

Platooning Connected, Autonomous, and Human-Driven Vehicles: A Deep Reinforcement Learning-based Approach

Zhen Qin^a, Dong-Fan Xie^{a,*}, Heng Ma^{a,*}, Xiaomei Zhao^a, Zhengbing He^{b,*}

^a*School of Systems Science, Beijing Jiaotong University, Beijing 100044, China*

^b*Faculty of Science and Engineering, University of Nottingham Ningbo China*

Abstract

Conventionally, existing vehicle platooning approaches are designed for connected vehicles, typically including connected autonomous vehicles and connected human-driven vehicles. Non-connected vehicles, such as non-connected autonomous or human-driven vehicles, are not incorporated. As a result, these platooning approaches may not properly reflect real-world mixed traffic conditions at the current stage. To address this limitation, this study proposes a hybrid platooning pattern that conditionally permits non-connected vehicles to join platoons, thereby enhancing platooning diversity and flexibility. However, it was found that the unregulated integration of non-connected vehicles can trigger rapid platoon expansion, significantly amplifying the risk of disturbance propagation in traffic flow. This, in turn, exacerbates the inherent conflict between traffic throughput and stability. To mitigate these challenges, this paper further develops a hybrid platooning control strategy based on deep reinforcement learning (DRL). This strategy integrates vehicle dynamics, platoon topology, and traffic flow states through a multi-level state representation network, enabling a dynamic trade-off between traffic capacity and stability. Numerical simulations demonstrate that the proposed strategy effectively suppresses velocity disturbance propagation by dynamically optimizing platoon structures, thereby significantly enhancing the stability and safety of mixed traffic while reducing fuel consumption and emissions.

Keywords: Mixed traffic, vehicle platooning, deep reinforcement learning, connected and autonomous vehicle, stability

*Corresponding authors

Email addresses: dfxie@bjtu.edu.cn (Dong-Fan Xie), HengMa1024@outlook.com (Heng Ma), he.zb@hotmail.com (Zhengbing He)

1. Introduction

The rapid progress of new technologies such as the Internet of Vehicles (IoV) and Artificial Intelligence (AI) has facilitated the development of vehicles. Many existing studies categorize these vehicles into the following four types: (i) **Human-driven Vehicles (HVs)**, i.e., traditional vehicles without automation or connection. (ii) **Autonomous Vehicles (AVs)**, which provide driving automation but no connection. (iii) **Connected Human-driven Vehicles (CHVs)**, which offer communication capabilities but no automation. (iv) **Connected Autonomous Vehicles (CAVs)**, which integrate both automation and connection. Obviously, HVs/AVs are non-connected vehicles, while CHVs/CAVs are connected vehicles. Those connected and non-connected vehicles form a new mixed traffic flow that is expected to become a defining feature of future traffic systems.

In mixed traffic, CHVs/CAVs can form platoons to enhance traffic throughput by enabling coordination among vehicles (Huang et al., 2018; Chang et al., 2020; Xin et al., 2021). Under optimal platooning conditions, road capacity can exceed even 12,000 vehicles per hour per lane, representing a 5.3-fold increase compared with scenarios without platooning (Sala and Soriguera, 2020). Furthermore, reduced inter-vehicle spacing within platoons decreases aerodynamic drag, thereby lowering fuel consumption and emissions (Shi et al., 2025a).

Nevertheless, most existing platooning strategies rely on homogeneous and continuously distributed CAVs, which limits platooning scalability (Guan et al., 2023). Notably, CAVs are equipped with multi-source perception systems, such as LiDAR, radar, and vision sensors, enabling them to accurately perceive the states of HVs/AVs. By broadcasting processed environmental information through communication networks, CAVs can incorporate HVs/AVs into platoons, thereby relaxing the requirement for continuous CAV distribution and enabling scalable platoon deployment.

In hybrid platooning strategies, CAVs assume dual roles as information relays and motion coordinators. Optimizing the vehicle interaction network through topological restructuring enables the conditional integration of HVs/AVs into platoons. This approach enhances platoon scenario diversity and flexibility, thereby better aligning with real-world traffic demands. However, the unregulated incorporation of HVs/AVs can cause rapid expansion of platoon size, resulting in excessively large formations. This substantially amplifies the risk of traffic flow disturbance propagation and increases the likelihood of congestion arising from systemic instability. Moreover, conventional maximum platoon length constraints, originally developed for homogeneous platoons, are no longer adequate to accommodate the requirements of these emerging platooning strategies.

To address the aforementioned challenges, this study introduces a hybrid platooning pattern that permits the integration of HVs/AVs into platoons. Building on this framework, we further develop a Deep Reinforcement Learning (DRL)-based platooning strategy to optimize multiple objectives in such mixed traffic flow. The proposed strategy employs a multi-tier state representation network to integrate vehicle dynamics, platoon topology, and traffic flow states, enabling a dynamic trade-off between traffic throughput and flow stability. Furthermore, the strategy adopts the Proximal Policy Optimization (PPO) algorithm within a Centralized Training with Decentralized Execution (CTDE) architecture, which helps improve policy optimization while enabling decentralized platooning decisions based on local observations. Simulation results demonstrate that the proposed DRL-based hybrid platooning strategy effectively improves mixed traffic flow stability by dynamically optimizing platoon structures and suppressing velocity disturbance propagation. Evaluation of the trained policy confirms its robustness across different CHV/CAV penetration rates, with a 51.08% improvement in the stability discriminant when $P_{CAV} > 0.1$. The trained policy also reduces peak average TTCI by 52.95% under low CAV penetration and lowers fuel consumption and CO₂ and NO_x emissions in disturbance-propagation scenarios.

The remainder of this study is organized as follows. Section 2 presents a systematic review of existing platooning strategies, including both homogeneous and heterogeneous approaches. Section 3 describes the mixed traffic environment and introduces the proposed hybrid platooning pattern incorporating HVs/AVs. Section 4 develops the DRL-based hybrid platooning strategy and details its training and testing procedures. Section 5 evaluates the proposed strategy through comprehensive simulations, assessing its generalization capability, traffic stability, safety, energy consumption, and emissions. Finally, Section 6 summarizes the main findings and outlines potential directions for future research.

2. Literature review

2.1. Platooning strategies

Although extensive research has been conducted on cooperative control strategies to exploit the benefits of vehicle platooning, many existing studies primarily focus on platooning strategies for a single vehicle type. While advanced optimization- and learning-based algorithms have been widely explored, rule-based platooning strategies remain among the simplest and most computationally efficient solutions. For example, Wu et al. (2024) and Maiti et al. (2023) introduced rule-based platoon control strategies that offer high computational efficiency but exhibit limited adaptability in complex mixed traffic environments.

Vehicle platooning fundamentally relies on the sharing of vehicle state information and aims to enhance collective performance, which naturally motivates its formulation as an optimal control problem. Within this framework, optimal control techniques, such as Model Predictive Control (MPC) and Linear Quadratic Regulator (LQR), have been extensively adopted due to their effectiveness in handling complex and dynamic traffic conditions (Wu et al., 2023; Wang et al., 2025; Yang et al., 2023; Sun et al., 2024). Specifically, Lan et al. (2023) investigated cooperative adaptive cruise control (CACC) for mixed traffic platoons by developing a data-driven MPC algorithm that explicitly accounts for noisy measurements. Jin et al. (2021a) proposed a continuum model for mixed traffic flow and provided analytical methodologies for the design of cooperative platoon control strategies. Zhang et al. (2025) developed a self-organized cooperative control strategy for platoon formation based on distributed MPC. Zong et al. (2025) introduced an adaptive control framework that allows CAVs to dynamically switch between platooning and individual operation, improving traffic stability and reducing emissions, while incorporating human driver behavior through transposition prediction. In addition, Liu et al. (2025) proposed a cooperative control strategy that combines an LQR controller with an active disturbance rejection controller (ADRC) to manage platoon state transitions under varying time headways. From a decentralized perspective, Wijnbergen et al. (2025) developed a non-linear spacing policy that guarantees string stability when each vehicle follows its immediate predecessor.

As platooning research has expanded from control under prescribed vehicle formations to coordination problems involving mixed traffic, heterogeneous vehicle behaviors, and more flexible platoon interactions, it has become increasingly difficult to rely solely on handcrafted rules or explicitly formulated optimization models. This shift has motivated the use of RL and DRL, which offer a practical way to derive adaptive control strategies from traffic interactions. For example, Song et al. (2026) integrated global perception prediction for CHVs/CAVs with lane-specific reinforcement-learning-based variable speed limit control, showing how richer traffic-state information can support adaptive decision-making in mixed traffic environments. Han et al. (2025) proposed a data-driven finite-time adaptive controller that achieves model-free tracking of vehicle position and velocity. Lin et al. (2024) presented an enhanced state representation for multi-agent reinforcement learning to improve car-following performance. By integrating heuristic learning with distributed modeling, Du et al. (2024) developed a multi-layer predictive control framework for autonomous vehicle platoons. For heterogeneous non-linear platoons, Ma et al. (2024) proposed a model-free distributed control strategy that relies only on local information from neighboring vehicles and guarantees stability via linear matrix

inequalities without pre-training. Addressing communication efficiency in multi-agent reinforcement learning, [Hua et al. \(2025\)](#) developed a fully decentralized CACC framework that incorporates quantified stochastic gradient descent and binary differential consensus mechanisms.

