Reactive-Voltage Coordinated Control for Offshore Wind Farm Based on Multi-Agent Deep Reinforcement Learning

Introduction:

In this research, my work is using a multi-agent deep reinforcement learning-based approach on distributed voltage regulation of offshore wind turbine farm; the details are:

(1) Established an optimization model of voltage which takes reactive power of each generator as variables and the power limitation and power flow as limits. Moreover, the optimization model is formulated as Markov Decision Process (MDP).

(2) Using Deep Deterministic Policy Gradient (DDPG) to solve proposed MDP, where the policy network is trained in an offline method and examined online. As a result, the average voltage deviation is decreased to 0.00054 p.u. after adding the forward difference of active power into observation. (using real-world wind farm active power data)

(3) To have better performance, the distributed control problem of each sub-network is modeled as Markov games and solved by the formulated Multi-Agent DDPG (MADDPG) algorithm, where each sub-network is modeled as an adaptive agent. All agents are centrally trained to learn the optimal coordinated voltage regulation strategy and executed it in a distributed manner while making decisions only based on local information. In this part, the average voltage deviation is decreased to 0.00036 p.u. which enhances the performance by 33.33%.

Table 1 Voltage distribution of each model (voltage statistical dispersion)

Model Number	Target Voltage (p.u.)	Average Voltage (p.u.)	Standard Deviation ($\times 10^{-4}$ p.u.)	Average Voltage Deviation ($\times 10^{-4}$ p.u.)	Solve Time (ms)	Algorithm
No Control	/	1.0767	26.9	/	/	/
1	1.03	1.0310	11.0	10.0	/	DDPG
2	1.03	1.0296	8.717	6.1	/	DDPG
3	1.03	1.0299	6.798	5.4	3.3	DDPG
4	1.03	1.0300	4.990	3.6	2.9	MADDPG
5	1.00	1.0001	21.45	5.2	/	DDPG
6	1.00	1.0000	5.120	3.3	/	MADDPG

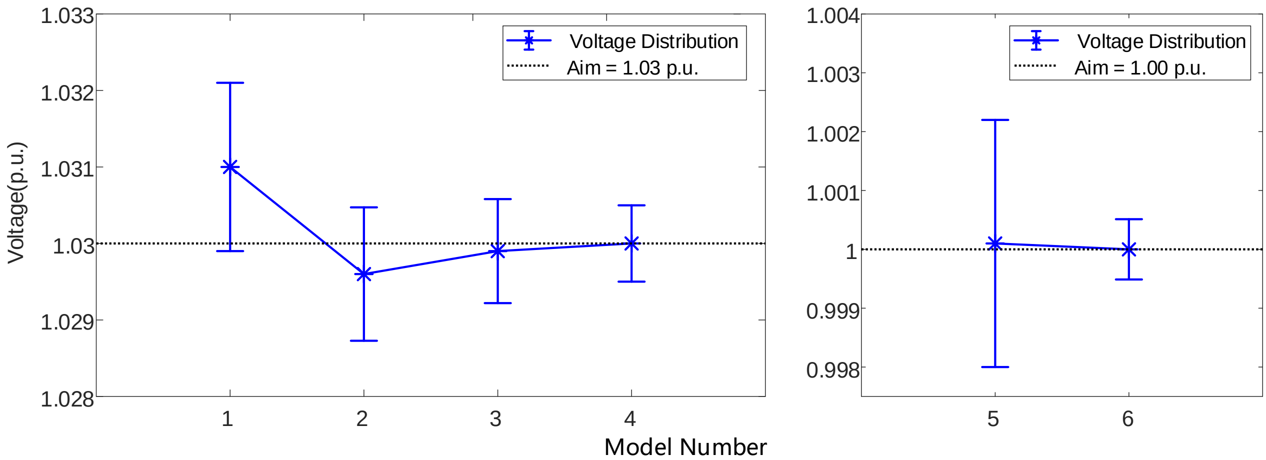


Fig 1 Voltage distribution of each model

Structure of the Wind Farm (left) & Using DDPG to Train the Model (right):

