



## Article

# Multi-Fault-Tolerant Operation of Grid-Interfaced Photovoltaic Inverters Using Twin Delayed Deep Deterministic Policy Gradient Agent

Shyamal S. Chand <sup>1,\*</sup>, Branislav Hredzak <sup>1,\*</sup> and Maurizio Cirrincione <sup>2,†</sup>

<sup>1</sup> School of Electrical Engineering and Telecommunications (EET), University of New South Wales (UNSW), Sydney, NSW 2052, Australia

<sup>2</sup> Université Marie et Louis Pasteur, UTBM, CNRS, Institut FEMTO-ST, F-90010 Belfort, France; maurizio.cirrincione@usp.ac.fj

\* Correspondence: shyamal.chand@unsw.edu.au (S.S.C.); b.hredzak@unsw.edu.au (B.H.)

† School of Information Technology, Engineering, Mathematics, and Physics (STEMP), University of the South Pacific (USP), Laucala Campus, Suva, Fiji.

**Abstract:** The elevated penetration of renewable energy has seen a significant increase in the integration of inverter-based resources (IBRs) into the electricity network. According to various industrial standards on interconnection and interoperability, IBRs should be able to withstand variability in grid conditions. Positive sequence voltage-oriented control (PSVOC) with a feed-forward decoupling approach is often adopted to ensure closed-loop control of inverters. However, the dynamic response of this control scheme deteriorates during fluctuations in the grid voltage due to the sensitivity of proportional–integral controllers, the presence of the direct- and quadrature-axis voltage terms in the cross-coupling, and predefined saturation limits. As such, a twin delayed deep deterministic policy gradient-based voltage-oriented control (TD3VOC) is formulated and trained to provide effective current control of inverter-based resources under various dynamic conditions of the grid through transfer learning. The actor–critic-based reinforcement learning agent is designed and trained using the model-free Markov decision process through interaction with a grid-connected photovoltaic inverter environment developed in MATLAB/Simulink<sup>®</sup> 2023b. Using the standard PSVOC method results in inverter input voltage overshoots of up to 2.50 p.u., with post-fault current restoration times of as high as 0.55 s during asymmetrical faults. The designed TD3VOC technique confines the DC link voltage overshoot to 1.05 p.u. and achieves a low current recovery duration of 0.01 s after fault clearance. In the event of a severe symmetric fault, the conventional control method is unable to restore the inverter operation, leading to integral-time absolute errors of 0.60 and 0.32 for the currents of the  $d$  and  $q$  axes, respectively. The newly proposed agent-based control strategy restricts cumulative errors to 0.03 and 0.09 for the  $d$  and  $q$  axes, respectively, thus improving inverter regulation. The results indicate the superior performance of the proposed control scheme in maintaining the stability of the inverter DC link bus voltage, reducing post-fault system recovery time, and limiting negative sequence currents during severe asymmetrical and symmetrical grid faults compared with the conventional PSVOC approach.

**Keywords:** reinforcement learning; twin delayed deep deterministic policy gradient; data-driven control; grid fault ride-through; voltage-oriented control; asymmetrical and symmetrical faults



Academic Editors: Marcel Nicola, Petre-Marlin Nicolae and Dan Selisteanu

Received: 20 November 2024

Revised: 16 December 2024

Accepted: 21 December 2024

Published: 26 December 2024

**Citation:** Chand, S.S.; Hredzak, B.; Cirrincione, M. Multi-Fault-Tolerant Operation of Grid-Interfaced Photovoltaic Inverters Using Twin Delayed Deep Deterministic Policy Gradient Agent. *Energies* **2025**, *18*, 44. <https://doi.org/10.3390/en18010044>

**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The revolution in renewable energy is evolving substantially with the advancement of energy policies, the development of infrastructure, and the integration of continually improving power technologies. Renewable energy sources (RES), such as photovoltaics (PVs), wind energy conversion systems (WECSs), hydroelectric power stations (HPSs), and fuel cells (FCs), combined with various forms of energy storage devices, such as batteries and flywheels, are becoming increasingly dominant components of the power network. Power generation using distributed energy resources (DER) requires optimal deployment of extraction and control techniques for large-scale integration and transmission. Improving the stability and increasing the operational reliability of these microgrids is a topic of significance among many researchers around the world today [1,2].

Microgrids (MGs) generally operate in two main modes of operation: islanded or standalone and grid-connected. In islanded mode, central or decentralized control can be used to provide set voltage and frequency points, such as droop control or master–slave methods [3]. However, when coupled to the upstream electricity network, the grid-connected inverter is programmed to follow the voltage and frequency of the grid using control methods such as voltage-oriented control [2]. Faults in power systems can generally be categorized as symmetrical and asymmetrical faults. An asymmetric fault occurs when the voltage phases of the grid are unequally affected, creating an imbalance in the system, resulting in different fault currents in each phase. Due to the unbalanced nature of asymmetrical faults, they are more challenging to model and analyze. However, a symmetric fault affects the three phases of the grid voltage equally, causing the fault current in all three phases to be identical in magnitude and phase angle. Symmetrical faults are typically the most severe type and generate the highest fault currents, but occur less frequently than asymmetrical faults.

Most MGs connected to the grid are not of special concern for the conventional power network, as they cannot influence its operating parameters (voltage and frequency). However, when several MGs are integrated into a low- or medium-voltage network, the stability and reliability of the larger grid can be compromised [4]. To ensure standard operation, virtual inertias are often incorporated into droop control loops, allowing the voltage and frequency magnitudes to be adjusted according to the active and reactive power generated by the inverter using low-bandwidth communication. This approach can effectively prevent voltage sagging while giving flexibility to transition between island-connected and grid-connected operations through a phase-locked loop (PLL). Droop control has a simple implementation; however, there are a few downfalls, such as a slow transient response, circulating current between converters, and an imprecise power distribution during grid faults [5]. Furthermore, most standalone industrial grid-interfaced power converters are now required to have multiple fault ride-through (FRT) capabilities [6]. For instance, IEEE Std. 2800-2022—“IEEE Standard for Interconnection and Interoperability of Inverter-Based Resources (IBRs) Interconnecting with Associated Transmission Electric Power Systems”—clearly specifies performance requirements for dynamic active and reactive power support under abnormal frequency or voltage, negative sequence current injection, and low voltage or high voltage ride-through operation. As such, the IBR plant must be able to switch between multiple modes of operation when required [7].

Consequently, in the event of voltage imbalance faults at a particular node, the effect is transmitted to power transformers, converters, and the larger microgrid, eventually affecting the synchronization of inverters and sensitive loads. The protection system of MGs, whether in grid-connected or islanded operation, must respond to abnormal grid conditions. In comparison, the fault current in the connected mode of the grid is 10–50 times the full load per unit current due to the low impedance of the utility network,

but in the islanded mode it is only 1.2–5 times higher [8]. As an ancillary service, negative sequence currents need to be injected to compensate for any voltage imbalances at the node. A complex proportional resonant controller can be designed to achieve this by sharing negative sequence currents between participating inverters connected to an AC microgrid using a communication link that transmits at 10 Hz [9]. However, the communication link is prone to failure and low-bandwidth data transfer can increase communication delays, resulting in deteriorated robustness during faults. In [10], an improved inverter control technique based on a grid impedance estimation technique is proposed using a Newton–Raphson algorithm implemented in the stationary reference frame that takes into account positive sequence phasors as input. This method iteratively adjusts the virtual impedance and stabilizes the inverter under unbalanced and harmonic-inflicted grid voltages. Similarly, another method involves a positive–negative sequence synchronverter capable of limiting output currents to reduce power oscillations when subjected to disturbances in the grid [11]. When the grid voltages abruptly differ from the nominal operating set points, these control approaches have poor transients and a long settling time.

Moreover, to leverage the operation of smart inverters, low-voltage ride-through (LVRT) capability is essential. An autonomous model predictive controller (AMPC), which is an improvement over the classical MPC, can be adopted for adjustment of the injected active and reactive powers of a grid-connected quasi-Z-source inverter during such scenarios. Compared with conventional MPC, the trial-and-error design stage of the weight factor is replaced with an auto-tuning objective normalization function. An additional benefit of this method is the provision for a transition between maximum power point tracking (MPPT) and FRT operation with ease based on the command of the grid operator or the grid condition [12]. Intrinsically, it can sometimes be difficult to reduce inverter currents during LVRT, as it adjusts the reactive current to support the voltage of the grid [13]. As such, [14] proposes a method to reduce limit fault currents using phase angle adjustment and achieve improved voltage support by embedding the network impedance in the current control loop. In other works, a bridge-type fault current limiter (BFCL) is connected between the distributed energy resources and the main grid at the point of common coupling (PCC) to limit fault currents and provide reactive power support through cooperative control with the voltage source converter [15]. For a switched resistance wind generator (SRWG), an optimal proportional–integral (PI) voltage controller tuning method using the elephant herd algorithm (EHA) is another way to provide LVRT capabilities according to the grid code [16]. Further analysis of other existing grid fault-ride-through techniques is presented in Table 1.

**Table 1.** Critical analysis of existing fault-ride-through methods.

Ref.	System Architecture	Type of Fault(s)	Effects	Adopted Mitigation Technique	Advantages	Remarks/Limitation
[17]	Grid-connected multi-modular converter	Voltage swell	Cause over-modulation of multi-modular converter leading to distortions in current waveforms and instability	Fundamental zero-sequence voltage injection	Gurrantees even modulation margins in all three phases. Improves multi-modular converter tolerance during high voltage ride-through	No provision for compensating other major types of fault in the network
[18]	Grid-connected three-phase four-wire inverter	Voltage sag (symmetrical and asymmetrical) and short-circuit ratio variation	Significant increase in current harmonics. Significant drop in inverter power at severe sags. Oscillation in DC-link voltage	Constant average active power/constant active current/constant peak current injection	Constant average active power injection can assure operated at rated conditions during faults	Does not offer smooth transition to pre-fault condition. Multiple current references must be specified for fault ride-through

Table 1. Cont.

Ref.	System Architecture	Type of Fault(s)	Effects	Adopted Mitigation Technique	Advantages	Remarks/Limitation
[19]	Three-phase energy storage converter	Unbalanced voltage sags	Causes drastic changes in power angle. Disruptions in power supplied to the grid.	Current droop control	Provide reactive power support. Ability to limit power angle deviation. Does not require power calculations in the outer control loop.	Modified control hinders commanding in terms of required power. The battery state of charge is not considered during fault ride-through.
[20]	Modified IEEE 9-bus network with solar, wind, and battery storage	Asymmetrical and symmetrical	Affects stability and reliability of the entire network.	Sliding mode linearization recurrent high-order neural controller trained using an extended Kalman filter	Allows mitigation of nonlinear dynamics due to parameter variation or faults. Reduced divergence in the DC link voltage and active and reactive power during abnormal operating conditions.	The approach does not consider nonlinear loading (inductive or capacitive). The use of a recurrent high-order neural network, as well as a sliding mode controller, complicates the design process.
[21]	Doubly-fed induction generator-based wind turbine systems	Voltage sags, asymmetrical and symmetrical	Increase in stator currents which leads to transient overshoots in rotor currents due to magnetic coupling.	Nonlinear sliding mode controller for bridge-type fault current limiter	To overcome the chattering phenomenon, an exponential reaching law is proposed for the sliding surface design. Current limiting is achieved during the fault period.	There is a lack of stability analysis as operating margins of the proposed control technique are not specified. Does not account for scalability issues that may arise for larger wind farms.
[22]	Standalone inverter	Asymmetrical and symmetrical	Can damage the semiconductor switches. May cause the inverter to shut down, depriving the power supply to critical loads.	Finite control set model predictive control	High-speed fault detection. Ability to limit total harmonic distortion of currents during faults. Fast recovery to nominal mode of operation.	Requires operation in dual modes, nominal and fault conditions. The objective functions need to be well defined to suit the system. Relies on accurate fault modeling and system parameters.
[23]	Doubly-fed induction generator-based wind turbine systems	Asymmetrical and symmetrical	Current surge in stator and armature winding. Compromises drive-train integrity due to induced torque transients.	Event-triggered sliding mode control combined with supplementary energy storage support interfaced via an alternate active bridge converter	Effectively regulates the DC link voltage. Low use of communication channels.	The use of supercapacitor banks, disturbance observer, and additional dual active bridge converters complicates the design process. Requires supplementary energy sources.

Recent advances in computational power density have led to the rise of numerous applications of machine learning (ML) or artificial intelligence (AI) in demand-side energy management, distribution networks, and improving the stability of power converters in MGs [24]. In an effort to advance the application of the Markovian jump stochastic neural network, a framework is particularly necessary to model abrupt variation of the system architecture and parameters in the context of stability and immunity to input disturbances [25]. The analysis of the Cohen–Grossberg bidirectional associative memory neural network comprising time delay can be reconfigured to inverter control under uncertainties [26]. In addition, synchronization of control systems techniques with sampled data prior to additive delay has also been investigated in detail [27]. Such approaches can be applied to improve the reliability and synchronization of multiple grid-connected inverters for optimal power distribution. In addition, fast and accurate fault detection often relies on appropriate data selection and feature extraction methods. High-quality hybrid feature

selection based on correlated items applied in the context of classification can be used to identify critical aspects of inverter performance by detecting voltage variation or unpredictable changes in load [28–30]. Through careful evaluation and reduction of the feature space with a focus on the most relevant data points, the computational cost for training and deploying sophisticated algorithms such as the adaptive neural fuzzy interference system for strategic decision-making in the energy sector has also been progressively advanced due to improved overall reliability in dynamic environments [31,32].

In hierarchical control, machine learning-based approaches have induced significant improvements in power system performance at all levels (primary, secondary, and tertiary) and have been shown to be effective in load forecasting, generation prediction, and fault diagnosis [33]. In a low-voltage grid, LVRT and constant reactive current injection during voltage sags can be achieved through online detection of grid abnormalities based on wavelet transforms and shallow neural networks and the resultant adjustments in the power references. The fault identification technique reported a classification precision of 98.3% with PCC voltage and frequency as input data [34]. Terminal voltage regulation is crucial during transitions between grid-connected and standalone operations. A DG inverter can be controlled using a deep neural network (DNN) trained offline in conjunction with a proportional integral derivative (PID) controller to ensure a smooth transition. Additionally, a feedforward loop can be incorporated to mitigate harmonics for nonlinear loads [35]. Battery energy storage systems (BESS) in MGs are commonly used for voltage support and fault compensation. To address the issue of short circuit-induced voltage sags in power lines using BESS in hybrid MG, a recurrent wavelet petri fuzzy neural network (RWPFNN) can be used in the control loop for LVRT and fast voltage restoration [36]. Other studies have employed neural networks with an extended Kalman filter to handle the control of doubly fed induction generators during sensor failure on controlled variables of back-to-back power converters [37].

