q-values are calculated using Bellman's equation, as expressed in Equation (7). The mean squared error (MSE) is then computed based on the target and actual Q-values. As the two Q-networks are initialized with random values, consideration is given to Q_1 and Q_2. The critic network parameters are updated through gradient descent by back-propagating the MSE.

$$y(r, s', d) = r + \gamma(1-d)\sum_{i=1}^{2^{\min}} Q_{\phi_{\text{target},i}}(s', a'(s')) \tag{7}$$

$$\nabla_{\phi_i}\frac{1}{|B|}\sum_{(s,a,r,s',d)\in B}(Q_{\phi_i}(s,a) - y(r,s',d))^2 for \quad i=1,2 \tag{8}$$

Given the absence of an explicit mathematical function for the calculation of the expected return for the actor-network, the expectation of Q-values from the critic network is regarded as the expected return. Q-function values are updated using Equation (8), and actor-network parameters are updated using the gradient ascent method by maximizing the expected return, as expressed in Equation (9).

$$\nabla_\theta \frac{1}{|B|}\sum_{s\in B}((s,a) - y(r,s',d))^2 Q_{\phi_i}(s, \mu_\theta(s)) \tag{9}$$

Finally, target network parameters are updated through the moving average principle, with the critic network being updated less frequently than the actor-network. This process continues until the maximum number of episodes (J), after which the network converges to the desired values, and the actor-network is saved as the optimal policy. Equations (10) and (11) illustrate the updating of target networks using the smoothing factor ρ in the moving average principle.

$$\phi_{target,i} \leftarrow \rho\phi_{target,i} + (1-\rho)\,\phi_i, \quad i \in \{1,\ 2\} \tag{10}$$

$$\theta_{\text{target}} \leftarrow \rho\theta_{\text{target}} + \ (1-\rho)\,\theta \tag{11}$$

IV. RESULTS AND DISCUSSION

In this section, three main cases are presented. In the first case (Case-I), the results of the output voltage regulation of the DC/DC buck converter feeding CPLs using the DRL-based TD3 method are showcased under both training and testing conditions. In the second case (Case-II), the proposed method, along with other DRL-based methods (DQN and DDPG)) found in the literature, is validated in real-time using the OPAL-RT experimental setup. The agent is trained with DRL-based modified DQN and modified DDPG methods, exhibiting suitable dynamic performance as detailed in [32]

FIGURE 6. Real-time hardware implementation setup.

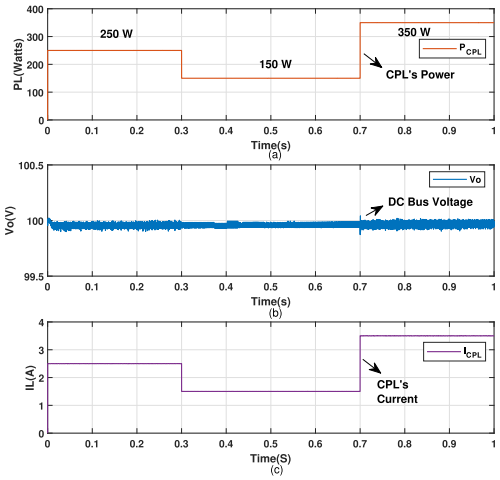

FIGURE 7. Assessing the response with the TD3 method under training conditions for (a) load power variations, (b) output voltage, and (c) inductor current.

and [36]. Moving to the next case (Case-III), the dynamic response of the system is compared among three DRL methods to determine superiority.

The DC/DC Buck converter parameters, as specified in [36], are considered in the experimentation for reproducible and effective comparison. The parameters of the DC/DC buck converter feeding CPLs are as follows: Input voltage (V_{Input}) = 200 V, Output voltage (V_{Output}) = 100 V, Inductance (L) = 2 mH, Capacitance (C) = 150 μF, and Switching frequency (f) = 10 kHz. The simulations are conducted in the MATLAB 2021b environment on a Dell Precision 5820 workstation with 32GB RAM and a 16GB NVIDIA RTX A4000 GPU. Real-Time validation is performed on the OPAL-RT OP4512 target simulator. The complete Real-Time implemented hardware setup is illustrated in Figure. 6.