$$reward = -\sum_{i=1}^{N_s}[w_i(U_i - U_{i-ref})^2 + k_i(S_i - S_{i\max})]$$

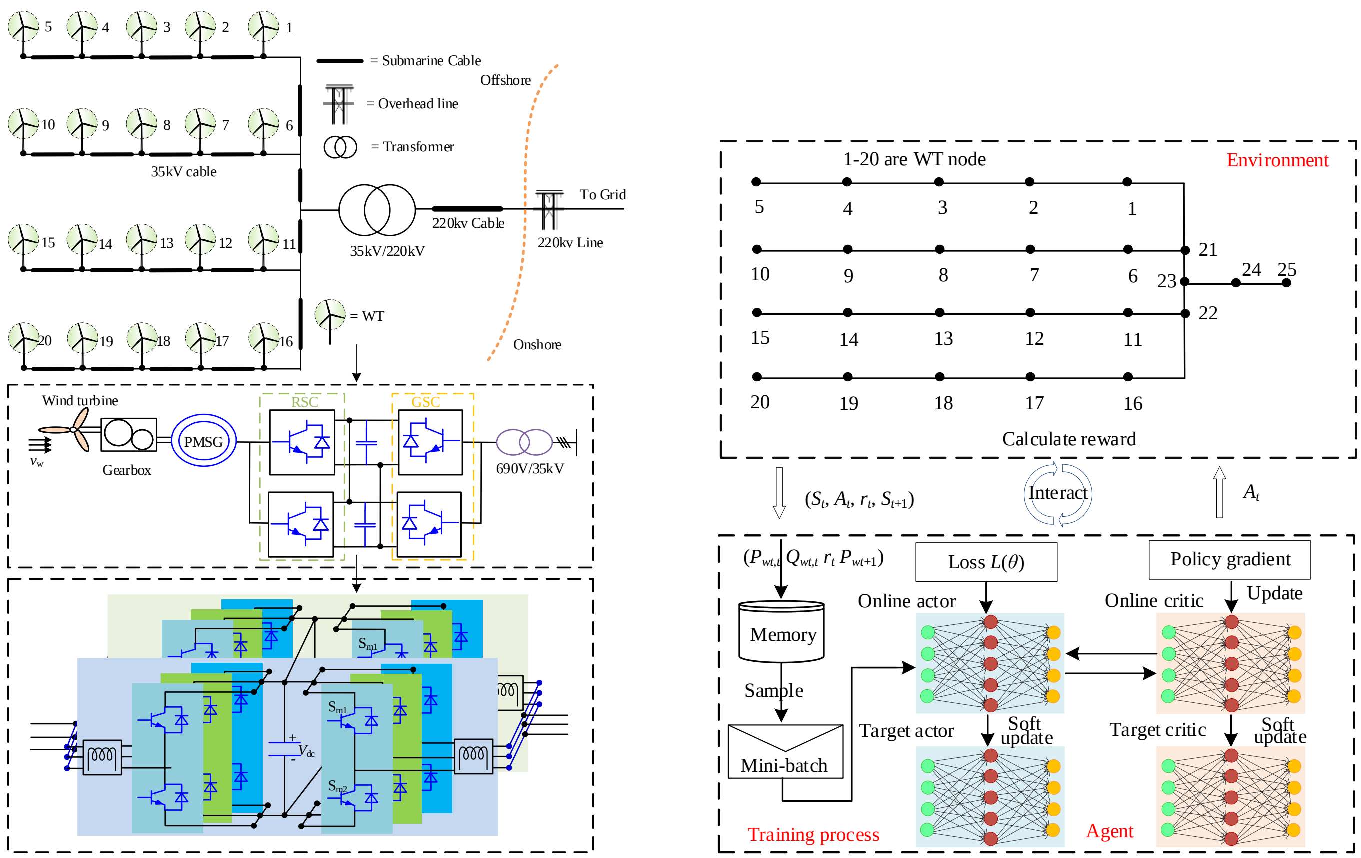
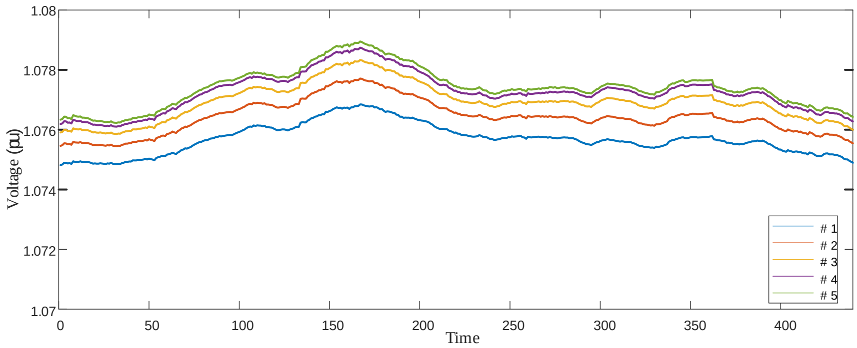