More closely related to this study, recent works have extended platoon control to mixed or heterogeneous vehicle settings. These studies typically take the platoon composition as a given setting and concentrate on the associated control problem, such as robust control of mixed platoons under HV prediction uncertainty ([Feng et al., 2021](#)), distributed MPC coordination of heterogeneous platoons with spacing constraints ([Qiang et al., 2023](#)), and hierarchical cooperative control of mixed platoons for oscillation mitigation ([Liang et al., 2025](#)). This tendency is also evident in recent DRL-based studies. [Shi et al. \(2023\)](#) developed a distributed DRL strategy for CAV longitudinal control in mixed platoons by aggregating consecutive HVs and incorporating their traffic effects into the state representation. [Yang et al. \(2025\)](#) compared single-agent and multi-agent RL methods for eco-CACC in predefined mixed-platoon scenarios, where the learned policies are used to determine CAV acceleration actions. [Liu et al. \(2026\)](#) considered mixed platoons containing HVs under a prescribed CAV–HV–CAV subplatoon structure and proposed a DRL-based cooperative controller with guaranteed string stability. [Shi et al. \(2025b\)](#) moved one step closer to platoon organization by studying mixed-vehicle platoon forming; however, their framework first generates a feasible target formation through rule-based design and then applies MARL to control CAV acceleration, deceleration, cruising, and lane-changing actions to realize that formation. Overall, these studies mainly address how vehicles should be controlled once the platoon composition or target formation has been assumed, rather than how heterogeneous vehicles, including HVs and AVs, should be organized into feasible platoons in mixed traffic. This leaves an important gap between within-platoon vehicle control and hybrid platooning strategy design, which is the focus of this study.

2.2. Impacting factors of vehicle platooning on traffic flow

The construction of platoons in mixed traffic requires more than identifying which vehicles are technically able to connect. It demands a clear understanding of what impacts different platoon arrangements may bring. Existing studies suggest that platoon formation pattern, maximum platoon length, and CAV clustering intensity are among the factors most closely related to these impacts.

Platoon formation pattern and CAV clustering intensity are closely related to how vehicles are spatially organized in mixed traffic. [Jin et al. \(2021b\)](#) evaluated different platoon formation patterns and showed that the arrangement of platoons can affect traffic efficiency and vehicle

energy consumption. Similarly, [Li et al. \(2022\)](#) compared uniform, random, and platoon-based vehicle distributions from a set-function optimization perspective. Beyond formation patterns, CAV clustering intensity has also been used to characterize the spatial concentration of CAVs in mixed traffic. Related studies suggest that clustering intensity is associated with platooning opportunities, macroscopic traffic characteristics, and safety performance ([Guan et al., 2023](#); [Zhou and Zhu, 2021](#); [Jiang et al., 2023](#); [Ghiasi et al., 2017, 2020](#)).

Maximum platoon length is another key consideration, as platoon expansion may bring both traffic benefits and operational risks. [Van Arem et al. \(2006\)](#) pointed out that unrestricted platoon growth may adversely affect traffic safety, while [Liu et al. \(2018a\)](#) argued that platoon splitting becomes necessary once platoon length exceeds an allowable threshold. From a traffic-flow perspective, [Jin et al. \(2021a\)](#) showed that road capacity can improve with platoon formation, but may decline when platoon size exceeds a critical level. Considering capacity, safety, and environmental impacts jointly, [Seraj et al. \(2018\)](#) and [Zhou and Zhu \(2021\)](#) further showed that larger platoons may improve throughput and environmental performance, although these gains can be accompanied by safety or stability concerns. Accordingly, recommended platoon lengths vary across studies depending on the target objective ([Liu et al., 2018b](#); [Seraj et al., 2018](#); [Zhou et al., 2021](#); [Yao et al., 2022](#)).

3. Platooning scenarios in mixed traffic

For clarity, a platoon is defined as a coordinated group of adjacent vehicles traveling longitudinally with a shared platooning status. A hybrid platoon may include HVs and AVs if their states can be sensed by adjacent CAVs and made available to CHVs and CAVs, subject to the restriction that HVs and AVs cannot act as platoon leaders or tail vehicles.

3.1. Mixed traffic scenario

The mixed traffic flow considered in this study consists of four categories of vehicles: CAVs, CHVs, AVs, HVs. The vehicle type set is defined as:

$$\mathcal{V} = \{\text{CAV}, \text{CHV}, \text{AV}, \text{HV}\} \quad (1)$$

From a communication perspective, CHVs/CAVs are capable of vehicle-to-vehicle (V2V) communication through dedicated short-range communication (DSRC) or cellular vehicle-to-everything (C-V2X) technologies. In contrast, HVs/AVs do not possess communication capabilities and therefore cannot exchange information with other vehicles. Accordingly, vehicles

can be classified into two categories based on their communication capability: CHVs/CAVs ($\mathcal{V}_C$) and HVs/AVs ($\mathcal{V}_H$):

$$\begin{cases} \mathcal{V}_C = \{\text{CAV}, \text{CHV}\}, \\ \mathcal{V}_H = \{\text{AV}, \text{HV}\}. \end{cases} \quad (2)$$

While CHVs and HVs are operated by human drivers, decision-making and control in CAVs and AVs are fully executed by autonomous driving systems, which typically employ control algorithms based on frameworks such as MPC or DRL.

This study focuses on mixed traffic scenarios on straight road segments, excluding external influences such as on- and off-ramps and lane-changing maneuvers. Vehicles are assumed to be spatially distributed with randomized type configurations. The penetration rates of CAVs, CHVs, AVs, and HVs are denoted by P_{CAV} , P_{CHV} , P_{AV} , and P_{HV} , respectively, and satisfy:

$$P_{\text{CAV}} + P_{\text{CHV}} + P_{\text{AV}} + P_{\text{HV}} = 1 \quad (3)$$

Based on the above description of the mixed traffic scenario, the following assumptions are made: (1) Vehicles move on a single-lane road without considering lane-changing behavior; (2) The four types of vehicles, i.e., CAVs, CHVs, AVs, HVs, are randomly distributed along the road; (3) V2V communication delays and sensor data processing latencies are neglected; (4) CAVs and AVs operate fully under automated control without human intervention, whereas CHVs and HVs rely entirely on human drivers; (5) CAVs are regarded as either (i) communication-enhanced AVs, namely, AVs equipped with additional V2V communication capabilities for state information sharing; or (ii) control-enhanced CHVs, namely, CHVs equipped with autonomous driving systems that replace human drivers; (6) Vehicle platoon formation and dissolution are assumed to occur instantaneously, without considering transient effects induced by dynamic reorganization processes; (7) All vehicles unconditionally participate in platooning operations whenever platoon size constraints and communication connectivity conditions are satisfied.

3.2. *Heterogeneous platooning behavior*

Based on whether HVs/AVs are permitted to participate in platoons, this study investigates vehicle platooning behaviors under two distinct scenarios.

3.2.1. *Platooning scenario for CHVs and CAVs*

In the traditional platooning scenario, only CHVs/CAVs, are allowed to form platoons.

When a CHV/CAV attempts to join a platoon, it first identifies a communication target vehicle (CTV) traveling ahead and seeks to establish a vehicle-to-vehicle (V2V) communication link. The vehicle can join the CTV's platoon if and only if the inter-vehicle distance is within the maximum communication range of both vehicles.

Although any member of the target platoon could theoretically serve as a CTV, this study focuses on the most practical and commonly adopted case in which vehicles communicate exclusively with their immediately preceding neighbors. As illustrated in Fig. 1 (a), all platoon members, except for the leader, maintain communication solely with their direct predecessor. For notational simplicity, this platooning scheme is referred to as PS1.

The probability that an ego vehicle successfully joins the platoon of its immediately preceding vehicle is denoted as $P_{\text{pla}, Y_{\text{ego}}, Y_{\text{front}}, u}^{\text{PS1}}$, where u represents the position index of the preceding vehicle within the platoon, and Y_{ego} and Y_{front} denote the vehicle types of the ego vehicle and the communication target vehicle, respectively. Based on the steady-state velocity v and the time headway ΔT , the corresponding probability model is given by:

$$P_{\text{pla}}^{\text{PS1}}(Y_{\text{ego}}, Y_{\text{front}}, u) = \begin{cases} \Pr(v \Delta T_{\text{ego}} < \min\{d_{Y_{\text{ego}}}, d_{Y_{\text{front}}}\}), & u \leq l_{\text{max}} - 1, \\ 0, & u = l_{\text{max}}. \end{cases} \quad (4)$$

where $d_{Y_{\text{ego}}}$ and $d_{Y_{\text{front}}}$ denote the maximum effective communication ranges of the ego vehicle and the communication target vehicle, respectively. This study assumes identical maximum communication ranges for CHVs/CAVs. Since HVs/AVs lack communication capabilities, their maximum effective communication ranges are set to zero. ΔT_{ego} represents the time headway after the ego vehicle joins the platoon, and l_{max} denotes the maximum allowable platoon size, i.e., the maximum number of vehicles that a single platoon can accommodate.

3.2.2. Platooning scenario for all vehicles

Conventional homogeneous platooning requires all vehicles to possess communication capabilities. However, under randomly distributed mixed traffic, HVs/AVs that lack communication modules impose stringent constraints, as platoon formation becomes strictly dependent on the contiguous spatial distribution of CHVs/CAVs. This dependence severely limits the scalability of platooning technology.

To address this limitation, this study proposes a heterogeneous platooning strategy. The approach comprises the following two steps. (1) Multi-sensor perception: CAVs utilize onboard multimodal sensors (lidar, mmWave radar, and cameras) to acquire real-time states (velocity,

acceleration, and spacing) of adjacent HVs/AVs. (2) Information propagation: CAVs broadcast the perceived state information of HVs/AVs via V2X communication, establishing an intra-platoon communication topology. This renders the states of all HVs/AVs observable to every CHV/CAV within the platoon.

This strategy enforces strict platoon integrity conditions: if the state of any vehicle becomes unobservable to a connected platoon member, immediate platoon dissolution and dynamic reorganization are triggered.

This platooning strategy conditionally allows HVs/AVs to join platoons, forming heterogeneous structures such as CAV–HV–HV–CAV. This enables more diverse and flexible platooning scenarios that better align with real-world traffic conditions. For subsequent reference, this strategy is referred to as platooning scheme PS2.

