



## Deep Q-Network–Based Power Sharing Control for Islanded Microgrids

Hoang Dinh Co , Nguyen Cao Cuong\* 

Faculty of Electrical Engineering - Automation, University of Economics – Technology for Industries, Ha Noi 100000, Vietnam

Corresponding Author Email: [nccuong@uneti.edu.vn](mailto:nccuong@uneti.edu.vn)

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/jesa.590303>

### ABSTRACT

**Received:** 2 January 2026

**Revised:** 19 March 2026

**Accepted:** 27 March 2026

**Available online:** 31 March 2026

#### Keywords:

*Deep Q-Network, deep reinforcement learning, islanded microgrid, power sharing, parallel inverters*

This study presents a data-driven power sharing control problem for independent microgrids using Deep Q-Network (DQN) enhancement learning. Unlike conventional change-based methods that rely on predefined control laws and precise knowledge of line parameters, this proposed approach allows the inverter controller to learn optimal control actions through continuous interaction with the microgrid environment. A neural network is used to approximate the value function in the action space, enabling adaptive tuning of inverter voltage reference data to improve reactive power performance. The controller only requires locally measured electrical variables, such as voltage, current, and power signals, and therefore eliminates the need for explicit line impedance estimation. As a result, the proposed strategy provides enhanced adaptability under load variations and network uncertainties. The performance of the control scheme proposed by the authors was evaluated using simulation results performed in the MATLAB/Simulink environment. These results show that the DQN controller significantly improves the accuracy of reactive power distribution while maintaining voltage stability at the common connection point.

## 1. INTRODUCTION

The rapid integration of distributed renewable energy sources, including photovoltaic (PV) systems, wind generators, and battery energy storage units, has significantly accelerated the development of modern microgrid technologies. A microgrid can integrate multiple distributed generation units through power electronic converters and operate either in grid-connected mode or in islanded mode. In particular, islanded microgrids play an important role in supplying reliable electricity to remote areas, isolated communities, and critical infrastructures where the main utility grid is unavailable or unreliable [1, 2].

One of the major challenges in islanded microgrids is achieving accurate power sharing among multiple parallel-connected inverters while maintaining acceptable voltage and frequency levels at the point of common coupling (PCC). In practical installations, distributed generators are connected to the PCC through distribution feeders with different electrical characteristics. The mismatch of feeder impedances can cause unequal voltage drops and circulating reactive power among inverters, which degrades voltage quality and overall system performance [2, 3].

Droop control is one of the most widely adopted decentralized control strategies for inverter-based microgrids. In this method, active and reactive power outputs are regulated through predefined frequency–power and voltage–reactive power relationships, allowing distributed generators to share loads without requiring communication among controllers [1, 2]. However, numerous studies have shown that conventional

droop control often leads to significant reactive power sharing errors when line impedances are mismatched or when inverter ratings are different [3, 4].

To address these limitations, various improved control strategies have been proposed in the literature. These approaches include adaptive droop control methods, virtual impedance techniques, and hierarchical control frameworks that aim to enhance power sharing accuracy and voltage regulation in microgrids [4-6]. Although these techniques can improve system performance, many of them rely on accurate parameter estimation or communication links between distributed generators, which may increase implementation complexity and reduce system robustness in uncertain operating conditions [5, 6].

Recently, the application of artificial intelligence (AI) and machine learning techniques has been developing, offering promising solutions for controlling complex power systems, even in Vietnam. In particular, reinforcement learning (RL) has demonstrated strong capability in solving control problems involving nonlinear dynamics and uncertain environments [7, 8]. Unlike traditional model-based control strategies, RL algorithms enable controllers to learn optimal policies through interactions with the environment without requiring an explicit mathematical model of the system [9].

Among different RL algorithms, the Deep Q-Network (DQN) has attracted significant attention due to its ability to approximate the action–value function using deep neural networks [7]. The incorporation of techniques such as experience replay and target networks improves training stability and enables DQN to handle high-dimensional state



spaces effectively [8, 9]. As a result, deep RL methods have been successfully applied to a wide range of energy system applications, including energy management, voltage regulation, and optimal control of distributed generation units [10-12].

Recently, several studies have explored the use of RL in microgrid control problems. These works demonstrate that learning-based controllers can improve system adaptability and robustness compared with traditional control methods [13, 14]. Furthermore, multi-agent RL approaches have been investigated to coordinate multiple distributed generators in microgrids, enabling decentralized decision-making and improved system scalability [15-18].

In addition, recent research has focused on intelligent inverter control and data-driven strategies for inverter-dominated power systems. These studies highlight the potential of AI techniques to improve voltage regulation, power sharing accuracy, and overall microgrid stability [19-21]. Data-driven control approaches are particularly attractive for modern microgrids where system parameters may be uncertain or time-varying [21, 22].

Moreover, AI has been increasingly integrated into power electronic converter control and smart grid operation, enabling more flexible and adaptive control architectures for future energy systems [23, 24]. Recent surveys also indicate that RL-based control strategies have strong potential for autonomous microgrid operation and intelligent energy management [24, 25].

