Deep Reinforcement Learning based Adaptive Control Strategy for Virtual Synchronous Generator

Xiaoke Ding
Department of Automation
Tsinghua University
Beijing China
ding_xiaoke@126.com

Junwei Cao*
Beijing National Research Center for Information Science and Technology
Tsinghua University
Beijing China
jcao@tsinghua.edu.cn

Xiawei Wang
Beijing Dahua Radio Instrument Co., Ltd.
Hai Dian District
Beijing China
wangxw@dhelec.com.cn

Abstract—Virtual Synchronous Generator (VSG) control strategy is a research hotspot at present, which simulates the operating characteristics of conventional generators and brings stability to renewable energy power system. To solve the difficulties in parameters tuning for a VSG, this paper presents a data-driven approach to meet achieve better frequency response against disturbances. In the proposed approach, the virtual moment of inertia and virtual Damping factor are provided by a reinforcement learning agent. In the reward function, the deviation of frequency, the rate of change of frequency (ROCOF) and the settling time are all considered simultaneously. To maximum the reward, this paper employs the Deep Deterministic Policy Gradient (DDPG) algorithm, whose action space is continuous. Finally, numerical validation in MATLAB/Simulink confirms the validity of the algorithm.

Key Words: virtual synchronous generator, deep reinforcement learning, deep deterministic policy gradient, virtual inertia, virtual damping, rate of change of frequency.

INTRODUCTION

With the extensive application of renewable energy generation and its characteristics of randomness and volatility, the stable operation of the power system is confronted with great challenges [1], [2]. The improve the stability of the system, the grid-forming control draws more and more attention both in scientific literature and industry-oriented research. The voltage-source converters (VSCs), which connect renewable energy and power systems, are able to form the system voltage and frequency under grid-forming control strategy. Various strategies are proposed, among whom virtual synchronous generator (VSG) is the hottest research spot.

As the name shows, VSG aims to mimic the inertia and damping characteristics of synchronous generators (SGs) [3], [4]. Not only that, an advantage of VSG compared to SG is its flexibility. In SGs, since the rotor mass is fixed, the available inertia is also fixed. While in VSG, since it is literally a control algorithm, which is executed by a software, its parameters can be made adaptive to the system design [5]. Since the dynamic response of the VSG is directly related to its parameters, research on parameter self-tuning has been an active area in recent years. Current work can be divided into two main categories: model-based and model-free methods.

The mode-based method depends on the accurate system model to develop a certain adaptive law for tuning. In [6], an adaptive virtual inertia control method is proposed to improve the dynamic performance of the system. In [7], both virtual inertia and damping are taken into consideration. The effect of control parameters on output power and frequency have been intuitively studied in [8] and a strategy with the features of short response time and small overshoot is proposed.

The mathematical relationship and the interaction between a VSG and other part of a grid are often complicated. And the exact model may not work well if the structure is changed. Model-free methods, relying on data measurements instead of the model or structure of the system, have become an attractive alternative approach. An RBF network is applied to adjust the parameters in [9], which shows a good dynamic performance and verifies the powerful ability of machine learning (ML). An effective and hot research spot among all model-free methods is reinforcement learning (RL). In [10], Q-learning is adopted to adjust the VSG parameters when frequency changes. However, the state space and action space are discrete in Q-learning, which relies on a lookup table to store Q-value for each state-action pair. As a result, when state space and action space expands, its performance tends to degrade significantly. A DQN algorithm is proposed in [11], which replaces the Q-table with a neural network. But the action space remains discrete. A DDPG algorithm is applied in [12] but the reward function is designed without considering the rate of change of frequency.

Motivated by the above, this paper proposed a RL-based parameters self-tuning strategy for a VSG. The main contributions of the work are summarized below:

- A DRL based adaptive controller is designed, which adjusts the virtual inertia and damping flexibly when operating condition changes;
- Compared with traditional VSG control method, the proposed DDPG method shows better dynamic performance.
- The reward function of DDPG takes ROCOF into consideration and the results are compared with method in [12].

