Dear reviewers:

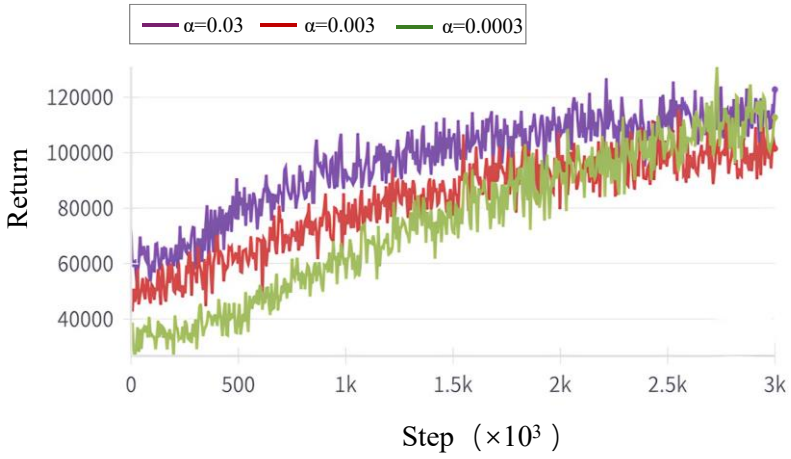
I sincerely thank you for your careful review and valuable comments on this paper! The issues you have pointed out are extremely constructive and crucial to improving the quality of the paper. We have carefully revised the paper one by one, and now we would like to report the revisions as follows:

1.The differences of the MAMQPPO algorithm from previous studies should be clearly stated.

1)Algorithmic differences: we describe the structure of the algorithm in detail in section 3.3.1 in blue, as well as the technical route flow, compared to the previous algorithm, our algorithm is designed with a two-layer network structure, while the proposed MAMQPPO algorithm consists of two Actor and two Critic networks, where the AV-Actor and Platoon-Actor represent the two Actor networks. The AV-Actor is a seven-layer network, including an input layer, three hidden layers (with 64, 128, and 64 neurons, respectively), and an output layer, utilizing ReLU and Tanh activation functions to enhance nonlinear processing. The Platoon-Actor incorporates the Max-Value function (MaxQ) from the DQN network at the output layer to select the maximum-value platoon gap and actions。

2.Hyperparameter selections and the training process should be explained in more detail.

We designed multiple sets of experiments and set the learning rate hyperparameter to 0.03 as detailed below:



learning rate α: When α = 0.03, the model reaches peak returns (>1.2 × 10^5) in 1.5 × 10^3 steps, with the fastest convergence, but with ±8% late fluctuations; α = 0.003 requires 2 × 10^3 steps to stabilise to 1.0 × 10^5, with a 37% lower speed of convergence, but with better stability; and α =0.0003 is the least efficient, reaching only 8 × 10^4 in 3 × 10^3 steps. Experiments show that a high learning rate (α = 0.03) reinforces early exploration capability through fast gradient updating and is suitable for real-time demands of queue

control (e.g., emergency lane changing), although stability risks need to be balanced, while a low learning rate ($\alpha = 0.003$) is suitable for long-term strategy optimisation scenarios. In this paper, we choose $\alpha = 0.03$ to prioritise the real-time control objective.

3.Experimental results should be analyzed in more depth, especially in terms of success rate, energy consumption and the effect of the number of vehicles in the convoy.

3)We have restructured the experimental analysis graphs and language, as shown in (Sections 4.1 and 4.2 in blue), and have added a new experiment with the quintic polynomial + PID from an existing method to compare with our proposed model, and to analyse the data in detail.

4.A discussion on how to validate the model with real-world data should be added.

4)Thank you for your suggestion. Our research focuses on models and algorithms for merging autonomous vehicles into platoon, and we have mainly adopted the Multi-agent Deep Reinforcement Learning (MADRL) method and conducted simulations to validate it. Currently, since there is no publicly available dataset of real fleets, we constructed a simulation environment and evaluated the effectiveness of the algorithms based on it. In order to better fit the real traffic situation and sumo has a significant advantage in simulating macroscopic traffic flow, we therefore utilise sumo for the construction, details of which are given in Section 4.1 in blue.

5.The obtained results should be discussed in detail. In addition, the limitations and deficiencies of the study should be emphasized and suggestions for future studies should be made.

5)We analyse the experimental results in more depth, and their limitations are detailed in the last paragraph of Section 5 in blue

6.All figure and table references should be corrected and missing explanations should be completed.

6)We made careful adjustments to the figures and tables, as well as the references, to ensure that the submission was formatted correctly.

Thank you once again for your professional advice!

The above modifications have been verified through simulation experiments and updated in the manuscript, significantly enhancing the paper's completeness and practical value. If you have any other suggestions, we are willing to make further adjustments.