Case-I: The agent is trained with TD3 algorithm on DC/DC buck converter for load power variations given by equation (12).

$$P(t) = \begin{cases} 250\ W & 0 \le t \le 0.3 \\ 150\ W & 0.3 \le t \le 0.7 \\ 350\ W & 0.7 \le t \le 1 \end{cases} \tag{12}$$

After successful training, the agent is tested with other set of load power variations (not covered in the training) given

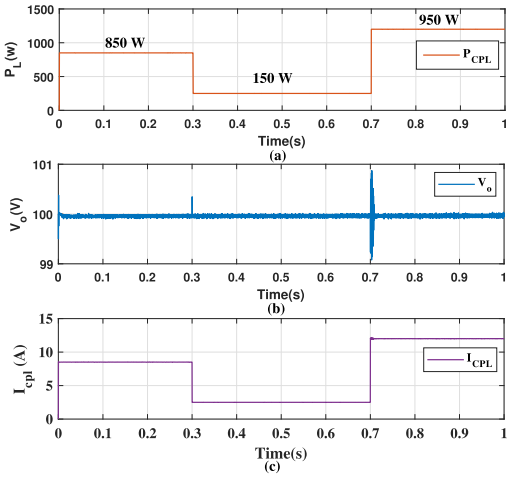

FIGURE 8. Evaluating the response with the TD3 method under testing conditions for (a) load power variations, (b) output voltage, and (c) inductor current.

by equation (13).

$$P(t) = \begin{cases} 850\ W & 0 \le t \le 0.3 \\ 150\ W & 0.3 \le t \le 0.7 \\ 950\ W & 0.7 \le t \le 1 \end{cases} \tag{13}$$

The waveforms depicting CPL power variations, output voltage, and the CPL current of the DC/DC buck converter feeding CPLs are presented in Figure. 7(a)-Figure. 7(c) under training conditions and also in Figure. 8(a)-Figure. 8(c) under testing conditions. From these graphs, it is evident that the agent has successfully learned the underlying patterns and relationships in the training data and can generalize that knowledge to unseen data. For the small-signal stability evaluation, 5% parameter variations are applied to the inductor and capacitor, i.e., (L= 2.1 mH, C=157.5 μF). The results of small-signal stability for load power variations, output voltage, and inductor current under the same load power variations are illustrated in Figure. 9(a)-Figure. 9(c). To analyze the robustness of the proposed method, large load variations are applied as disturbances (load power transients), as given by equation (14).

$$P(t) = \begin{cases} 200\ W & 0 \le t \le 0.14 \\ 1600\ W & 0.14 \le t \le 0.2 \\ 200\ W & 0.2 \le t \le 0.4 \end{cases} \tag{14}$$

The large power variations, output voltage, and inductor current are depicted in Figure. 10(a)-Figure. 10(c). These figures show that the controller has achieved satisfactory results with good dynamic performance (lesser overshoot) and steady-state performance (steady-state error below 2%, considered from a power engineering aspect [32]).

Case-II: For Real-Time validation, the three DRL algorithms are implemented as controllers and simulated on the OPAL-RT platform. The output voltage regulation, inductor current, and load power variations with the TD3 method are depicted in Figure. 11. The results corresponding to

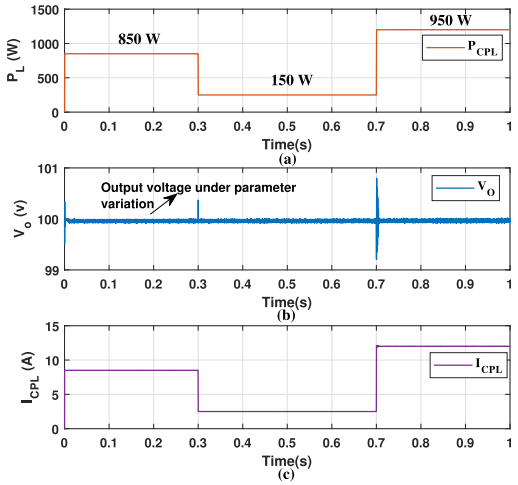

FIGURE 9. Assessing the response with the TD3 method under small signal conditions for (a) load power variations, (b) output voltage, and (c) inductor current.