The rest of the paper is organized as follows. The structure and principle of VSG are introduced in section II. Section III provides a brief introduction of reinforcement learning and DDPG algorithm. In section IV, the results of simulation carried out in Simulink using proposed method is discussed, which verifies the effectiveness of the proposed strategy.

PRINCIPLE OF VIRTUAL SYNCHRONOUS GENERATOR

A. Distributed System with VSG Control

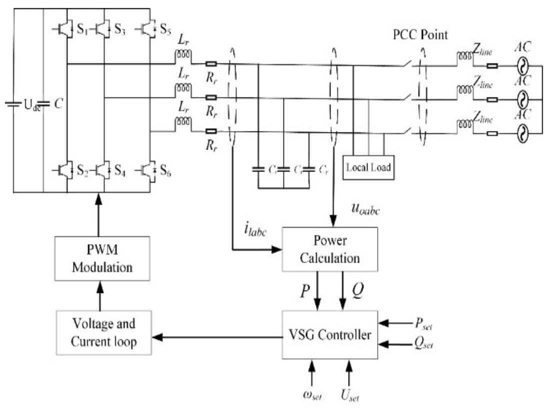


Fig. 2. VSG control system

The basic VSG control system is shown in Fig. 1, where L_r and C_r are respectively the filter inductance and filter capacitance of the inverter, Z_{line} is the line impedance of the inverter output to the AC bus, P and Q are respectively the active power and reactive power output. The power supply on the DC side represents distributed power supplies such as photovoltaic and energy storage system. And the AC side can either operate in off-grid mode to support local loads, or provide inertia to the grid in grid-connected mode.

The basic VSG control typically consists of an active power loop and a reactive power loop. The controller collects voltage and current information from the AC side and calculates them to obtain the reference signals required by PWM generation module to control the converter on and off.

B. Active and Reactive Power Loop

The virtual inertia is derived from the use of swing equation of a SG into the active power loop, which can be expressed as:

$$\begin{cases} P_m - P_e = J\omega_0 \frac{d\omega}{dt} + D\omega_0(\omega_0 - \omega) \\ \omega = \frac{d\theta}{dt} \end{cases} \quad (1)$$

where P_m and P_e are mechanic power and output active power, J is the moment of inertia, D is the damping coefficient, ω_0 and ω are rated angular frequency and output angular frequency of the VSG, respectively. By simulating the second-order model of the rotor, converters under VSG control can simulate the virtual moment of inertia and virtual damping when disturbance occurs. The control module established according to Formula (1) is shown in Fig. 2.

An important function of SG is to participate in the primary frequency modulation of the power grid, which is achieved by a governor. Fig. 3 shows the basic workflow of a traditional SG governor. When the frequency deviates from the rated frequency, the signal is transmitted to the governor, and the governor action changes the mechanical power input by the prime mover, thus adjusting the generator speed to track the rated frequency of the power grid.

The static regulation equation of SG can be expressed as:

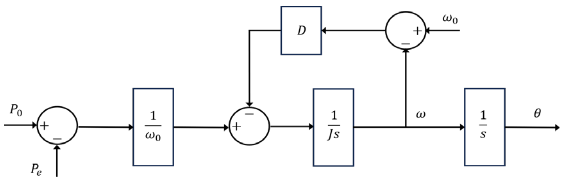


Fig. 3. Control model of VSG

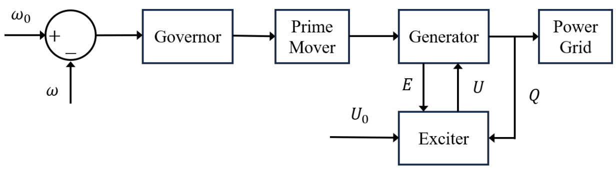


Fig. 1. SG governor workflow

$$K_f(\omega_0 - \omega) + (P_m - P_0) = 0 \quad (2)$$

where P_0 is active power reference, K_f is the frequency adjustment coefficient, which can be also regarded as the P-f droop coefficient.