Fig 2 Structure of wind farm and training the agent

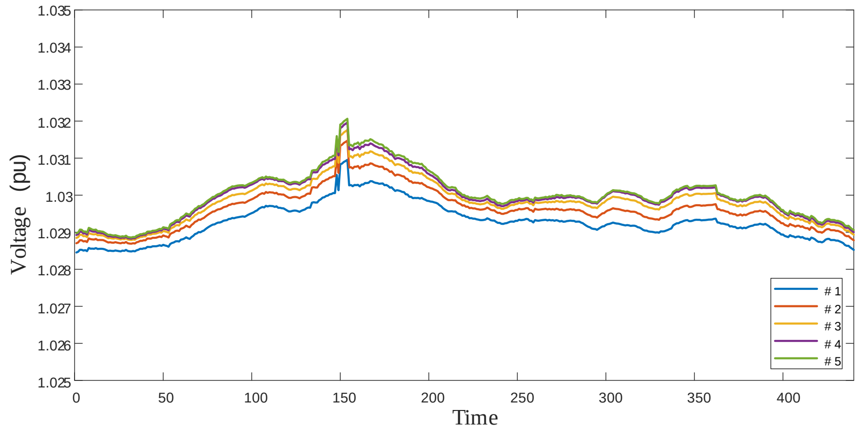
Comparison of each model:

The model trained by MADDPG has a voltage that can change more smooth and has less deviation

the voltage without control



control with model trained by DDPG (model 3)



control with model trained by MADDPG (model 4)

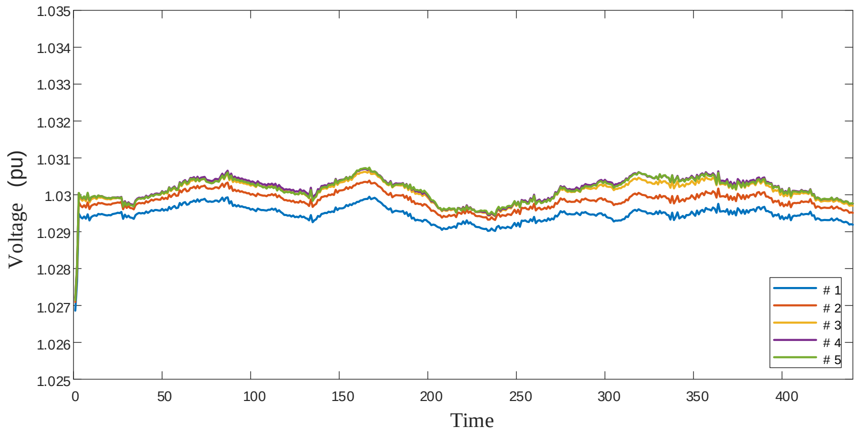


Fig. 3 Voltage of the node 1 to 5

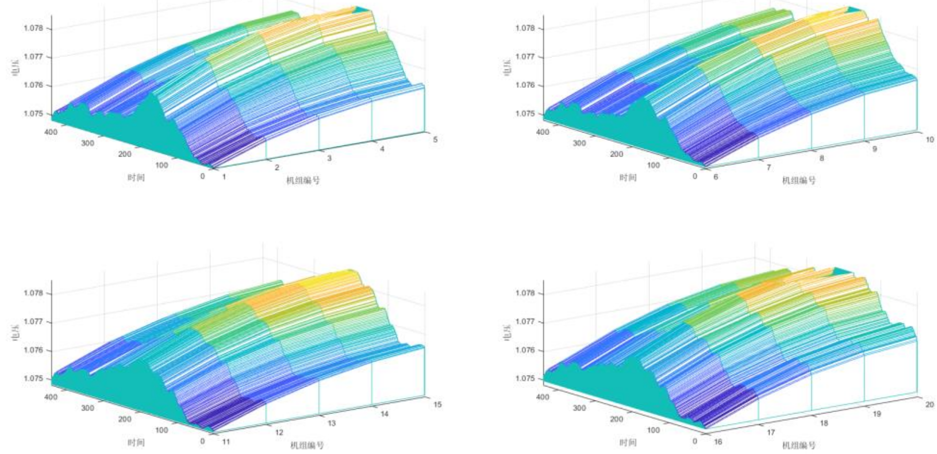


Fig. 4 Voltage distribution of the wind farm (No Control)