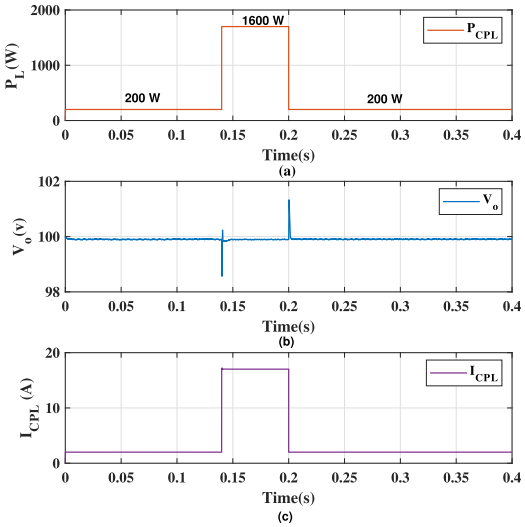

FIGURE 10. Evaluating the response with the TD3 method under large signal conditions for (a) load power variations, (b) output voltage, and (c) inductor current.

the DDPG method and the DQN method are illustrated in Figure. 12 and Figure. 13, respectively. The OPAL-RT target simulator (OP4512) has analog output pins with a maximum voltage of 16V. Therefore, the experimental results are scaled down to 16V, and the corresponding scaling factor is shown in the waveforms. While all three controllers demonstrate appropriate steady-state and dynamic performance under normal operating conditions, their effective training and robustness in performance are compared in the next case.

Case-III: Here, the proposed TD3 method is compared with similar state-of-the-art DRL methods in the literature, focusing on aspects of robustness, effectiveness in static, dynamic, and robust performance. The output voltage regulation with DQN, DDPG, and the TD3 method, along with their comparison under normal conditions, is shown in Figure. 14. Small-signal stability is explored under parameter variations for these methods, and the corresponding results are presented in Figure. 15. These figures indicate that under

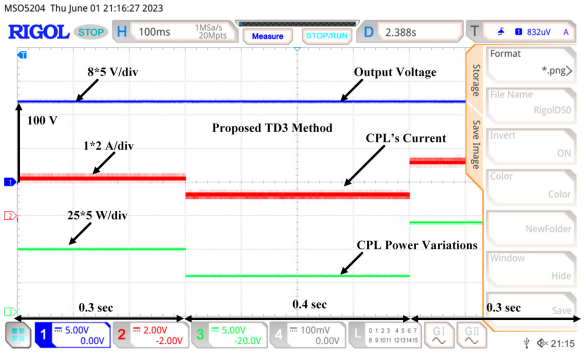

FIGURE 11. Real-time validation of the TD3 method under typical conditions involves assessing variations in output voltage, inductor current, and load power.

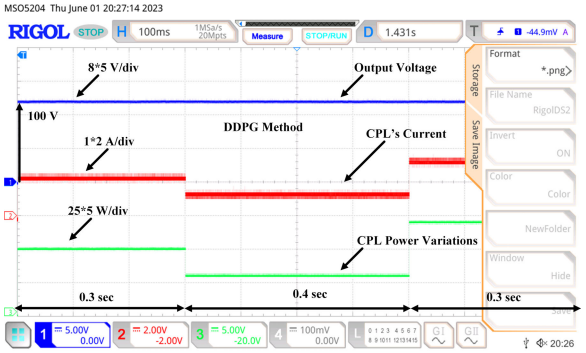

FIGURE 12. Real-time validation of the DDPG method under normal conditions includes assessing variations in output voltage, inductor current, and load power.