By simulating the speed regulation of SG, VSG can participate in the primary frequency modulation of power system. And it is also conducive to the parallel operation of multiple virtual synchronous generators and the realization of active power distribution. The virtual governor module is shown in Fig. 4.

Regarding the simulation of the electromagnetic equation of the stator of synchronous generator, the researchers have different opinions. This paper focuses on the effects of virtual moment of inertia and virtual damping coefficient on active power and frequency. In order to reduce the complexity of the problem, electromagnetic equations are not considered. Therefore, the reactive power loop only adopts the reactive power-voltage droop control equation, which simplifies the model and ensures the ability of primary voltage regulation. It can be expressed as:

$$K_u(Q_0 - Q) + (U_0 - E) = 0 \quad (3)$$

where Q_0 and Q are the reactive power reference and reactive power output, K_u is the voltage adjustment coefficient as well as the *Q-V* droop coefficient. The reactive power control module is shown in Fig. 5.

Deep Deterministic Policy Gradient Algorithm

In this paper, an agent trained by DDPG algorithm is applied for parameter adaptive control of VSG. The key parameters J and D in VSG control are automatically adjusted in the operation by the agent.

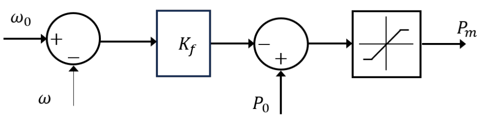


Fig. 4. Virtual governor module

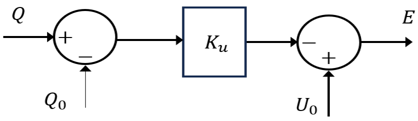


Fig. 5. Reactive power control module

A reinforcement learning problem can be described as a Markov Decision Process (MDP), which consists of a set of interacting objects, the agent and environment. A reinforcement learning agent can make decisions by sensing the state of external environment, take actions and adjust decisions through the feedback of the environment. The environment refers to all things outside the agent in the MDP model, the state of whom is changed by the actions of the agent. The changes can be fully or partially perceived by the agent. A positive change may give a reward to the agent while a negative one can bring a punishment, causing the agent to adjust accordingly.

MDP can be described as a tuple $\langle S, A, P, R\rangle$, where:

- $S = \{s_1, s_2, \dots, s_t\}$ stands for the state space, which can be either discrete or continuous. It refers to the state of environment observed by the agent. In this paper, the state S at each time step t can be described as $s_t = \left(P_t, \omega_t, \frac{d\omega_t}{dt}\right)$, where P_t is the active power output, ω_t is the output angular frequency, $\frac{d\omega_t}{dt}$ is the ROCOF.
- $A = \{a_1, a_2, \dots, a_t\}$ stands for the action space taken by agent, which can be either discrete or continuous. a_t refers to the action produced by agent decision at each time step t. After each action, the environment will enter the next state. The action A in this paper consists of the virtual moment of inertia J and the virtual damping D.
- P refers to the transfer probability function, which represents the possibility that the system moves to the next state after the agent takes a certain action in the current state.
- R refers to the reward function. By interacting with the environment, the agent will receive a reward for its action. A positive reward indicates the action is effective while a negative reward, also known as a punishment, indicates a wrong action. The goal of an agent is to maximize the expected future rewards by optimizing the policy. In this paper, the deviation of output angular frequency from rating and ROCOF are taken into consideration.