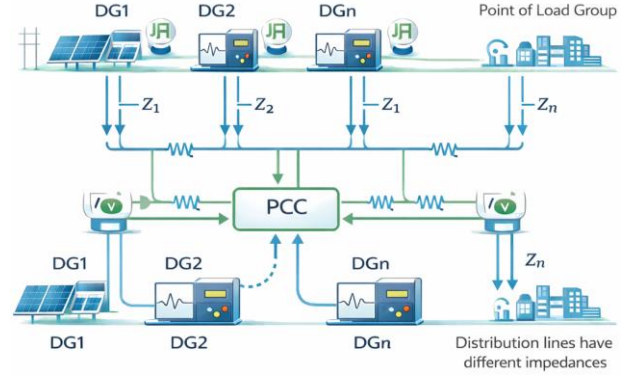
Motivated by these developments, this paper proposes a power sharing control strategy for islanded microgrids based on the DQN reinforcement learning algorithm. In the proposed framework, each inverter is modeled as an intelligent agent capable of learning an optimal control policy to minimize reactive power sharing errors while maintaining voltage stability at the PCC. The proposed controller relies only on locally measured electrical variables and does not require explicit knowledge of line impedance parameters. Simulation studies are conducted to evaluate the effectiveness of the proposed approach and to compare its performance with that of conventional droop-based control strategies.

## 2. MICROGRID MODELING AND REACTIVE POWER SHARING PROBLEM

Consider an islanded microgrid composed of several three-phase voltage source inverters connected in parallel to a common bus through distribution feeders, as illustrated in Figure 1 [1-3]. Each inverter represents a distributed generation unit interfaced through power electronic converters. In practical installations, the electrical characteristics of the feeders are generally different due to variations in cable length, conductor type, and installation conditions. As a consequence, the line impedances between individual inverters and the PCC are not identical.

This impedance mismatch introduces unequal voltage drops along the feeders, which directly affects the reactive power exchanged between inverters and the PCC. As a result, conventional decentralized control strategies may lead to inaccurate reactive power sharing and circulating currents among distributed generators.

Each inverter is controlled by a hierarchical two-loop control structure:



**Figure 1.** Cavity diagram of the three-phase voltage source inverters connected to the point of common coupling (PCC) in a microgrid

### Inverter Control Structure

- **Inner control loop:**

Within the proposed control architecture, the inner loop is responsible for controlling the inverter output current and voltage using the synchronous  $dq$  coordinate system. This structure allows the system to achieve fast transient performance while preserving local stability. The voltage dynamics of the  $i$ -th inverter in the  $dq$  frame can therefore be written as:

$$\begin{cases} v_{di} = R_i i_{di} + L_i \frac{di_{di}}{dt} - \omega L_i i_{qi} + u_{di} \\ v_{qi} = R_i i_{qi} + L_i \frac{di_{qi}}{dt} + \omega L_i i_{di} + u_{qi} \end{cases} \quad (1)$$

where,  $i_{di}$  and  $i_{qi}$  denote the  $dq$ -axis current components, while  $u_{di}$  and  $u_{qi}$  represent the associated control signals. The parameters  $R_i$  and  $L_i$  indicate the resistance and inductance of the inverter output filter, respectively.

- **Outer control loop:**

The outer loop performs the task of power regulation while ensuring proper load sharing among the parallel inverters in the microgrid.

### Power Model and Load Sharing

The active and reactive power delivered by the  $i$ -th inverter to the PCC are defined as follows:

$$\begin{cases} P_i = \frac{V_i V_{PCC}}{X_i} \sin(\delta_i) \\ Q_i = \frac{V_i}{X_i} (V_i - V_{PCC}) \end{cases} \quad (2)$$

where,  $V_i$  is the magnitude of the inverter terminal voltage,  $V_{PCC}$  denotes the voltage magnitude at the PCC,  $X_i$  corresponds to the feeder reactance, and  $\delta_i$  represents the phase angle difference between the inverter and the PCC.

Under the small-angle assumption ( $\delta_i \approx 0$ ), the reactive power can be approximated as a linear function of voltage:

$$Q_i \approx \frac{V_i}{X_i} (V_i - V_{PCC}). \quad (3)$$

The reactive power reference is allocated proportionally to the rated capacity of each inverter:

$$Q_i^* = \frac{S_i}{\sum_{j=1}^n S_j} Q_{\text{load}} \quad (4)$$

where,  $S_i$  denotes the apparent power capacity assigned to the  $i$ -th inverter.

Accordingly, the reactive power sharing error is defined as the difference between the actual and reference reactive powers, i.e.,  $e_{Q,i} = Q_i - Q_i^*$ .

### Drawbacks of Conventional Droop Control

According to the classical  $Q - V$  droop mechanism, the voltage reference signal of each inverter is defined as:

$$V_i^* = V_0 - n_Q Q_i \quad (5)$$

where,  $n_Q$  is a fixed droop coefficient. Due to unequal line reactances  $X_i$ , using identical droop coefficients results in unequal reactive power sharing:  $Q_1 \neq Q_2 \neq \dots \neq Q_n$ , which leads to circulating reactive power and degraded voltage regulation [5, 6].

### Motivation for DQN-Based Control

To overcome the above limitations, the proposed approach replaces the fixed  $Q - V$  droop characteristic with a DQN-based controller. The voltage reference of the  $i$ -th inverter is modified as:

$$V_i^* = V_0 + \Delta V_{DQN,i} \quad (6)$$

where,  $\Delta V_{DQN,i}$  is the voltage correction signal generated by the DQN agent.

Through iterative interactions with the microgrid system, the DQN agent is trained to obtain an optimal policy that determines the appropriate adjustment of  $\Delta V_{DQN,i}$  such that:

$$\lim_{t \rightarrow \infty} e_{Q,i}(t) \rightarrow 0, \forall i = 1, \dots, n, \quad (7)$$

while maintaining the PCC voltage close to its nominal value:  $V_{PCC} \approx V_0$ .

This learning-based control strategy enables accurate reactive power sharing without requiring prior knowledge of line parameters and provides strong adaptability to load variations and network uncertainties.