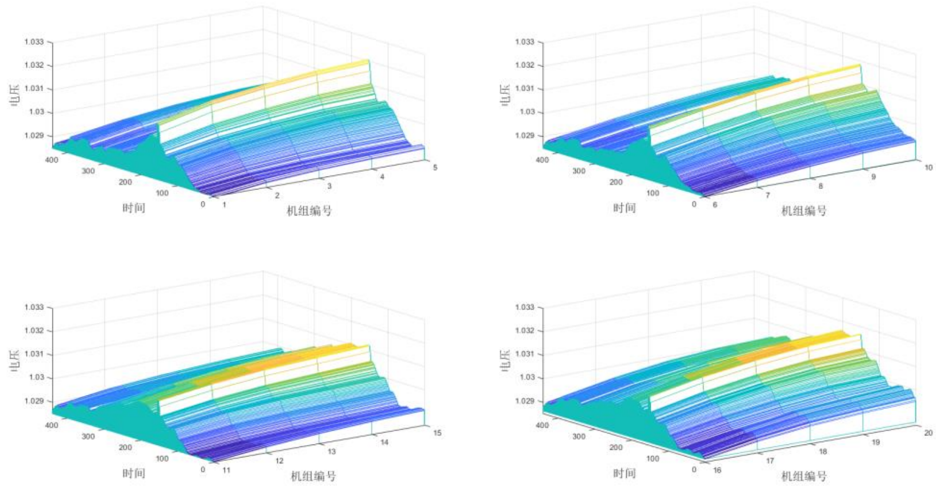


Fig. 5 Voltage distribution of the wind farm (DDPG)

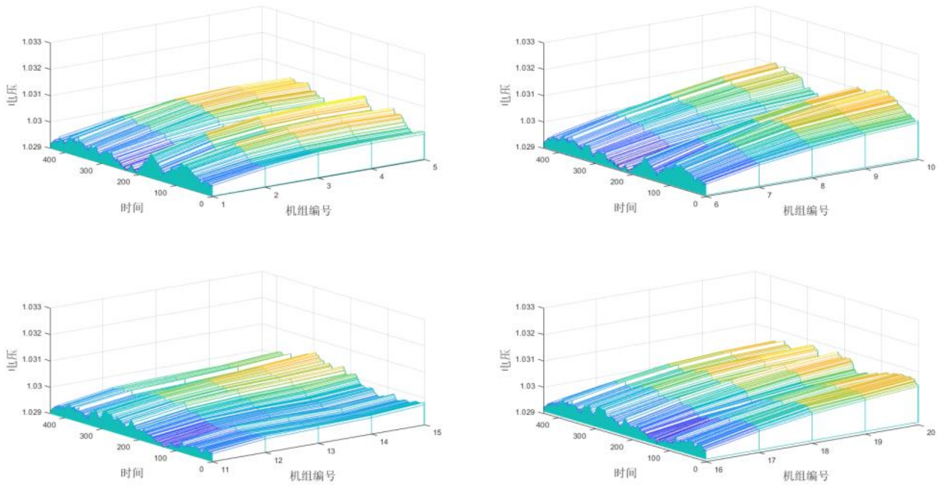


Fig. 6 Voltage distribution of the wind farm (MADDPG)

Comparison of MADDPG and DDPG:

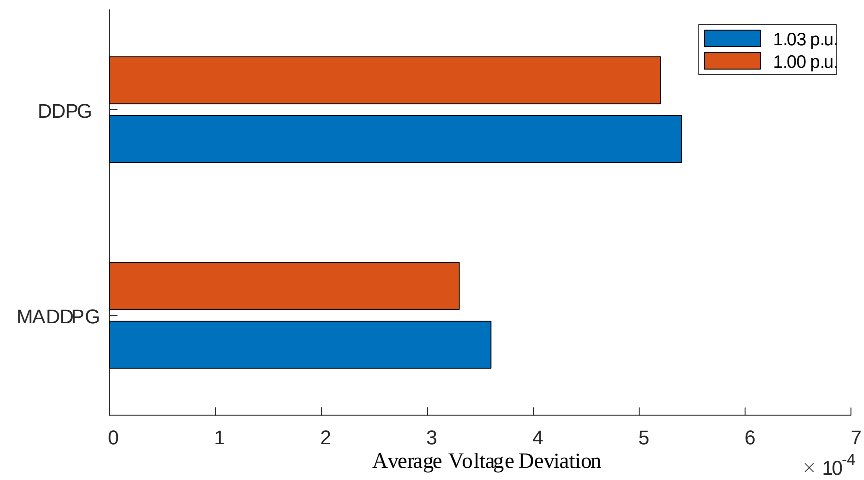


Fig 7 Average voltage deviation using MADDPG and DDPG (the less the better)

