



Adaptive Control of VSG Inertia Damping Based on MADDPG

Demu Zhang ¹, Jing Zhang ¹ , Yu He ^{1,*} , Tao Shen ¹ and Xingyan Liu ²¹ College of Electrical Engineering, Guizhou University, Guiyang 550025, China;

gs.dmzhang22@gzu.edu.cn (D.Z.); zhangjing@gzu.edu.cn (J.Z.); gs.tshen22@gzu.edu.cn (T.S.)

² Power Grid Planning and Research Center of Guizhou Power Grid Co., Ltd., Guiyang 550002, China;

liuxy0916@gz.csg.cn

* Correspondence: yhe7@gzu.edu.cn

Abstract: As renewable energy sources become more integrated into the power grid, traditional virtual synchronous generator (VSG) control strategies have become inadequate for the current low-damping, low-inertia power systems. Therefore, this paper proposes a VSG inertia and damping adaptive control method based on multi-agent deep deterministic policy gradient (MADDPG). The paper first introduces the working principles of virtual synchronous generators and establishes a corresponding VSG model. Based on this model, the influence of variations in virtual inertia (J) and damping (D) coefficients on fluctuations in active power output is examined, defining the action space for J and D . The proposed method is mainly divided into two phases: “centralized training and decentralized execution”. In the centralized training phase, each agent’s critic network shares global observation and action information to guide the actor network in policy optimization. In the decentralized execution phase, agents observe frequency deviations and the rate at which angular frequency changes, using reinforcement learning algorithms to adjust the virtual inertia J and damping coefficient D in real time. Finally, the effectiveness of the proposed MADDPG control strategy is validated through comparison with adaptive control and DDPG control methods.

Keywords: VSG; multi-agent; deep deterministic policy gradient; frequency control



Citation: Zhang, D.; Zhang, J.; He, Y.; Shen, T.; Liu, X. Adaptive Control of VSG Inertia Damping Based on MADDPG. *Energies* **2024**, *17*, 6421. <https://doi.org/10.3390/en17246421>

Academic Editor: Frede Blaabjerg

Received: 4 November 2024

Revised: 18 December 2024

Accepted: 18 December 2024

Published: 20 December 2024



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

As the transition to green energy continues to deepen, the proportion of renewable energy sources and power electronic devices in the power system is increasing. Consequently, the disadvantages of low inertia and low damping in the system have become more pronounced [1,2]. The VSG technology, since its inception, has been widely employed to address the issues of low inertia and low damping in renewable energy power systems [3,4]. However, due to the strong intermittency and uncertainty of new energy generation, coupled with its increasing proportion, the parameters of traditional VSG control strategies, which are fixed, may not be suitable for all operational scenarios of renewable energy systems with a high proportion of renewables. Therefore, researching effective adaptive control methods for VSG is crucial for the development of systems with a high proportion of renewable energy [5,6].

Scholars have achieved remarkable results in the research on VSG parameter adaptation. Ref. [7] constructs a VSG control strategy with adaptive droop coefficients that enhance transient stability and frequency support. Additionally, some scholars have divided the oscillation period into four regions based on the active power–frequency curve and have constructed different adaptive control functions to accommodate the distinct power and angle change characteristics of each region [8]. Refs. [9,10] propose an improved virtual inertia-damping adaptive control method that does not rely on phase-locked loops for measurement. This approach helps avoid stability issues in weak grids and enhances the method’s adaptability within the power grid. In Ref. [11], by analyzing the response characteristics of power and angular frequency, and based on the principles of parameter design,

scholars have determined the range of parameter action and proposed a dual-adaptive inertia control strategy. Traditional adaptive control strategies are constructed based on the analysis of the variation rules of J and D and are function-based for adaptive adjustment; although this adaptive control design rule is flexible and variable, capable of appropriately adjusting parameters for different operational scenarios, the complexity arises from the multitude of control parameters and their intricate interrelationships. Additionally, the setting of certain parameters necessitates regulatory experience, which leads to difficulties in the design of adaptive control methods. To address this issue, some scholars have explored the use of intelligent algorithms to dynamically improve VSG parameter control, such as the whale optimization algorithm [12], particle swarm optimization [13], and radial basis function neural network algorithms [14]. The methods design adaptive control rules that are flexible and variable, capable of adjusting parameters in real time for VSG control based on target function design. Experience in parameter setting and selection is not required; however, the high degree of uncertainty and intermittency inherent in renewable energy sources makes precise modeling of renewable energy power systems challenging. Yet, the effectiveness of VSG control methods based on intelligent algorithms heavily relies on the accurate modeling of the power system, a dependency that does not align well with the inherent uncertainty of renewable energy systems. Reinforcement learning, on the other hand, does not depend on an exact system model and learns the optimal strategy through interaction with the environment [15,16]. Refs. [17,18] have investigated a novel twin delayed deep deterministic policy gradient (TD3)-based adaptive control strategy for VSG. This strategy is capable of stabilizing and rapidly compensating for uncertainties in the system, effectively enhancing the robustness and interference resistance of the system. Ref. [19] utilizes a DDPG agent to output active power compensation actions. This method is capable of enhancing the control performance of VSG, even under the traditional fixed-parameter control of VSG. In Ref. [20], by employing TD3, the study manages to optimize the reactive power reference for each VSG, minimizing the reactive-power-sharing error and ensuring voltage stability, which is crucial for the reliable operation of power systems. Refs. [21,22] apply reinforcement learning to VSG control strategies, where the former optimizes the maximum frequency deviation and the rate of frequency change using the proposed approach, and the latter applies the proposed method to multi-VSG systems, enhancing the transient stability of the power system and suppressing fault currents. In Ref. [23], an agent-based DDPG algorithm is designed to enable VSG to cooperatively adjust the moment of J and D of the system. In Ref. [24], a novel deep reinforcement learning control strategy is proposed for simultaneously regulating the frequency of parallel VSGs in an isolated microgrid while minimizing their active power sharing errors. The aforementioned references introduce reinforcement learning, where the agent decides the next action based on the current state and updates the strategy based on the reward after executing the action. Through the two stages of offline training and online application, they solve the problems of model dependence and the complexity of design rules. However, when a single agent in reinforcement learning outputs both J and D simultaneously, the computational complexity is high, and the coordination between control parameters is weak. In Ref. [25], the feasibility of integrating multi-agents into virtual power plants was studied. The paper indicates that in a multi-agent environment, the control and operation of virtual synchronous generators still need to be refined. Moreover, the existing research on multi-agents concerning VSGs is mostly reflected in distributed control of multiple VSGs [26,27], and there is still a lack of research on the application of multi-agents in the adaptive output control of a single VSG with multiple parameters.

Building upon this, the present paper proposes a VSG inertia and damping adaptive control method based on MADDPG. This method, through a multi-agent cooperative mechanism, achieves adaptive parameter adjustment of VSG in a multi-agent system, significantly enhancing the dynamic response capability and stability of the power grid. The online learning characteristic of the MADDPG algorithm enables it to effectively train agents and optimize control strategies in the absence of an exact power grid model through

experience replay and gradient backpropagation mechanisms. This approach is suitable for addressing the continuous variation issues of frequency and power in VSG control, thereby enhancing the grid's robustness to active power fluctuations. Through simulation verification in MATLAB/Simulink, the VSG control algorithm based on MADDPG demonstrates its effectiveness in improving the grid's frequency and active power output, making a significant contribution to the enhancement of power system stability and reliability.