## 3. DEEP Q-NETWORK-BASED POWER SHARING CONTROL

### 3.1 Reinforcement learning problem formulation

In the proposed control framework, the power sharing problem is formulated as an RL task in which each inverter is represented by an intelligent agent. The agent interacts with the microgrid environment by observing system variables, selecting control actions, and receiving feedback in the form of reward signals. Through repeated interactions, the agent gradually learns a control policy that optimizes system performance.

The state vector is constructed from locally measured electrical variables that reflect the operating condition of the inverter and the microgrid. Typical state variables include inverter active power, reactive power, PCC voltage magnitude,

and the reactive power sharing error. Based on the observed state, the DQN agent selects a discrete control action that modifies the inverter control variable, such as the voltage reference or the droop coefficient [4, 12].

#### 3.1.1 State

The state vector of the agent at time step  $k$  is defined as:

$$s_k = [P_k, Q_k, V_{PCC,k}, \Delta Q_k] \quad (8)$$

where:

$P_k$  and  $Q_k$  correspond to the inverter active and reactive power components, respectively;

$V_{PCC,k}$  refers to the voltage magnitude at the PCC;

$\Delta Q_k = Q_k - Q_k^*$  expresses the mismatch between the actual and reference reactive powers.

#### 3.1.2 Action

Since DQN handles a discrete action space, the action set is constructed by quantizing the control signal, for example:

$$a_k \in \{-\Delta V, 0, +\Delta V\} \quad (9)$$

The selected action is used to adjust one of the following control variables: the voltage reference  $V_{ref}$ ; or the  $Q-V$  droop coefficient; or the equivalent virtual impedance value.

#### 3.1.3 Reward function

The reward function is designed to minimize power sharing errors while maintaining voltage stability:

$$r_k = -\alpha |\Delta Q_k| - \beta |V_{PCC,k} - V_0| \quad (10)$$

where,  $V_0$  is the rated voltage, and  $\alpha$  and  $\beta$  are weighting coefficients.

The weighting factors  $\alpha$  and  $\beta$  are selected based on a sensitivity analysis to balance voltage regulation and reactive power sharing. In this study,  $\alpha = 1$  and  $\beta = 1.5$  are chosen as they provide a good trade-off between the two objectives.

### 3.2 Deep Q-Network algorithm

The DQN algorithm is a deep RL approach that integrates traditional Q-learning with deep neural networks in order to handle control problems involving high-dimensional and nonlinear state spaces [7, 12]. In the DQN framework, the action-value function  $Q(s, a)$  is not represented using a tabular form as in classical Q-learning. Instead, it is approximated by a deep neural network parameterized by  $\theta$ .

In the proposed microgrid control framework, each inverter is treated as an autonomous RL agent. At every discrete time instant  $k$ , the agent observes the current system state  $s_k$ , selects a control action  $a_k$ , and subsequently receives a reward  $r_k$  together with the next system state  $s_{k+1}$ . Through this interaction process, the agent gradually learns a control strategy that maximizes the cumulative discounted reward over time.

The action-value function used in the DQN algorithm can be expressed as [9]:

$$Q(s_k, a_k) = E \left[ \sum_{t=k}^{\infty} \gamma^{t-k} r_t \right] \quad (11)$$

where,  $\gamma \in (0,1)$  denotes the discount factor that determines the relative weighting between immediate rewards and future rewards.

To improve the stability of the training process, the DQN algorithm incorporates two important mechanisms: experience replay and a target network. The experience replay buffer stores transition samples in the form of tuples  $(s_k, a_k, r_k, s_{k+1})$  obtained from the agent's interaction with the environment. During training, mini-batches are randomly sampled from this buffer, which helps eliminate the strong temporal correlations between successive data and consequently enhances the convergence behavior of the neural network.

Furthermore, a separate target network with parameters  $\theta^-$  is introduced to generate the target Q-values during the learning process. The parameters of this target network are periodically synchronized with those of the primary network. This mechanism mitigates learning oscillations and contributes to a more stable training procedure. Accordingly, the loss function of the DQN algorithm can be expressed as:

$$L(\theta) = E \left[ \left( r_k + \gamma \max_a Q(s_{k+1}, a'; \theta^-) - Q(s_k, a_k; \theta) \right)^2 \right] \quad (12)$$

Throughout the learning process, an  $\epsilon$ -greedy policy is used to maintain a trade-off between exploration and exploitation. More specifically, the agent selects a random action with probability  $\epsilon$  to investigate new possibilities, while with probability  $1 - \epsilon$ , it chooses the action that maximizes the estimated Q-value provided by the neural network.

In the proposed control strategy, the DQN algorithm is integrated into the outer control loop of each inverter to adjust reference signals, such as the voltage reference or the equivalent droop coefficient. Owing to its ability to learn directly from operational data, the DQN-based controller does not require prior knowledge of the mathematical model or line impedance parameters, and it exhibits strong adaptability to load variations and changes in microgrid topology.

### 3.3 Integration of the Deep Q-Network algorithm into the microgrid control structure

In the proposed control strategy, the DQN algorithm is integrated into the outer control loop of each inverter, while the inner control loop continues to employ conventional linear controllers to ensure fast stability of current and voltage dynamics, as illustrated in Figure 2 [16-18].