Notably, this study stipulates that HVs/AVs are not allowed to serve as platoon leaders or tail vehicles. Due to the absence of communication capabilities, HVs/AVs cannot obtain platoon-level information nor disseminate control or status signals to define platoon boundaries. Moreover, their car-following behaviors remain identical regardless of whether they are inside or outside a platoon. Consequently, positioning HVs/AVs at the head or tail of a platoon fails to establish effective platoon boundaries and does not contribute to improvements in traffic efficiency or stability through cooperative driving. The following sections systematically analyze and elaborate on all feasible platooning scenarios under the proposed framework.

As shown in Fig. 1, when two adjacent vehicles are both CHVs/CAVs, the follower selects its immediate predecessor as the communication target vehicle (CTV). The follower is allowed to platoon with the preceding vehicle if the following conditions are satisfied: (1) the current platoon size is smaller than the maximum allowable platoon length; and (2) the communication distance constraint is satisfied. The joining probability is given by

$$p_{\text{pla}, Y_{\text{ego}}, Y_{\text{front}}, Z_m}^{\text{PS2}} = \begin{cases} \Pr(v \Delta T_{\text{ego}} < \min(d_{Y_{\text{ego}}}, d_{Y_{\text{front}}})) , & Z_m \leq l_{\text{max}} - 1, \\ 0, & Z_m = l_{\text{max}}. \end{cases} \quad (5)$$

where Z_m is the position of vehicle $m \in \{1, 2, \dots, l_{\text{max}}\}$ in the platoon; and $d_{Y_{\text{ego}}}$ and $d_{Y_{\text{front}}}$ represent the corresponding maximum effective communication ranges.

Figs. 1 (b) and (c) illustrate the scenario in which a single HVs/AVs vehicle joins a platoon. As shown in Fig. 1(b), the m -th vehicle is a CHV, the $(m + 1)$ -th vehicle is a HV/AV, and the $(m + 2)$ -th vehicle is a CAV. In this case, the HV/AV $(m + 1)$ can rely on the CAV to

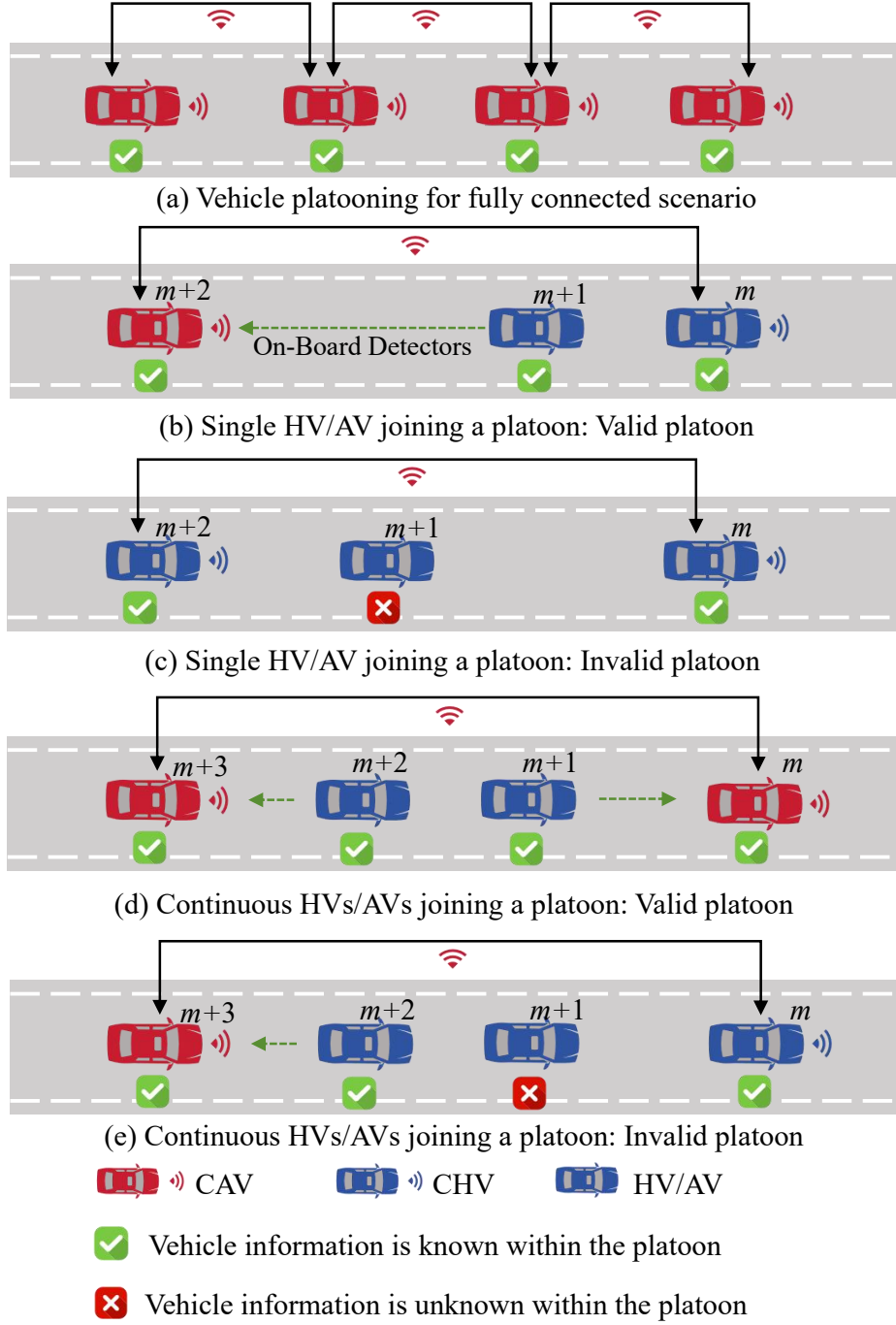


Figure 1: Vehicle platooning scenarios.

acquire its motion state information via onboard sensing units, which is then relayed to other CHVs/CAVs within the platoon. Consequently, the motion states of all vehicles within the platoon become observable to all CHVs/CAVs, enabling the formation of a functional platoon. However, under conventional connected platooning strategies, this scenario would prevent effective platoon formation between vehicle m and vehicle $(m + 2)$ due to the presence of the HV/AV $(m + 1)$.

Furthermore, since HVs/AVs cannot serve as the lead vehicle or the tail vehicle of a platoon, the preceding vehicle m and the following vehicle $(m + 2)$ of the HV/AV $(m + 1)$ must join the same platoon as vehicle $(m + 1)$. Therefore, the conditions for vehicle $(m + 1)$ to join the platoon are: (1) the platoon length constraint remains satisfied after the following vehicle $(m + 2)$ joins; and (2) the communication distance constraint is satisfied between vehicle $(m + 2)$ and its target communication vehicle m . The probability of vehicle $(m + 1)$ joining the platoon can be expressed as:

$$\begin{aligned} P_{\text{pla}, Y_m, Y_{m+2}, Z_m}^{\text{PS2}} \\ = Pr\{Z_m \leq l_{\max} - 2 \cap v(\Delta T_{m+1} + \Delta T_{m+2}) \\ < \min(d_{Y_m}, d_{Y_{m+2}})\} \end{aligned} \quad (6)$$

where ΔT_i is the time headway between vehicle i and its preceding vehicle.

Notably, in this scenario, the absence of communication capability in the HV/AV necessitates the presence of a CAV as a communication intermediary among its adjacent vehicles for it to join the platoon. As illustrated in Fig. 1 (c), the lack of a CAV near the HV/AV $(m + 1)$ implies that its motion state information cannot be reliably acquired and forwarded to other CHVs/CAVs. Consequently, it is unable to form a platoon with its preceding and following vehicles.

As shown in Fig. 1 (d), vehicle m is a CAV, vehicles $(m + 1)$ and $(m + 2)$ are HVs/AVs, and vehicle $(m + 3)$ is a CHV. In this configuration, the motion state information of vehicle $(m + 1)$ must be acquired and relayed through its preceding vehicle m , while the motion state information of vehicle $(m + 2)$ requires acquisition and relaying via its following vehicle $(m + 3)$. Furthermore, since HVs/AVs cannot serve as the lead or tail vehicle of a platoon, vehicles m and $(m + 3)$ must form the same platoon encompassing the HVs/AVs. Consequently, the conditions for the HVs/AVs $(m + 1)$ and $(m + 2)$ to join the platoon are: (1) the platoon length constraint remains satisfied after vehicle $(m + 3)$ joins; (2) the communication distance constraint is met between vehicle $(m + 3)$ and its target communication vehicle m . The probability

Table 1: Platoon structures under hybrid platooning strategy

Platooning structure	Platooning scenario
C-C	Continuous CHV/CAV platooning
CAV-H-CHV CHV-H-CAV CAV-H-CAV	Single HV/AV join a platoon
CAV-H-H-CAV	Two HVs/AVs join a platoon

Note: C denotes CHVs/CAVs and H represents HVs/AVs.

of the HVs/AVs $(m + 1)$ and $(m + 2)$ joining the platoon can then be expressed as:

$$P_{\text{pla}, Y_m, Y_{m+3}, Z_m}^{\text{PS2}} = \Pr\{Z_m \leq l_{\max} - 3 \cap v \cdot \sum_{i=m+1}^{m+3} \Delta T_i < \min(d_{Y_m}, d_{Y_{m+3}})\} \quad (7)$$

In this scenario, any HVs/AVs must have at least one CAV among its adjacent vehicles. Consequently, the maximum number of consecutive HVs/AVs within a platoon is limited to two. As illustrated in Fig. 1(e), due to the absence of a CAV to probe and relay its motion state, the state information of the HV/AV $(m + 1)$ remains unknown to other CHVs/CAVs. Thus, a platoon cannot be formed.

In summary, Table 1 presents the three distinct platoon structures under the hybrid platooning strategy.