In addition, reinforcement learning (RL) techniques have been extensively utilized to design better secondary control (power and frequency) of power inverters. For unstructured MGs with tidal power units and vehicle-to-grid connections, load frequency control is extremely important to restore stability. A fractional gradient descent based on a fuzzy logic controller can be implemented as the main frequency regulator with a deep deterministic policy gradient (DDPG) composed of an actor–critic architecture to produce additional control signals for frequency stabilization [38]. In [39], a quantum neural network (QNN) is formulated to cooperatively control the frequency of an isolated MG. It combines DRL and quantum machine learning to minimize the number of parameters and minimize training requirements for a 13-bus network. A data set containing the frequency response of unoptimized linear controllers is initially required to train the network through supervised learning. In [40], a frequency regulation approach is proposed using deep reinforcement learning (DRL), considering both the frequency performance and the economic operational limits of the MG. The twin delayed deep-deterministic policy gradient (TD3PG) algorithm is implemented based on historical data for adaptive selection of the best commands to ensure frequency stability.

Moreover, a virtual inertia emulation technique is established in [41] using the twin delayed deep deterministic policy gradient (TD3PG) reinforcement learning algorithm for frequency regulation in weak grids with reduced inertia and damping capability. Verification shows the ability of such controllers to better stabilize frequency in grid-connected and standalone MGs. In another study, a Kuramoto-based consensus algorithm is used to implement multi-agent reinforcement learning to handle frequency deviations. In [42], an adaptive controller based on TD3PG is proposed to replace virtual synchronous generators (VSG) in multilevel modular converter control. This approach reported better system

strength and higher resistance to disturbances than conventional VSGs. Reinforcement learning algorithms can also be used in designing parameters of existing controllers such as proportional–integral, proportional–resonant, and fractional-order proportional–integral regulators, etc. In [43], a deep Q network is adopted to adjust the parameters of the proportional–integral controllers for a dual active bridge converter according to the modulation phase shifted angles. This allows real-time mapping of phase-shift angles to the necessary control parameters for stability.

Furthermore, reinforcement learning is applied to design Q-learning and TD3PG current controllers in variable-speed doubly fed induction generators for the rotor and grid-side inverters [44]. The performance of both agents is evaluated under dynamic conditions, leading to the conclusion that Q-learning agents may not perform well under significant disturbances compared to TD3 agents. In another study, the TD3PG agent is utilized again to attain superior performance of permanent magnet synchronous machines (PMSM) governed using the classical field-oriented control scheme (FOC). The agent works in tandem with ESO-type speed observers, which allows sensorless FOC control through online estimation of the machine speed. Multiple neural networks are used to obtain the load torque with the TD3 agent performing a real-time estimated speed correction. The results indicate a better reference tracking capability with the use of the RL TD3PG agent [45]. DDPG agents have also been implemented in a similar way and have been trained to work in parallel with the outer loop speed controller designed using sliding mode control (SMC) to improve the load torque rejection capability in PMSM [46]. In this work, it is shown that RL agents have the ability to directly manipulate control signals to obtain the desired performance without a high computational or mathematical burden. Although there are extensive applications of RL agents and other intelligent control methods in machine control [47], there is limited literature that focuses on designing robust and adaptive artificial-based controllers for inverters to withstand adverse grid disturbances.

With the increasing penetration of renewable energy into the electrical network and the increasing demand for resilient inverter-based technology, there is a growing need to develop advanced control methods to stabilize MGs under various dynamic conditions. The expanding nature of the incorporation of DERs with utility and standard ancillary practices makes the operation and control of voltage, frequency, and FRT mechanisms more challenging. AI has the capability to bring about a new frontier in power engineering with promising advances in processing density and improving deep learning algorithms [48]. If implemented appropriately, a self-governing, highly intelligent controller capable of withstanding various grid abnormalities is definitely a possibility in MG applications. Although the above-mentioned model-free control methods deliver promising results, most depend on the type of data used during the training process. As such, considering deep learning methods that do not require an exact MG model for efficient FRT and have provision for transfer learning is desirable [49]. Evidently, during grid faults, such as voltage swells, dips, unbalances, frequency disturbances, harmonics, asymmetrical, and symmetrical faults, the current control loop on IBRs is important in regulating power quality and ensuring stability. Most of the methods available in the literature often target compensating for one or a few types of fault using a particular approach, while some techniques require switching between multiple control modes depending on the induced fault. Some classical FRT controllers perform additional computations to generate negative sequence components for fault compensation; however, this requires exact plant modeling, which makes the design process monotonous.

In addition, the current loop with cross-coupling terms in conventional voltage-oriented control in inverter-based MGs causes instabilities during abnormal grid conditions, which are attributed to poor proportional–integral controller dynamics and saturation lim-

its. Based on the aforementioned constraints and inspired by reinforcement learning techniques, a novel twin delayed deep deterministic policy gradient-based reinforcement learning control approach is proposed to provide effective FRT operation and ensure optimal current injection from RES. The main contributions of this research article are as follows.

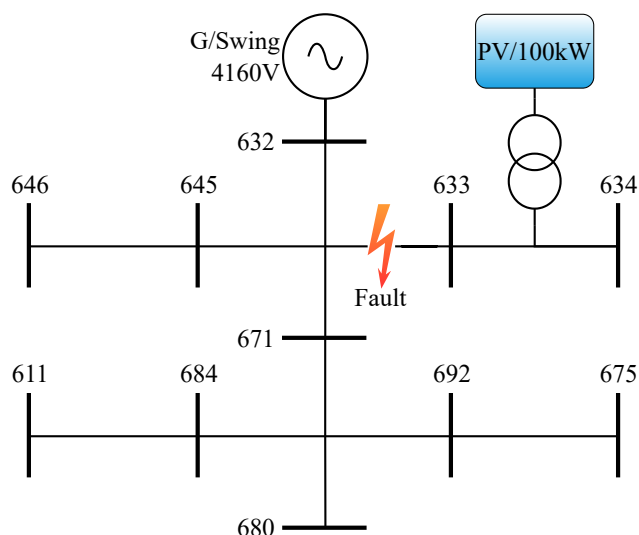
- A novel twin delayed deep-deterministic policy gradient agent is formulated and trained to generate the voltages of the direct and quadrature axes in real time for resilient operation of the inverter considering abnormal conditions of the grid based on a set of observed states from the environment developed in MATLAB/Simulink®.
- A framework for training and deploying fully AI-based current controllers has been established to achieve the optimal response of grid-interfaced inverters using a model-free Markov decision process guided through a suitable reward function, thus eliminating the need to design controllers based on uncertain operating parameters and an accurate system model.
- Initially, the agent is trained for a single asymmetric phase-to-phase fault and then applied to improve the inverter response for other asymmetric and symmetric grid faults using transfer learning. This feature allows the controller to continuously adapt to variations in operating conditions and be resilient despite deviations in grid voltages without requiring complex deep neural network architectures and minimizes computational or training costs.

Such control approaches have not been studied in depth in the field of inverter fault-ride operation. Therefore, this article focuses on developing the actor–critic-based data-driven control strategy, which has been thoroughly verified under asymmetric faults (L-L and L-L-G) and symmetric faults (L-L-L). The proposed fault ride-through approach is also compared with the standard positive sequence voltage-oriented control technique.

The organization of this paper is as follows. Section 1 provides a detailed background and an extensive review of the literature on the subject. Section 2 covers the description of the power conversion system, while Section 3 describes the traditional positive sequence voltage-oriented control method. The novel design process for the twin delayed deep deterministic policy gradient agent, intended for grid fault ride-through, is detailed in Section 4. This includes specifics on the architecture of the deep neural network, the formulation of the reward function, and the training procedure. Section 5 presents a comparative evaluation of the proposed fault ride-through approach against the conventional control method. The research considers three distinct fault scenarios, with the DC link voltage response, post-fault recovery time, and cumulative integral-time error analysis serving as primary evaluation metrics. Section 6 discusses the limitations of the study and future research opportunities, and the article is concluded in Section 7.

## 2. System Description

The IEEE 13-bus network is a benchmark test case system used for research to implement new concepts in electricity generation, transmission, and distribution. For modeling purposes, single AC generators are represented as the swing voltage source with a fixed X/R ratio of 10. The classical 13 bus system has multiple predetermined single- and three-phase loads connected to all buses. For this study, an inverter is connected to an existing 4160/415 V transformer between buses 633 and 634 to facilitate the integration of a 100 kW solar system as shown in Figure 1.



**Figure 1.** Single line diagram of the IEEE 13 bus network with an interconnected solar system.

In the primary power conversion circuit (Figure 2), a three-phase two-level inverter is connected to the IEEE 13 bus network between buses 633 and 634 with an LCL-type current harmonic filter. A DC/DC boost converter is used to increase the voltage of the PV panels to charge the DC link capacitors. The scheme follows a centralized inverter topology in which a single inverter encompasses the entire power conversion circuit and controls the power using a single DC link [50]. The incremental conductance (INC) maximum power point tracking algorithm is used for optimum power extraction. INC requires measurement of the array voltage and current to calculate the instantaneous conductance and update the reference duty cycle accordingly. It has the benefit of faster adaptation to rapidly changing climate conditions and fewer oscillations when attaining the MPP. The size of the perturbation voltage step is set to  $1 \times 10^{-3}$  V to minimize tracking error after reaching the desired operating point. Other parameters of the power conversion circuit are stated in Table 2.

**Table 2.** Power conversion circuit and controller specifications.

Parameter	Symbol	Value	Unit
Grid frequency	$\omega$	60	Hz
Rated PV power	$P_n$	100	kW
Utility line voltage	$E_{LL}$	415	V
Rated DC bus voltage	$V_{dc}$	700	V
Inverter switching frequency	$F_{sw}$	15	kHz
Input capacitor	$C_{in}$	10	$\mu$ F
Boost Inductance	$L$	2.4	mH
DC link capacitor	$C_{dc}$	2000	$\mu$ F
Boost switching frequency	$F_b$	5	kHz
Damping resistor	$R_d$	0.4	$\Omega$
Filter capacitance	$C_f$	77	$\mu$ F
Internal resistance of $L_g$	$R_g$	18	m $\Omega$
Internal resistance of $L_i$	$R_i$	28	m $\Omega$
Grid side filter inductance	$L_g$	177	$\mu$ H
Inverter side filter inductance	$L_i$	280	$\mu$ H
Current controller gains		$K_p = 0.6$ $K_i = 243$	
DC link controller gains		$K_p = 1.35$ $K_i = 395$	



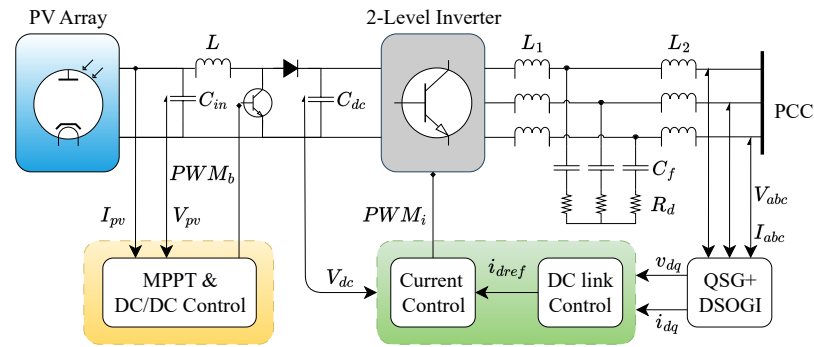


Figure 2. Dual-stage photovoltaic power conversion circuit.

For this study, as a prerequisite, voltages and currents are normalized to per unit values. The use of per-unit (p.u.) measurements significantly enhances scalability, making it easier to design, analyze, and operate systems of varying sizes and complexities with various control techniques. Using a base apparent power ( $S_{base}$ ) of 100 kVA and a nominal AC phase-to-phase voltage ( $V_{AC}$ ) of 415 V, the voltage and current bases can be specified as

$$V_{AC(base)} = \sqrt{\frac{2}{3}} \times V_{AC} \quad I_{AC(base)} = \sqrt{\frac{2}{3}} \times \frac{S_{base}}{V_{AC}} \quad (1)$$

Thus, the instantaneous voltages and currents of the grid at each phase measured at PCC are independently converted to their equivalents per unit using Equation (2). The transformed voltages and currents are then fed to a phased-locked loop to extract their  $d$  and  $q$  counterparts in normalized form.

$$V_{abc}(p.u.) = \frac{V_{PCC}}{V_{AC(base)}} \quad I_{abc}(p.u.) = \frac{I_{PCC}}{I_{AC(base)}} \quad (2)$$

Grid synchronization is crucial to maintaining quality power injection into the grid despite the volatile nature of the grid. As such, the double second-order generalized integrator phase-locked loop (DSOGI-PLL) is often used to synchronize the inverter voltages with the main grid. DSOGI-PLL has established superior performance in stabilizing the inverter under numerous disturbances of the grid [51,52]. It is based on an instantaneous symmetric component method with two inbuilt quadrature signal generators (QSG) to filter out additional harmonics present in the grid voltages in the first stage. Each QSG produces two voltages on the quadratic axis ( $90^\circ$  apart) that can be recombined to form positive and negative stationary reference frame grid components using Clarke's transformation of the grid voltage at PCC ( $v_\alpha$  and  $v_\beta$ ). Since in the case of a grid, the frequency is constant under normal operating conditions, the filter bandwidth is essentially dependent on the damping factor ( $K_\gamma$ ), which is generally chosen as  $\sqrt{2}$ . This makes it also suitable for variable-frequency applications [53].