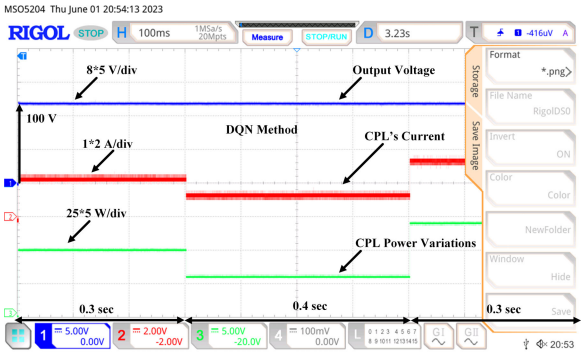

FIGURE 13. Real-time validation of the DQN method under typical conditions involves assessing variations in output voltage, inductor current, and load power.

normal conditions and in the context of small-signal stability, the TD3 method stands out with its relatively improved steady-state and transient performance.

The robustness of the controller is also studied under external disturbances such as large load power variations for these methods, and their results are illustrated in Figure. 16. From Figure. 16, it is indicated that the percentage overshoot in output voltage exceeds the limits under large-signal stability with the DQN method compared to DDPG & TD3. Additionally, there are small oscillations sustained in the DDPG method. The dynamic and steady-state performance parameters, i.e., peak overshoot, settling time, and steady-state error, are listed in Table 2 for all three methods under

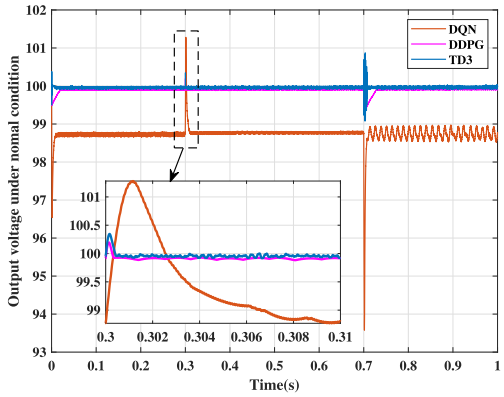

FIGURE 14. Under typical conditions, the voltage response to CPL variations is examined for the DQN, DDPG, and TD3 methods.

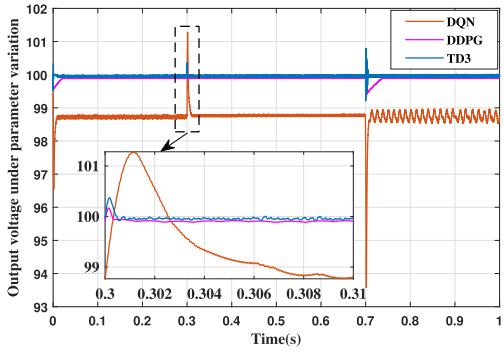

FIGURE 15. Assessing voltage response under small-signal stability (parameter variation) conditions for the DQN, DDPG, and TD3 methods.

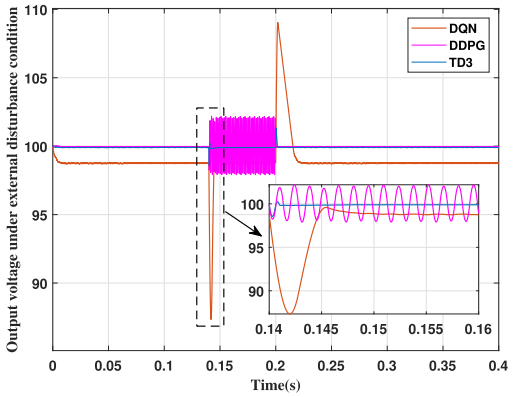

FIGURE 16. Evaluating voltage response to CPL variations under large signal stability conditions (load transients) for the DQN, DDPG, and TD3 methods.

three conditions (normal, parameter variation, & external disturbances). It is evident from Table 2 that the TD3 method performs well compared to the DQN and DDPG methods.