Besides the elements above, a policy $u(a_t|s_t)$ is a function that maps the state s_t to the action a_t. The ultimate goal of a DDPG algorithm is to find out a certain policy to maximize the benefits. The reward function is designed as:

$$r_t = -\alpha(\Delta\omega_t)^2 - \beta\left(\frac{d\omega_t}{dt}\right)^2 \tag{4}$$

$Q^u(s_t, a_t)$ is the action value function, which is used to evaluate the long-term benefits of taking action a_t in the current state s_t. It can be expressed as:

$$Q^u(s_t, a_t) = \mathbb{E}_u\left[R(s_t, a_t) + \gamma \max_{a_{t+1}} Q^u(s_{t+1}, a_{t+1})\right] \tag{5}$$

The implementation process of the DDPG algorithm in this paper is shown in Fig.6. The Actor network is a neural network with parameters θ^u that works as the policy $u(s_t)$ to decide how the network should act in current state. The Critic network is another neural network with parameters θ^Q to evaluate the action value function $Q^u(s_t, a_t)$.

Additionally, target actor network and target critic network are used in the algorithm to stabilize the process. It has been verified in [13] that learning with target networks performs better in many reinforcement learning tasks. The updating for parameters in target networks, denoted as $\theta^{u'}$ and $\theta^{Q'}$, slowly tracks the parameters in actor and critic network.

Algorithm 1 shows how the DDPG agent is trained. During the training process, the agent has no experience in how to act in the environment at the beginning. To prevent the agent from acting only within the explored action space, a noise decaying with time is added to encourage exploration. The predicted action a_t based on the current state s_t is applied to the environment and then the system goes to a new state s_{t+1}. The reward r_t is calculated by the consequence of taking the action a_t to measure how good the action is. Such sequence of events can be denoted as a tuple (s_t, a_t, r_t, s_{t+1}), which can be saved in a replay buffer $\boldsymbol{B}$ as an experience. Experiences are randomly sampled and used during the training process.

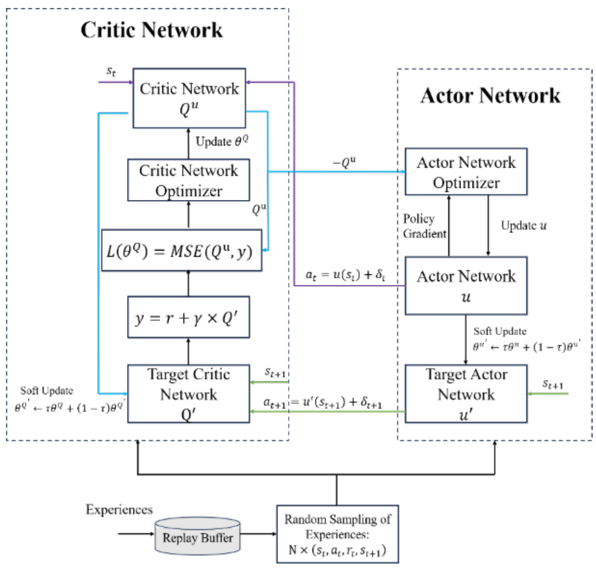


Fig. 6. Detailed DDPG agent

Algorithm 1: Deep Deterministic Policy Gradient

Initialize actor network $u(s_t|\theta^u)$ and critic network $Q^u(s_t, a_t|\theta^Q)$ with random weights θ^u and θ^Q.

Initialize target networks u' and Q' with $\theta^{u'} \leftarrow \theta^u$, $\theta^{Q'} \leftarrow \theta^Q$.

Initialize replay buffer $\boldsymbol{B}$.

for *episode = 1* to *M* **do**

Receive initial observation state s_1.

for *t=1* to *T* **do**

Select action $a_t = u(s_t|\theta^u) + \delta_t$ according to policy and exploration disturbance.

Calculate reward r_t. Observe new state s_{t+1}.

Store transition (s_t, a_t, r_t, s_{t+1}) in $\boldsymbol{B}$.

Randomly sample mini-batch of *N* transitions (s_i, a_i, r_i, s_{i+1}) from $\boldsymbol{B}$.

Set $y_i = r_i + \gamma Q'(s_{i+1}, u'(s_{i+1}|\theta^{u'})|\theta^{Q'})$.