Taking  $v_\alpha'$ ,  $v_\beta'$ ,  $qv_\alpha'$ , and  $qv_\beta'$  as QSG output, the decoupled positive and negative sequence components of the grid voltage can be obtained by using the sequence calculation in Equation (3). Thus, the corresponding positive and negative voltages on the  $d$  and  $q$  axes can be easily obtained by Park's transformation of the respective stationary reference frame voltages with  $\omega t$  as the estimated grid angle according to Equations (4) and (5).  $v_d^+$  gives the amplitude of the input voltage, while  $v_q^+$  gives information on the phase error. A PI controller is used as a loop filter to maintain  $v_d^+$  at zero and send it to the voltage-controlled oscillator represented by the integrator, thus allowing the phase angle of the input signal to be estimated [54]. Transformation of three-phase voltage and currents from their natural

states (*abc*) into (*dq*) synchronous rotating frame DC quantities allows decoupled control of grid-connected inverters in both positive and negative domains.

$$\begin{aligned} v_{\alpha}^+ &= \frac{1}{2}(v_{\alpha}' - qv_{\beta}') & v_{\beta}^+ &= \frac{1}{2}(v_{\beta}' + qv_{\alpha}') \\ v_{\alpha}^- &= \frac{1}{2}(v_{\alpha}' + qv_{\beta}') & v_{\beta}^- &= \frac{1}{2}(v_{\beta}' - qv_{\alpha}') \end{aligned} \quad (3)$$

$$v_d^+ = v_{\alpha}^+ \sin(\omega t) - v_{\beta}^+ \cos(\omega t) \quad \& \quad v_q^+ = v_{\alpha}^+ \cos(\omega t) + v_{\beta}^+ \sin(\omega t) \quad (4)$$

$$v_d^- = v_{\alpha}^+ \sin(-\omega t) - v_{\beta}^+ \cos(-\omega t) \quad \& \quad v_q^- = v_{\alpha}^+ \cos(-\omega t) + v_{\beta}^+ \sin(-\omega t) \quad (5)$$

The approach can also be applied to the measured grid currents at PCC to obtain its positive and negative sequence components. With  $i_{\alpha}$  and  $i_{\beta}$  as inputs to the dual QSG, the corresponding outputs  $i_{\alpha}'$ ,  $i_{\beta}'$ ,  $qi_{\alpha}'$ , and  $qi_{\beta}'$  can be recombined and produce the positive and negative current sequence components required for control purposes.

$$\begin{aligned} i_{\alpha}^+ &= \frac{1}{2}(i_{\alpha}' - qi_{\beta}') & i_{\beta}^+ &= \frac{1}{2}(i_{\beta}' + qi_{\alpha}') \\ i_{\alpha}^- &= \frac{1}{2}(i_{\alpha}' + qi_{\beta}') & i_{\beta}^- &= \frac{1}{2}(i_{\beta}' - qi_{\alpha}') \end{aligned} \quad (6)$$

$$i_d^+ = i_{\alpha}^+ \sin(\omega t) - i_{\beta}^+ \cos(\omega t) \quad \& \quad i_q^+ = i_{\alpha}^+ \cos(\omega t) + i_{\beta}^+ \sin(\omega t) \quad (7)$$

$$i_d^- = i_{\alpha}^+ \sin(-\omega t) - i_{\beta}^+ \cos(-\omega t) \quad \& \quad i_q^- = i_{\alpha}^+ \cos(-\omega t) + i_{\beta}^+ \sin(-\omega t) \quad (8)$$

### 3. Conventional Voltage-Oriented Control

The positive sequence voltage-oriented control (PSVOC) is most famously used to interface photovoltaic inverters with the grid. It has two cascaded control loops; the inner loop regulates the inverter currents on the *d* and *q* axes, and the outer loop ensures the stability of the DC link voltage, as shown in Figure 3. Normally, the DC link voltage reference is kept constant, and a PI controller is used to generate the direct axis current reference that corresponds to the net active power to be injected by the inverter into the grid [55]. Transient conditions and variations in the DC bus voltage are regulated by the capacitor’s charge-and-discharge process. In a grid-connected situation, the voltage can fluctuate due to changes in solar irradiation or temperature levels of the photovoltaic array, as well as oscillations in the AC power caused by grid imbalances.

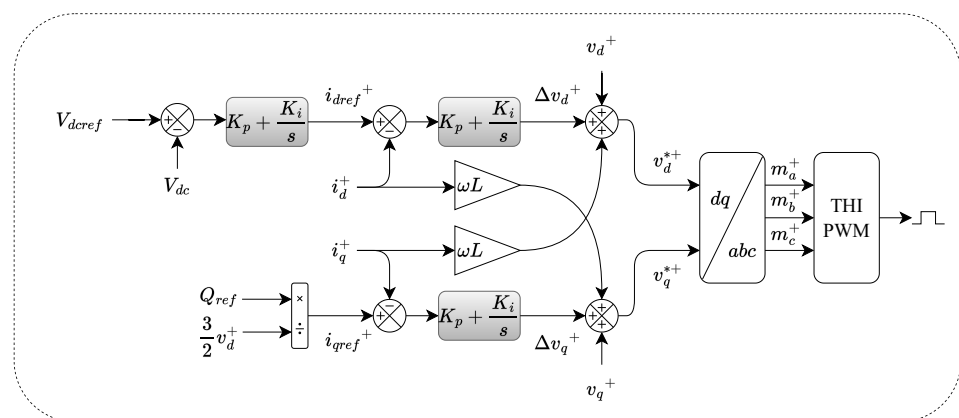


Figure 3. Conventional positive sequence voltage oriented control (PSVOC).

Injection of active power into the grid can be indicated by an increase in three-phase currents at the PCC, assuming that the grid voltage remains steady during normal operation. The change in DC link voltage is determined by the power balance between the solar panels and the currents injected into the grid. The reactive power supply can be commanded directly by specifying the required amount and calculating the current on

the quadrature axis. The error for both currents can be processed by inner-loop dual PI controllers independently. Typically, the interconnecting filter transfer function is used to tune the proportional ( $K_p$ ) and integral ( $K_i$ ) controller gains. The final inverter reference voltage is the combined output of the current PI controllers, the grid voltages, and the cross-coupling terms [56].

$$\begin{aligned} v_d^{*+} &= \left( K_p + \frac{K_i}{s} \right) \left( i_{dref}^+ - i_d^+ \right) + \omega L i_q^+ + v_d^+ \\ v_q^{*+} &= \left( K_p + \frac{K_i}{s} \right) \left( i_{qref}^+ - i_q^+ \right) + \omega L i_d^+ + v_q^+ \end{aligned} \quad (9)$$

Decoupling allows independent control of active and reactive powers for grid-connected inverters. According to instantaneous power theory, the active and reactive powers on the AC side per unit can be evaluated using  $P^+ = (v_d^+ i_d^+ + v_q^+ i_q^+)$  and  $Q^+ = (v_q^+ i_d^+ - v_d^+ i_q^+)$ , respectively. It is evident that the active and reactive power produced by the inverter is directly proportional to the direct- and quadrature-axis currents in the synchronous reference frame. Therefore, for the unity power factor, the reference reactive power ( $Q_{ref}$ ) can be kept zero. The PI controllers employ a fixed frequency free-running carrier signal, resulting in a well-bounded current total harmonic distortion (THD) with constant frequency inverter operation. This aids in instantaneous takeover of current control and has the advantage of eliminating the effect of DC side ripple on the phase currents of the inverter.

The third harmonic injection pulse width modulation (THIPWM) strategy is highly desirable for producing IGBT gate pulses, as it can help improve inverter performance under conditions of low DC link voltage. In THIPWM, a sine wave with a higher oscillating frequency is added to the modulating signals, leading to a reduction in the peak of the resulting modulating signal ( $m_a^+$ ,  $m_b^+$ , and  $m_c^+$ ). With  $v_a^+$ ,  $v_b^+$ , and  $v_c^+$  as the amplitude of the modulating signals of each phase and  $v_3$  as the amplitude of the third harmonic, the positive sequence inverter modulating signals with the embedded third harmonic are given in Equation (10).  $v_3$  is the amplitude of the third harmonic component, which is usually between 0.15 and 0.2 [57]. The resulting modulating signals are compared with a higher-frequency triangular wave generator to produce logic pulses for semiconductor devices.

$$\begin{aligned} m_a^+ &= v_a^+ \sin(\omega t) + v_3 \sin(3\omega t) \\ m_b^+ &= v_b^+ \sin\left(\omega t - \frac{2\pi}{3}\right) + v_3 \sin(3\omega t) \\ m_c^+ &= v_c^+ \sin\left(\omega t - \frac{4\pi}{3}\right) + v_3 \sin(3\omega t) \end{aligned} \quad (10)$$

The gains of the PI controller in such control approaches directly affect the stability of the inverter and the response to abrupt changes in operating conditions. Although there are many methods for designing these controller parameters, such as Ziegler–Nichols and pole placement, many cannot guarantee the best performance during grid disturbances and nonlinear loading conditions. For a more robust operation, metaheuristic-based algorithms inspired by nature have gained popularity in inverter control applications [58–61]. For this article, the widely used genetic algorithm (GA) optimization algorithm is implemented to iteratively determine the best set of PI values for the regulation of DC link voltage and inverter current. GA has been adopted because it guarantees convergence given certain constraint bands and does not require knowledge of derivatives, as the algorithm can process with input–output mapping [62]. For the inner current loop, the PI values depend on the transfer function of the LCL filter. With a cost function derived based on the integral time absolute error (ITAE), a population size of 10, and lower/upper limits for PI values as [0 500], 52 generations are sufficient for convergence to optimal values. Similarly, the DC link transfer function established in [63] is utilized for optimization with the established parameters. In this case, a total of 107 generations are required to achieve convergence

to the optimum values. Thus, the PI values of the best fit for the current controller are determined as  $K_p = 0.60$  and  $K_i = 243$ , while for the DC link PI controller,  $K_p = 1.35$  and  $K_i = 395$ . All controllers are equipped with anti-windups to reduce the effect of saturation and prevent large overshoots or oscillations.

#### 4. Reinforcement Learning Agent Design for Grid Fault Ride-Through

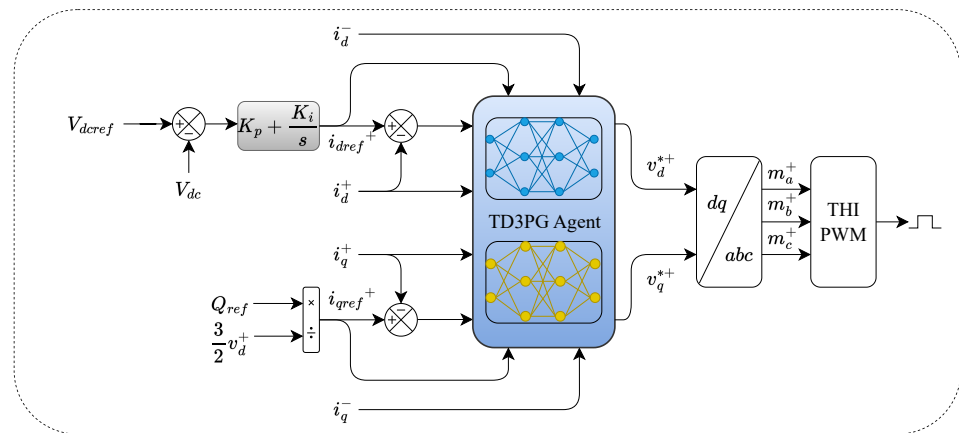
Fault ride-through requirements state that all IBRs must remain connected to the grid for a certain duration and wait for clearance to confirm that the fault is temporary. From the inverter's perspective, low voltage is simply a dip/sag in the voltage profile at PCC causing a drop in the computed  $d$  and  $q$  axes components. Depending on the percentage of dip and duration, the currents injected from the inverter may increase considerably. During faults, conventional PI controllers suffer from poor performance and are unable to track the designated current references in the synchronous reference frames. Prolonged symmetric or asymmetric faults can cause DC link spikes, overcurrents, and deviations of active and reactive powers. The dynamic response of the feedforward decoupled control approach also deteriorates during fluctuations in grid voltages due to the presence of voltage terms of the  $d$  and  $q$  axes in cross-coupling.

Although controller gains can be optimized, obtaining knowledge of the grid condition and incorporating it into the process is difficult due to the grid's dynamic nature. Additionally, PI controllers are insensitive to changes in grid operating conditions and pose a difficulty in maintaining inverter stability during abnormal grid conditions due to predefined saturation limits. Due to the presence of the grid voltage in the cross-coupling in the inner control loop, the PI controller tends to increase the inverter modulating voltages, causing an increase in the supplied currents. Prolonged divergence of the inverter currents results in the output of the PI controller reaching maximum thresholds, and thus the inverter goes into overmodulation mode.

Most other types of controller, such as proportional resonant, deadbeat, and fractional-order PI controllers, are inherently designed to provide a steady-state response and may not respond rapidly to sudden variations during grid faults. Nonlinearities are often introduced in the system due to disturbances in grid voltage and frequency and cause responses that standard controllers struggle to handle effectively. Other advanced control techniques, such as sliding mode and model predictive control, also encounter issues during grid transients due to the chattering phenomenon, sensitivity to uncertain parameter variation, tuning complexities, and dependency on accurate system models. As such, a twin delayed deep-determined policy gradient (TD3PG) current controller is formulated and trained to provide effective current control of inverter-based resources under various severities of grid disturbances. The improved voltage-oriented control based on reinforcement learning for effective fault management of grid-connected inverters is depicted in Figure 4.

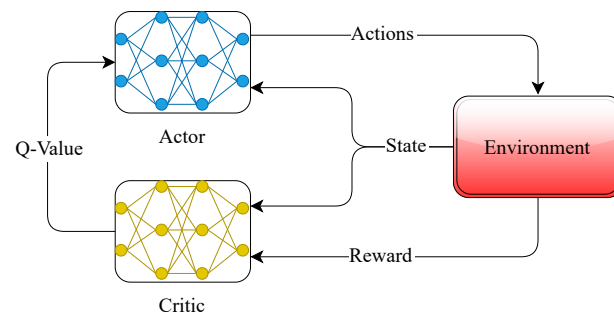
Reinforcement learning is an effective tool that can greatly improve the design process of power converter controllers. It provides various benefits, particularly in intricate and dynamic environments where conventional control methods may face difficulties. RL algorithms are capable of continually learning and adapting to changing operational conditions, such as load variations or fluctuations in input voltage. This is particularly advantageous in the control of grid-feeding inverters, where the power network can be highly unstable. Inverters frequently demonstrate non-linear behavior, especially under fault conditions. RL agents are anticipated to perform well in handling such non-linearity without requiring explicit modeling, enabling more robust and flexible control mechanisms. Traditional control techniques typically depend on precise mathematical models of the inverter and its operational context. This model-free approach eliminates the need for an explicit system model. Instead, it derives the control strategy directly through interaction

with the environment, making it ideal for systems where precise models are hard to develop, enhancing fault resilience.