The action signal selection of the three control algorithms is shown in Figure. 17. The duty ratio plot between the two control techniques infers that a wider duty ratio range indicates higher flexibility in adjusting the converter's output voltage. It allows for better adaptation to varying load conditions and greater control over the converter's operation. A stable control technique will maintain a consistent duty ratio around the desired value without excessive variations or instability. From Figure. 17, it is clear that the DQN method

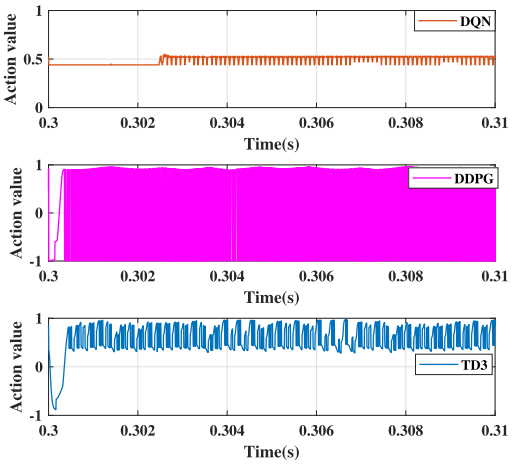

FIGURE 17. Duty ratio comparison among the DQN, DDPG, and TD3 methods.

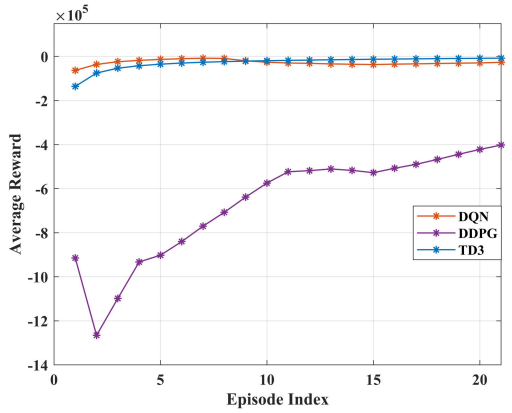

FIGURE 18. Training statistics for the DQN, DDPG, and TD3 methods, illustrating the learning curves.

& TD3 have a wider range of duty ratios which hold on the switch for a long duration to limit the transient current. It is also evident that the TD3 method maintains a consistent duty ratio about the operating point, whereas DDPG shows irregular behavior. Furthermore, it is understood from the literature [37] that the training complexity of the DQN algorithm escalates with a large action space. So, the values of the action space are curtailed based on domain experts for faster convergence and effectiveness in agent learning. Additionally, the identification of limits on the action space is not practical for various tasks. For continuous action space, the DDPG and TD3 performances are acceptable under all operating conditions. However, the TD3 method is preferable to DDPG for the following reasons.

The training statistics of the three algorithms are illustrated in Figure. 18, presenting the learning curve (Average reward vs. Episode number) of the algorithm. It is observed from Figure. 18 that the sample efficiency, defined as achieving a higher average reward using a minimal number of interactions, is higher for the TD3 algorithm compared to the other two algorithms. This indicates that TD3 can learn more from each interaction. The figure. 18 serves as evidence that DDPG exhibits unstable training due to the overestimation

TABLE 2. Analyzing transient and steady-state performance characteristics of the DQN, DDPG, and TD3 control algorithms.

Condition	Parameter	DQN	DDPG	TD3
Normal condition	% overshoot	2.5	1.3	1
	settling time	12 msec	8 msec	1 msec
	steady state error	1.2	0.5	0.2
Parameter variation	% overshoot	2.9	1.5	1.2
	settling time	15 msec	11 msec	3 msec
	steady state error	1.5	0.4	0.2
External disturbances	% overshoot	10	2.5	1.7
	settling time	250 msec	200 msec	160 msec
	steady state error	2	1.6	0.3

TABLE 3. Error performance metrics of the DQN, DDPG, and TD3 algorithms.