The following Table 1 shows a performance comparison of the various VSG control methods mentioned in this paper. Table 2 below illustrates and explains the variables used in this paper.

Table 1. Control method difference table.

Control Method	Frequency Nadir	Settling Time	Scenario Applicability	Model Dependency	Design Complexity	Synergy	References
MADDPG	✓	✓	✓	✓	✓	✓	[26,27]
DDPG	✓	✓	✓	✓	✓		[21,23]
Intelligent algorithm optimization	✓	✓	✓			✓	[12,13]
Function-based adaptive control	✓	✓	✓				[6–8]
Fixed-parameter control				✓	✓		[3]

Table 2. Symbol description.

Nomenclature		Nomenclature	
U_{dc}	DC input power source	L_S, R_S	stator armature resistance
R_S	synchronous inductance of the synchronous generator	u_{oa}, u_{ob}, u_{oc}	induced electromotive force of the VSG
e_a, e_b, e_c	three-phase output terminal voltage	i_a, i_b, i_c	grid-connected current
J	virtual moment of inertia	D	damping coefficient
D	damping coefficient	P_e	VSG output power
ω	output angular frequency	ω_0	rated angular frequency
K_q	reactive power voltage droop coefficient	K_u	voltage regulation coefficient
Z	impedance of the filter circuit	T_m	mechanical torque
T_e	electromagnetic torque	T_D	damping torque
t_s	settling time	$\sigma\%$	overshoot percentage
θ'	parameters of the target critic network	ϕ'	target actor network
∇_θ	gradient for the parameters θ	s_i	state at the current time step
a_i	action at the current time step	y_i	Q-value estimated via the Bellman equation
γ	discount factor	∇_ϕ	gradient for the parameters ϕ
ΔP_{\max}	maximum active power deviation	$\Delta\omega_{\max}$	maximum angular frequency deviation

2. Introduction to the Principles

2.1. VSG Overall Topology

VSG technology mimics the active power frequency and reactive power voltage regulation traits of synchronous generators. This compensates for the lack of rotational inertia and damping in power electronic devices, thereby providing crucial support to the system. A VSG system is primarily composed of a main circuit and a control circuit. The main circuit includes distributed power sources, a three-phase inverter, and a filter circuit. For simplified analysis, the distributed power source is represented by an ideal voltage source. The control circuit mainly consists of several components, such as a virtual

governor, an excitation controller, and measurement and computation units. The system diagram of the VSG is shown in Figure 1.

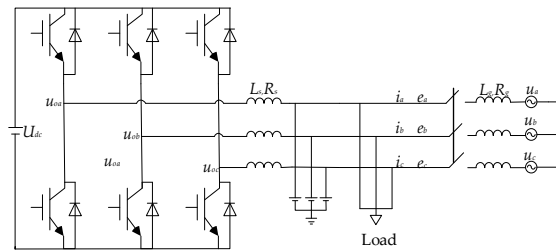


Figure 1. Topology diagram of the VSG principle.

The circuit topology and control of the virtual synchronous machine are illustrated in Figure 1. In this setup, U_{dc} represents the DC input power source, where L_s and R_s simulate the stator armature resistance and synchronous inductance of the synchronous generator, respectively. u_{0a}, u_{0b}, u_{0c} represent the induced electromotive force of the VSG. e_a, e_b, e_c denote the three-phase output terminal voltage, and i_a, i_b, i_c represent the grid-connected current. This figure illustrates the topology diagram of the VSG principle.

2.2. Control Strategies for VSG

VSG control mainly includes active power–frequency controllers and reactive power–voltage controllers. The active power–frequency controller is primarily based on the generator rotor equation, and its formula is as follows:

$$J \frac{d\omega}{dt} = \frac{1}{\omega} (P_T - P_e) - D(\omega - \omega_0). \tag{1}$$

where J represents the virtual moment of inertia; D represents the damping coefficient; P_T and P_e are the input mechanical power and VSG output power, respectively; ω and ω_0 are the output angular frequency and the rated angular frequency, respectively.

The reactive power–voltage controller is primarily based on the generator rotor equation, and its formula is as follows:

$$E = (U_{set} - U)K_u + (Q_{set} - Q)K_q + E_0. \tag{2}$$

where K_q and K_u are the reactive power voltage droop coefficient and the voltage regulation coefficient, respectively.

According to Figure 1, the output power of the VSG is calculated as

$$\begin{cases} P_e = \frac{EU \cos(\varphi - \sigma)}{Z} - \frac{U^2 \cos \varphi}{Z} \\ Q = \frac{EU \sin(\varphi - \sigma)}{Z} - \frac{U^2 \sin \varphi}{Z} \end{cases}. \tag{3}$$

where Z represents the impedance of the filter circuit.

The equivalent circuit model of the VSG grid-connected system is shown in Figure 2.

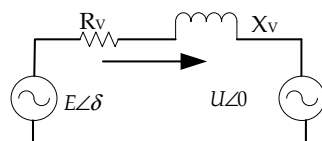


Figure 2. Equivalent circuit model.

In the equivalent circuit, the resistance and reactance are represented by R_v and X_v , respectively. Since the simulation environment is a high-voltage power grid, the resistance in the circuit is considered negligible, represented by $U \approx E$ and $\sin \sigma = \sigma$. The formula for the active power output of the VSG can be expressed as follows:

$$P = \frac{EU}{X_V} \sin \sigma \approx \frac{EU}{X_V} \sigma, \quad (4)$$

By combining Equations (1) and (2), the closed-loop second-order transfer function expression for the VSG active power output loop can be obtained.

$$G(s) = \frac{P(s)}{P_{ref}(s)} = \frac{EU}{JX_V\omega_0s^2 + (k_p + D\omega_0)X_Vs + EU} \quad (5)$$

Based on the previous equation, the natural angular frequency and the damping ratio are determined as follows:

$$\begin{cases} \omega_n = \sqrt{\frac{EU}{J\omega_0X_V}} \\ \zeta = \frac{D}{2} \sqrt{\frac{\omega_0X_V}{JEU}} + \frac{k_p}{2} \sqrt{\frac{X_V}{J\omega_0EU}} \end{cases}, \quad (6)$$

2.3. The Impact of Rotational Inertia and Damping Coefficient on the Dynamic Characteristics of the VSG

According to Equation (1), it can be derived that

$$\Delta\omega = \frac{\frac{P_T - P_e}{\omega} - J \frac{d\omega}{dt}}{D} = \frac{T_m - T_e - J \frac{d\omega}{dt}}{D} \quad (7)$$

$$T_D = D(\omega - \omega_0) \quad (8)$$

where T_m , T_e , and T_D represent the mechanical torque, electromagnetic torque, and damping torque of the synchronous generator, respectively. In the formula, if the value of $T_m - T_e - J \frac{d\omega}{dt}$ remains constant, the larger the value of J , the smaller the rate of change in the angular frequency. Based on this principle, the VSG can adjust J and D to change the angular frequency deviation and the rate of change in the angular frequency, thereby achieving control effects.