**Figure 4.** Twin delayed deep deterministic voltage-oriented control (TD3VOC) for inverter.

Twin delayed deep deterministic policy gradient is an advanced reinforcement learning algorithm that serves as an upgrade to the deep deterministic policy gradient algorithm. RL can be applied to control systems to train deep neural networks by interacting with an environment, enabling the achievement of tasks that are difficult to achieve using traditional control methods. The actor–critic agent is widely used in these applications, where the actor neural network determines the actions to be taken and the critic network assesses the effectiveness of those actions based on the rewards received from the environment (Figure 5).



**Figure 5.** General reinforcement learning workflow.

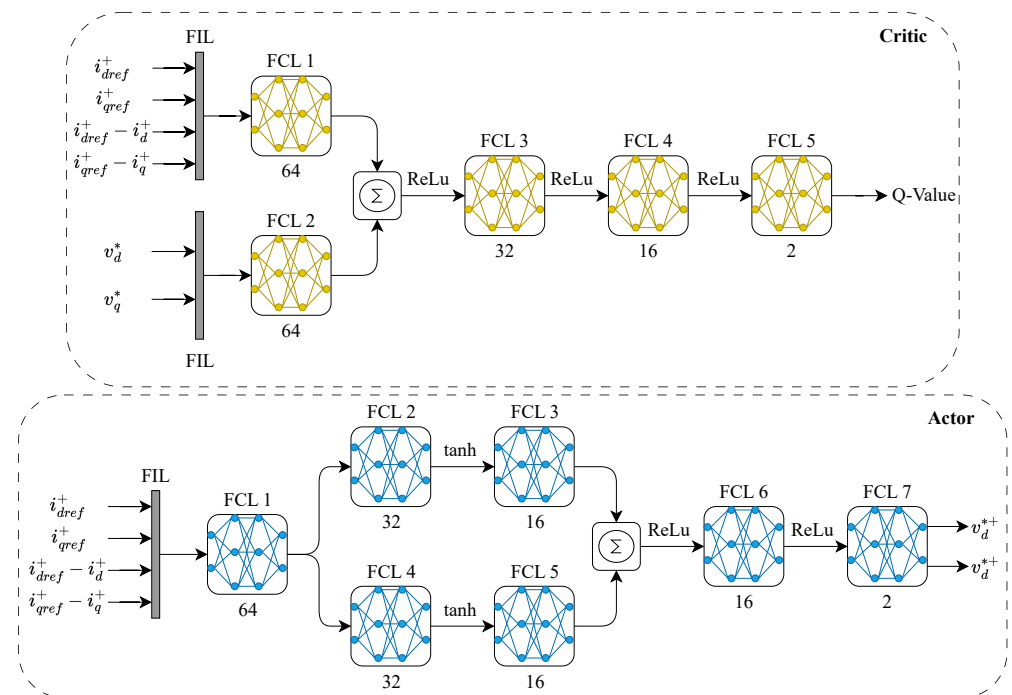
Markov decision process (MDP) is a mathematical modeling tool that is used to address decision-making problems involving an RL agent and its interaction with the environment. It provides a framework for modeling sequential decision processes or a continuous action space in uncertain states. Optimizing an agent involves finding the best policy that maximizes the expected cumulative rewards for the agent using a suitable algorithm such as Q-learning or policy gradient. In MDP, states ( $s$ ) represent all the information an agent receives from the environment, actions ( $a$ ) are the array of numeric input (continuous or discrete) applied in the environment determined by the actor, and ( $r$ ) is the cumulative reward obtained for the applied action. During training, the actor maps the states and actions, eventually converging to a deterministic policy ( $\pi$ ) after numerous iterations [64].

#### 4.1. Deep Neural Networks

To denote the primary indicators for selecting various actions in the subspace,  $[i_{dref}^+, i_{qref}^+, (i_{dref}^+ - i_d^+), (i_{qref}^+ - i_q^+)]$  are relevant states that are chosen as observations for the agent,

while the modulation voltages of the inverter  $[v_d^{*+}, v_q^{*+}]$  are two actions required from the actor network to provide enhanced stability to the inverter. The critic network has two feature input layers (FILs): one for the states and the other for the corresponding actions. Information measured via sensors is passed through two fully connected layers (FCLs), and the output is summed together, forming a single feedforward path. The rectified linear unit (ReLU) activation function performs a threshold operation before the data is inputted to the next FCL, where all negative values are set to zero. The final FCL outputs the required Q-value at each time step.

In a similar manner, the states are fed to the actor network using a FIL and passed through an FCL with 64 outputs, which are given two separate but identical paths with two FCLs. The outputs are summed and propagated through two more FCLs, and the final modulation voltages are produced. In the actor network, the hyperbolic tangent activation function (tanh) is employed between FCL due to its high performance and naturally normalizes inputs in the range  $[-1, 1]$ , which act as saturation limits for direct and quadrature voltages. The deep neural networks of critics and actors are shown in Figure 6. The 'Adam' learning algorithm is used to train dual critic and actor networks, each with a total of 3100 and 5800 learnable parameters, respectively. In general, both networks follow a similar pattern in which the number of outputs in the FCL is progressively reduced up to the output layer. This prevents overfitting and ensures a more stable training convergence.



**Figure 6.** Critic and actor deep neural network architecture.

#### 4.2. Reward Function

It is necessary to use a reward function that will allow the agent to learn the best possible policy by encapsulating a function that minimizes the inverter current deviation from the nominal set points at each discrete time step during abrupt disturbances on the grid side. The reward function is derived on the premise of effectively controlling the direct- and quadrature-axes currents based on the observed states. It is a combination of the measured  $d$  and  $q$  current errors with respect to the reference, the integral square error, and the negative sequence currents of the inverter, which must be minimized with any imbalance induced in the grid voltage. The total episode reward is the sum of all instantaneous rewards at each time step until the execution is terminated, given as

$$\mathfrak{R} = \sum_{t=0}^T \left( \begin{matrix} M \left( \int (i_{dref}^+ - i_d^+)^2 dt + \int (i_{qref}^+ - i_q^+)^2 dt \right) + \\ N \left( (i_{dref}^+ - i_d^+)^2 + (i_{qref}^+ - i_q^+)^2 \right) + \\ O \left( (i_d^-)^2 + (i_q^-)^2 \right) \end{matrix} \right) \quad (11)$$

The reward function assigns different weights to different aspects of the reward, represented by the arbitrary gain values  $M$ ,  $N$ , and  $O$ . In this case,  $N$  and  $O$  are set to  $-1$  and  $M$  to  $-0.5$ . The weights are selected to guarantee quick convergence during training. The adjustable gains permit customization of the significance assigned to each part of the reward function. For example, the squared errors of the positive and negative sequence currents are assigned the same gains so that both errors are equally minimized as the injected inverter currents deviate during faults. In contrast, the cumulative error increase over time is assigned a lower weight to ensure that it does not affect the real-time inverter regulation under the usual grid conditions. In addition, an average reward is calculated based on the length of the score averaging window. To simulate the agent’s adaptation, a short-circuit fault is introduced between phases B and C at bus 633. The agent is expected to adjust its behavior based on the defined reward function to minimize current deviations in the inverter and improve the resilience of the inverter. All measurements are taken as per unit values.

### 4.3. TD3PG Training Process

The agent training process involves systematically updating the parameters of the actor–critic network. In TD3PG, two critic networks are present that help mitigate the issue of overestimation bias of the DDPG agent and also incorporate delayed updates of the target networks to stabilize convergence during training. The output of the critic networks, known as the Q-value, is an important component of the training process and acts as an indicator for the estimated cumulative reward given a set of observations and actions data influenced by certain policies. The complete agent training process is shown in Figure 7, and the general agent training pseudocode is presented as Algorithm 1.

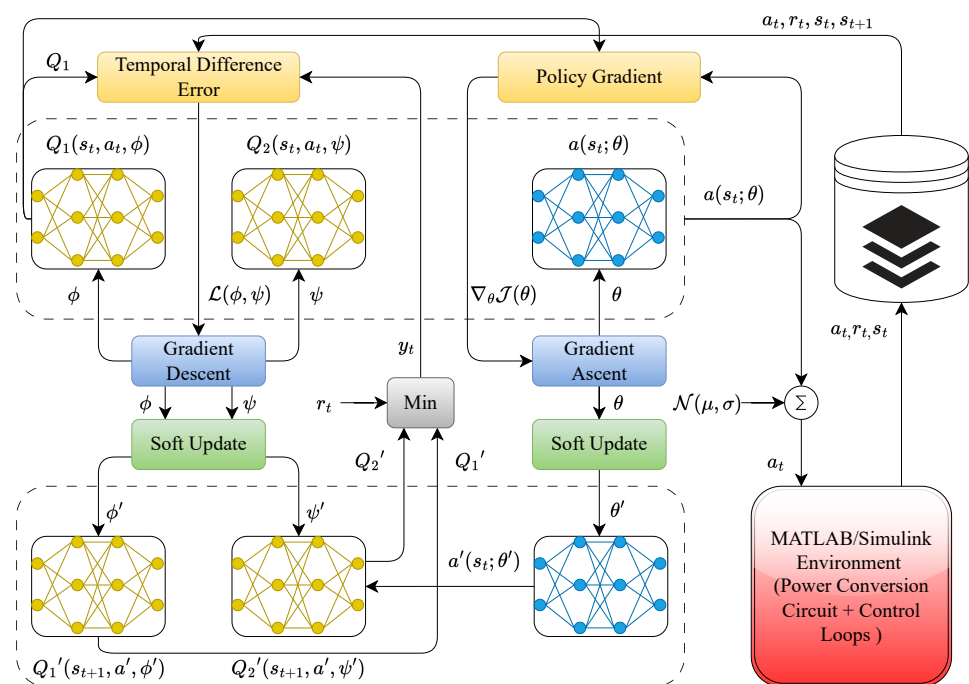


Figure 7. Training process of TD3PG RL agent incorporating data flow and environment interactions.

**Algorithm 1:** Pseudocode for training the TD3PG agent for inverter regulation

---

```

1 Setup the agent in the MATLAB/Simulink environment containing the
  grid-following inverter topology with appropriate grid fault models
2 Initialize critic networks  $Q_1, Q_2$  with random parameters  $\phi, \psi$ 
3 Initialize actor network  $\mu$  with random parameters  $\theta$ 
4 Initialize target networks  $Q'_1, Q'_2, \mu'$  with  $\phi'$  and  $\psi'$  and  $\theta'$ 
5 Set replay buffer  $\mathcal{R}$  and exploration noise  $\mathcal{N}$ 
6 Define hyperparameters: discount factor, learning rates, target update rate,
  regularisation, etc
7 for each episode do
8   Initialize environment and get initial state ( $s_t$ )
9   for each time step ( $t$ ) in the episode do
10    Select action + exploration noise:  $a'_t = a(s_t; \theta) + \mathcal{N}(\mu, \sigma)$ 
11    Execute action  $a$  in the environment and observe reward  $r$  and next state
       $s_{t+1}$ 
12    Store transition  $(s_t, a_t, r_t, s'_t)$  in replay buffer  $\mathcal{R}$ 
13    if time to update then
14      for each update step do
15        Sample a random mini-batch of  $\mathcal{M}$  transitions  $(s_i, a_i, r_i, s_{t+1})$  from
           $\mathcal{R}$ 
16        Compute target Q-values using target networks and clipped
          double Q-learning:
17        Compute target Q-values:
           $y_t = r_t + \gamma \min(Q'_1(s_{t+1}, a'(s_{t+1}; \theta'); \phi'), (Q'_2(s_{t+1}, a'(s_{t+1}; \theta'); \psi'))$ 
18        Update the critic networks by minimising the loss:
19         $\mathcal{L}(\phi) = \frac{1}{\mathcal{M}} \sum_{i=1}^n (Q_1(s_t, a_t, \phi) - y_t)^2$  and
           $\mathcal{L}(\psi) = \frac{1}{\mathcal{M}} \sum_{i=1}^n (Q_2(s_t, a_t, \psi) - y_t)^2$ 
20        Update critic parameters  $\phi$  and  $\psi$  using gradient descent
21         $\phi \rightarrow \phi + \alpha_{critic} \nabla_{\phi} \mathcal{L}(\phi, \psi)$  and  $\psi \rightarrow \psi + \alpha_{critic} \nabla_{\psi} \mathcal{L}(\phi, \psi)$ 
22        if update step policy delay == 0 then
23          Update actor network by maximizing the Q-value:
24           $\nabla J(\theta) \approx E_{s_t \sim \mathcal{R}} [\nabla_{\theta} Q_1(s_t, a_t; \phi) |_{a_t=a(s_t; \theta)} \nabla_{\theta} a(s_t; \theta)]$ 
25          Update actor parameters  $\theta$  using gradient ascent:
           $\theta \rightarrow \theta + \alpha_{actor} \nabla_{\theta} J(\theta)$ 
26          Update target networks:
27           $\phi' \rightarrow \phi'(1 - \tau) + \tau\phi, \psi' \rightarrow \psi'(1 - \tau) + \tau\psi$  and
           $\theta' \rightarrow \theta'(1 - \tau) + \tau\theta$ 
28        end
29      end
30    end
31    Update the state:  $s_{t+1} \rightarrow s_t$ 
32    if the episode is done then
33      Break the loop
34    end
35  end
36 end
37 return The trained actor network or the final optimised policy ( $\pi$ )

```

---



Denoting the parameters of the actor network as  $\theta$ , the weight of the first critic network ( $Q_1$ ) as  $\phi$ , the weights of the second critic network ( $Q_2$ ) as  $\psi$ , and the parameters of the corresponding target network as  $\theta'$ ,  $\phi'$ , and  $\psi'$ . The actor is expected to maximize the expected cumulative long-term reward by minimizing the negative Q-value function given as

$$\mathcal{J}(\theta) = -E_{s_t, a_t \sim \mathcal{R}}[Q_1(s_t, a_t; \phi)] \quad (12)$$

$E$  represents the expectation operator that is associated with the stochastic property of the transition probability and policy,  $\pi$ . Assuming  $a(s_t; \theta)$  is the output of the actor network given a state in a time interval ( $s_t$ ) and specific weights and  $\nabla_{\theta} Q_1(s_t, a_t; \phi)$  is the calculated gradient of the first critic network with respect to the action, the gradient ascent can be used to update the parameters of the actor network using the chain rule applied to the expected return objective according to the actor parameters [65].