Performance metric	DQN	DDPG	TD3
IAE	1.241	0.091	0.0422
ISE	1.541	0.0085	0.0021
ITAE	0.619	0.0045	0.0041
ITSE	0.767	0.0041	0.0009

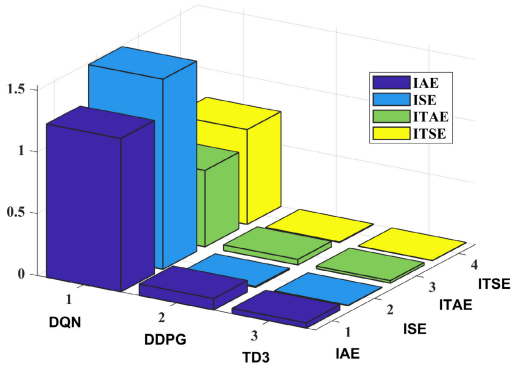

FIGURE 19. Performance metrics, including IAE, ISE, ITAE, and ITSE, for the DQN, DDPG, and TD3 algorithms.

of its Q-function value during training, a challenge overcome by TD3 with its twin critic networks, resulting in a smoother training curve. Error performance indices such as Integral Time Square Error (ITSE), Integral Time Absolute Error (ITAE), Integral Square Error (ISE), and Integral Absolute Error (IAE) for all three methods are provided in Table 3. It is evident from Figure. 19 that TD3 outperforms all the other methods.

V. CONCLUSION

A DRL-based modified Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm has been employed as an *end-to-end* controller for voltage regulation of a DC/DC buck converter feeding Constant Power Loads (CPLs). The TD3 method has demonstrated improved dynamic and steady-state results under untrained data of load power variations and parameter fluctuations. The OPAL-RT results indicate that the recommended method has achieved superior performance compared to other DRL controllers, particularly in scenarios with large CPL variations and model uncertainty in real-time. The dynamic and steady-state performance of the TD3 method is smoother, and the training complexity is lower compared to DQN and DDPG. Future work could involve exploring reward shaping using a non-heuristic approach. Additionally, the agent could be investigated with other neural network architectures such as Recurrent Neural Networks (RNN), Long Short Term Memory (LSTM), etc.

ACKNOWLEDGMENT

The authors would like to acknowledge the Common Research Technological Development Hub (CRTDH), NIT Andhra Pradesh sponsored by DSIR, Ministry of Science and Technology, Government of India for equipment support. They also thank Dr. V. Sandeep and Dr. P. Shankar, National Institute of Technology Andhra Pradesh for data collection.

REFERENCES

[1] S. K. Sahoo, A. K. Sinha, and N. K. Kishore, "Control techniques in AC, DC, and hybrid AC–DC microgrid: A review," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 6, no. 2, pp. 738–759, Jun. 2018.

[2] X. Li, X. Zhang, W. Jiang, J. Wang, P. Wang, and X. Wu, "A novel assorted nonlinear stabilizer for DC–DC multilevel boost converter with constant power load in DC microgrid," *IEEE Trans. Power Electron.*, vol. 35, no. 10, pp. 11181–11192, Oct. 2020.

[3] A. Emadi, B. Fahimi, and M. Ehsani, "On the concept of negative impedance instability in the more electric aircraft power systems with constant power loads," *SAE Trans.*, vol. 108, pp. 689–699, Jan. 1999.

[4] M. Wu and D. D. Lu, "A novel stabilization method of LC input filter with constant power loads without load performance compromise in DC microgrids," *IEEE Trans. Ind. Electron.*, vol. 62, no. 7, pp. 4552–4562, Jul. 2015.

[5] S. Kapat and P. T. Krein, "A tutorial and review discussion of modulation, control and tuning of high-performance DC–DC converters based on small-signal and large-signal approaches," *IEEE Open J. Power Electron.*, vol. 1, pp. 339–371, 2020.