In the control theory, it is known that the settling time t_s and the overshoot percentage $\sigma\%$ are important indicators for measuring the dynamic performance of a system. When the system is underdamped ($0 < \zeta < 1$), and an error band $\Delta = 2\%$ is chosen, then

$$\begin{cases} \sigma\% = e^{\frac{-\pi\zeta}{\sqrt{1-\zeta^2}}} \times 100\% \\ t_s = \frac{4}{\zeta\omega_n} \end{cases}, \quad (9)$$

Based on Equations (5), (6), and (9), this paper will further analyze the impact of J and D on the system and construct an appropriate reward function accordingly. The simulations in this paper tested the dynamic response characteristics of the active power output of the VSG under an output power of 6 kW, with different J and D . The simulation results are shown in Figures 3 and 4, and the analysis will help evaluate the impact of these parameters on the system's dynamic performance.

Based on the simulation results, it can be observed that as the virtual inertia increases and the damping coefficient decreases, the system's damping ratio declines. This leads to an increase in the frequency peak and a longer settling time, which is detrimental to the stable operation of the power system. As the damping coefficient gradually increases, the system's damping ratio correspondingly improves, which reduces overshoot and shortens the control time. The overshoot of active power diminishes gradually, whereas the rise time lengthens. This leads to the conclusion that J has a significant impact on the

frequency of system oscillations, while D mainly affects the decay rate of these oscillations. Adjusting either J or D individually cannot fully optimize active power, frequency, and the overall system response speed. Therefore, VSG adaptive control needs to manage both J and D simultaneously. A single agent may struggle to manage this complexity, as it must independently handle all parameters and make decisions, potentially leading to inefficiencies. As a result, this paper utilizes the MADDPG algorithm to improve VSG control and seek an optimal control strategy.

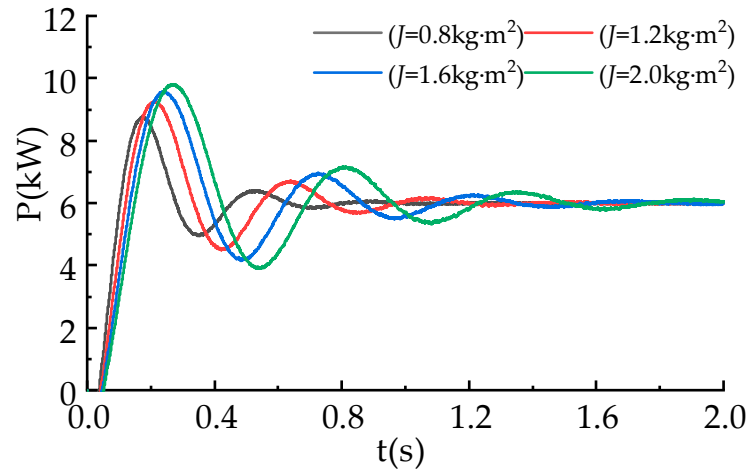


Figure 3. Active power response curve (J varies).

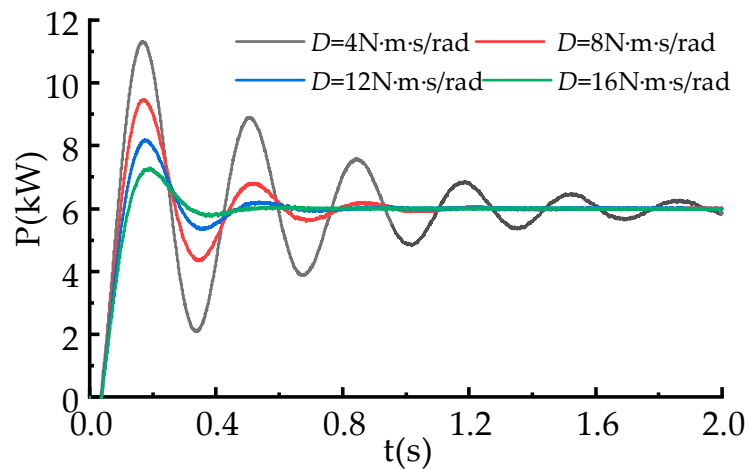


Figure 4. Active power response curve (D varies).

To analyze the impact of J and D on system stability, the linearization of Equations (1), (2), and (4) yields the small-signal model of the VSG power controller as follows:

$$\begin{cases} \hat{\omega} = (\hat{P}_{set} - \hat{P}_e) / [\omega_0(Js + D)] \\ \hat{\sigma} = \hat{\omega} / s \\ \hat{E} = (\hat{Q}_{set} - \hat{Q})K_q / [\sqrt{2}(1 + K_u)] \\ \hat{P}_e = (EU\hat{\sigma} + U\sigma\hat{\sigma}) / X_V \\ \hat{Q} = (U\hat{E} - EU\sigma\hat{\sigma}) / X_V \end{cases} \quad (10)$$

Based on Equation (10), the small-signal model of the power controller in the s -domain can be depicted. The active and reactive power control loops of the VSG are shown in Figure 5 as follows:

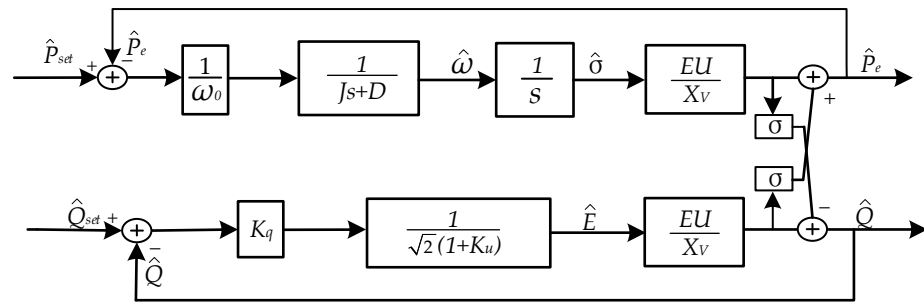


Figure 5. Active and reactive power control loop.

In the active power control loop, the system calculates the frequency deviation by comparing the set active power with the actual output. It then converts this frequency deviation into angular velocity deviation using inertia and damping elements, along with integrators. The reactive power control loop adjusts the reactive power through a proportional gain. After processing using a low-pass filter, the loop also controls the reactive power output via voltage regulation. Together, these two control loops work in concert to achieve precise control over both the active and reactive power in the power grid.

By approximately decoupling the active power loop from the reactive power loop in the aforementioned figure, the closed-loop characteristic equation for the active power response can be obtained as follows:

$$1 + G(s)H(s) = s^2 + \frac{D}{J}s + \frac{EU}{J\omega_0 X_v}, \quad (11)$$

3. Optimization of VSG Parameters Based on MADDPG

3.1. Principles of MADDPG

To eliminate the dependency of controllers on the accuracy of power grid models and past regulatory experience, this paper adopts a VSG parameter of adaptive control strategy based on multi-agent reinforcement learning. Agents learn through interaction with a dynamic environment and by engaging in a process of trial and error. During each training episode, the agent observes the current state of the environment and chooses a suitable action, which then leads to a transition to a new state. The agent receives feedback and evaluates the impact of its actions on state changes, thereby gradually optimizing its policy.