Update critic network by minimizing loss: $L = \frac{1}{N}\sum_i \left(y_i - Q^u(s_i, a_i|\theta^Q)\right)^2$.

Update actor network by deterministic policy gradient:

$\nabla_{\theta^u} J = \frac{1}{N}\sum \nabla_a Q^u(s, \mathrm{a}|\theta^Q)|_{a=u(s_i)} \times \nabla_{\theta^u} u(s|\theta^u)\,|_{s=s_i}$.

Update the target networks: $\theta^{u'} \leftarrow \tau\theta^u + (1-\tau)\theta^{u'}$, $\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'}$.

end for

end for

PROPOSED METHODOLOGY

To verify the effectiveness of proposed method, a simulation built on MATLAB/Simulink platform is carried out in this section. The power system parameters are depicted in Table.1.

At the beginning, the system operates in a stable state with active power reference $P_0 = 13\ kW$ and reactive power reference $Q_0 = 0$. At $5s$, the active power reference is changed to $8\ kW$, causing the system enter a transient process. In order to verify the performance of proposed DDPG method, fixed parameter VSG control, model-based adaptive parameter VSG control and RBF method in [9] are all employed as comparison cases. Fig. 7 shows the active power response and frequency response with different controller when disturbance occurs. Compared with other methods, the DDPG algorithm proposed in this paper provides the best performance with the smallest overshoot and the least settling time.

Fig. 8 shows the different dynamic performance when ROCOF is taken into consideration or not. It can be derived from the active power response and frequency response that when ROCOF is considered, the curve changes more gently and smoothly, leading to a smaller overshoot and a relatively longer settling time.

TABLE I. PARAMETERS AOR A VSG SYSTEM

Parameters	Values
DC bus Voltage	1000 V
Filter inductance	2 mH
Filter capacitance	50 μF
Filter resistance	0.02 Ω
Rated frequency	50 Hz
Grid-side inductance	1.2 mH
Grid-side resistance	0.05 Ω

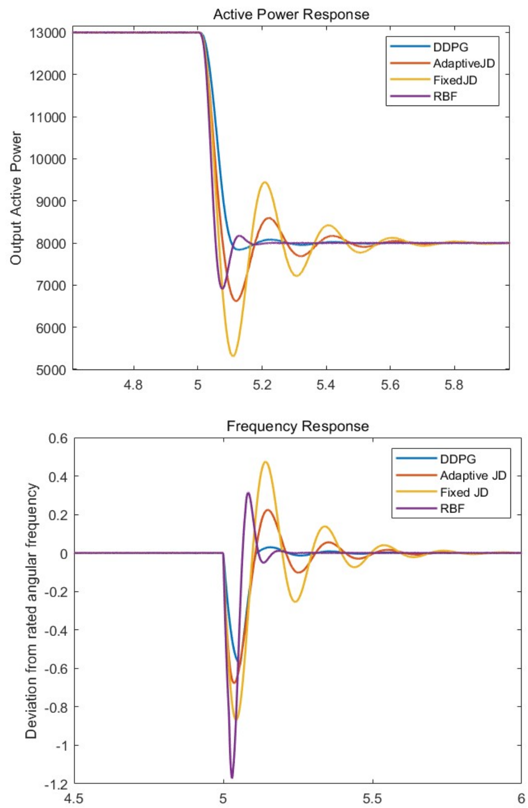
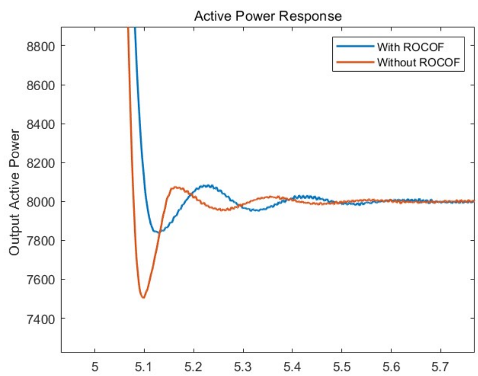