3.3. Heterogeneous car-following behavior and models

Based on the vehicle interaction space of Cartesian product $\mathcal{V} \times \mathcal{V}$, car-following behaviors in mixed traffic flow can be classified into two categories according to platoon status:

(1) Fundamental car-following types: Sixteen elementary car-following pairs $s - r \in \mathcal{V}^2$ formed by pairwise combinations of four vehicle types, where s denotes the leading vehicle type and r the following vehicle type. This classification reveals fundamental interaction patterns between heterogeneous vehicles at the microscopic level.

(2) Platoon-based car-following types: Considering the dynamic influence of platooning on control modes, when CAVs, CHVs, AVs, and HVs overcome communication barriers to form hybrid platoons, their intra-platoon and inter-platoon car-following behaviors exhibit significant heterogeneity. Thus, we extend the fundamental pairs into 16 intra-platoon relationships, e.g., $\text{CAV}_{\text{pla}}\text{-CHV}_{\text{pla}}$ and $\text{CHV}_{\text{pla}}\text{-HV}_{\text{pla}}$.

In summary, mixed traffic flow contains 16 fundamental non-platoon car-following types and 16 platoon-based car-following types, totaling 32 car-following configurations.

However, for HVs/AVs, their inability to acquire real-time platoon status via communication implies their time headway remains independent of instantaneous platoon conditions. That is, their car-following behavior is identical both inside and outside platoons. Applying the established CHV/CAV demotion criteria, these 32 car-following types can be consolidated into six distinct time headway patterns, called as fundamental autonomous headway (ΔT_A), manual driving headway (ΔT_H), connected autonomous headway outside platoons (ΔT_{CA}), connected manual headway outside platoons (ΔT_{CH}), intra-platoon cooperative autonomous headway (ΔT_{PA}), and intra-platoon cooperative manual headway (ΔT_{PH}), respectively. For conciseness, the definitions of the six consolidated time headway patterns and their detailed mappings to the original car-following configurations are provided in Appendix [Appendix A](#). Generally speaking, autonomous systems are capable of operating at shorter time headways than human drivers. Communication enables cooperative driving that further reduces time headways, while platooning, as an advanced form of cooperative driving, achieves even smaller headways. Consequently, these time headways exhibit the following hierarchical relationship:

$$\Delta T_H > \Delta T_A > \Delta T_{CH} > \Delta T_{CA} > \Delta T_{PH} > \Delta T_{PA} \quad (8)$$

Based on the above classification of heterogeneous car-following behaviors, the car-following process can be further represented through acceleration functions. Given the headway, velocity difference, and velocity of the ego vehicle, the car-following dynamics can be expressed by the following ordinary differential equations:

$$\dot{x}_n(t) = v_n(t) \quad (9)$$

$$\dot{v}_n(t) = f(s_n(t), v_n(t), \Delta v_n(t)) \quad (10)$$

where $x_n(t)$ and $v_n(t)$ denote the position and velocity of vehicle n at time t ; $s_n(t) = x_{n-1}(t) - x_n(t)$ represents the headway between vehicle n and the preceding vehicle $n - 1$; and $\Delta v_n(t) = v_{n-1}(t) - v_n(t)$ is the velocity difference. To characterize the heterogeneous driving behaviors associated with different vehicle types and connectivity conditions, three car-following models are adopted in this study. The detailed model formulations, parameter definitions, and parameter settings are provided in Appendix [Appendix B](#).

4. Platooning strategy using deep reinforcement learning

The design of the hybrid platooning strategy should comprehensively integrate multiple factors including heterogeneous vehicle dynamics, communication topology constraints, and traffic flow stability. Formulated within a Partially Observable Markov Decision Process (POMDP) framework, we develop a distributed platooning strategy for hybrid platooning in mixed traffic, that comprises state space, action space, multi-objective reward functions, and training architecture. The primary structure of the DRL-based platooning strategy is illustrated in Fig.2.

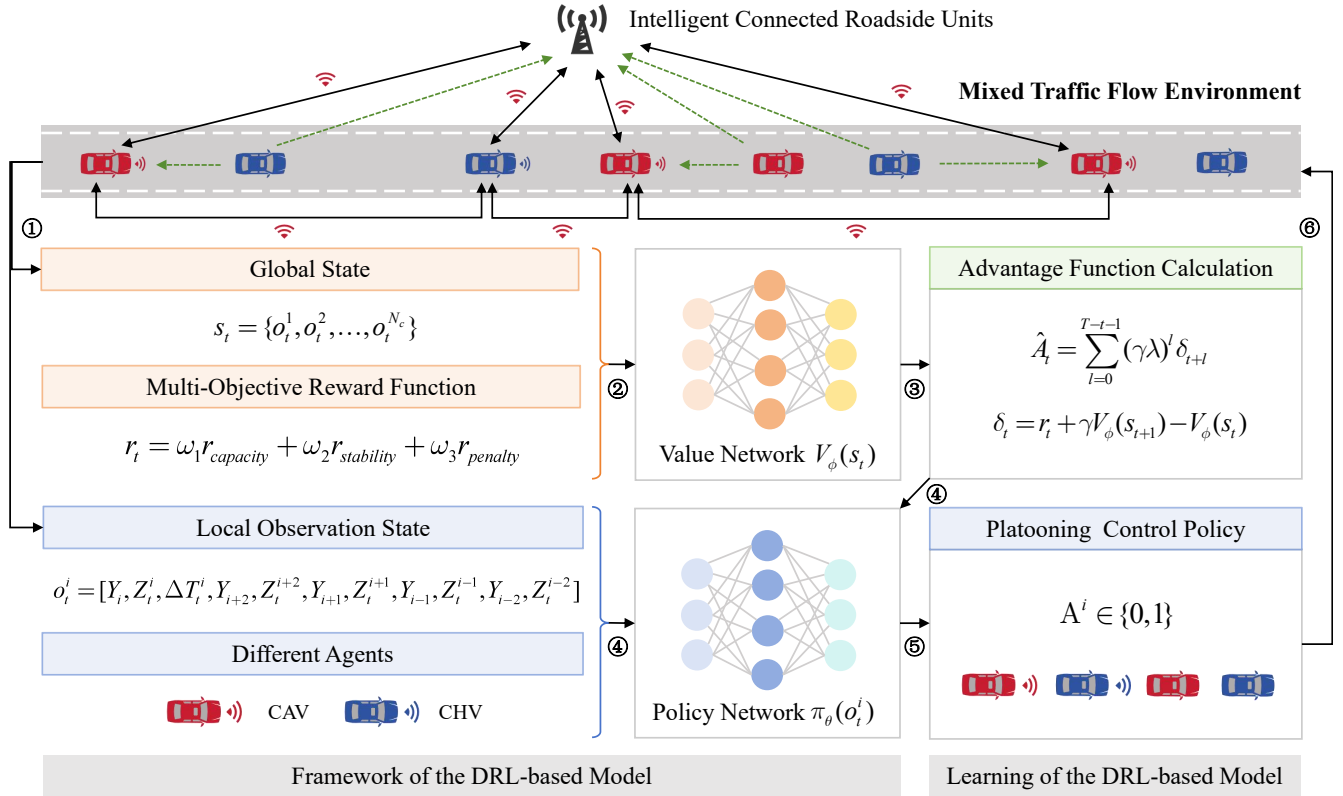


Figure 2: Structure of the DRL-based hybrid platooning model considering multi-type vehicles.

4.1. State space

Connected vehicles (CHVs/CAVs) possess communication capabilities that enable them to acquire partial state information of surrounding vehicles from vehicle-to-everything (V2X) links. To enable multi-dimensional environmental perception and coordinated platooning decisions in mixed traffic flows, the state space integrates multi-modal data, including vehicle dynamics parameters, platoon topological relationships, and traffic environment characteristics. This integration ensures real-time capture of dynamic traffic conditions, instantaneous platoon structure control, and optimized traffic efficiency/stability. This integration lays the

groundwork for establishing spatiotemporal coupling relationships between microscopic vehicle states and macroscopic traffic flow. The state-space vector $o_t^i \in \mathbb{R}^{11}$, where i denotes the index of the ego vehicle, and t denotes the discrete time step, is constructed by concatenating ego-vehicle states with ambient environment states derived from multi-source heterogeneous perception data.

$$o_t^i = \{s_t^{\text{ego}}, s_t^{\text{env}}\} \quad (11)$$

where s_t^{ego} denotes the ego-vehicle state; and s_t^{env} represents the ambient environment state.

The ego-vehicle state s_t^{ego} comprises three elements, including vehicle type, position within the platoon, and time headway.

$$s_t^{\text{ego}} = \{Y_i, Z_t^i, \Delta T_t^i\} \quad (12)$$

Vehicle type and platoon position serve as critical inputs for platooning decisions. The introduction of HVs/AVs significantly increases platoon length. To capture essential queue state information while avoiding the curse of dimensionality from expanding input features, which would escalate model training complexity, this study selects vehicle types and positions from the two immediately preceding and following vehicles, respectively, to construct the state s_t^{env} .

$$s_t^{\text{env}} = \{Y_{i+2}, Z_t^{i+2}, Y_{i+1}, Z_t^{i+1}, Y_{i-1}, Z_t^{i-1}, Y_{i-2}, Z_t^{i-2}\} \quad (13)$$

Leveraging local observations, including ego-vehicle platoon position, time headway, and surrounding vehicle states, CHVs/CAVs dynamically generate platoon control decisions through policy networks. These decisions enable real-time structural adjustments to platoon formations, achieving communication efficiency and traffic flow stability in mixed traffic scenarios.

4.2. Action space

The platooning decision states for vehicles in mixed traffic flow comprise three categories: join platoon, leave platoon, maintain independent driving. The platooning decision action space $\mathcal{A}^i$ can be formulated as,

$$\mathcal{A}^i = \{a_t^i \in \{0, 1\}\} \quad (14)$$

where $a_t^i = 1$ denotes vehicle i executing a joining-platoon action at timestep t ; while $a_t^i = 0$ denotes vehicle i executing either a leaving-platoon action or maintaining independent driving at t .