$$\nabla \mathcal{J}(\theta) \approx E_{s_t \sim \mathbf{R}}[\nabla_{\theta} Q_1(s_t, a_t; \phi)|_{a_t=a(s_t; \theta)} \nabla_{\theta} a(s_t; \theta)] \quad (13)$$

Therefore, the new weights of the actor network can be calculated using the return objective and the actor learning rate ( $\alpha_{actor}$ ) as follows.

$$\theta \rightarrow \theta + \alpha_{actor} \nabla_{\theta} \mathcal{J}(\theta) \quad (14)$$

Furthermore, critic networks are trained to reduce the Q-value output of critic networks ( $Q_1$  and  $Q_2$ ) and target critic networks  $Q'_1$  and  $Q'_2$ . The targeted Q-value ( $y_t$ ) is obtained by evaluating the minimum of the two Q-values of the target networks, which also include the discounted reward ( $r_t$ ) and the predicted Q-values in the next state ( $s_{t+1}$ ), commonly calculated using the Bellman equation. Minimum operation ensures reduced overestimation bias. A discount factor ( $\gamma$ ), generally between 0 and 1, determines the weight given to future rewards. High  $\gamma$  values give more importance to compounded rewards, and a small value prioritizes immediate rewards, creating a trade-off between past and instant rewards.

$$y_t = r_t + \gamma \min(Q'_1(s_{t+1}, a'(s_{t+1}; \theta'); \phi'), (Q'_2(s_{t+1}, a'(s_{t+1}; \theta'); \psi'))) \quad (15)$$

The critic tends to reduce the difference between the target Q-value (from the Bellman equation) and the estimated Q-values, known as the temporal difference error (TDE), using a formulated loss function stated in Equation (8). Given  $M$  as the size of the mini-batch and the set of state-action data from the stored experiences in the memory at each time step, the critic networks obtain a better approximation for the true Q-value via minimization of this loss function.

$$\mathcal{L}(\phi, \psi) = \frac{1}{M} \sum_{i=1}^M (Q_1(s_t, a_t, (\phi, \psi)) - y_t)^2 \quad (16)$$

Eventually, both parameters of the critic network are updated using gradient descent considering a reasonable learning rate ( $\alpha_{critic}$ ).

$$\phi \rightarrow \phi + \alpha_{critic} \nabla_{\phi} \mathcal{L}(\phi, \psi) \quad (17)$$

$$\psi \rightarrow \psi + \alpha_{critic} \nabla_{\psi} \mathcal{L}(\phi, \psi) \quad (18)$$

The target networks ( $Q'_1$ ,  $Q'_2$ , and  $a'$ ) are updated through a soft update process that involves slowly evolving the network parameters with the current policy network parameters. A predefined target smooth factor ( $\tau$ ) is used to provide a more stable learning process.

$$\theta' \rightarrow \theta'(1 - \tau) + \tau\theta \quad (19)$$

$$\phi' \rightarrow \phi'(1 - \tau) + \tau\phi \quad (20)$$

$$\psi' \rightarrow \psi'(1 - \tau) + \tau\psi \quad (21)$$

During the training process, noise exploration data are often added to actions to allow for more robust policy development. Sampled Gaussian noise  $\mathcal{N}$  characterized by a normal distribution with reasonable mean ( $\mu$ ) and standard deviation ( $\sigma$ ) can be used to modify the selected actions before being applied.

$$a'_t = a(s_t; \theta) + \mathcal{N}(\mu, \sigma) \quad (22)$$

In this way, the agent can explore various regions of the action space via controlled randomness. It prevents it from obtaining a suboptimal policy, which is especially crucial in discovering effective continuous-deterministic actions. However, high values of  $\mu$  and  $\sigma$  lead to more exploration but may cause overfitting, resulting in less deterministic behavior. Therefore, it is important to maintain balance without compromising training or policy stability. In some cases, a standard deviation decay rate ( $\epsilon$ ) is used, allowing controlled reduction in the amount of noise that is added as agents take progressive steps and helping to shift from exploration to exploitation.

Furthermore, the TDP3PG agent offers several notable advantages that make it highly suitable for improving the effectiveness of such control tasks. It is specifically designed to handle problems with continuous action spaces, such as the regulation of power converters. This design choice eliminates the need for sampling from a probability distribution, as the agent can directly output continuous actions. Consequently, the complexity of decision making is reduced. Additionally, the use of the minimum operation in the Bellman equation combined with clipped double Q-learning critic networks ensures a more stable training process by mitigating issues related to overestimated Q-values and training divergence. Moreover, delayed updates of critic networks, in comparison to the actor, result in decorrelated target Q-values. This contributes to reducing variance and facilitates efficient training [41].

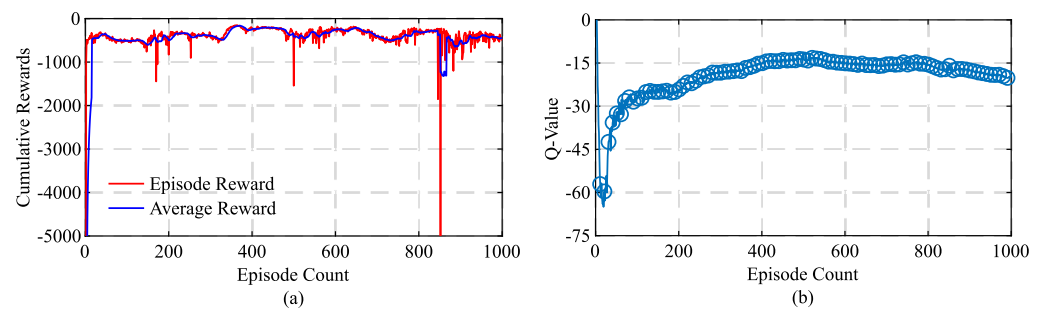
## 5. Results and Discussion

Training is performed using the high-performance shared computational cluster “Katana” with 64 GB of memory using a single CPU core. It takes approximately 22 h for the training to be completed, with 3000 steps per episode. All training hyperparameter specifications are provided in Table 3.

As a secondary training criterion, all agents that achieve an average of  $-500$  reward points or more are saved for further evaluation, resulting in a total of 833 being saved. Training progress with the episode reward obtained, the average reward, and the Q-value is depicted in Figure 8. Initially, the agent starts with a low reward ( $-15,000$ ), but slowly progresses to higher values as the actor neural network parameters become accustomed to the objective based on the updated weights. It is observed that after 840 episodes there is a sharp drop in the episode reward, after which a decline is noted in the training behavior. This can be an indication of overfitting, as the actor may already have reached its optimum policy. It may also be an indication that the actor is stuck in a local suboptimal policy that causes a temporary decrease in reward due to too much or insufficient exploration.

**Table 3.** TD3 agent training parameters.

Hyper Parameter	Symbol	Value
Sample time	$T_s$	100 $\mu$ s
Mini batch size	$\mathcal{M}$	64
Discount factor	$\gamma$	0.99
Maximum episodes	$\mathcal{E}_{max}$	1000
Actor learning rate	$\alpha_{actor}$	$1 \times 10^{-3}$
Critic learning rate	$\alpha_{critic}$	$1 \times 10^{-3}$
Target smooth factor	$\tau$	$1 \times 10^{-3}$
Policy update frequency	$\zeta$	1
L2 regularisation (actor)	$\rho_{actor}$	$1 \times 10^{-4}$
L2 regularisation (critic)	$\rho_{critic}$	$1 \times 10^{-4}$
Experience buffer length	$\mathcal{R}$	$1 \times 10^6$
Number of steps to look ahead	$n$	1
Score averaging window length	$\mathcal{K}$	15
Standard deviation	$\sigma$	0.1
Minimum standard deviation	$\sigma_{min}$	0.025
Standard deviation decay rate	$\sigma_{decay}$	$1 \times 10^{-6}$

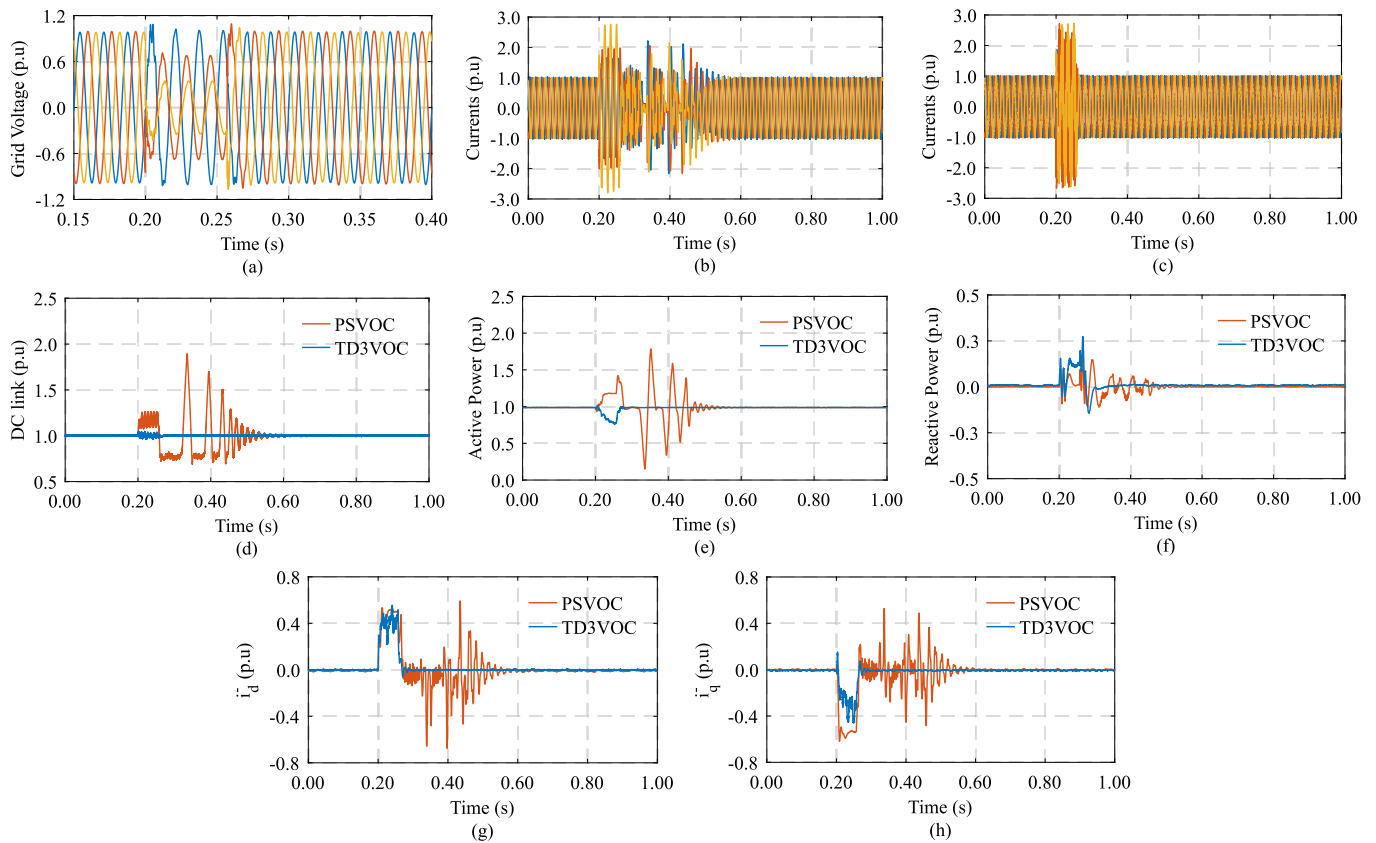
**Figure 8.** (a) Cumulative reward obtained and (b) Q-value progression at each successive iteration.

Since the actor–critic agent is deployed in a detailed time-varying environment containing feedback signals with inverter switching noise, observations and action pairs in the experience buffer length can also cause ambiguity, thus reducing training effectiveness and reward accumulation. If this pattern continues, the final agent at the end of maximum episodes may not perform well even if the learning rate is slow with a small regularization parameter. Since the approach deals with inverter current control at the primary level, to select the finest possible policy, the agent that attains the highest average reward is chosen as the best contender for controlling the inverter during faults. Using this criterion, the agent in episode 364 is adopted, which attains the highest reward of  $-151$  with a Q-value of  $-16.70$ .

Under all test conditions, the DC link reference is 1 p.u. while the reactive power reference is kept at 0 p.u. It should be noted that the TD3PG agent is trained only for a single asymmetrical fault between phases B and C across a  $0.25 \Omega$  resistor. However, its effectiveness in mitigating other types of symmetrical and asymmetrical faults through transfer learning has been verified. Using leverage of the pre-trained model significantly reduces the training time, especially for various types of grid faults. Training the agent for only one fault also allows the actor neural network to learn more generalized actions and leads to better accuracy. In addition, applying transfer learning reduces the need for computational resources, making the approach more efficient. The performance of the positive sequence voltage-oriented control and the designed twin delayed deep deterministic policy gradient voltage-oriented control is evaluated under both asymmetric and symmetric grid faults based on the adaptation to faults and operation restoration time.

### 5.1. Asymmetrical Fault: Short Circuit Between Phases B and C

In this test case, a phase-to-phase fault is induced between phases B and C (Figure 9a) over a fault resistance of 75 mΩ between time 0.20 and 0.25 s on bus 633. In particular, the trained agent has significantly reduced the post-fault recovery time of the inverter currents with immediate settling after fault clearance. With PSVOC, the current waveform settles at around 0.60 s. However, the proposed control approach takes only 0.06 s for the current to return to nominal values, and the waveform is only distorted for the duration of the fault without any significant increase in peak values, as shown in Figure 9b,c. This also helps to maintain the integrity of the grid voltages, as prolonged injection of distorted currents adds unwanted harmonics and may cause the inverter to isolate from the grid.