The DDPG algorithm combines policy gradient methods with Q-learning, effectively addressing challenges in high-dimensional continuous action spaces and expanding the application domain of reinforcement learning [28]. In the DDPG algorithm, the policy function is represented by the actor network, while the Q-function is realized by the critic network. These networks update their parameters through continuous interaction with the environment. To enhance training stability, the DDPG algorithm introduces the concept of main and target networks. The main network computes the actions or Q-values at the current time step, while the target network computes the actions and Q-values for the next time step. The parameters of the target network gradually approach those of the main network through a mechanism known as soft updating, with the parameter update formula calculated by Equation (11).

$$\begin{aligned} \theta' &= \tau\theta + (1 - \tau)\theta' \\ \phi' &= \tau\phi + (1 - \tau)\phi' \end{aligned} \quad (12)$$

where θ' , ϕ' represent the parameters of the target critic network and the target actor network, respectively. θ and ϕ denote the parameters of the main critic network and the main actor network, respectively.

The output of the primary critic network is the state-action Q-value, and its parameter update objective is to minimize the loss function, with parameters updated using the

gradient descent method. The loss function of the primary critic network is calculated using Equation (12), and its gradient can be computed using Equation (13):

$$L(\theta) = \frac{1}{N} \sum_i (Q_\theta(s_i, a_i) - y_i)^2 \quad (13)$$

$$\nabla_\theta L(\theta) = \frac{1}{N} \sum_i 2(Q_\theta(s_i, a_i) - y_i) \nabla_\theta Q_\theta(s_i, a_i). \quad (14)$$

where N represents the sampled mini-batch size; s_i and a_i denote the state and action at the current time step; and ∇_θ represents the gradient for the parameters θ . $Q_\theta(s_i, a_i)$ is the Q-value from the primary network at the current time step, and y_i is the Q-value estimated via the Bellman equation, which can be calculated using Equation (13):

$$y_i = r(s_i, a_i) + \gamma Q_{\theta'}(s_{i+1}, \mu_{\phi'}(s_{i+1})) \quad (15)$$

where $r(s_i, a_i)$ represents the reward at the current time step, and γ is the discount factor. Additionally, $Q_{\theta'}(s_{i+1}, \mu_{\phi'}(s_{i+1}))$, s_{i+1} , and $\mu_{\phi'}(s_{i+1})$ denote the target network Q-value, state, and action at the next time step, respectively.

The primary actor network functions as the policy network for the agent, generating action commands based on environmental observations. The goal of optimizing the neural network parameters is to maximize the Q-value, which is achieved by updating the parameters through the gradient ascent method. The objective function of the primary actor network can be computed as specified in Equation (15).

$$J(\phi) = \frac{1}{N} \sum_i Q(s_i, a_i) \Big|_{a_i = \mu_\phi(s_i)} \quad (16)$$

Its gradient is computed using Equation (16):

$$\nabla_\phi J(\phi) = \frac{1}{N} \sum_i \nabla_{a_i} Q(s_i, a_i) \Big|_{a_i = \mu_\phi(s_i)} \nabla_\phi \mu_\phi(s_i) \quad (17)$$

where $\mu_\phi(s_i)$ represents the action value produced by the primary actor network based on the current state, and ∇_ϕ denotes the gradient for the parameters ϕ .

MADDPG extends the actor–critic framework of DDPG to multi-agent environments. In these environments, the behavior of each agent is influenced not only by the state of the environment but also by the policies of other agents [29]. This interdependence and interaction render the learning and optimization of multi-agent systems complex and challenging. To effectively handle this complexity, MADDPG introduces a “centralized training and decentralized execution” framework. During the training phase, MADDPG utilizes a centralized training approach. Each agent’s critic network relies not only on its own state and actions but also takes into account the joint states and actions of all other agents. Through this centralized critic network, agents are better equipped to capture the complex interactions within the multi-agent system, thereby providing more accurate value estimations.

MADDPG achieves decentralized decision making by implementing an independent actor–critic architecture for each agent. Each agent possesses its own actor and critic networks, enabling it to make decisions based on local observations. This decentralized execution ensures that each agent relies exclusively on its own observational information during operation, thereby enhancing the system’s flexibility and efficiency in practical applications.

The parameter update formulae for each agent’s neural network in MADDPG are essentially identical to those in DDPG. However, due to the incorporation of a centralized critic network that considers not only the agent’s own state and actions but also those of other agents, the equation for computing the Q-value is modified as follows:

$$Q_i^j = Q_{\theta_j}(x_i, a_i^1, \dots, a_i^j, \dots, a_i^m). \quad (18)$$

where $a_i^1, \dots, a_i^j, \dots, a_i^m$ represents the actions of all agents at time i ; m denotes the number of agents; and x_i is the set of observations from all agents.

$$x_i = (s_i^1, \dots, s_i^j, \dots, s_i^m) \quad (19)$$

Therefore, with the modification of the Q-value calculation formula, the target functions and gradient calculation formulae for updating the parameters of the critic and actor networks in MADDPG also change accordingly. Under the modified formulae

$$\begin{aligned} L(\theta_j) &= \frac{1}{N} \sum_i (Q_{\theta_j}(x_i, a_i^1, \dots, a_i^j, \dots, a_i^m) - y_i)^2 \\ \nabla_{\theta_j} L(\theta_j) &= \frac{1}{N} \sum_i 2(Q_{\theta_j}(x_i, a_i^1, \dots, a_i^j, \dots, a_i^m) - y_i) \nabla_{\theta_j} Q_{\theta_j}(x_i, a_i^1, \dots, a_i^j, \dots, a_i^m) \\ J(\phi_j) &= \frac{1}{N} \sum_i Q_{\theta_j}(x_i, a_i^1, \dots, a_i^j, \dots, a_i^m) \Big|_{a_i^j = \mu_{\phi_j}(s_i^j)} \\ \nabla_{\phi_j} J(\phi_j) &= \frac{1}{N} \sum_i \nabla_{a_j} Q_{\theta_j}(x_i, a_i^1, \dots, a_i^j, \dots, a_i^m) \Big|_{a_i^j = \mu_{\phi_j}(s_i^j)} \nabla_{\phi_j} \mu_{\phi_j}(s_i^j) \end{aligned} \quad (20)$$

where θ_j and ϕ_j represent the parameters of the critic and actor networks of the agent j , respectively. s_i^j denotes the observation of the agent j at time step i .

3.2. Adaptive Control Strategy Based on MADDPG-VSG

MADDPG is categorized into completely cooperative, mixed-task, and completely competitive categories based on the relationships among agents. In the completely cooperative category, the core idea is that all agents share the same maximization reward as their goal, which is particularly applicable in VSG parameter tuning.

In the VSG parameter optimization based on the DDPG algorithm, the core objective is to train the intelligent agent's neural network to learn how to adaptively adjust the J and D to achieve frequency control. This process primarily consists of four components: state space, action space, reward function, and neural network architecture.

(1) State Space: The state space encompasses the state information of the power system, including voltage, current, frequency, and other relevant metrics at sampling points. This state information determines the subsequent action. To achieve adaptive tuning of VSG parameters, it is essential to gather information closely related to frequency, frequency variations, and other pertinent factors.