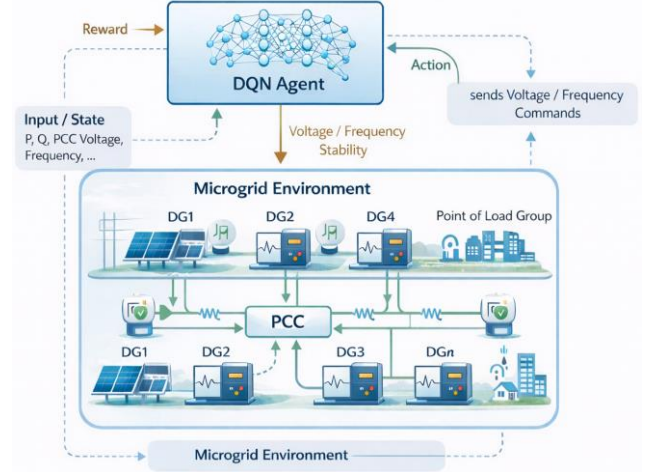
Specifically, the DQN controller generates a voltage correction signal  $\Delta V_{DQN,i}$  to adjust the reference voltage of the  $i$ -th inverter as follows:

$$V_i^* = V_0 + \Delta V_{DQN,i} \quad (13)$$

where,  $V_0$  denotes the rated voltage of the microgrid. The signal  $\Delta V_{DQN,i}$  is updated according to the action selected by the DQN agent at each discrete time step. As a result, the inverter can autonomously regulate its output voltage to compensate for reactive power sharing errors caused by mismatched line impedances.

The incorporation of the DQN controller into the outer control loop enables the system to operate without relying on line impedance parameters, adapt effectively to load variations,

and significantly reduce circulating reactive power among parallel inverters.



**Figure 2.** Deep Q-Network (DQN)-based microgrid control structure

To estimate the action-value function, a deep neural network is employed in the proposed DQN framework. The network takes the system state vector as the input and outputs the estimated Q-values corresponding to all possible control actions. In this study, a fully connected feedforward neural network is adopted due to its simplicity and effectiveness for nonlinear control problems.

The neural network consists of an input layer, two hidden layers, and an output layer. The hidden layers contain 64 and 32 neurons, respectively, and use the Rectified Linear Unit (ReLU) activation function to enhance nonlinear representation capability. The output layer employs a linear activation function to generate the Q-values associated with each action.

During training, the network parameters are optimized using the Adam optimizer with a learning rate of  $10^{-3}$ . The discount factor is set to  $\gamma = 0.95$ , and the experience replay buffer size is chosen as 10,000 samples. Mini-batches of size 64 are randomly sampled from the replay buffer to update the network parameters. The target network parameters are periodically updated every 200 training steps to improve learning stability. The architecture of the neural network used in the DQN controller and the corresponding training parameters are summarized in Table 1.

**Table 1.** Deep Q-Network (DQN) neural network architecture and training parameters

Parameter	Value
Network type	Fully connected feedforward network
Input layer	State vector $s_k$
Hidden layer 1	64 neurons, ReLU
Hidden layer 2	32 neurons, ReLU
Output layer	Q-values for each action (linear activation)
Optimizer	Adam
Learning rate	0.001
Discount factor $\gamma$	0.95
Replay memory size	10000
Mini-batch size	64
Target network update	Every 200 steps

**Algorithm:** DQN-based power sharing control for islanded microgrids

**Input:**

- Rated voltage  $V_0$
- Discount factor  $\gamma$
- Reward weighting coefficients  $\alpha, \beta$
- Voltage adjustment step  $\Delta V$
- Initial exploration rate  $\varepsilon$

**Output:**

- Voltage correction signal  $\Delta V_{DQN,i}$  for each inverter

**Initialization**

1. Initialize the main Q-network with parameters  $\theta$
2. Initialize the target network with parameters  $\theta^- \leftarrow \theta$
3. Initialize the experience replay memory  $D$
4. Set  $\Delta V_{DQN,i}(0) = 0$  for all inverters

**Training and Control Process**

5. For each training episode do
6. Observe the initial state
 
$$s_{i,0} = [P_{i,0}, Q_{i,0}, V_{PCC,0}, e_{Q,i,0}]$$
7. For each time step  $k = 0, 1, 2, \dots$  do
8. Select action  $a_{i,k}$  according to the  $\varepsilon$ -greedy policy:
 
$$a_{i,k} = \begin{cases} \text{random action, with probability } \varepsilon \\ \arg \max_a Q(s_{i,k}, a; \theta), \text{ with probability } 1 - \varepsilon \end{cases}$$
9. Update the voltage correction signal:
 
$$\Delta V_{DQN,i}(k+1) = \Delta V_{DQN,i}(k) + a_{i,k}$$
10. Apply the reference voltage:
 
$$V_i^* = V_0 + \Delta V_{DQN,i}(k+1)$$
11. Measure the new state  $s_{i,k+1}$  and the reward:
 
$$r_{i,k} = -\alpha |e_{Q,i,k}| - \beta |V_{PCC,k} - V_0|$$
12. Store the transition
 
$$(s_{i,k}, a_{i,k}, r_{i,k}, s_{i,k+1})$$

into the replay memory  $D$

13. Sample a random minibatch from  $D$
14. Compute the target value:
 
$$y_k = r_{i,k} + \gamma \max_a Q(s_{i,k+1}, a'; \theta^-)$$
15. Update the main network parameters  $\theta$  by minimizing:
 
$$L(\theta) = E \left[ \left( y_k - Q(s_{i,k}, a_{i,k}; \theta) \right)^2 \right]$$
16. Gradually decrease the exploration rate  $\varepsilon$
17. If  $k \bmod N_{\text{target}} = 0$  then
18. Update the target network:  $\theta^- \leftarrow \theta$
19. **End For**
20. **End For**

## 4. STABILITY ANALYSIS OF THE DEEP Q-NETWORK-BASED MICROGRID SYSTEM

### 4.1 Reactive power sharing error model

Consider the reactive power sharing error of the  $i$ -th inverter defined as:

$$e_{Q,i} = Q_i - Q_i^* \quad (14)$$

where,  $Q_i^*$  is the reference reactive power allocated according to the rated capacity of each inverter.