[6] A. Kwasinski and C. N. Onwuchekwa, "Dynamic behavior and stabilization of DC microgrids with instantaneous constant-power loads," *IEEE Trans. Power Electron.*, vol. 26, no. 3, pp. 822–834, Mar. 2011.

[7] A. Riccobono and E. Santi, "Comprehensive review of stability criteria for DC power distribution systems," *IEEE Trans. Ind. Appl.*, vol. 50, no. 5, pp. 3525–3535, Sep. 2014.

[8] M. Céspedes, T. Beechner, L. Xing, and J. Sun, "Stabilization of constant-power loads by passive impedance damping," in *Proc. 24th Annu. IEEE Appl. Power Electron. Conf. Expo. (APEC)*, Feb. 2010, pp. 2174–2180.

[9] X. Lu, K. Sun, J. M. Guerrero, J. C. Vasquez, L. Huang, and J. Wang, "Stability enhancement based on virtual impedance for DC microgrids with constant power loads," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2770–2783, Nov. 2015.

[10] A. Emadi and M. Ehsani, "Negative impedance stabilizing controls for PWM DC–DC converters using feedback linearization techniques," in *Proc. Collection Tech. Papers., 35th Intersociety Energy Convers. Eng. Conf. Exhib. (IECEC)*, 2000, pp. 613–620.

[11] V. Grigore, J. Hatonen, J. Kyyra, and T. Suntio, "Dynamics of a buck converter with a constant power load," in *Proc. 29th Annu. IEEE Power Electron. Specialists Conf.*, vol. 1, May 1998, pp. 72–78.

[12] S. Pang, B. Nahid-Mobarakeh, S. Pierfederici, Y. Huangfu, G. Luo, and F. Gao, "Large-signal stable nonlinear control of DC/DC power converter with online estimation of uncertainties," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 9, no. 6, pp. 7355–7368, Dec. 2021.

[13] A. Griffo, J. Wang, and D. Howe, "Large signal stability analysis of DC power systems with constant power loads," in *Proc. IEEE Vehicle Power Propuls. Conf.*, Sep. 2008, pp. 1–6.

[14] N. Sarrafan, J. Zarei, R. Razavi-Far, M. Saif, and M.-H. Khooban, "A novel on-board DC/DC converter controller feeding uncertain constant power loads," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 9, no. 2, pp. 1233–1240, Apr. 2021.