(2) Action Space Design: The DDPG algorithm has been widely applied to solve continuous action problems. In VSG parameter optimization, the J and D are selected as action variables. However, improper settings of VSG parameters may result in issues such as decreased system stability and fluctuations in power output. Therefore, it is necessary to impose restrictions on the action space.

Without considering the output limitation of energy storage, inverters can only operate continuously within a frequency variation range of ± 0.5 Hz. At this time, the active power required for frequency regulation by the inverter should be 40–100% of its rated capacity.

$$\frac{0.4\Delta P_{\max}}{\omega_0 \cdot \Delta\omega_{\max}} \leq D \leq \frac{\Delta T_{\max}}{\Delta\omega_{\max}} = \frac{\Delta P_{\max}}{\omega_0 \cdot \Delta\omega_{\max}} \quad (21)$$

where ΔP_{\max} and $\Delta\omega_{\max}$ represent the maximum active power deviation and the maximum angular frequency deviation of the inverter output, respectively. In this paper, the selected rated capacity is 30 kV·A; therefore, the range of D values is [12.16, 30.4]. Considering the overshoot and adjustment time comprehensively, the damping ratio range is set to [0.7, 1].

$$\frac{D^2\omega_0}{2^2S_E} \leq J \leq \frac{D^2\omega_0}{1.4^2S_E} \quad (22)$$

The range of J values can be obtained as [0.03, 0.58].

(3) Reward Function Design: The fundamental concept of reinforcement learning is to identify the optimal policy by maximizing the reward; thus, the quality of the reward function directly impacts the training outcomes. In this paper, the reward function primarily consists of two components.

1. Frequency Deviation

To ensure control effectiveness and minimize the maximum frequency deviation during VSG control, the reward function is defined as follows:

$$r_f = \begin{cases} 0 & 0 < |\Delta f| \leq 0.005 \\ -\lambda_1|\Delta f| & 0.005 < |\Delta f| \leq 0.02 \\ -\lambda_2|\Delta f| & 0.02 < |\Delta f| \leq 0.05 \\ -\lambda_3|\Delta f| & 0.05 < |\Delta f| \leq 0.08 \\ -\lambda_4|\Delta f| & 0.08 < |\Delta f| \leq 0.1 \\ -\lambda_5|\Delta f| & |\Delta f| > 0.1 \end{cases} \quad (23)$$

where $\lambda_1 \sim \lambda_5$ is the frequency reward coefficient, and in this paper, it is selected to be 10, 20, 30, 40, and 50.

2. Steady-State Time

To achieve a rapid response time and stable control performance of the VSG, the steady-state time and response time are incorporated as indicators in the VSG reward function.

$$r_k = -5 \int_{t_0}^{t_f} \Delta f dt \quad (24)$$

By combining the rewards of both indicators, the reward function presented in this paper is defined as follows:

$$R = r_f + r_k \quad (25)$$

(4) Neural Network Structure

In the design of the VSG controller based on MADDPG, the neural network architecture plays a critical role in deep reinforcement learning. Neural networks can capture complex environmental features and behavioral patterns through hierarchical nonlinear transformations, thereby enhancing the decision-making capabilities of the agent. The neural network architecture proposed in this paper utilizes fully connected layers and the ReLU activation function. The advantage of fully connected layers lies in their generality and strong representational power, as they can effectively learn the high-dimensional mapping relationships between input data and output results. Additionally, the introduction of the ReLU activation function helps mitigate the vanishing gradient problem, improves training efficiency, and increases the likelihood of the network converging to the global optimum. This combination ensures network complexity while enhancing both training stability and performance.

When employing MADDPG to configure the parameters for the VSG, two agents output the J and D to the VSG controller. Each agent observes the environmental state and provides feedback to the other agent through the reward function. This paper introduces a multi-agent reinforcement learning method, based on traditional VSG control, to adaptively adjust the J and D to deal with the complexities and variability of the power system. The improved VSG control block diagram is shown below.

In Figure 6, it is clear that the VSG environment is composed of electromagnetic equations, virtual excitation control, virtual governor control, power calculation, proportional-resonant control, and a pulse-width modulation waveform generator. The VSG environ-

ment adjusts its J and D based on the actions selected by the agent through the actor network. These adjustments affect the output of the VSG, which in turn influences the frequency and voltage of the power grid.

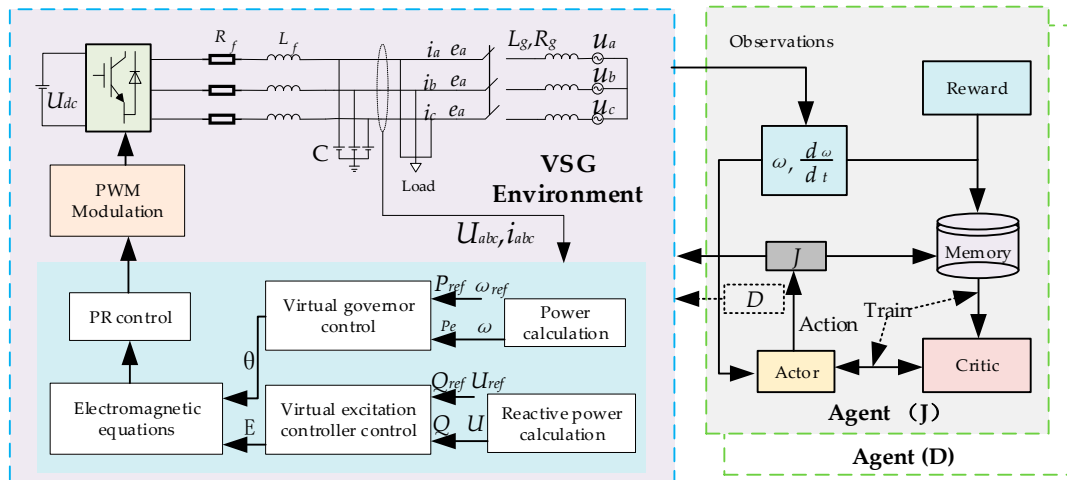


Figure 6. MDDPG parameters are self-adapting.

The agent observes the rate of change and the angular frequency within the VSG environment. It receives reward signals and assesses the value of these responses through the critic network. Using this information, the agent updates its policy to optimize control actions. The control actions of the agent influence the VSG environment, which then provides feedback to the agent. This feedback allows the agent to adjust its strategy, thus forming a closed-loop control system.

3.3. Pre-Training Stage

In deep reinforcement learning, the agent must undergo a pre-training phase characterized by random fluctuations before it can be used effectively. By continuously accumulating experience, the optimal value function Q-network can be developed. Based on this, this paper generates 40 s of random active power fluctuations using white noise and utilizes this dataset for training. The pre-training active power fluctuations are illustrated in Figure 7.