Under the assumptions of small phase angle differences and predominantly inductive line impedances, the reactive power can be approximated as a linear function of the voltage, yielding [11, 21]:

$$e_{Q,i} = K_i \quad (15)$$

where,  $K_i > 0$  is a positive constant dependent on system parameters.

Since

$$V_i = V_0 + \Delta V_{DQN,i} \quad (16)$$

it follows that:

$$e_{Q,i} = K_i \Delta V_{DQN,i} \quad (17)$$

### 4.2 Lyapunov stability analysis

A candidate Lyapunov function for the overall microgrid is chosen as [4, 23]:

$$V = \frac{1}{2} \sum_{i=1}^n e_{Q,i}^2 \quad (18)$$

Taking the time derivative of  $V$ , we obtain:

$$\dot{V} = \sum_{i=1}^n e_{Q,i} \dot{e}_{Q,i} = \sum_{i=1}^n K_i e_{Q,i} \Delta V_{DQN,i} \quad (19)$$

Assume that the DQN agent learns an equivalent control policy of the form:

$$\Delta V_{DQN,i} = -k_i e_{Q,i} + \varepsilon_i \quad (20)$$

where,  $k_i > 0$  and  $\varepsilon_i$  represents the approximation error of the neural network, which is assumed to be bounded as:  $|\varepsilon_i| \leq \bar{\varepsilon}_i$ .

Substituting into  $\dot{V}$  yields:

$$\dot{V} = - \sum_{i=1}^n k_i K_i e_{Q,i}^2 + \sum_{i=1}^n K_i e_{Q,i} \varepsilon_i \quad (21)$$

Applying Young's inequality, we obtain:

$$\dot{V} \leq - \sum_{i=1}^n \frac{k_i K_i}{2} e_{Q,i}^2 + \sum_{i=1}^n \frac{K_i}{2k_i} \bar{\varepsilon}_i^2 \quad (22)$$

Therefore, the microgrid system is uniformly ultimately bounded (UUB) in the Lyapunov sense. When the learning process converges and  $\varepsilon_i \rightarrow 0$ , the reactive power sharing error asymptotically converges to zero.

## 5. SIMULATION RESULTS

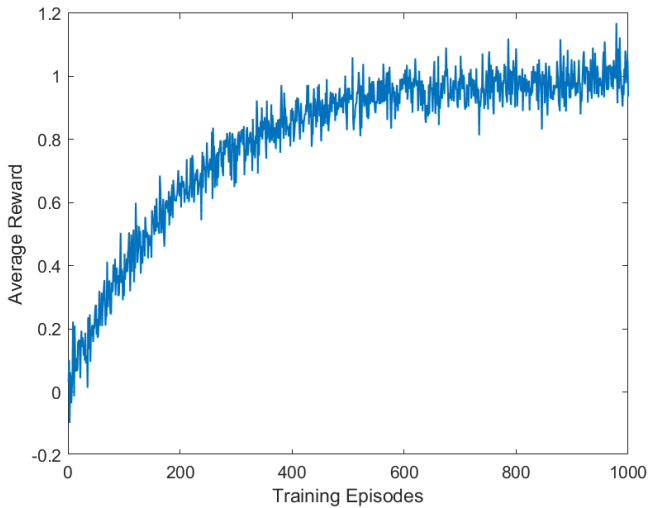
To evaluate the effectiveness of the proposed DQN-based power sharing control strategy, comprehensive simulations were carried out in the MATLAB/Simulink environment. The simulated system consists of an islanded microgrid with multiple three-phase inverters connected in parallel to a PCC through distribution lines with different impedances. The system parameters and controller settings were selected in accordance with typical microgrid studies reported in the literature. The system parameters used in the simulation are listed in Table 2.

In the simulation scenarios, the microgrid operates in islanded mode with time-varying loads to examine the adaptability and robustness of the proposed controller. The inverters have different rated power capacities and are connected to the PCC through lines with unequal reactances, which leads to reactive power sharing inaccuracies when conventional droop control is applied. The performance of the proposed method is directly compared with that of the conventional droop control to highlight the advantages of the DQN-based controller.

**Table 2.** System parameters of the studied microgrid

Parameter	Symbol	Value
Rated apparent power (Inverter 1)	$S_1$	8 kVA
Rated apparent power (Inverter 2)	$S_2$	4 kVA
Nominal voltage	$V_{nom}$	380 V
Nominal frequency	$f_{nom}$	50 Hz
Line resistance (Inv1-PCC)	$R_1$	0.2 $\Omega$
Line reactance (Inv1-PCC)	$X_1$	0.4 $\Omega$
Line resistance (Inv2-PCC)	$R_2$	0.3 $\Omega$
Line reactance (Inv2-PCC)	$X_2$	0.6 $\Omega$
Filter inductance	$L_f$	2 mH
Filter resistance	$R_f$	0.1 $\Omega$
Load active power	$P_L$	10 kW
Load reactive power	$Q_L$	5 kVar
Sampling time	$T_s$	0.0001 s

To evaluate the training stability of the proposed method, a reward-versus-episode curve is presented, which is a standard metric for assessing convergence in RL. Figure 3 shows the evolution of the average reward over training episodes.