4.3. Multi-objective reward function

Guided by multi-objective optimization principles, the reward function r_t is formulated to achieve a dynamic trade-off between traffic efficiency and stability, and is defined as:

$$r_t = \omega_1 r_{\text{capacity}} + \omega_2 r_{\text{stability}} + \omega_3 r_{\text{penalty}} \quad (15)$$

where r_{capacity} , $r_{\text{stability}}$ and r_{penalty} denote the traffic efficiency reward, traffic flow stability reward, and capacity constraint penalty, respectively; ω_i , $i \in \{1, 2, 3\}$ represent their corresponding weight coefficients.

(1) The traffic efficiency reward is,

$$r_{\text{capacity}} = \frac{C_{RL} - C_{\text{base}}}{C_{\text{base}}} \times 100\% \quad (16)$$

where C_{RL} is the lane capacity under the DRL-based platooning strategy. Without loss of generality, the baseline capacity C_{base} is defined as the lane capacity under the PS2 hybrid platooning scheme.

(2) The traffic flow stability reward is,

$$r_{\text{stability}} = \frac{G_h^{RL} - G_h^{\text{base}}}{G_h^{\text{base}}} \times 100\% \quad (17)$$

where G_h^{RL} is the hybrid traffic flow stability discriminant value under the DRL-based platooning strategy; and G_h^{base} is the corresponding value under the PS2 hybrid platooning scheme.

Following the mixed traffic stability criterion established by [Ngoduy \(2013\)](#) and [Xie et al. \(2019\)](#), G_h is evaluated under the DRL-based and PS2 platooning strategies to compare their impacts on mixed traffic flow stability:

$$G_h = \sum_{j=1}^J \left[P_{H_j} \frac{(f_v^j)^2 - 2f_s^j - 2f_v^j f_{\Delta v}^j}{(f_s^j)^2} \right] \quad (18)$$

where $j \in J = \{1, 2, \dots, 6\}$ corresponds to the six car-following behavior types in Section 3.3, and P_{H_j} is the corresponding proportion. The terms f_v^j , f_s^j , and $f_{\Delta v}^j$ represent the partial derivatives of the car-following model for type j with respect to velocity, spacing, and velocity difference, respectively.

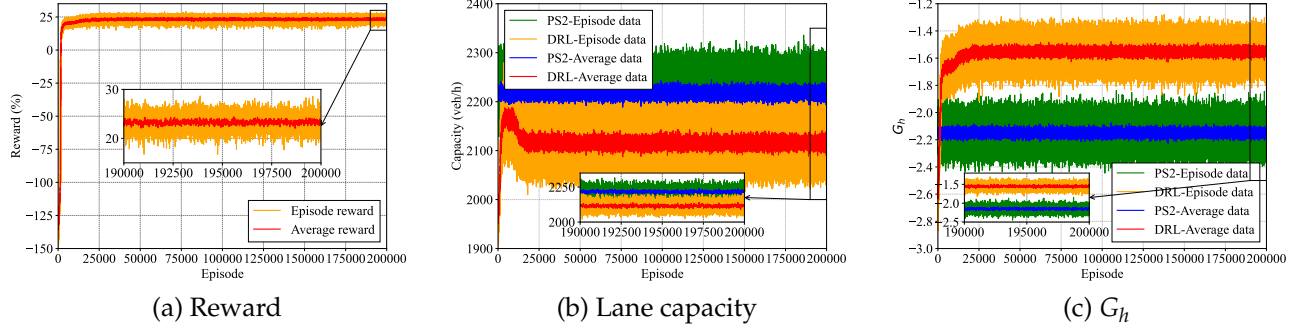


Figure 3: Evolution of reward, lane capacity, and G_h during the training process.

(3) The capacity constraint penalty term is,

$$r_{\text{penalty}} = \begin{cases} -1, & \frac{C_{\text{RL}} - C_{\text{base}}}{C_{\text{base}}} < 0.05 \\ 0, & \text{else} \end{cases} \quad (19)$$

This reward term penalizes significant lane capacity losses.

To prevent single-objective dominance in policy updates and ensure the model improves traffic flow stability while maintaining traffic efficiency, this study assigns equal weights to the traffic efficiency and stability reward terms. Furthermore, since the impact of platoon structure changes on traffic efficiency and flow stability exhibits coupling characteristics, identical weights compel the agents (CHVs/CAVs) to dynamically trade off between these objectives and avoid extreme strategies. To prioritize preventing significant capacity degradation and avoid sacrificing traffic efficiency for excessive stability optimization, a fixed efficiency loss threshold is established as a boundary condition. This is reinforced by assigning a larger weight coefficient to amplify the impact of excessive efficiency loss.

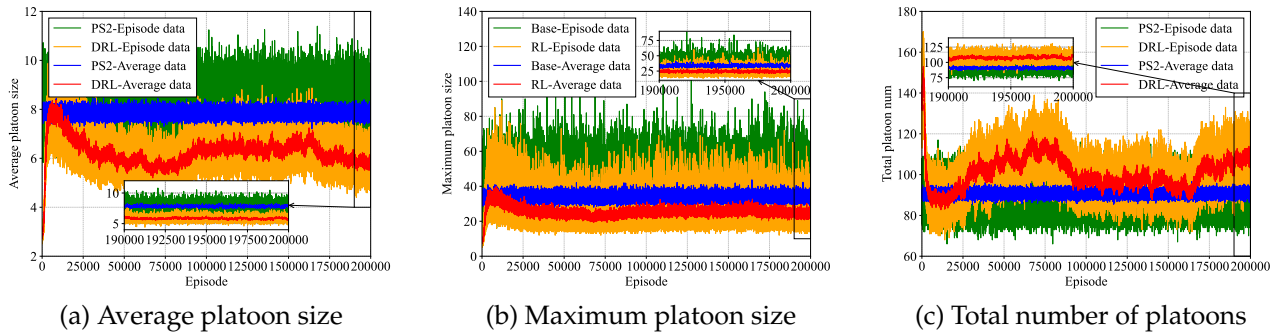


Figure 4: Evolution of platoon characteristics during the training process.

4.4. Training framework

To address platooning behavior decision-making in large-scale mixed traffic flows, this study adopts the Centralized Training with Decentralized Execution (CTDE) framework (Bernstein et al., 2000). This architecture obtains globally optimal policies through centralized training, while during decentralized execution, each vehicle makes real-time decisions relying solely on local observations, effectively mitigating the curse of dimensionality in large-scale traffic flows.

(1) Policy network π_θ : The policy network employs a dual-network structure. The feature extraction layer consists of bilinear fully-connected layers and ReLU activation functions to capture nonlinear vehicle features. The final layer outputs action probability distributions through a Sigmoid activation function.

(2) Value network V_ϕ : The value network serves as a global state value estimator. It fuses multi-agent state information via mean pooling, with linear layers outputting value estimates, thereby establishing a mapping between global states and values.

(3) PPO algorithm: The PPO algorithm ensures policy update stability through a clipped objective function and experience replay (Lin, 1991).

(4) Generalized advantage estimation (GAE) algorithm (Schulman et al., 2015): The GAE algorithm combines temporal difference errors and multi-step prediction concepts, computing advantage estimates via exponentially weighted averaging to enhance the stability and efficiency of policy gradient methods.

The detailed training procedure is outlined in Algorithm 1.

4.5. Training of the DRL-based platooning model

This subsection systematically analyzes the training process of the DRL-based hybrid platooning strategy, examining policy optimization dynamics through the convergence of the reward function, lane capacity, traffic flow stability, and platoon state evolution. The parameters for this training procedure are provided in Table 2.

Fig. 3(a) illustrates the reward progression during training, with a moving average (window size: 20 episodes) computed to mitigate stochasticity from vehicle sequencing. During initial training (<5000 episodes), the reward exhibits significant volatility with negative moving averages. This originates from the DRL-based platooning strategy’s initial inability to effectively coordinate platoon formation with traffic efficiency. As evidenced in Figs. 4(a)-(b), both average and maximum platoon lengths remain suboptimal during this phase, resulting in insufficient vehicles maintaining small time headways for platooning. Consequently, lane capacity experiences significant degradation (Fig. 3(b)).

Algorithm 1 PPO-based platooning algorithm with CTDE architecture

Inputs: Vehicle number N , observation dimension d_s , training epochs K , discount factor γ , GAE parameter λ , learning rate α_{ppo} , clipping parameter ϵ , loss weights c_1 , c_2 , reward weights $\omega_1, \omega_2, \omega_3$

Outputs: Optimal policy π_θ^*

```
1: Initialize policy network  $\pi_\theta$  and value network  $V_\phi$ 
2: Initialize experience buffer  $\mathcal{B}$ 
3: for  $k = 1$  to  $K$  do
4:   for  $t = 0$  to  $T - 1$  do
5:     Observe global state  $\mathbf{s}_t = \{o_t^i\}_{i=1}^N$ 
6:     for  $i = 1$  to  $N$  do
7:       Sample action  $a_t^i \sim \pi_\theta(o_t^i)$ 
8:     end for
9:     Execute joint action  $\mathbf{a}_t$ , observe  $(\mathbf{s}_{t+1}, r_t)$ 
10:    Store  $(\mathbf{s}_t, \mathbf{a}_t, r_t, V_\phi(\mathbf{s}_t))$  in  $\mathcal{B}$ 
11:  end for
12:  Compute GAE advantage estimates  $\hat{A}_t$ 
13:  for  $m = 1$  to  $M$  do
14:    Sample minibatch from  $\mathcal{B}$ 
15:    Compute importance ratio  $\rho_t$ 
16:    Update policy  $\pi_\theta$  using clipped PPO objective
17:    Update value function  $V_\phi$  with squared error loss
18:  end for
19: end for
```

Table 2: Parameters for training the DRL-based platooning model

Parameter	Value
Total vehicle number, N	1000
Local observation state dimension, d_s	11
Total training epochs, K	2×10^5
Discount factor, γ	0.99
GAE parameter, λ	0.95
PPO learning rate, α_{ppo}	3×10^{-4}
Policy loss clipping hyperparameter, ϵ	0.2
Value loss weight coefficient, c_1	0.5
Entropy regularization weight coefficient, c_2	0.01
Traffic efficiency reward weight, ω_1	0.5
Traffic flow stability reward weight, ω_2	0.5
Traffic capacity constraint penalty weight, ω_3	100

As training progresses, the value network progressively learns the mapping between platoon formation and lane capacity through experience replay. Concurrently, the policy network

increasingly favors multi-vehicle cooperative platooning actions. The average reward subsequently demonstrates rapid improvement, ultimately stabilizing within a narrow band centered at 23.30%, confirming the stability of the DRL-based platooning strategy.