Fig. 7. System responses to active power reference change with different methods:(a) Active power response (b) Frequency



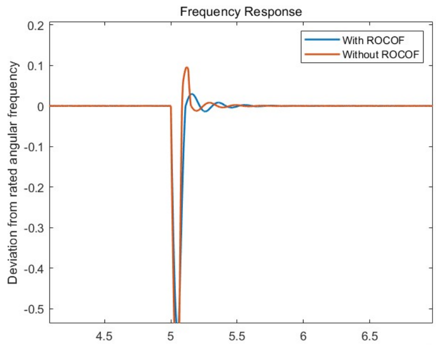


Fig. 8. System responses to active power reference change with different reward functions:(a) Active power response (b) Frequency response

CONCLUSION

This paper presents a DDPG algorithm based VSG controller to achieve parameters tuning and better dynamic performance. The superiority of proposed method is shown by comparing its response with other methods. Besides, ROCOF is taken into account when designing the reward function and the performance shows smaller overshoot in dynamic process compared with DDPG without taking ROCOF into consideration. In the future work, DDPG algorithm based VSG in islanded-mode and multiple VSGs operating in parallel should be further studied. The effects of different elements with different weights in reward function should be also explored.

[1] Lasseter, R.H., Z. Chen, and D. Pattabiraman, Grid-Forming Inverters: A Critical Asset for the Power Grid. IEEE Journal of Emerging and Selected Topics in Power Electronics, 2020. 8(2): p. 925-935.

[2] Chen, M., D. Zhou, and F. Blaabjerg, Enhanced Transient Angle Stability Control of Grid-Forming Converter Based on Virtual Synchronous Generator. IEEE Transactions on Industrial Electronics, 2022. 69(9): p. 9133-9144.

[3] Vasilakis, A., et al., The Evolution of Research in Microgrids Control. IEEE Open Access Journal of Power and Energy, 2020. 7: p. 331-343.

[4] Cheema, K.M., A comprehensive review of virtual synchronous generator. International Journal of Electrical Power & Energy Systems, 2020. 120.

[5] She, B., et al., Virtual inertia scheduling for power systems with high penetration of inverter-based resources. arXiv preprint arXiv:2209.06677, 2022.

[6] Alipoor, J., Y. Miura, and T. Ise, Power System Stabilization Using Virtual Synchronous Generator With Alternating Moment of Inertia. IEEE Journal of Emerging and Selected Topics in Power Electronics, 2015. 3(2): p. 451-458.

[7] Li, D., et al., A Self-Adaptive Inertia and Damping Combination Control of VSG to Support Frequency Stability. IEEE Transactions on Energy Conversion, 2017. 32(1): p. 397-398.

[8] Wang, F., et al., An Adaptive Control Strategy for Virtual Synchronous Generator. IEEE Transactions on Industry Applications, 2018. 54(5): p. 5124-5133.

[9] Yao, F., et al., RBF Neural Network Based Virtual Synchronous Generator Control With Improved Frequency Stability. IEEE Transactions on Industrial Informatics, 2021. 17(6): p. 4014-4024.

[10] 1Zhang, K., et al. A Virtual Synchronous Generator Control Strategy with Q-Learning to Damp Low Frequency Oscillation. in 2020 Asia Energy and Electrical Engineering Symposium (AEEES). 2020.

[11] Wu, W., et al., Deep Q-Network based Adaptive Robustness Parameters for Virtual Synchronous Generator, in 2022 IEEE Transportation Electrification Conference and Expo, Asia-Pacific (ITEC Asia-Pacific). 2022. p. 1-4.

[12] Xiong, K., et al., Deep reinforcement learning based parameter self-tuning control strategy for VSG. Energy Reports, 2022. 8: p. 219-226.

[13] Wang, X., et al., Deep Reinforcement Learning: A Survey. IEEE Trans Neural Netw Learn Syst, 2022. PP.