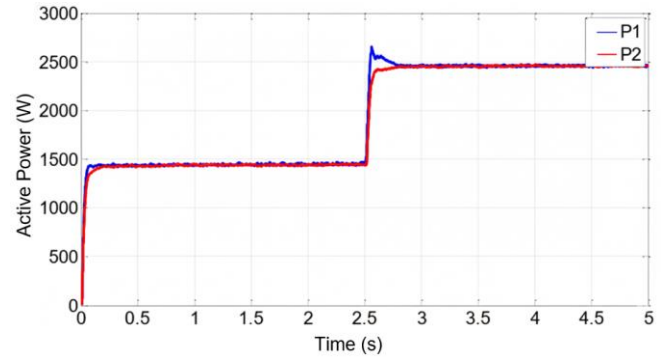


**Figure 3.** Training convergence of the Deep Q-Network (DQN) controller

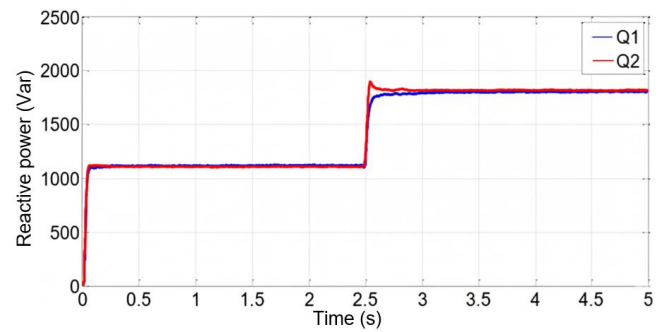
Figure 4 presents the active power distribution among the inverters. Because of the inherent properties of droop-based control, active power is shared with satisfactory accuracy under both control schemes. The results show that the proposed DQN-based controller achieves a level of active power sharing similar to that of the conventional droop approach, without introducing any noticeable degradation in microgrid frequency stability.

Figure 5 presents the reactive power responses of the inverters in the microgrid. With conventional droop control, the unequal line impedances lead to significant reactive power sharing errors, resulting in circulating reactive power among

the inverters. In contrast, the DQN-based control method significantly improves the accuracy of reactive power sharing. The inverters distribute reactive power according to their rated power capacities, and the sharing error is substantially reduced after a short learning period.



**Figure 4.** Active power sharing performance



**Figure 5.** Reactive power sharing performance

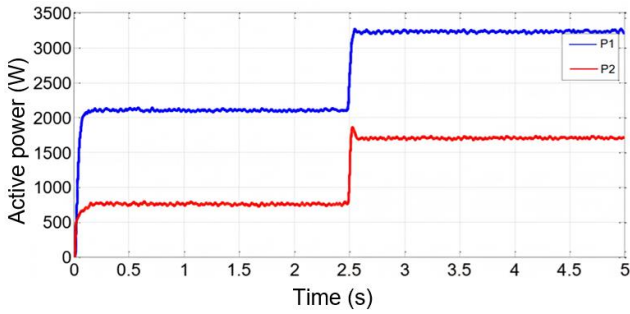
Figures 6 and 7 show the simulated power-sharing behavior in an islanded microgrid with two parallel-connected inverters, where the intended sharing ratio is set to 2:1 using the proposed DQN-based control strategy. The results demonstrate that, following a short transient interval, the output power of the inverters converges to the desired 2:1 distribution. During the initial stage, due to startup conditions and the learning process of the DQN controller, small power oscillations may occur. However, as the learning process gradually converges, these oscillations are rapidly attenuated, and the system reaches a stable steady state.

Specifically, the first inverter supplies approximately twice the power of the second inverter, which is consistent with the prescribed power-sharing ratio. The power-sharing error is significantly reduced and converges to nearly zero, demonstrating the high accuracy of the proposed control strategy. Moreover, the adjustment process is smooth, without inducing large oscillations or adversely affecting the overall stability of the microgrid.

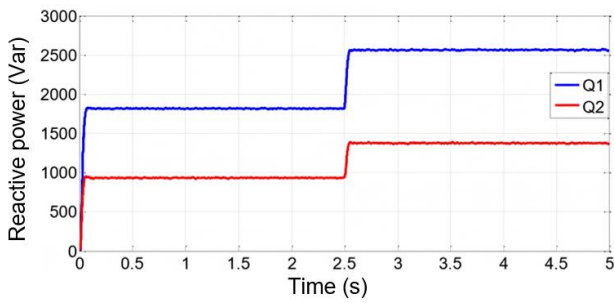
The results presented in Figures 6 and 7 demonstrate that the proposed DQN-based controller is capable of adaptively regulating the power-sharing process to achieve the specified ratio, even without prior information about the line impedance parameters. These results highlight the effectiveness of the proposed approach in enabling flexible and accurate power distribution among parallel inverters operating in islanded microgrids.

Figure 8 depicts the voltage profile at the PCC when the proposed DQN-based power-sharing controller is implemented. It can be seen that the PCC voltage remains

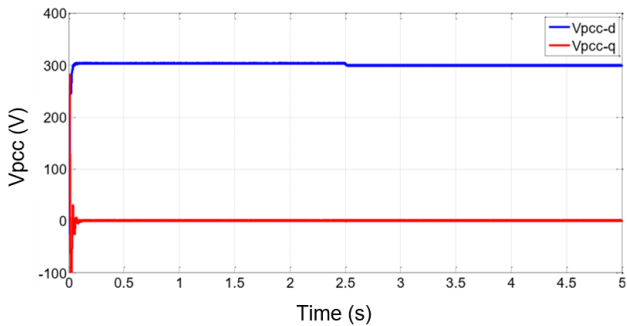
stable and close to its nominal level during the entire operation, despite the presence of line impedance mismatches and the interactions among parallel inverters. At the initial stage, minor voltage variations may occur due to the initialization and training process of the DQN controller; however, these fluctuations are quickly attenuated as the learning algorithm gradually converges.