Fig. 3 illustrates the mechanistic impact of DRL-based platooning strategy on macroscopic traffic flow characteristics. During the training phase with episodes ≥ 25000 , the single-lane capacity under the DRL-based platooning strategy converges to 2115veh/h, representing a 4.6% reduction compared to the hybrid platooning strategy (2217veh/h). This capacity compromise, however, achieves a substantial stability enhancement: the hybrid traffic flow stability discriminant value increases from -2.15 (hybrid strategy PS2) to -1.55 , marking a 27.9% improvement. This performance trade-off empirically validates the inherent efficiency-stability contradiction in mixed traffic flows. The DRL strategy effectively balances these competing objectives through persistent exploratory training, demonstrating its capability to navigate complex traffic optimization landscapes.

Fig. 4 further validates the microscopic evolutionary mechanisms of DRL-based platooning strategy optimization through hybrid traffic flow platoon state evolution. The DRL strategy regulates the average platoon length within a steady-state band centered at 5.84, representing a 25.5% reduction compared to the hybrid platooning scheme PS2. The maximum platoon length is reduced by 28.3% to 24.3, effectively mitigating stability risks from large-scale platoon aggregation. Notably, the number of platoons increases by 19.3% to 108.3 under the DRL-based platooning strategy, indicating optimized platoon structure through enhanced distribution density of medium-scale platoons while maintaining stable total platooned vehicles. This platooning approach simultaneously mitigates stability degradation from oversized platoons via optimized length parameters and alleviates capacity loss from reduced platoon sizes through increased platoon count.

To further examine the suitability of the adopted PPO-based training framework, we compare it with two representative alternatives, namely IPPO (Shi et al., 2025b) and MADDPG (Shi et al., 2023). The comparison includes: reward, lane capacity, the traffic-flow stability criterion G_h , average platoon size, maximum platoon size, and the total number of platoons. As shown in Fig. 5 (a)–(c), PPO achieves the highest and most stable reward and converges to better macroscopic traffic performance, with higher lane capacity and a more favorable traffic-flow stability criterion G_h . This improvement can be attributed to the clipped policy update in PPO, which constrains abrupt policy changes and thus helps maintain a more consistent platooning strategy during training. The platoon-level results in Fig. 5 (d)–(f) further show that PPO forms platoons with sufficient scale while avoiding excessive platoon expansion. In con-

trast, IPPO relies on more independent agent-wise learning and tends to produce smaller and more fragmented platoons, which benefits stability but limits capacity improvement. MADDPG forms platoons with sizes comparable to those of PPO, but its lower final reward and G_h suggest less effective coordination in the proposed hybrid platooning task.

In summary, experimental results from the training process validate the effectiveness of the DRL-based platooning strategy in addressing multi-objective optimization challenges in mixed traffic flows. The policy network achieves a nonlinear equilibrium between lane capacity and traffic flow stability. This capability fundamentally stems from optimizing hybrid traffic flow platoon structures through the policy network, whereby regulating platoon scale distributions enables nonlinear balancing of traffic efficiency and flow stability. Training data quantitatively demonstrate that compared to traditional fixed hybrid-platooning schemes, the DRL approach significantly enhances traffic flow stability within acceptable traffic efficiency margins.

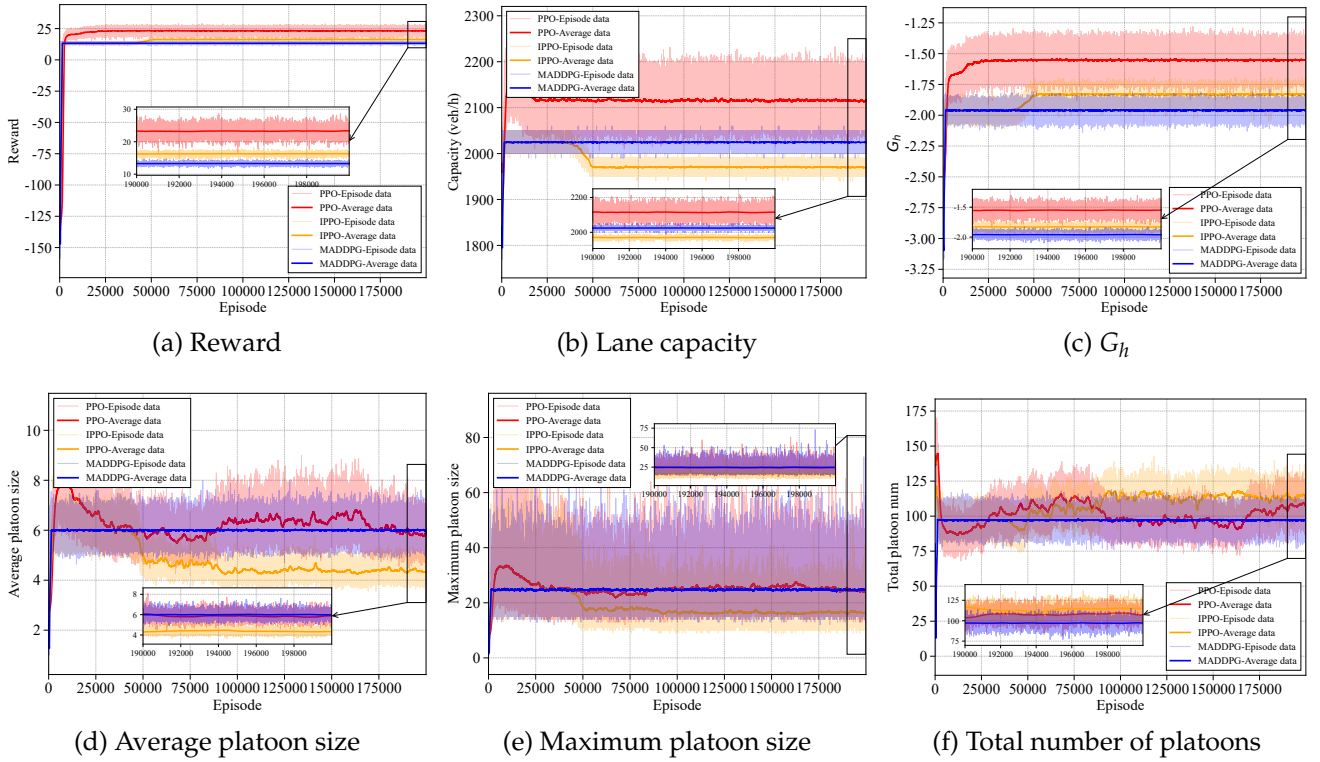


Figure 5: Comparison of training performance among PPO, IPPO, and MADDPG.

5. Performance evaluation of the DRL-based platooning strategy

To validate the effectiveness of the DRL-based platooning strategy, this section conducts numerical simulations across four critical dimensions: model generalization capability, traf-

fic flow stability, traffic safety, and energy consumption/emissions. Performance of the DRL-based strategy is comparatively evaluated against hybrid platooning scheme PS2 and traditional connected vehicle platooning scheme PS1.

5.1. Generalization capability of the strategy

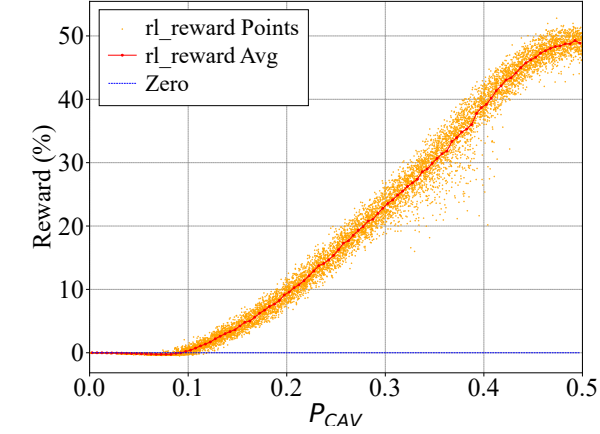
Through rigorous numerical experiments, this subsection validates the generalization capability of the proposed DRL-based platooning strategy in terms of penetration rates of CHVs/CAVs. Employing a controlled-variable approach, the experimental framework incorporates 10000 independent randomized trials to mitigate stochastic variations.

The experimental framework configures vehicle penetration parameters as $P_{CAV} = P_{CHV}$, $P_{AV} = P_{HV}$, with P_{CAV} following uniform distribution over $[0.0, 0.5]$. Resultant performance metrics are illustrated in Fig. 6.