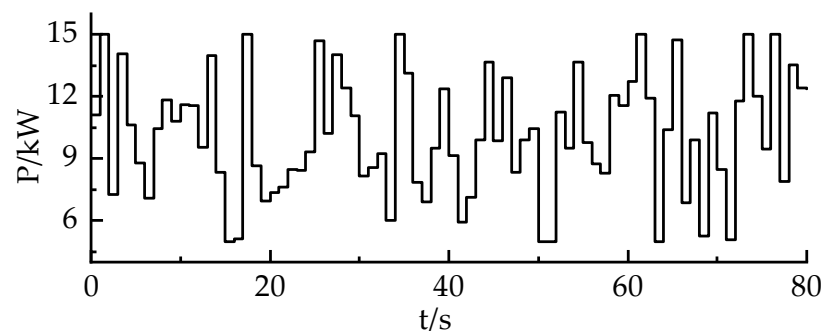


Figure 7. Pre-training active power fluctuation chart.

The results of the control performance after iterative training are presented in Figure 8. Based on the training results, it can be observed that when the agent encounters intermittent active power fluctuations over 80 s, the baseline active power is 10 kW, with an upper and lower fluctuation range of 5 kW, and the maximum frequency deviation is 0.147 Hz.

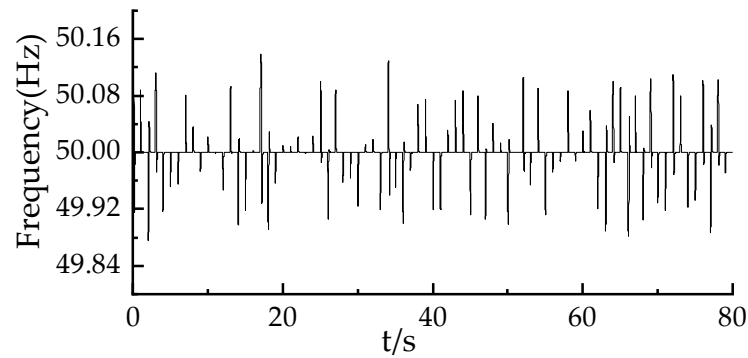


Figure 8. Training results chart.

4. Case Study

To verify the theoretical analysis presented earlier, as well as the correctness and superiority of the proposed control method, a VSG grid-connection model was constructed on the MATLAB/Simulink simulation platform for experimental validation [30]. Table 3 presents the parameters of the system settings.

Table 3. System parameters.

Parameter	Value	Parameter	Value
DC voltage	800 V	Filter capacitor	30
Line voltage	310 V	Filter inductor	3
Rated frequency	50	Resistance	0.1

To demonstrate the superiority of the control strategy presented in this paper, the same initial environment is established for the subsequent three control strategies. The maximum deviation from the reference power and the maximum deviation from the set frequency output are utilized as indicators for analyzing oscillation amplitude. Smaller values of these indicators indicate a stronger ability of the strategy to suppress oscillations.

Operating Condition One: The system is configured with an active power of 10 kW, which suddenly drops to 6 kW at 0.5 s, increases to 13 kW at 1 s, and then decreases to 8 kW at 1.6 s. The reactive power remains constant, and the simulation time is set to 2.4 s. The simulation results are presented in Figures 9 and 10 below.

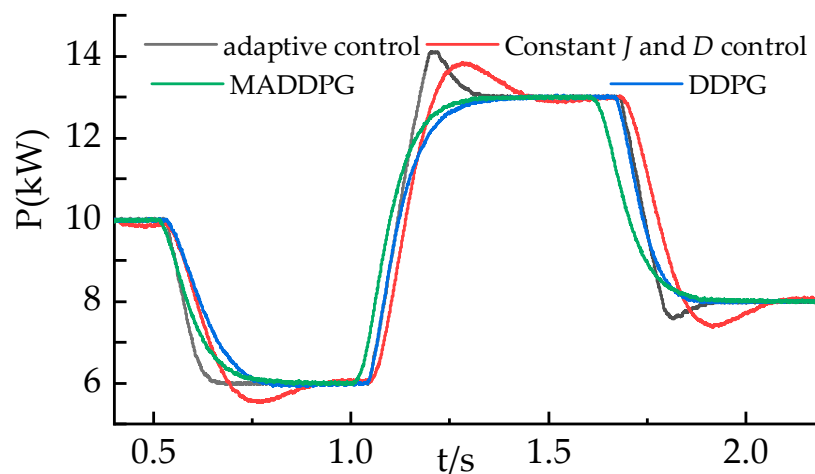


Figure 9. Case 1 frequency response curve chart.

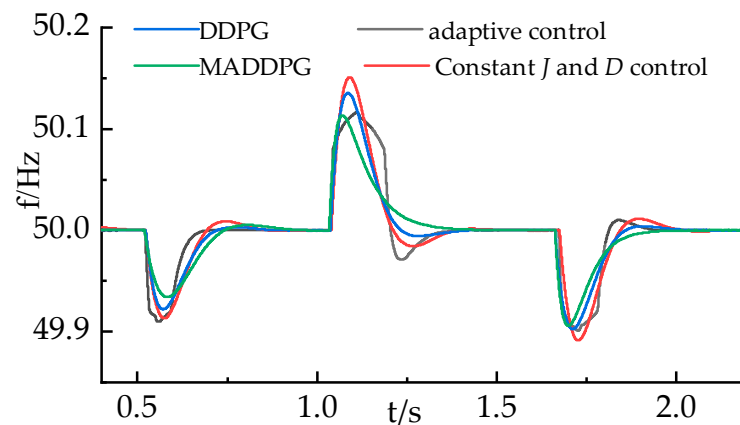


Figure 10. Case 1 power response curve chart.

To display the control effect more clearly and intuitively, the control performance indicators under different control modes are shown in Table 4.

Table 4. Case1 control performance comparison.

t/s	Control Strategy	Frequency Nadir	Settling Time	Overshoot
0.5	MADDPG	49.938	0.406	12.40%
	DDPG	49.921	0.383	15.80%
	adaptive control	49.913	0.260	17.40%
	Constant <i>J</i> and <i>D</i> control	49.909	0.464	18.20%
1	MADDPG	50.112	0.364	−22.40%
	DDPG	50.135	0.398	−27.00%
	adaptive control	50.117	0.375	−23.40%
	Constant <i>J</i> and <i>D</i> control	50.151	0.446	−30.20%
1.5	MADDPG	49.907	0.364	18.60%
	DDPG	49.903	0.390	19.40%
	adaptive control	49.898	0.358	20.40%
	Constant <i>J</i> and <i>D</i> control	49.891	0.407	21.80%

The results presented in the table indicate that when the active power input increases from 6 kW to 13 kW and then decreases from 13 kW to 8 kW, the MADDPG control strategy yields a smaller maximum frequency deviation compared to the other three control methods. This suggests that this control method has the potential to reduce frequency deviations, thereby minimizing the impact on power equipment and systems and enhancing overall system robustness. When confronted with an increase in active power fluctuations, the MADDPG method demonstrates smaller active power overshoot and shorter settling times compared to the other three control methods, thereby helping to reduce operational risks in the system.

The experimental results demonstrate that the proposed MADDPG control method significantly outperforms other control methods in reducing maximum frequency deviation when resolving active power fluctuations in VSGs. The frequency settling time is comparable to that of the adaptive control and DDPG control methods, indicating that the MADDPG-based VSG parameter optimization method effectively optimizes both the frequency and active response curves, thereby enhancing system robustness.