**Figure 6.** Active power sharing performance



**Figure 7.** Reactive power sharing performance



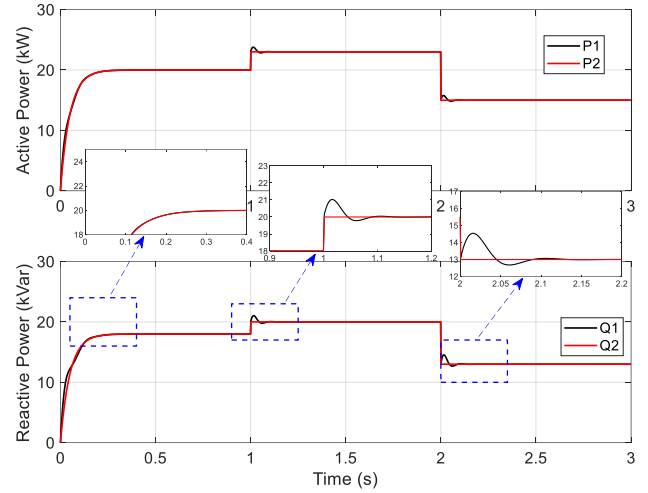
**Figure 8.** Voltage at the point of common coupling (PCC)

These results indicate that the DQN-based controller not only improves reactive power sharing accuracy but also significantly enhances voltage quality at the PCC, effectively overcoming the voltage deviation and fluctuation issues commonly associated with conventional droop control. Therefore, Figure 8 clearly demonstrates the effectiveness and adaptability of the method in ensuring PCC voltage stability for islanded microgrid operation.

Figure 9 illustrates the system response under a sudden load change. The simulation results demonstrate that the DQN-based controller rapidly adapts to new operating conditions, maintaining PCC voltage stability and ensuring accurate reactive power sharing among the inverters. In contrast, the conventional droop control exhibits a slower dynamic response and larger reactive power errors during the transient period.

From the simulation results, it is evident that the proposed

DQN-based controller provides improved performance over the traditional droop control method, particularly with respect to reactive power sharing precision and PCC voltage quality. In addition, the proposed strategy operates without requiring prior information about line impedances and shows robust adaptability to variations in load demand and operating conditions. This indicates a promising capability for deployment in islanded microgrids subject to significant uncertainties.



**Figure 9.** System response under sudden load change

## 6. CONCLUSIONS

This paper has proposed a reactive power sharing control strategy for islanded microgrids based on the DQN algorithm. By integrating DQN into the outer control loop of inverters, the proposed method enables the system to adapt to line impedance mismatches and load variations, thereby overcoming the limitations of conventional droop-based control. The DQN algorithm is designed with appropriate state variables and a reward function, which effectively reduces reactive power sharing errors while maintaining the PCC voltage close to its nominal value. Lyapunov-based stability analysis demonstrates that the microgrid system achieves UUB stability, and the reactive power sharing error asymptotically converges to zero as the learning process converges. The results indicate that the proposed approach provides an effective and flexible solution for microgrid control, particularly suitable for systems with uncertain parameters and varying operating conditions.

Future work will focus on real-time implementation of the proposed DQN-based controller on embedded platforms such as DSPs and FPGAs, with particular attention to latency and computational efficiency. In addition, the framework will be extended to large-scale multi-inverter microgrids under highly dynamic operating conditions, while incorporating advanced RL techniques to further enhance convergence speed and robustness.

## ACKNOWLEDGMENT

This study was supported by the University of Economics-Technology for Industries, Ha Noi - Vietnam;