[15] J. Solsona, S. Gomez Jorge, and C. Busada, "Nonlinear control of a buck converter feeding a constant power load," *IEEE Latin Amer. Trans.*, vol. 12, no. 5, pp. 899–903, Aug. 2014.
[16] J. A. Solsona, S. G. Jorge, and C. A. Busada, "Nonlinear control of a buck converter which feeds a constant power load," *IEEE Trans. Power Electron.*, vol. 30, no. 12, pp. 7193–7201, Dec. 2015.
[17] A. M. Rahimi and A. Emadi, "Active damping in DC/DC power electronic converters: A novel method to overcome the problems of constant power loads," *IEEE Trans. Ind. Electron.*, vol. 56, no. 5, pp. 1428–1439, May 2009.
[18] Q. Xu, Y. Yan, C. Zhang, T. Dragicevic, and F. Blaabjerg, "An offset-free composite model predictive control strategy for DC/DC buck converter feeding constant power loads," *IEEE Trans. Power Electron.*, vol. 35, no. 5, pp. 5331–5342, May 2020.
[19] O. Andrés-Martínez, A. Flores-Tlacuahuac, O. F. Ruiz-Martinez, and J. C. Mayo-Maldonado, "Nonlinear model predictive stabilization of DC–DC boost converters with constant power loads," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 9, no. 1, pp. 822–830, Feb. 2021.
[20] S.-C. Tan, Y. M. Lai, and C. K. Tse, "An evaluation of the practicality of sliding mode controllers in DC–DC converters and their general design issues," in *Proc. 37th IEEE Power Electron. Specialists Conf.*, Jun. 2006, pp. 1–7.
[21] Y. He and F. Luo, "Design and analysis of adaptive sliding-mode-like controller for DC–DC converters," *IEE Proc.-Electr. Power Appl.*, vol. 153, no. 3, pp. 401–410, 2006.
[22] Z. Wang, S. Li, and Q. Li, "Discrete-time fast terminal sliding mode control design for DC–DC buck converters with mismatched disturbances," *IEEE Trans. Ind. Informat.*, vol. 16, no. 2, pp. 1204–1213, Feb. 2020.
[23] M. Asghar, A. Khattak, and M. M. Rafiq, "Comparison of integer and fractional order robust controllers for DC/DC converter feeding constant power load in a DC microgrid," *Sustain. Energy, Grids Netw.*, vol. 12, pp. 1–9, Dec. 2017.
[24] B. A. Martínez-Treviño, R. Jammes, A. E. Aroudi, and L. Martínez-Salamero, "Sliding-mode control of a boost converter supplying a constant power load," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 7807–7812, Jul. 2017.
[25] N. Vafamand, M. H. Khooban, T. Dragicevic, F. Blaabjerg, and J. Boudjadar, "Robust non-fragile fuzzy control of uncertain DC microgrids feeding constant power loads," *IEEE Trans. Power Electron.*, vol. 34, no. 11, pp. 11300–11308, Nov. 2019.
[26] M. M. Mardani, N. Vafamand, M. H. Khooban, T. Dragicevic, and F. Blaabjerg, "Design of quadratic D-stable fuzzy controller for DC microgrids with multiple CPLs," *IEEE Trans. Ind. Electron.*, vol. 66, no. 6, pp. 4805–4812, Jun. 2019.
[27] Y. Yuan, C. Chang, Z. Zhou, X. Huang, and Y. Xu, "Design of a single-input fuzzy PID controller based on genetic optimization scheme for DC–DC buck converter," in *Proc. Int. Symp. Next-Generation Electron. (ISNE)*, May 2015, pp. 1–4.
[28] H. Farsizadeh, M. Gheisarnejad, M. Mosayebi, M. Rafiei, and M. H. Khooban, "An intelligent and fast controller for DC/DC converter feeding CPL in a DC microgrid," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 67, no. 6, pp. 1104–1108, Jun. 2020.
[29] E. Mohammadi, M. Alizadeh, M. Asgarimoghaddam, X. Wang, and M. G. Simoes, "A review on application of artificial intelligence techniques in microgrids," *IEEE J. Emerg. Sel. Topics Ind. Electron.*, vol. 3, no. 4, pp. 878–890, Oct. 2022.
[30] 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.
[31] 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.
[32] 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.
[33] 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.
[34] 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
[35] 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
[36] 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.
[37] M. Zhan, J. Chen, C. Du, and Y. Duan, "Twin delayed multi-agent deep deterministic policy gradient," in *Proc. IEEE Int. Conf. Prog. Informat. Comput.*, Dec. 2021, pp. 48–52.

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 the National Institute of Technology 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 continues his research in machine learning implementation in dc and ac microgrids, poised for further contributions to artificial intelligence in power electronics applications. He remains 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 the Ph.D. degree, he expanded his expertise during a Postdoctoral Fellowship with IIT, from 2017 to 2018. He is currently an Assistant Professor with the Electrical Engineering Department, National Institute of Technology 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.

YANNAM RAVI SHANKAR was born in July 1986. He received the B.Tech. degree in electrical and electronics engineering from the Prakasam Engineering College, JNTU, Hyderabad, Andhra Pradesh, India, in 2008, and the M.Tech. degree in power and industrial drives from JNTU, Kakinada, India, in 2011. He worked as an Assistant Professor with the SSN Engineering College, Ongole, Andhra Pradesh, for two years. He is currently an Assistant Professor with Dire Dawa University, Dire Dawa, Ethiopia. His research interests include power electronics, power systems, advanced control systems, and electrical machines.