Operating Condition Two: The simulation grid-connection model is configured to experience a 0.1 Hz frequency fluctuation in the main power grid starting at 1 s and ending at 1.4 s. The total experimental simulation time is set to 2 s. The frequency fluctuation response results of the simulation experiment are presented in Figure 11 and Table 5.

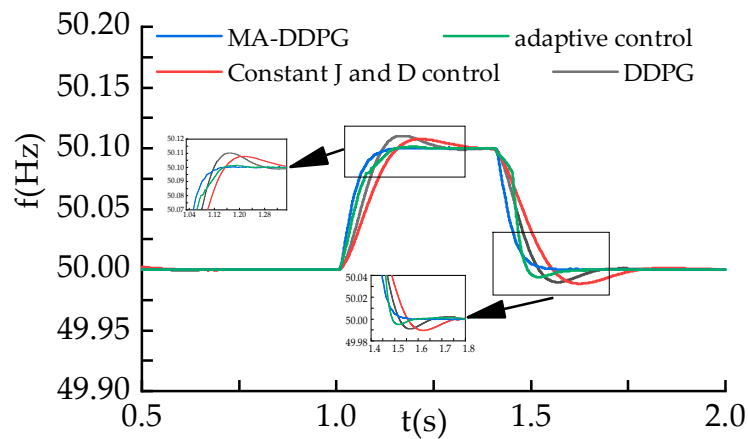


Figure 11. Case 2 frequency response curve chart.

Table 5. Case 2 control performance comparison.

t/s	Control Strategy	Frequency Nadir	Settling Time	Overshoot
1	MADDPG	50.101	0.140	0.2%
	DDPG	50.110	0.295	2%
	adaptive control	50.102	0.197	0.4%
	Constant <i>J</i> and <i>D</i> control	50.107	0.367	2.14%
1.4	MADDPG	49.998	0.172	−0.4%
	DDPG	49.989	0.187	−2.2%
	adaptive control	49.993	0.272	−1.4%
	Constant <i>J</i> and <i>D</i> control	49.987	0.384	−2.6%

By comparing the frequency response curves in the graph, it is evident that the control method proposed in this study outperforms the DDPG control, adaptive control, and constant *J* and *D* control methods in terms of frequency overshoot and settling time when addressing large grid frequency fluctuations.

Taking Case Study 1 as an example, the variation in *J* and *D* under the MADDPG control strategy is shown in Figure 12.

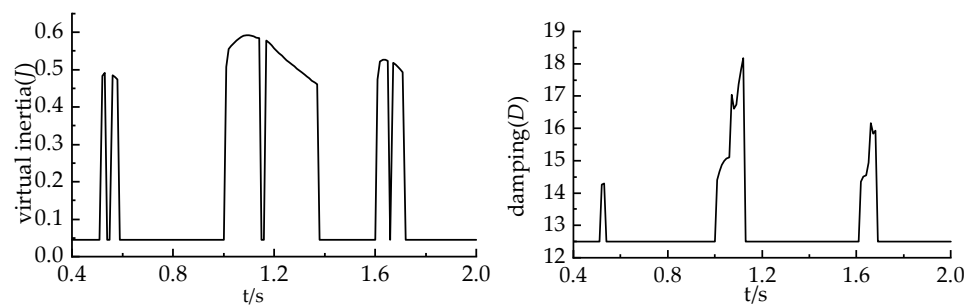


Figure 12. *J* and *D* variation chart.

As can be seen in Figure 12, by simulating the static stability regulation principle of a SG and using the MADDPG method for adaptive control of *J* and *D*, the VSG can significantly adjust in response to changes in observed quantities. This enables the VSG to truly achieve the effect of virtual governor control.

5. Conclusions

This paper presents an adaptive control method for VSG parameters based on MADDPG. The proposed method reduces the reliance on past control experience and eliminates the need for repetitive parameter settings in adaptive control. In contrast to intelligent algorithm-based control methods that require precise modeling of the power system, this

approach offers the advantage of not needing an accurate system model. The utilization of VSG in lower level control endows the equipment with inertia response and frequency regulation capabilities. The controller employs intelligent agents to monitor frequency deviation and the rate of frequency change and, based on the MADDPG algorithm, outputs J and D values in real time. The simulation experiments yield the following results:

- (1) The analysis examined the impact of J and D on active power. As the damping ratio increases, the system exhibits reduced active power overshoot, faster response speed, and longer control time.
- (2) Compared to traditional constant J and D control, as well as adaptive control methods and DDPG control, the MADDPG control method proposed in this study exhibits smaller frequency amplitudes and shorter settling times when tackling active power fluctuations. This effectively enhances the response curves of active power and frequency compared to traditional methods.
- (3) When confronted with large power grid fluctuations, the frequency deviation produced by the MADDPG control method proposed in this paper is smaller than that of traditional methods, and the return-to-steady time is also shorter.

Author Contributions: Methodology, D.Z. and Y.H.; Software, D.Z.; Resources, J.Z. and Y.H.; Writing—Original draft, D.Z.; Writing—Review and editing, D.Z., J.Z., Y.H., T.S. and X.L.; Supervision, J.Z. and Y.H.; Project administration, J.Z. and Y.H.; Funding acquisition, J.Z. and Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Foundation of Guizhou Province, grant numbers ([2022] general013, [2022] general014).

Data Availability Statement: The data are not available to the public.

Conflicts of Interest: Author Xingyan Liu was employed by the company Power Grid Planning and Research Center of Guizhou Power Grid Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Liu, J.; Miura, Y.; Bevrani, H.; Ise, T. A Unified Modeling Method of Virtual Synchronous Generator for Multi-Operation-Mode Analyses. *IEEE J. Emerg. Sel. Top. Power Electron.* **2021**, *9*, 2394–2409. [\[CrossRef\]](#)
2. Wu, Q.H.; Zhang, C.J.; Zhao, X.J.; Lin, H.W.; Zhang, X.Y.; Wang, F.X. Comparative Analysis and Improvement of Generalized Droop Control and Virtual Synchronous Generator for Rate of Change of Frequency Constraint and Transient Power Suppression. *Int. J. Circuit Theory Appl.* **2024**. [\[CrossRef\]](#)
3. Chen, Y.; Wang, K.; Tang, H.L.; Qi, Z.Y.; Tang, H.Y. Energy storage quasi-Z source photovoltaic grid-connected virtual impedance VSG control strategy considering secondary frequency regulation. *J. Power Electron.* **2024**. [\[CrossRef\]](#)
4. Zhang, Y.; Li, S.C.; Liu, Y.H.; Tang, Z.J.; Luo, B. Unconstrained optimization MPC method for qZSI-VSG grid-connected wind power system. *Int. J. Electr. Power Energy Syst.* **2024**, *162*, 110276. [\[CrossRef\]](#)
5. Li, K.X.; Wei, Y.Q.; Zhang, J.R. Adaptive Virtual Synchronous Generator Control Strategy Based on Frequency Integral Compensation. *Electronics* **2024**, *13*, 4318. [\[CrossRef\]](#)
6. Xing, D.F.; Tian, M.X. Analysis and Fuzzy Neural Networks-Based Inertia Coefficient Adjustment Strategy of Power Converters. *Concurr. Comput. Pract. Exp.* **2024**, *37*, e8311.
7. Liu, J.; Liu, X.Y.; Liu, J.J.; Li, X.J.; Wang, J.S. Adaptive-Droop-Coefficient VSG Control for Cost-Efficient Grid Frequency Support. *IEEE Trans. Power Syst.* **2024**, *39*, 6768–6771. [\[CrossRef\]](#)
8. Wang, M.J.; Cen, H.L.; Lu, M.; Li, J.B.; Zhao, M.; Li, J.Y. An Adaptive Control Strategy for VSG Parameters With Perturbation Segmentation. *Int. J. Circuit Theory Appl.* **2024**, *52*. [\[CrossRef\]](#)
9. Gurski, E.; Kuiava, R.; Perez, F.; Benedito, R.; Damm, G. A Novel VSG with Adaptive Virtual Inertia and Adaptive Damping Coefficient to Improve Transient Frequency Response of Microgrids. *Energies* **2024**, *17*, 4370. [\[CrossRef\]](#)
10. Behera, S.K.; Panda, A.K.; Naik, N. An Adaptive Control Approach for Improved Power Quality and Power Ripple Mitigation in a Self-Synchronized Grid-Tied VSG. *IEEE Trans. Ind. Electron.* **2024**, 1–12. [\[CrossRef\]](#)
11. Li, M.Y.; Huang, W.T.; Tai, N.L.; Yang, L.Q.; Duan, D.L.; Ma, Z.J. A Dual-Adaptivity Inertia Control Strategy for Virtual Synchronous Generator. *IEEE Trans. Power Syst.* **2020**, *35*, 594–604. [\[CrossRef\]](#)
12. Li, L.L.; Li, H.Y.; Tseng, M.L.; Feng, H.; Chiu, A. Renewable Energy System on Frequency Stability Control Strategy Using Virtual Synchronous Generator. *Symmetry* **2020**, *12*, 1697. [\[CrossRef\]](#)