## REFERENCES

- [1] Shahgholian, G., Moradian, M., Fathollahi, A. (2025). Droop control strategy in inverter-based microgrids: A brief review on analysis and application in islanded mode of operation. *IET Renewable Power Generation*. <https://doi.org/10.1049/rpg2.13186>
- [2] Guerrero, J.M., Vasquez, J.C., Matas, J., De Vicuña, L.G., Castilla, M. (2011). Hierarchical control of droop-controlled AC and DC microgrids—A general approach toward standardization. *IEEE Transactions on Industrial Electronics*, 58(1): 158-172. <https://doi.org/10.1109/TIE.2010.2066534>
- [3] Wang, K., Huang, X., Fan, B., Yang, Q., Li, G., Crow, M.L. (2018). Decentralized Power Sharing Control for Parallel-Connected Inverters in Islanded Single-Phase Micro-Grids. *IEEE Transactions on Smart Grid*, 9(6): 6721-6730. <https://doi.org/10.1109/TSG.2017.2720683>
- [4] Lewis, F.L., Vrabie, D. (2009). Reinforcement learning and adaptive dynamic programming for feedback control. *IEEE Circuits and Systems Magazine*, 9(3): 32-50. <https://doi.org/10.1109/MCAS.2009.933854>
- [5] Zhang, Z., Zhang, D., Qiu, R.C. (2020). Deep reinforcement learning for power system applications: An overview. *CSEE Journal of Power and Energy Systems*, 6(1): 213-225. <https://doi.org/10.17775/CSEEJPES.2019.00920>
- [6] Vázquez-Canteli, J.R., Nagy, Z. (2019). Reinforcement learning for demand response: A review of algorithms and modeling techniques. *Applied Energy*, 235: 1072-1089. <https://doi.org/10.1016/j.apenergy.2018.11.002>
- [7] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540): 529-533. <https://doi.org/10.1038/nature14236>
- [8] Van Hasselt, H., Guez, A., Silver, D. (2016). Deep reinforcement learning with double Q-learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Phoenix, Arizona, USA, 30(1): 2094-2100. <https://doi.org/10.1609/aaai.v30i1.10295>
- [9] Zhang, S., Sutton, R.S. (2017). A deeper look at experience replay. *arXiv preprint arXiv:1712.01275*. <https://doi.org/10.48550/arXiv.1712.01275>
- [10] Cui, Y., Xu, Y., Li, Y., Wang, Y., Zou, X. (2026). Deep reinforcement learning based optimal energy management of multi-energy microgrids with uncertainties. *CSEE Journal of Power and Energy Systems*, 12(2): 1002-1014. <https://doi.org/10.17775/CSEEJPES.2023.05120>
- [11] Li, S., Cao, D., Hu, W., Huang, Q., Chen, Z., Blaabjerg, F. (2023). Multi-energy management of interconnected multi-microgrid system using multi-agent deep reinforcement learning. *Journal of Modern Power Systems and Clean Energy*, 11(5): 1606-1617. <https://doi.org/10.35833/MPCE.2022.000473>
- [12] Vincent, F.L., Peter, H., Riashat, I., Marc, G.B., Joelle, P. (2018). An introduction to deep reinforcement learning. *Foundations and Trends® in Machine Learning*, 11(3-4): 219-354. <https://doi.org/10.1561/22000000071>
- [13] Barbalho, P.I.N., Lacerda, V.A., Fernandes, R.A.S., Coury, D.V. (2022). Deep reinforcement learning-based secondary control for microgrids in islanded mode. *Electric Power Systems Research*, 212: 108315. <https://doi.org/10.1016/j.epr.2022.108315>
- [14] Jeyaraj, P.R., Asokan, S.P., Kathiresan, A.C., Samuel Nadar, E.R. (2023). Deep reinforcement learning-based network for optimized power flow in islanded DC microgrid. *Electrical Engineering*, 105: 2805-2816. <https://doi.org/10.1007/s00202-023-01835-1>
- [15] Bidram, A., Davoudi, A. (2012). Hierarchical structure of microgrids control system. *IEEE Transactions on Smart Grid*, 3(4): 1963-1976. <https://doi.org/10.1109/TSG.2012.2197425>
- [16] Oboreh-Snapps, O., Strathman, S.A., Saelens, J., Fernandes, A., Kimball, J.W. (2024). Addressing reactive power sharing in parallel inverter islanded microgrid through deep reinforcement learning. In *2024 IEEE Applied Power Electronics Conference and Exposition (APEC)*, Long Beach, CA, USA, pp. 2946-2953. <https://doi.org/10.1109/APEC48139.2024.10509093>
- [17] Zhang, M., Guo, G., Magnússon, S., Pilawa-Podgurski, R.C.N., Xu, Q. (2024). Data driven decentralized control of inverter-based renewable energy sources using safe guaranteed multi-agent deep reinforcement learning. *IEEE Transactions on Sustainable Energy*, 5(2): 1288-1299. <https://doi.org/10.1109/TSTE.2023.3341632>
- [18] Lin, S.W., Chu, C.C., Tung, C.F. (2023). Data-driven distributed Q-learning droop control for frequency synchronization and voltage restoration in isolated AC micro-grids. *IEEE Transactions on Industry Applications*, 59(6): 7306-7317. <https://doi.org/10.1109/TIA.2023.3300290>
- [19] Chaturvedi, S., Bui, V.H., Su, W., Wang, M. (2024). Reinforcement learning-based integrated control to improve the efficiency of DC microgrids. *IEEE Transactions on Smart Grid*, 15(1): 149-159. <https://doi.org/10.1109/TSG.2023.3286801>
- [20] Jin, P., Qin, J., Ma, C., Sha, G., Cong, X. (2025). Comparing SAC and DQN in microgrid EMS: Baseline results under fixed electricity price. In *2025 10th International Conference on Power and Renewable Energy (ICPRE)*, Hangzhou, China, pp. 930-933. <https://doi.org/10.1109/ICPRE67300.2025.11274160>
- [21] Li, S., Wu, W., Guo, Z., Liu, H., Cheng, Y. (2025). Adaptive data-driven secondary control for renewables dominated islanded microgrids and field tests. *IEEE Transactions on Smart Grid*, 16(5): 3548-3562. <https://doi.org/10.1109/TSG.2025.3574483>
- [22] Fan, Z., Zhang, W., Liu, W. (2023). Multi-agent deep reinforcement learning-based distributed optimal generation control of DC microgrids. *IEEE Transactions on Smart Grid*, 14(5): 3337-3351. <https://doi.org/10.1109/TSG.2023.3237200>
- [23] Li, M., Wan, Z., Zou, T., Shen, Z., Li, M., Wang, C., Xiao, X. (2024). Artificial intelligence enabled self-powered wireless sensing for smart industry. *Chemical Engineering Journal*, 495: 152417. <https://doi.org/10.1016/j.cej.2024.152417>
- [24] Kwon, K.B., Mukherjee, S., Vu, T.L., Zhu, H. (2023). Risk-constrained reinforcement learning for inverter-

dominated power system controls. IEEE Control Systems Letters, 7: 3854-3859. <https://doi.org/10.1109/LCSYS.2023.3343948>

[25] Mohammadi, E., Alizadeh, M., Asgarimoghaddam, M., Wang, X., Simões, M.G. (2022). A Review on

application of artificial intelligence techniques in microgrids. IEEE Journal of Emerging and Selected Topics in Industrial Electronics, 3(4): 878-890. <https://doi.org/10.1109/JESTIE.2022.3198504>