Evolution of the Reward:

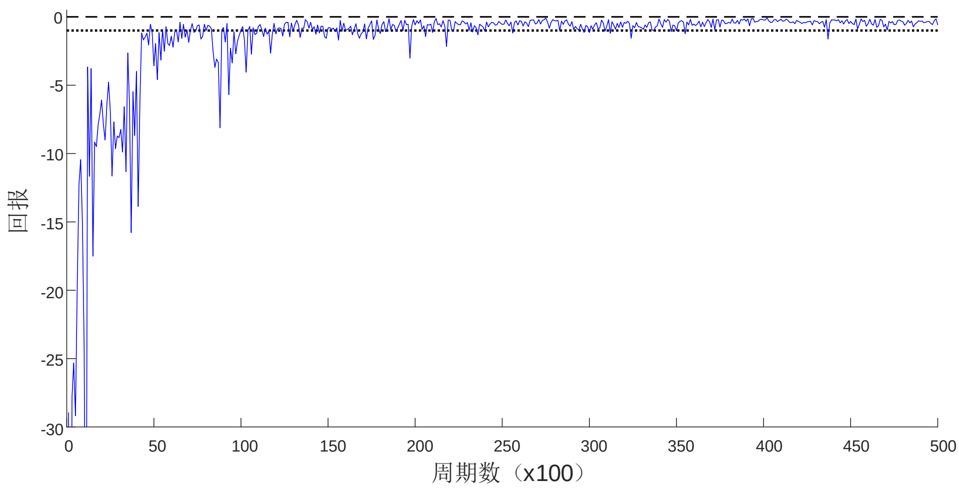


Fig. 8 The evolution of the reward during the training procedure (model 4, MADDPG)

MADDPG:

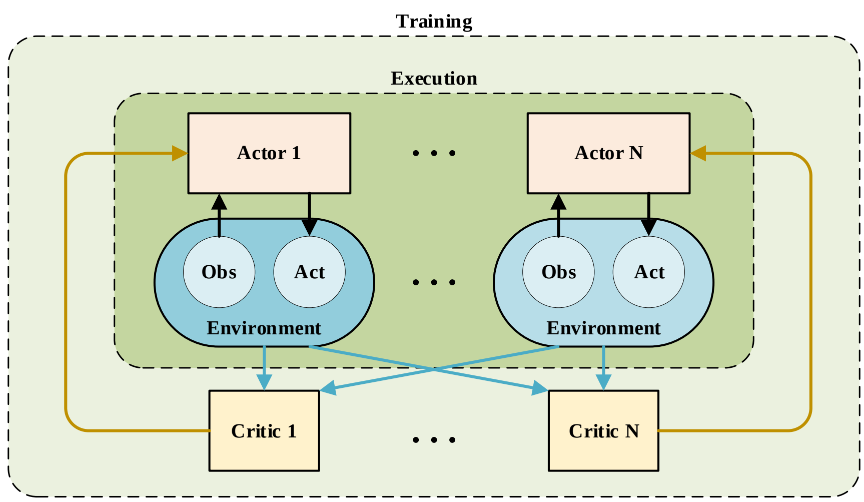


Fig. 9 Overview of multi-agent decentralized actor, centralized critic approach

Optimization Model:

1. Object:

$$\min F = \Delta U^2 = \sum (U_i - U_{i-ref})^2 \qquad i \in [1, N_s]$$

N_s : number of generators

U_i、U_{i-ref} : voltage of the wind turbine and its target

2. Constrains

power limitation

$$\begin{cases} Q_{i\min} \le Q_i \le Q_{i\max} \\ U_{i\min} \le U_i \le U_{i\max} \\ S_{i\min} \le S_i \le S_{i\max} \end{cases} \qquad i \in [1, N_s]$$

power flow constrains

$$\begin{cases} P_i = U_i \sum_{j=1}^{Ns} U_j (G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}) \\ Q_i = U_i \sum_{j=1}^{Ns} U_j (G_{ij} \sin\theta_{ij} - B_{ij} \cos\theta_{ij}) \end{cases} \qquad i \in [1, N_s]$$

3. Average Voltage Deviation

$$\Delta U = \frac{\sum_{i=1}^{N_s} | U_i - U_{ref} |}{N_s}$$