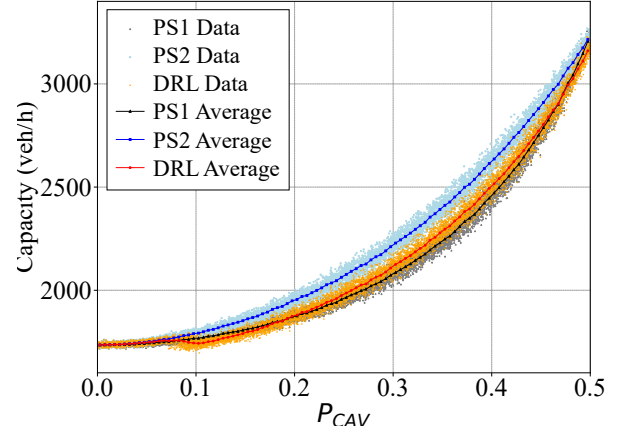
Fig. 6(a) illustrates the evolutionary trend of the reward function, indicating that the reward value approaches 0 when $P_{CAV} < 0.1$. Under this condition, the sparse distribution of CHVs and CAVs in mixed traffic flow impedes platoon formation, resulting in negligible performance differences among the three platooning strategies. As the CAV penetration rate increases within the range $P_{CAV} \in (0.1, 0.5)$, the DRL-based strategy demonstrates progressively enhanced optimization of platoon structure, yielding an approximately linear increase in reward value.

Fig. 6(b) reveals that lane capacity exhibits a distinct quadratic growth pattern with increasing CAV penetration. This phenomenon stems from CAVs' facilitation of platoon formation in mixed traffic flows. Notably, lane capacity shows insignificant variation across the three strategies when $P_{CAV} < 0.1$ or $P_{CAV} \rightarrow 0.5$. Within the interval $P_{CAV} \in (0.2, 0.45)$, the hybrid platooning strategy maximizes collaborative gains through CAVs' information relay and topology reconfiguration capabilities, achieving superior lane capacity compared to both DRL-based and conventional connected platooning strategies. However, the performance advantage comes at the cost of reduced traffic flow stability.

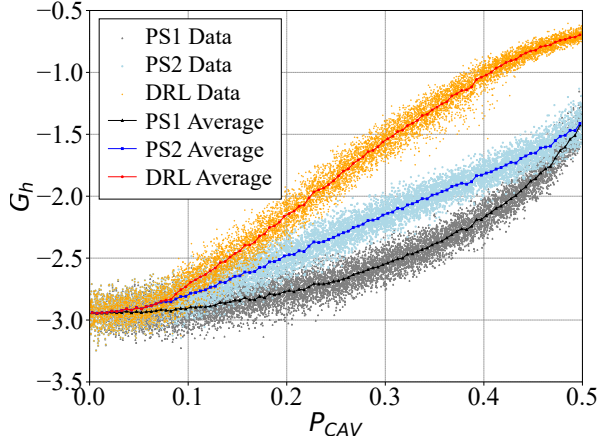
Fig. 6(c) demonstrates that the DRL-based platooning strategy delivers a 51.08% improvement in stability discriminant value over both hybrid and conventional connected strategies when $P_{CAV} > 0.1$. This enhancement is attributed to the strategy's online learning optimization of platoon structures in mixed traffic flows. In contrast, hybrid and conventional schemes (PS2 and PS1), constrained by fixed formation patterns, struggle to balance traffic efficiency with flow stability. Crucially, the DRL-based strategy maintains optimal traffic stability across all CAV penetration rates, fundamentally enabled by its integration of centralized optimization's global perspective and distributed control's local adaptability.



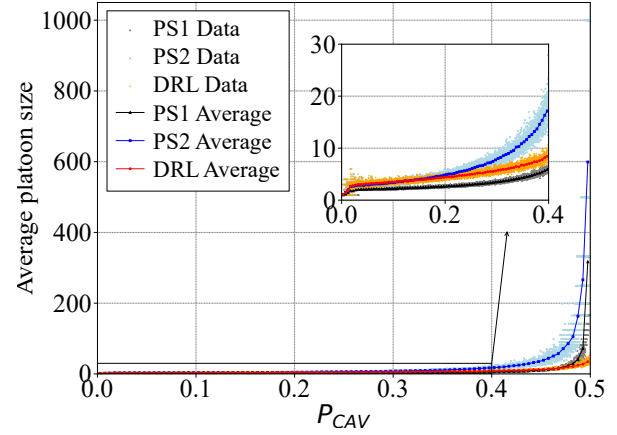
(a) Reward value



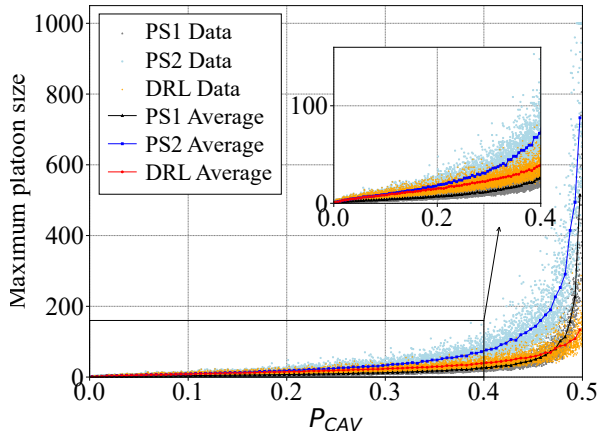
(b) Lane capacity



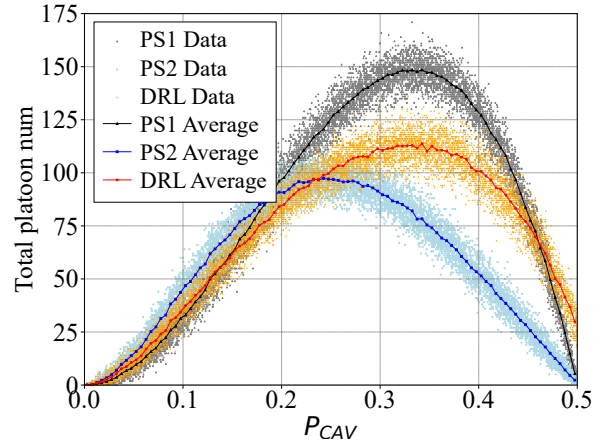
(c) G_h



(d) Average platoon size



(e) Maximum platoon size



(f) Total platoon number

Figure 6: Generalization performance of the proposed model under different P_{CAV} conditions.

Figs. 6(d)-(f) further illustrate the evolutionary mechanisms of platoon characteristics in mixed traffic flows in terms of platoon size and number. Fig. 6(d) demonstrates that under both hybrid and conventional platooning schemes (PS2 and PS1), increased CAV penetration at higher rates induces exponential growth in average platoon length. In contrast, the DRL-based platooning strategy maintains this metric at 35 vehicles through platoon optimization, effectively mitigating length inflation caused by rigid formation patterns in hybrid strategies. As shown in Fig. 6(e), maximum platoon length exhibits unbounded exponential growth under schemes PS2 and PS1, posing significant stability challenges to mixed traffic flows. The DRL-based platooning strategy, however, prevents irrational expansion of maximum length by incorporating stability reward constraints. Fig. 6(f) reveals a parabolic trend in platoon quantity, initially increasing then decreasing with rising P_{CAV} . Peak values occur at $P_{\text{CAV}} = 0.24$ for both DRL-based strategy and PS1 scheme (97 platoons), while the hybrid scheme PS2 peaks later at $P_{\text{CAV}} = 0.34$ (148 platoons).

Collectively, repeated independent experiments verify the generalizability and effectiveness of the DRL-based platooning strategy across diverse penetration scenarios. By optimizing platoon structures to reduce both average and maximum lengths, the DRL-based platooning strategy significantly enhances mixed traffic flow stability and achieves robust strategy transferability irrespective of vehicle penetration rates.

5.2. Stability of traffic flow

This study conducts traffic flow simulations under open boundary conditions to quantitatively analyze the impact of three platooning strategies on disturbance propagation characteristics in mixed traffic flows. The experimental setup involves 100 heterogeneous vehicles randomly distributed on a single-lane road, with all vehicles strictly adhering to their respective vehicle-type-specific car-following dynamics and platoon behavioral models.

Under fixed proportions $P_{\text{CAV}} = P_{\text{CHV}}$, $P_{\text{AV}} = P_{\text{HV}}$, four penetration scenarios with $P_{\text{CAV}} = 0.1, 0.2, 0.3, 0.4$ were tested, respectively. Following system initialization, traffic flow operates at steady-state velocity $v_e = 20$ m/s for 600s to achieve dynamic equilibrium. Subsequently, a disturbance is applied to the leading vehicle: a constant deceleration of -1 m/s² sustained for 3s, followed by acceleration at 1 m/s² until v_e is recovered. This disturbance velocity profile is

mathematically described as a piecewise linear function:

$$v_1(t) = \begin{cases} v_e, & 0 < t \leq 600 \\ v_e - (t - 600), & 600 < t \leq 603 \\ v_e + (t - 606), & 603 < t \leq 606 \\ v_e, & t > 606 \end{cases} \quad (20)$$

Fig. 7 illustrates the velocity propagation characteristics of each platooning strategy across four penetration scenarios. Experimental data reveal that under conventional connected platooning PS1, steady-state recovery times corresponding to P_{CAV} values of 0.1, 0.2, 0.3, and 0.4 measure 300s, 265s, 240s, and 190s respectively. Reduced recovery times at higher CAV penetration confirm a positive stability-penetration correlation.

As depicted in Figs. 7(c), (f), and (i), conventional platooning PS1 exhibits significant velocity oscillation amplification in low-penetration scenarios ($P_{CAV} \leq 0.3$). In contrast, both DRL-based strategy (Figs. 7(a), (d), and (g)) and hybrid scheme PS2 (Figs. 7(b), (e), and (h)) reduce average velocity fluctuations by approximately 1 m/s. Furthermore, the DRL-based platooning strategy achieves lower velocity fluctuation compared to the other strategies.