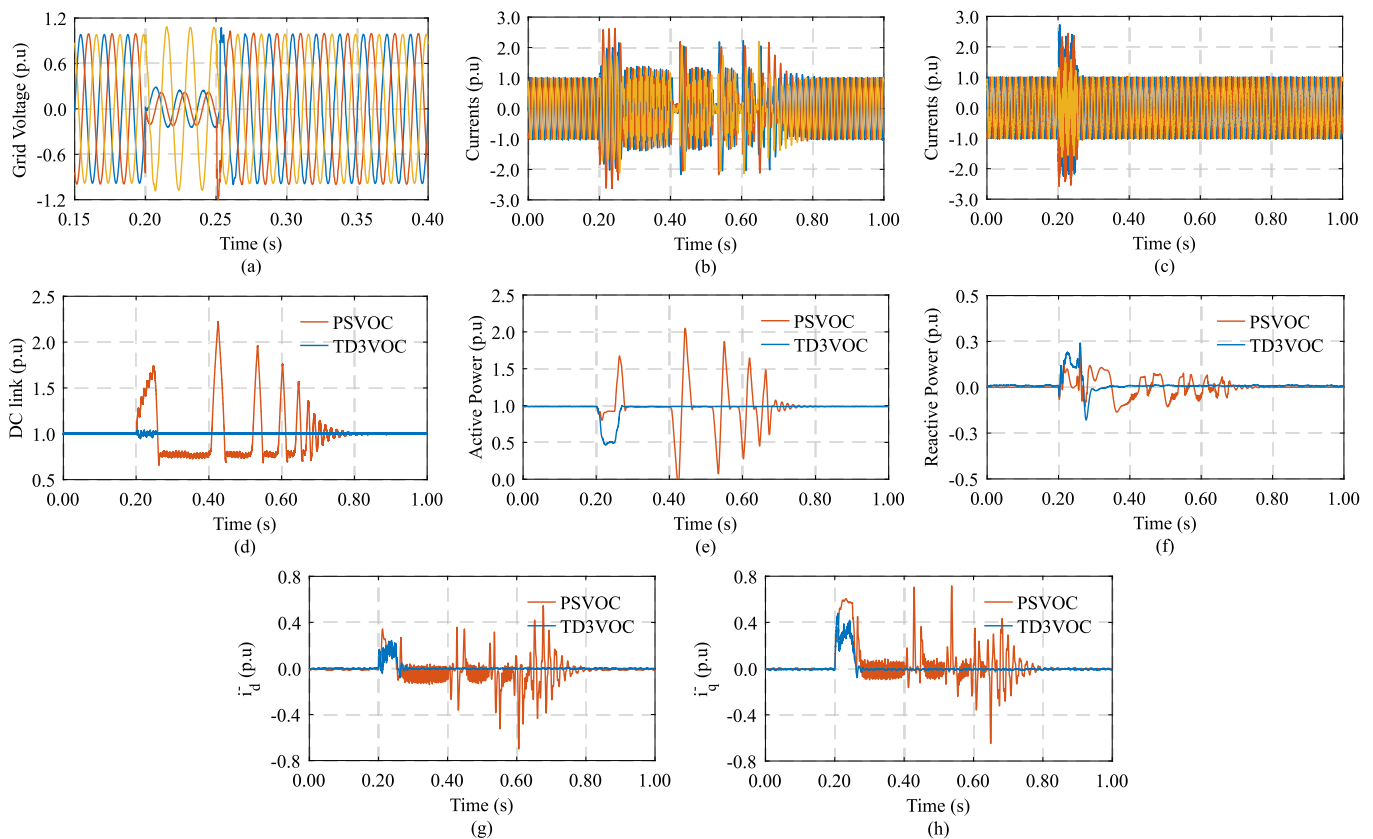
**Figure 9.** (a) Grid voltages, three-phase current waveform with (b) PSVOC and (c) TD3VOC, (d) DC link voltage, (e) active power, (f) reactive power, and negative sequence current on (g) direct and (h) quadrature axis response during asymmetrical short circuit fault between phases B and C.

As seen in Figure 9d, with PSVOC, an overshoot in the DC link voltage is observed during the fault. As the fault clears, an undershoot is seen as the maximum power point, and the DC-link PI controller attempts to recover from the disturbance. During the recovery period, the inverter consistently undergoes over- and undermodulation as a result of a destabilized DC link reaching a maximum voltage of 1.85 p.u. Following a period of 0.35 s after the fault has cleared, the nominal operating point is finally reached. However, the trained agent is far better at stabilizing the DC link voltage and shows immediate post-fault recovery. Consequent fluctuations in active and reactive powers (Figure 9e,f) are detected with large overshoots using the conventional controller, and although a slight drop in injected active power is observed with the TD3PG agent, the nominal operating point is quickly restored without any overshoot, showing a more robust response despite deviations during the fault. In addition, the TD3 agent performs well in minimizing negative sequence

direct and quadrature axis currents during the fault and immediate retention after the termination of the fault, as presented in Figure 9g,h.

### 5.2. Asymmetrical Fault: Phases A and B to Ground Fault

To further investigate the performance of the neural-based current controller, a double-phase-to-ground fault is applied at the PCC voltage for 50 ms (Figure 10a). The fault and ground resistance are assumed to be 50 mΩ. The three-phase currents experience oscillations and take about 0.55 s to settle after fault clearance when PI-based VOC is used. High current peaks are momentarily seen as the inverter attempts to recover from the fault, as shown in Figure 10b. This effect is completely eliminated with the RL agent. Although the current increases during the fault period, the inverter is quick to recover and attains stability without any post-fault transients, as depicted in Figure 10c. The superior response of the neural network is attributed to its fast input-to-output mapping. A similar effect is seen in the negative sequence currents on the direct and quadrature axes (Figure 10g,h). While the inverter is in overmodulation, the current oscillates and has several overshoots and undershoots during recovery with the conventional control approach. This problem is completely eliminated with the trained actor–critic neural networks ensuring a robust response.



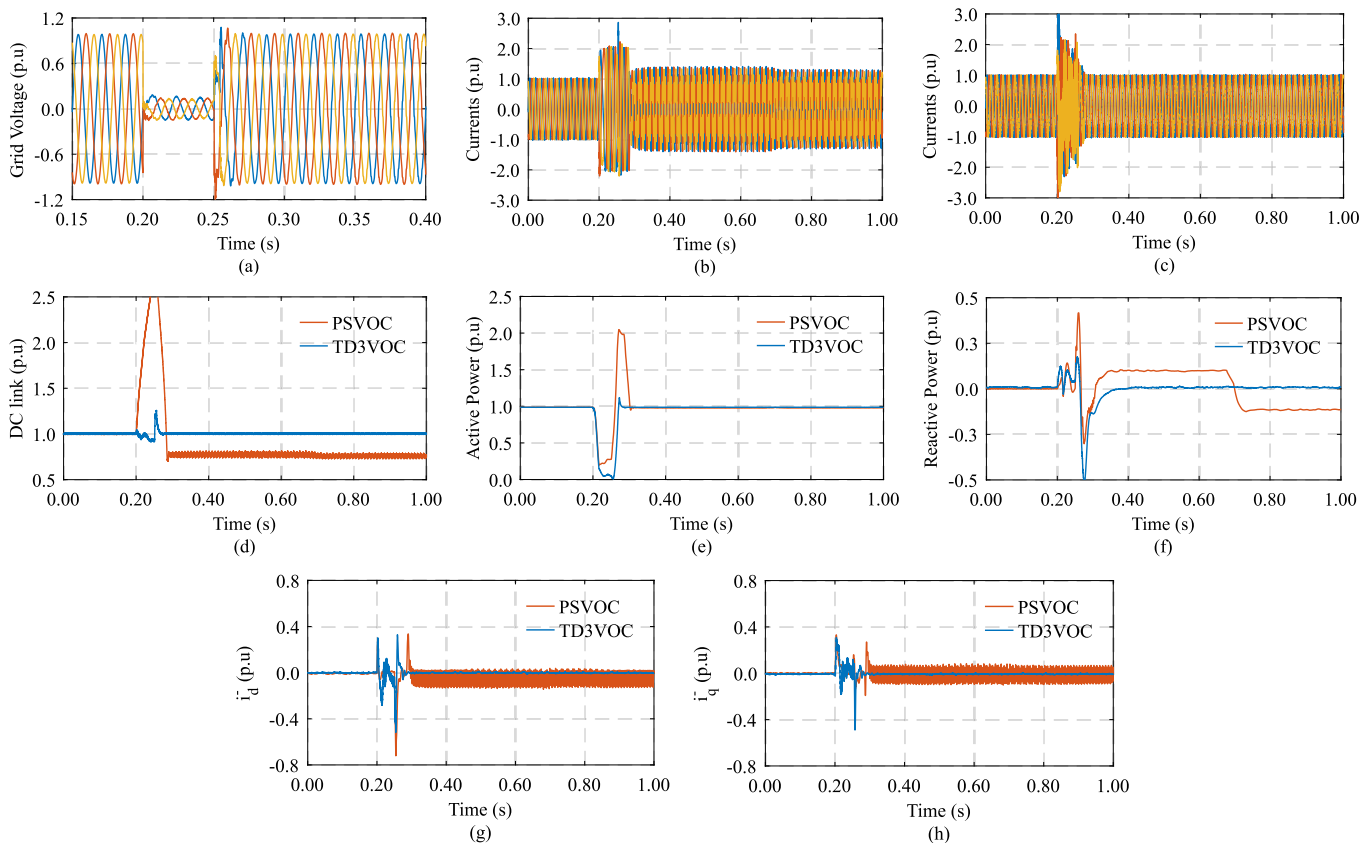
**Figure 10.** (a) Grid voltages, three-phase current waveform with (b) PSVOC and (c) TD3VOC, (d) DC link voltage, (e) active power, (f) reactive power, and negative sequence current on (g) direct and (h) quadrature axis response during phases A and B to ground fault.

During this period, a drastic response is observed for the DC link voltage compared with the previous case, reaching a peak of 2.25 p.u. It undergoes transients and settles at 0.80 s, as seen in Figure 10d when the inverter is controlled using PSVOC. With the TD3VOC, during the fault period, the voltage experiences slight oscillations but does not diverge from the reference value. When using cascaded PI controllers, inverter stability

depends on the accurate functioning of each loop. As such, a disturbance may cause the entire system to respond in an unpredictable manner, even if the controllers are properly tuned. This scenario is also seen in the active and reactive power response in Figure 10e,f, where the active and reactive powers diverge considerably due to the unstable DC link. Nevertheless, with the TD3 agent, immediate recovery is observed, as nominal operating points are restored without transients.

### 5.3. Symmetrical Fault: Phases A, B, and C Short Circuit Fault

A severe symmetric three-phase fault is applied between phases A, B, and C on bus 633 across a fault resistance of 35 m $\Omega$  (Figure 11a). The conventional PSVOC approach does not restore the inverter current waveform shown in Figure 11b due to the reaching of controller saturation limits. As a result, high oscillations and distortions in the currents are observed, even though they remain in phase with the grid voltage. Prolonged injection of highly distorted currents would eventually trigger protection devices, causing the inverter to disconnect from the grid, disrupting power transfer. In contrast, the RL agent can restore the inverter currents within 0.05 s after the fault is cleared, as shown in Figure 11c. The voltage diverges and reaches more than 2.5 times the nominal operating range, which will eventually activate the circuit tripping mechanisms. In contrast, the difference is significant when TD3VOC is used instead, leading to a much smoother response of the DC link voltage without much divergence (Figure 11d).



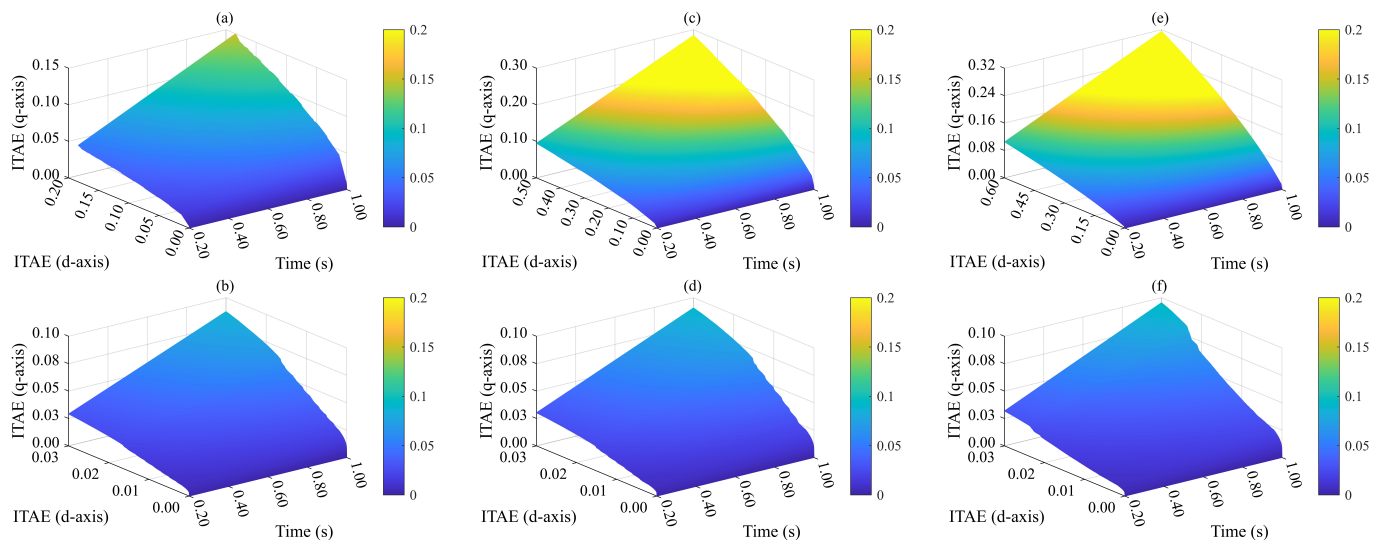
**Figure 11.** (a) Grid voltages, three-phase current waveform with (b) PSVOC and (c) TD3VOC, (d) DC link voltage, (e) active power, (f) reactive power, and negative sequence current on (g) direct and (h) quadrature axis response to a symmetrical short-circuit fault between phases A, B, and C.

Likewise, active power experiences disruptions during a fault, and a large overshoot occurs as the inverter tries to restore power post-fault. This recovery time is reduced with the proposed TD3PG current controller, as shown in Figure 11e,f. Although the active

power is restored to nominal values, it is at the cost of a lower DC link voltage and an overmodulating inverter. Due to not recovering from saturation and continuous operation with a lower DC link voltage, the negative sequence currents suffer from high oscillations after fault clearance when the inverter is controlled with PSVOC. However, the TD3VOC scheme performs quite well in restoring and reducing negative sequence currents even after severe disturbances (Figure 11g,h). It should be noted that although the agent was not specifically trained for symmetrical faults, it still demonstrates a good response to such severe disturbances, thus proving its adaptability through transfer learning.