13. Sun, X.S.; Cai, J.M.; Wang, D.S.; Lin, J.W.; Li, K. Small-disturbance stability analysis and control-parameter optimization of grid-connected virtual synchronous generator. *Front. Energy Res.* **2024**, *12*, 1428748. [[CrossRef](#)]
14. Yao, F.J.; Zhao, J.B.; Li, X.J.; Mao, L.; Qu, K.Q. RBF Neural Network Based Virtual Synchronous Generator Control With Improved Frequency Stability. *IEEE Trans. Ind. Inform.* **2021**, *17*, 4014–4024. [[CrossRef](#)]
15. Chen, P.; Zhao, J.F.; Liu, K.L.; Zhou, J.Y.; Dong, K.; Li, Y.F.; Guo, X.R.; Pan, X. A Review on the Applications of Reinforcement Learning Control for Power Electronic Converters. *IEEE Trans. Ind. Appl.* **2024**, *60*, 8430–8450. [[CrossRef](#)]
16. Egbomwan, O.E.; Liu, S.C.; 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](#)]
17. 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](#)]
18. Oboreh-Snapps, O.; She, B.; Fahad, S.; Chen, H.; Kimball, J.; Li, F.; Cui, H.; Bo, R. Virtual Synchronous Generator Control Using Twin Delayed Deep Deterministic Policy Gradient Method. *IEEE Trans. Energy Convers.* **2024**, *39*, 214–228. [[CrossRef](#)]
19. Li, C. Applications of Reinforcement Learning in Three-phase Grid-connected Inverter. In Proceedings of the 2023 IEEE 13th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), Qinhuangdao, China, 11–14 July 2023; pp. 580–585.
20. Oboreh-Snapps, O.; A Strathman, S.; Saelens, J.; Fernandes, A.; W Kimball, J. Addressing Reactive Power Sharing in Parallel Inverter Islanded Microgrid Through Deep Reinforcement Learning. In Proceedings of the 2024 IEEE Applied Power Electronics Conference and Exposition (APEC), Long Beach, CA, USA, 25–29 February 2024; pp. 2946–2953.
21. Benhmidouch, Z.; Moufid, S.; Ait-Omar, A.; Abbou, A.; Laabassi, H.; Kang, M.; Chatri, C.; Hammou Ou Ali, I.; Bouzekri, H.; Baek, J. A novel reinforcement learning policy optimization based adaptive VSG control technique for improved frequency stabilization in AC microgrids. *Electr. Power Syst. Res.* **2024**, *230*, 110269. [[CrossRef](#)]
22. Chen, L.; Tang, J.G.; Qiao, X.F.; Chen, H.K.; Zhu, J.H.; Jiang, Y.Q.; Zhao, Z.K.; Hu, R.Z.; Deng, X.Y. Investigation on Transient Stability Enhancement of Multi-VSG System Incorporating Resistive SFCLs Based on Deep Reinforcement Learning. *IEEE Trans. Ind. Appl.* **2024**, *60*, 1780–1793. [[CrossRef](#)]
23. Xiong, K.; Hu, W.H.; Zhang, G.Z.; Zhang, Z.Y.; Chen, Z. Deep reinforcement learning based parameter self-tuning control strategy for VSG. *Energy Rep.* **2022**, *8*, 219–226. [[CrossRef](#)]
24. Oboreh-Snapps, O.; Strathman, S.A.; Saelens, J.; Fernandes, A.; Morris, L.; Uddaraju, P.; Kimball, J.W. Simultaneous Frequency Regulation and Active Power Sharing in Islanded Microgrid Using Deep Reinforcement Learning. In Proceedings of the 2024 IEEE Kansas Power and Energy Conference (KPEC), Manhattan, KS, USA, 25–26 April 2024.
25. Xie, Y.J.; Zhang, Y.C.; Lee, W.J.; Lin, Z.L.; Shamash, Y.A. Virtual Power Plants for Grid Resilience: A Concise Overview of Research and Applications. *IEEE CAA J. Autom. Sin.* **2024**, *11*, 329–343. [[CrossRef](#)]
26. Yang, Q.F.; Yan, L.F.; Chen, X.; Chen, Y.; Wen, J.Y. A Distributed Dynamic Inertia-Droop Control Strategy Based on Multi-Agent Deep Reinforcement Learning for Multiple Paralleled VSGs. *IEEE Trans. Power Syst.* **2023**, *38*, 5598–5612. [[CrossRef](#)]
27. Hammad, E.; Farraj, A.; Kundur, D. On Effective Virtual Inertia of Storage-Based Distributed Control for Transient Stability. *IEEE Trans. Smart Grid* **2019**, *10*, 327–336. [[CrossRef](#)]
28. Fan, L.; Zhang, J.; He, Y.; Liu, Y.; Hu, T.; Zhang, H. Optimal Scheduling of Microgrid Based on Deep Deterministic Policy Gradient and Transfer Learning. *Energies* **2021**, *14*, 584. [[CrossRef](#)]
29. Wilk, P.; Wang, N.; Li, J. Multi-Agent Reinforcement Learning for Smart Community Energy Management. *Energies* **2024**, *17*, 5211. [[CrossRef](#)]
30. Gong, Z.; Su, Y.F.; Cai, R.B.; Rao, B.; Zhou, J. An adaptive control strategy for vsg based on energy storage capacity optimization of MMC-BESS. *Electr. Eng.* **2024**. [[CrossRef](#)]

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.