#### 5.4. Error Analysis

To accurately assess the performance of the proposed fault-tolerant agent-based control method, calculating the integral-time absolute error (ITAE) is crucial. The ITAE metric multiplies the absolute error by time, thereby giving greater importance to errors that occur later in the system compared to those at the start. Minimizing ITAE is important to achieve optimal performance during faults in inverter control applications. The error data for the  $d$  and  $q$  axes currents obtained from the evaluation models have been analyzed under the three types of faults, and the results are shown in Figure 12. With the conventional controller, ITAE on the  $d$  and  $q$  axes increases significantly as faults become more severe. During the phase-to-phase asymmetric fault between phases B and C, a maximum of 0.20 and 0.15 is reached for the  $d$  and  $q$  axes, respectively. However, this increases to 0.32 and 0.60 after a symmetric fault is induced in the system. With the proposed TD3PG current controller, it is observed that during all test cases, the error remains below 0.10 and 0.03 for the  $d$  and  $q$  axes, respectively. This metric authenticates the viability of the formulated agent in providing a reasonably better inverter response compared with the standard control approach.



**Figure 12.** Integral-time error visualization for  $d$  and  $q$  axes currents using (a,c,e) PSVOC and (b,d,f) TD3VOC for (a,b) asymmetric fault between phases B and C, (c,d) asymmetric phases A and B to ground fault, and (e,f) symmetric short-circuit fault between phases A, B and C.

#### 5.5. Summary

In summary, the key results obtained in this research for the existing voltage-oriented control scheme and the redesigned TD3PG voltage-oriented control architecture are stated in Tables 4 and 5. Key performance indicators were selected as the peak DC link voltage magnitude during the fault period with a nominal value of 1 p.u., the post-fault recovery time after fault clearance (PFRT), and ITAE on the  $d$  and  $q$  axes. In every scenario, the

suggested TD3VOC exhibits a robust fault-ride-through capability in all evaluations by the use of transfer learning. It exhibits minimized DC link voltage peaks during faults and reduced post-fault system recovery time. A significant attribute that has directly affected the performance of the TD3PG agent is its structural design. Bellman’s minimum equation paired with clipped double Q-learning mitigates the overestimation bias found in other continuous-time agents. The algorithm’s lower sensitivity to hyperparameter selection is due to postponed policy updates and target smoothing. These sophisticated features, along with its adaptability to unpredictable inverter operating conditions, make TD3PG an exceptionally effective option to improve the control of grid-feeding photovoltaic systems during grid disruptions. In addition, the per-unit system is beneficial in minimizing the alterations needed to apply this method to larger power systems. It ensures that the neural network generates modulation voltages in a consistent manner, regardless of the differing power ratings in various systems.

**Table 4.** Performance of conventional positive-sequence voltage-oriented control.

Type of Fault	Voltage (p.u.)	PFRT (s)	ITAE ( <i>d</i> -axis)	ITAE ( <i>q</i> -axis)
Phase-to-phase (B to C)	1.85	0.35	0.20	0.15
Phase-to-ground (A, B to G)	2.25	0.55	0.50	0.28
Phases A, B and C short circuit	>2.50	>0.80	0.60	0.32

**Table 5.** Performance of proposed twin delayed deep deterministic voltage-oriented control.

Type of Fault	Voltage (p.u.)	PFRT (s)	ITAE ( <i>d</i> -axis)	ITAE ( <i>q</i> -axis)
Phase-to-phase (B to C)	1.05	0.01	0.03	0.09
Phase-to-ground (A, B to G)	1.05	0.01	0.03	0.09
Phases A, B and C short circuit	1.25	0.02	0.03	0.09

## 6. Limitations and Future Works

Despite the impressive fault-resistant capabilities of the TD3VOC, it should be noted that achieving the optimal policy for effective real-time inverter control still requires approximately 20–25 h of episodic iteration. The duration of training is predominantly influenced by the processor power, the number of CPU cores, and the available memory. In addition, the environment plays a significant role in determining the required computational power. In this study, we developed a highly detailed grid-interfaced photovoltaic inverter environment in MATLAB/Simulink® with which the RL agent interacted. An average model could also be used to train the agent; however, it is uncertain whether the same agent would perform well when faced with a more elaborate model that includes switching noise and variability in feedback signals.

The results presented in this article are obtained using the twin delayed deep-deterministic policy gradient agent; however, there is the possibility of exploring other continuous-time agents such as soft actor–critic, trust region policy optimization, model-based policy optimization, etc. A comparative study can be conducted to determine the best candidate for developing data-driven controllers based on training time, number of hyperparameter selections, and adaptation of various grid conditions. It will also be interesting to compare the formulated fault ride-through approach with other established methods such as sliding mode and model reference adaptive controllers.

The proposed agent-based current controller can be trained and implemented for various other grid-feeding inverter control applications in weak microgrids that require a robust fault ride-through mechanism, such as wind energy conversion systems, grid-forming battery energy storage systems, fuel cell systems, and other inverter control schemes featuring an inner current loop. This direct design method can also be applied



to develop robust data-driven controllers for electrical machines that are controlled using cascaded field-oriented control or the direct torque control scheme to provide better torque handling capabilities, smoother acceleration, and increased energy efficiency.

In addition, the experimental validation of such advanced control techniques is still a subject of particular interest to numerous researchers today. Implementing RL agents for optimal regulation of the inverter during dynamic grid conditions will require an embedded system that involves a combination of advanced hardware and software features. It will be necessary to utilize high-performance microcontroller units (MCUs), digital signal processors (DSPs), or systems on chips (SoCs) capable of handling both RL algorithms and real-time inverter control with appropriate communication interfaces. It is suggested to use the Xilinx field-programmable gate array board, OPAL-RT technology, or DSPACE Microlabbox due to compatibility with MATLAB/Simulink<sup>®</sup>, customized software options, and expandable hardware. There is immense potential for RL-based current controllers to enhance the performance and reliability of inverter systems in various domains, particularly in the context of developing modern and intelligent energy systems.

## 7. Conclusions

In this paper, a robust data-driven grid fault ride-through approach is proposed for photovoltaic inverters. A twin delayed deep-deterministic policy gradient agent is embedded in the conventional voltage-oriented control scheme as an upgrade to the PI controllers of the  $d$  and  $q$  axes and cross-coupling terms. Classical current controllers are designed assuming a balanced and stable grid, which is not always the case due to the presence of unpredictable fault conditions, and also suffer from dependence on controller gains and predefined saturation limits. To address these issues, the reinforcement learning workflow is adopted to train actor–critic deep neural networks to improve the inverter current response under various grid conditions. A dual stage inverter environment with a suitable reward function was set up to guide the training process. Initially, the agent is trained for a single asymmetric fault between phases B and C. However, it is possible to use the same agent to improve inverter stability across a wide spectrum of grid faults via transfer learning. When using the PSVOC method, the DC link overshoots to 1.85 p.u. and 2.25 p.u. while a current post-fault restoration time of 0.35 s and 0.55 s for phase-to-phase and phase-to-ground faults, respectively. With the developed TD3VOC technique, voltage overshoot is limited to 1.05 p.u. with a current recovery time of 0.01 after fault clearance in both cases. For a severe symmetric fault, the conventional control approach cannot restore inverter operation with a voltage overshoot of more than 2.50 p.u. with integral-time absolute errors of 0.60 and 0.32 on the  $d$  and  $q$  axes, respectively. The proposed agent-based control method limits the voltage overshoot to 1.25 p.u. and achieves a cumulative error of 0.03 and 0.09 on the  $d$  and  $q$  axes, respectively. The proposed control approach is implemented using per-unit measurements to allow for scalability and can be extended to grid-connected wind turbine generators, grid-forming inverters, and other inverter-based resources that require robust fault ride through operation.

**Author Contributions:** Conceptualization, S.S.C.; methodology, S.S.C.; software, S.S.C.; validation, S.S.C.; formal analysis, S.S.C.; investigation, S.S.C.; resources, B.H.; writing—original draft preparation, S.S.C. and B.H.; writing—review and editing, S.S.C. and B.H.; supervision, B.H. and M.C.; project administration, B.H. and M.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding authors.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## List of Acronyms

MG	Microgrid
IBR	Inverter Based Resources
INC	Incremental Conductance
FRT	Fault-Ride Through
MDP	Markov Decision Process
PFRT	Post-fault Recovery Time
TD3PG	Twin Delay Deep Deterministic Policy Gradient
PSVOC	Positive Sequence Voltage-Oriented Control
TD3VOC	Twin Delay Deep Deterministic Voltage-Oriented Control
PV	Photovoltaic
DER	Distributed Energy Resource
QSG	Quadrature Signal Generator
PCC	Point of Common Coupling
DRL	Deep Reinforcement Learning
ITAE	Integral-time Absolute Error
DDPG	Deep Deterministic Policy Gradient
DSOGI	Double Second Order Generalized Integrator
THIPWM	Third Harmonic Injection Pulse Width Modulation

## References

- Panda, A.; Dauda, A.K.; Chua, H.; Tan, R.R.; Aviso, K.B. Recent advances in the integration of renewable energy sources and storage facilities with hybrid power systems. *Clean. Eng. Technol.* **2023**, *12*, 100598. [\[CrossRef\]](#)
- Shuai, Z.; Sun, Y.; Shen, Z.J.; Tian, W.; Tu, C.; Li, Y.; Yin, X. Microgrid stability: Classification and a review. *Renew. Sustain. Energy Rev.* **2016**, *58*, 167–179. [\[CrossRef\]](#)
- Rajesh, K.S.; Dash, S.S.; Rajagopal, R.; Sridhar, R. A review on control of ac microgrid. *Renew. Sustain. Energy Rev.* **2017**, *71*, 814–819. [\[CrossRef\]](#)
- Ahmed, M.; Meegahapola, L.; Vahidnia, A.; Datta, M. Stability and Control Aspects of Microgrid Architectures-A Comprehensive Review. *IEEE Access* **2020**, *8*, 144730–144766. [\[CrossRef\]](#)
- Nawaz, F.; Pashajavid, E.; Fan, Y.; Batool, M. A Comprehensive Review of the State-of-the-Art of Secondary Control Strategies for Microgrids. *IEEE Access* **2023**. [\[CrossRef\]](#)
- Qais, M.H.; Hasanien, H.M.; Alghuwainem, S. Whale optimization algorithm-based Sugeno fuzzy logic controller for fault ride-through improvement of grid-connected variable speed wind generators. *Eng. Appl. Artif. Intell.* **2020**, *87*, 103328. [\[CrossRef\]](#)
- IEEE Std 2800-2022; IEEE Standard for Interconnection and Interoperability of Inverter-Based Resources (IBRs) Interconnecting with Associated Transmission Electric Power Systems. IEEE: Piscataway, NJ, USA, 2012; pp. 1–180.
- Singh, M.; Basak, P. Q and frequency component-based fault and nature detection for hybrid microgrid. *Sustain. Energy Grids Netw.* **2021**, *28*, 100552. [\[CrossRef\]](#)
- Borrell, A.; Velasco, M.; Miret, J.; Camacho, A.; Marti, P.; Castilla, M. Collaborative Voltage Unbalance Elimination in Grid-Connected AC Microgrids with Grid-Feeding Inverters. *IEEE Trans. Power Electron.* **2021**, *36*, 7189–7201. [\[CrossRef\]](#)
- Gomes, H.M.; Carralero, L.L.; Suárez, J.H.; Tahim, A.N.; Pinheiro, J.R.; Costa, F.F. A Grid Impedance Estimation Method Robust Against Grid Voltages Unbalance and Harmonic Distortions. *J. Control Autom. Electr. Syst.* **2023**, *34*, 289–301. [\[CrossRef\]](#)
- Ferreira, R.V.; Silva, S.M.; Brandao, D.I. Positive–Negative Sequence Synchronverter for Unbalanced Voltage in AC Grids. *J. Control Autom. Electr. Syst.* **2021**, *32*, 711–720. [\[CrossRef\]](#)
- Easley, M.; Jain, S.; Shadmand, M.; Abu-Rub, H. Autonomous Model Predictive Controlled Smart Inverter with Proactive Grid Fault Ride-Through Capability. *IEEE Trans. Energy Convers.* **2020**, *35*, 1825–1836. [\[CrossRef\]](#)
- Nguyen, T.K.; Chang, G.W.; Liu, Y.J.; Li, G.Y. Improving Voltage Ride-Through Capability of Grid-Tied Microgrid With Harmonics Mitigation. *IEEE Trans. Power Deliv.* **2023**, *38*, 738–741. [\[CrossRef\]](#)
- Liu, X.; Chen, X.; Shahidepour, M.; Li, C.; Wu, Q.; Wu, Y.; Wen, J. Active Fault Current Limitation for Low-Voltage Ride-Through of Networked Microgrids. *IEEE Trans. Power Deliv.* **2022**, *37*, 980–992. [\[CrossRef\]](#)
- Bahramian-Habil, H.; Abyaneh, H.A.; Gharehpetian, G.B. Improving LVRT capability of microgrid by using bridge-type fault current limiter. *Electr. Power Syst. Res.* **2021**, *191*, 106872. [\[CrossRef\]](#)

16. El-Naggar, M.F.; Mosaad, M.I.; Hasanien, H.M.; AbdulFattah, T.A.; Bendary, A.F. Elephant herding algorithm-based optimal PI controller for LVRT enhancement of wind energy conversion systems. *Ain Shams Eng. J.* **2021**, *12*, 599–608. [[CrossRef](#)]
17. Xiao, Q.; Jin, Y.; Chen, L.; Yu, X.; Jia, H.; Liu, H.; Teodorescu, R.; Blaabjerg, F. A Simple Operation Approach for Modular Multilevel Converter under Grid Voltage Swell in Medium-Voltage Microgrids. *IEEE Access* **2019**, *7*, 147280–147291. [[CrossRef](#)]
18. Brandao, D.I.; Mendes, F.E.; Ferreira, R.V.; Silva, S.M.; Pires, I.A. Active and reactive power injection strategies for three-phase four-wire inverters during symmetrical/asymmetrical voltage sags. *IEEE Trans. Ind. Appl.* **2019**, *55*, 2347–2355. [[CrossRef](#)]
19. Wang, H.; Zhang, Q.; Wu, D.; Zhang, J. Advanced Current-Droop Control for Storage Converters for Fault Ride-Through Enhancement. *IEEE J. Emerg. Sel. Top. Power Electron.* **2020**, *8*, 2461–2474. [[CrossRef](#)]
20. Djilali, L.; Sanchez, E.N.; Ornelas-Tellez, F.; Avalos, A.; Belkheiri, M. Improving Microgrid Low-Voltage Ride-Through Capacity Using Neural Control. *IEEE Syst. J.* **2020**, *14*, 2825–2836. [[CrossRef](#)]
21. Firouzi, M.; Nasiri, M.; Mobayen, S.; Gharehpetian, G.B. Sliding Mode Controller-Based BFCL for Fault Ride-Through Performance Enhancement of DFIG-Based Wind Turbines. *Complexity* **2020**, *2020*, 1259539. [[CrossRef](#)]
22. Aboelsaud, R.; Ibrahim, A.; Aleksandrov, I.V.; Ali, Z.M. Model predictive control algorithm for fault ride-through of stand-alone microgrid inverter. *Int. J. Electr. Power Energy Syst.* **2022**, *135*, 107485. [[CrossRef](#)]
23. Musarrat, M.N.; Fekih, A.; Rahman, M.A.; Islam, M.R.; Muttaqi, K.M. An Event Triggered Sliding Mode Control-Based Fault Ride-Through Scheme to Improve the Transient Stability of Wind Energy Systems. *IEEE Trans. Ind. Appl.* **2024**, *60*, 876–886. [[CrossRef](#)]
24. Jia, H.; Hu, Y.; Jiang, T.; Li, G.; Tightiz, L.; Yoo, J. A Review on a Data-Driven Microgrid Management System Integrating an Active Distribution Network: Challenges, Issues, and New Trends. *Energies* **2022**, *15*, 8739. [[CrossRef](#)]
25. Cao, Y.; Chandrasekar, A.; Radhika, T.; Vijayakumar, V. Input-to-state stability of stochastic Markovian jump genetic regulatory networks. *Math. Comput. Simul.* **2024**, *222*, 174–187. [[CrossRef](#)]
26. Radhika, T.; Chandrasekar, A.; Vijayakumar, V.; Zhu, Q. Analysis of Markovian Jump Stochastic Cohen–Grossberg BAM Neural Networks with Time Delays for Exponential Input-to-State Stability. *Neural Process. Lett.* **2023**, *55*, 11055–11072. [[CrossRef](#)]
27. Tamil Thendral, M.; Ganesh Babu, T.R.; Chandrasekar, A.; Cao, Y. Synchronization of Markovian jump neural networks for sampled data control systems with additive delay components: Analysis of image encryption technique. *Math. Methods Appl. Sci.* **2022**. [[CrossRef](#)]
28. Mamdouh Farghaly, H.; Abd El-Hafeez, T. A new feature selection method based on frequent and associated itemsets for text classification. *Concurr. Comput. Pract. Exp.* **2022**, *34*, e7258. [[CrossRef](#)]
29. Mamdouh Farghaly, H.; Abd El-Hafeez, T. A high-quality feature selection method based on frequent and correlated items for text classification. *Soft Comput.* **2023**, *27*, 11259–11274. [[CrossRef](#)]
30. Farghaly, H.M.; Ali, A.A.; El-Hafeez, T.A. Developing an Efficient Method for Automatic Threshold Detection Based on Hybrid Feature Selection Approach. *Adv. Intell. Syst. Comput.* **2020**, *1225*, 56–72. [[CrossRef](#)]
31. Mostafa, G.; Mahmoud, H.; Abd El-Hafeez, T.; ElAraby, M.E. Feature reduction for hepatocellular carcinoma prediction using machine learning algorithms. *J. Big Data* **2024**, *11*, 1–27. [[CrossRef](#)]
32. Eliwa, E.H.I.; El Koshiry, A.M.; Abd El-Hafeez, T.; Omar, A. Optimal Gasoline Price Predictions: Leveraging the ANFIS Regression Model. *Int. J. Intell. Syst.* **2024**, *2024*, 8462056. [[CrossRef](#)]
33. Li, S.; Oshnoei, A.; Blaabjerg, F.; Anvari-Moghaddam, A. Hierarchical Control for Microgrids: A Survey on Classical and Machine Learning-Based Methods. *Sustainability* **2023**, *15*, 8952. [[CrossRef](#)]
34. Bai, K.; Sindhu, V.; Haque, A. Fault Ride Through approach for Grid-Connected Photovoltaic System. *e-Prime-Adv. Electr. Eng. Electron. Energy* **2023**, *5*, 100232. [[CrossRef](#)]
35. Benbouzid, M.; Elghali, B.; El, A.; Bouzid, M.; Zerrougui, M.; El-Ebiary, A.H.; Attia, M.A.; Marei, M.I.; Sameh, M.A. An Integrated Seamless Control Strategy for Distributed Generators Based on a Deep Learning Artificial Neural Network. *Sustainability* **2022**, *14*, 13506. [[CrossRef](#)]
36. Lin, F.J.; Liao, J.C.; Chen, C.I.; Chen, P.R.; Zhang, Y.M. Voltage Restoration Control for Microgrid with Recurrent Wavelet Petri Fuzzy Neural Network. *IEEE Access* **2022**, *10*, 12510–12529. [[CrossRef](#)]
37. Conchas, R.F.; Sanchez, E.N.; Ricalde, L.J.; Alvarez, J.G.; Alanis, A.Y. Sensor fault-tolerant control for a doubly fed induction generator in a smart grid. *Eng. Appl. Artif. Intell.* **2023**, *117*, 105527. [[CrossRef](#)]
38. Khooban, M.H.; Gheisarnejad, M. A Novel Deep Reinforcement Learning Controller Based Type-II Fuzzy System: Frequency Regulation in Microgrids. *IEEE Trans. Emerg. Top. Comput. Intell.* **2021**, *5*, 689–699. [[CrossRef](#)]
39. Yan, R.; Wang, Y.; Xu, Y.; Dai, J. A Multiagent Quantum Deep Reinforcement Learning Method for Distributed Frequency Control of Islanded Microgrids. *IEEE Trans. Control. Netw. Syst.* **2022**, *9*, 1622–1632. [[CrossRef](#)]
40. He, X.; Ge, S.; Liu, H.; Xu, Z.; Mi, Y.; Wang, C. Frequency regulation of multi-microgrid with shared energy storage based on deep reinforcement learning. *Electr. Power Syst. Res.* **2023**, *214*, 108962. [[CrossRef](#)]
41. Egbomwan, O.E.; Liu, S.; Chaoui, H. Twin Delayed Deep Deterministic Policy Gradient (TD3) Based Virtual Inertia Control for Inverter-Interfacing DGs in Microgrids. *IEEE Syst. J.* **2023**, *17*, 2122–2132. [[CrossRef](#)]

42. Yang, M.; Wu, X.; Loveth, M.C. A Deep Reinforcement Learning Design for Virtual Synchronous Generators Accommodating Modular Multilevel Converters. *Appl. Sci.* **2023**, *13*, 5879. [[CrossRef](#)]
43. You, W.; Yang, G.; Chu, J.; Ju, C. Deep reinforcement learning-based proportional–integral control for dual-active-bridge converter. *Neural Comput. Appl.* **2023**, *35*, 17953–17966. [[CrossRef](#)]
44. Kosuru, R.; Liu, S.; Shi, W. Deep Reinforcement Learning for Stability Enhancement of a Variable Wind Speed DFIG System. *Actuators* **2022**, *11*, 203. [[CrossRef](#)]
45. Nicola, M.; Nicola, C.I.; Ionete, C.; Şendrescu, D.; Roman, M. Improved Performance for PMSM Sensorless Control Based on Robust-Type Controller, ESO-Type Observer, Multiple Neural Networks, and RL-TD3 Agent. *Sensors* **2023**, *23*, 5799. [[CrossRef](#)]
46. Wang, H.; Vukosavic, S.N.; Katic, V.; Nicola, M.; Nicola, C.I.; Selis, D. Improvement of PMSM Sensorless Control Based on Synergetic and Sliding Mode Controllers Using a Reinforcement Learning Deep Deterministic Policy Gradient Agent. *Energies* **2022**, *15*, 2208. [[CrossRef](#)]
47. Prasad, R.; Kumar, D.; Chand, S.; Fagiolini, A.; Mudaliar, H.; Benedetto, M.; Cirrincione, M. Enhancing Speed Loop PI Controllers with Adaptive Feed-forward Neural Networks: Application to Induction Motor Drives. In Proceedings of the 2022 International Conference on Electrical Machines and Systems (ICEMS 2022), Chiang Mai, Thailand, 29 November–2 December 2022. [[CrossRef](#)]
48. Hu, J.; Shan, Y.; Cheng, K.W.; Islam, S. Overview of Power Converter Control in Microgrids - Challenges, Advances, and Future Trends. *IEEE Trans. Power Electron.* **2022**, *37*, 9907–9922. [[CrossRef](#)]
49. Mohammadi, F.; Mohammadi-Ivatloo, B.; Gharehpetian, G.B.; Ali, M.H.; Wei, W.; Erdinc, O.; Shirkhani, M. Robust Control Strategies for Microgrids: A Review. *IEEE Syst. J.* **2022**, *16*, 2401–2412. [[CrossRef](#)]
50. Chand, S.; Kumar, R.; Prasad, R.; Cirrincione, M.; Raj, K. Open Circuit (OC) and Short Circuit (SC) IGBT Switch Fault Detection in Three-Phase Standalone Photovoltaic Inverters Using Shallow Neural Networks. In Proceedings of the 2022 International Conference on Electrical Machines and Systems (ICEMS 2022), Chiang Mai, Thailand, 29 November–2 December 2022. [[CrossRef](#)]
51. Xu, J.; Qian, H.; Hu, Y.; Bian, S.; Xie, S. Overview of SOGI-Based Single-Phase Phase-Locked Loops for Grid Synchronization under Complex Grid Conditions. *IEEE Access* **2021**, *9*, 39275–39291. [[CrossRef](#)]
52. Li, K.; Bo, A.; Zheng, H.; Sun, N. Performance Analysis of Three-Phase Phase-Locked Loops for Distorted and Unbalanced Grids. *J. Power Electron.* **2017**, *17*, 262–271. [[CrossRef](#)]
53. Zhang, R.; Chand, S.S.; Hredzak, B.; Bie, Z. A Novel Synchronization Strategy for Distributed Energy Resources in Weak Grids Using Remote Strong Grid Sensing. *Energies* **2024**, *17*, 6135. [[CrossRef](#)]
54. Chand, S.; Prasad, R.; Mudaliar, H.; Kumar, D.; Fagiolini, A.; Benedetto, M.; Cirrincione, M. Improving Power Delivery of Grid-Connected Induction Machine Based Wind Generators under Dynamic Conditions Using Feedforward Linear Neural Networks. *IEEE Access* **2023**, *11*, 63550–63564. [[CrossRef](#)]
55. Khan, M.Y.A.; Liu, H.; Yang, Z.; Yuan, X. A Comprehensive Review on Grid Connected Photovoltaic Inverters, Their Modulation Techniques, and Control Strategies. *Energies* **2020**, *13*, 4185. [[CrossRef](#)]
56. Chand, S.; Prasad, R.; Mudaliar, H.; Kumar, D.; Fagiolini, A.; Di Benedetto, M.; Cirrincione, M. Enhanced Current Loop PI Controllers with Adaptive Feed-Forward Neural Network via Estimation of Grid Impedance: Application to Three-Phase Grid-Tied PV Inverters. In Proceedings of the 2022 IEEE Energy Conversion Congress and Exposition (ECCE 2022), Detroit, MI, USA, 9–13 October 2022. [[CrossRef](#)]
57. Subsingha, W. A Comparative Study of Sinusoidal PWM and Third Harmonic Injected PWM Reference Signal on Five Level Diode Clamp Inverter. *Energy Procedia* **2016**, *89*, 137–148. [[CrossRef](#)]
58. Tarequzzaman, M.; Ahmed, S.; Moznuzzaman, M. Performance Improvement of Direct Torque Control Induction Motor Drive using Genetic Algorithm Optimized PI Controller. In Proceedings of the 2018 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE 2018), Gazipur, Bangladesh, 22–24 November 2018. [[CrossRef](#)]
59. Yang, Q.; Dai, N. Genetic Algorithm for PI Controller Design of Grid-connected Inverter based on Multilayer Perceptron Model. In Proceedings of the 2021 IEEE Sustainable Power and Energy Conference: Energy Transition for Carbon Neutrality (iSPEC 2021), Nanjing, China, 23–25 December 2021; pp. 2770–2775. [[CrossRef](#)]
60. Debdouche, N.; Zarour, L.; Chebabhi, A.; Bessous, N.; Benbouhenni, H.; Colak, I. Genetic algorithm-super-twisting technique for grid-connected PV system associate with filter. *Energy Rep.* **2023**, *10*, 4231–4252. [[CrossRef](#)]
61. de Oliveira Evald, P.J.D.; Hollweg, G.V.; Borin, L.C.; Mattos, E.; Tambara, R.V.; Montagner, V.F.; Gründling, H.A. An optimal initialisation for robust model reference adaptive PI controller for grid-tied power systems under unbalanced grid conditions. *Eng. Appl. Artif. Intell.* **2023**, *124*, 106589. [[CrossRef](#)]
62. Ali, M.; Tariq, M.; Upadhyay, D.; Khan, S.A.; Satpathi, K.; Alamri, B.; Alahmadi, A.A. Genetic Algorithm Based PI Control with 12-Band Hysteresis Current Control of an Asymmetrical 13-Level Inverter. *Energies* **2021**, *14*, 6663. [[CrossRef](#)]
63. Yang, Y.; Kim, K.A.; Blaabjerg, F.; Sangwongwanich, A. Control of PV systems under normal grid conditions. In *Advances in Grid-Connected Photovoltaic Power Conversion Systems*; Woodhead Publishing Ltd.: Cambridge, UK, 2019; pp. 75–112. [[CrossRef](#)]

64. Neal, C.; Dagdougui, H.; Lodi, A.; Fernandez, J.M. Reinforcement Learning Based Penetration Testing of a Microgrid Control Algorithm. In Proceedings of the 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC 2021), Virtual, 27–30 January 2021; pp. 38–44. [[CrossRef](#)]
65. Fujimoto, S.; Van Hoof, H.; Meger, D. Addressing Function Approximation Error in Actor-Critic Methods. In Proceedings of the 35th International Conference on Machine Learning (ICML 2018), Stockholm Sweden, 10–15 July 2018; Volume 4, pp. 2587–2601.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.