When P_{CAV} reaches 0.4 (Figs. 7(j), (k), and (l)), the DRL-based strategy demonstrates notable platoon synergy effects, maintaining smaller velocity and acceleration fluctuations while enabling faster recovery to pre-disturbance steady-state operation.

Collectively, these experiments validate the effectiveness of DRL-based platooning in controlling disturbance propagation within mixed traffic environments through dual perspectives: velocity oscillation mitigation and acceleration response optimization.

5.3. Traffic safety

To evaluate the impact of platooning strategies on the safety of mixed traffic, this study employs the inverse time to collision (TTCI) and velocity standard deviation (SD) as safety metrics (Yao et al., 2023; Li et al., 2024; Hua et al., 2023; Luo et al., 2024). The TTCI dynamically quantifies longitudinal collision risk for vehicle i at time t , which is:

$$\text{TTCI}_i(t) = \begin{cases} \frac{v_i(t) - v_{i-1}(t)}{x_{i-1}(t) - x_i(t)}, & v_i(t) > v_{i-1}(t) \\ 0, & \text{Otherwise} \end{cases} \quad (21)$$

where $x_i(t)$ and $v_i(t)$ denote the position and velocity of vehicle i at time t , respectively. This dimensionally consistent metric assesses collision risk by calculating the velocity differential relative to inter-vehicle spacing, with higher values indicating elevated collision probability.

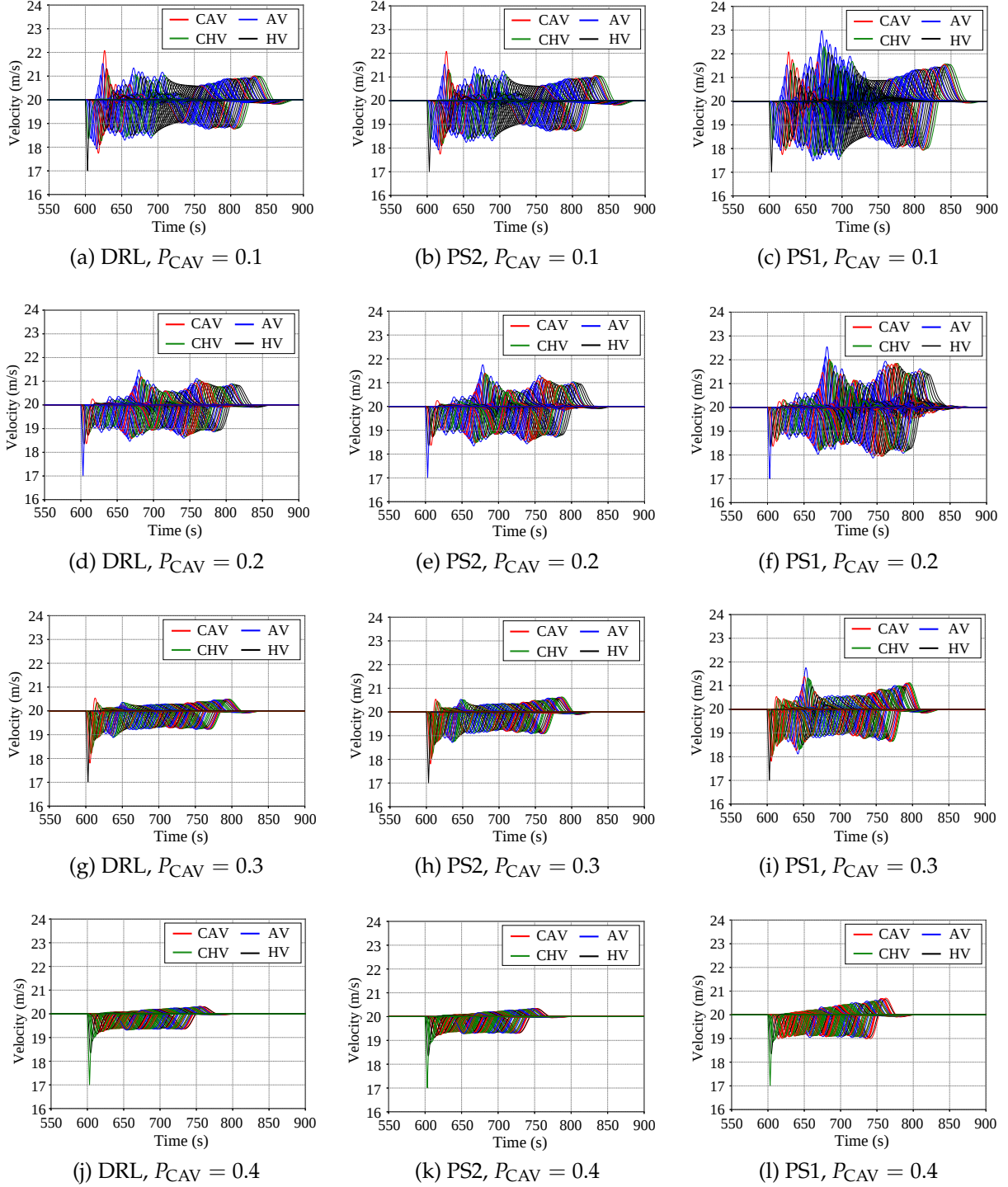


Figure 7: Velocity responses to head-vehicle perturbations under different penetration rates P_{CAV} . For $P_{CAV} = 0.1, 0.2, 0.3$, and 0.4 , the three subfigures from left to right correspond to the DRL-based strategy, hybrid strategy (PS2), and conventional strategy (PS1), respectively.

The velocity SD is mathematically defined as:

$$SD(t) = \sqrt{\frac{1}{N} \sum_{i=1}^N (v_i(t) - \bar{v}(t))^2} \quad (22)$$

where $\bar{v}(t)$ denotes the space-mean velocity of traffic flow at time t . This metric quantifies velocity dispersion within the traffic stream, with its temporal evolution capturing spatiotemporal patterns of velocity oscillations during disturbance propagation.

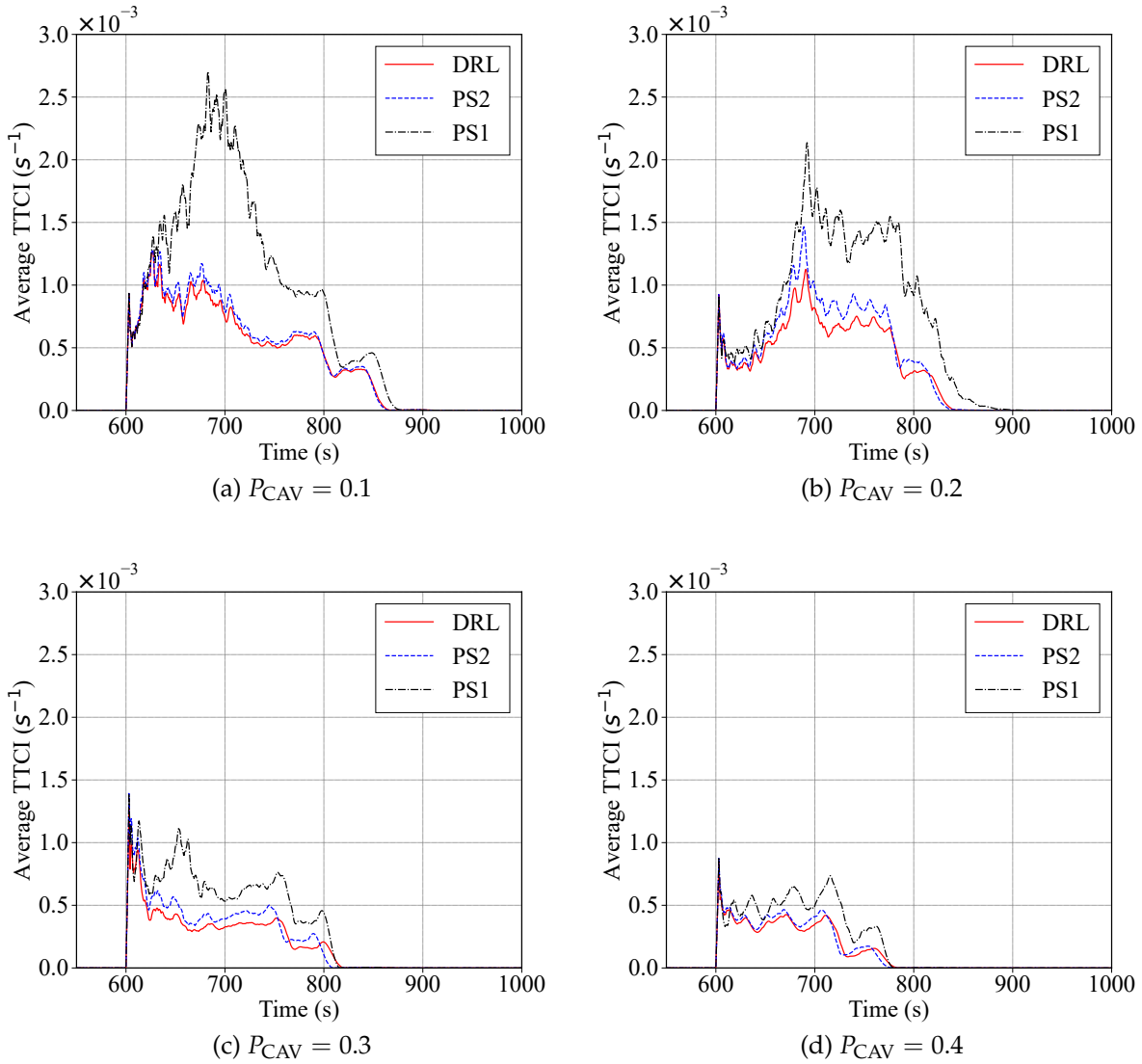


Figure 8: Average inverse time-to-collision (TTCI) under different penetration rates P_{CAV} .

Similar to the simulation scenarios in Section 5.2, Figs. 8 and 9 present the dynamic re-

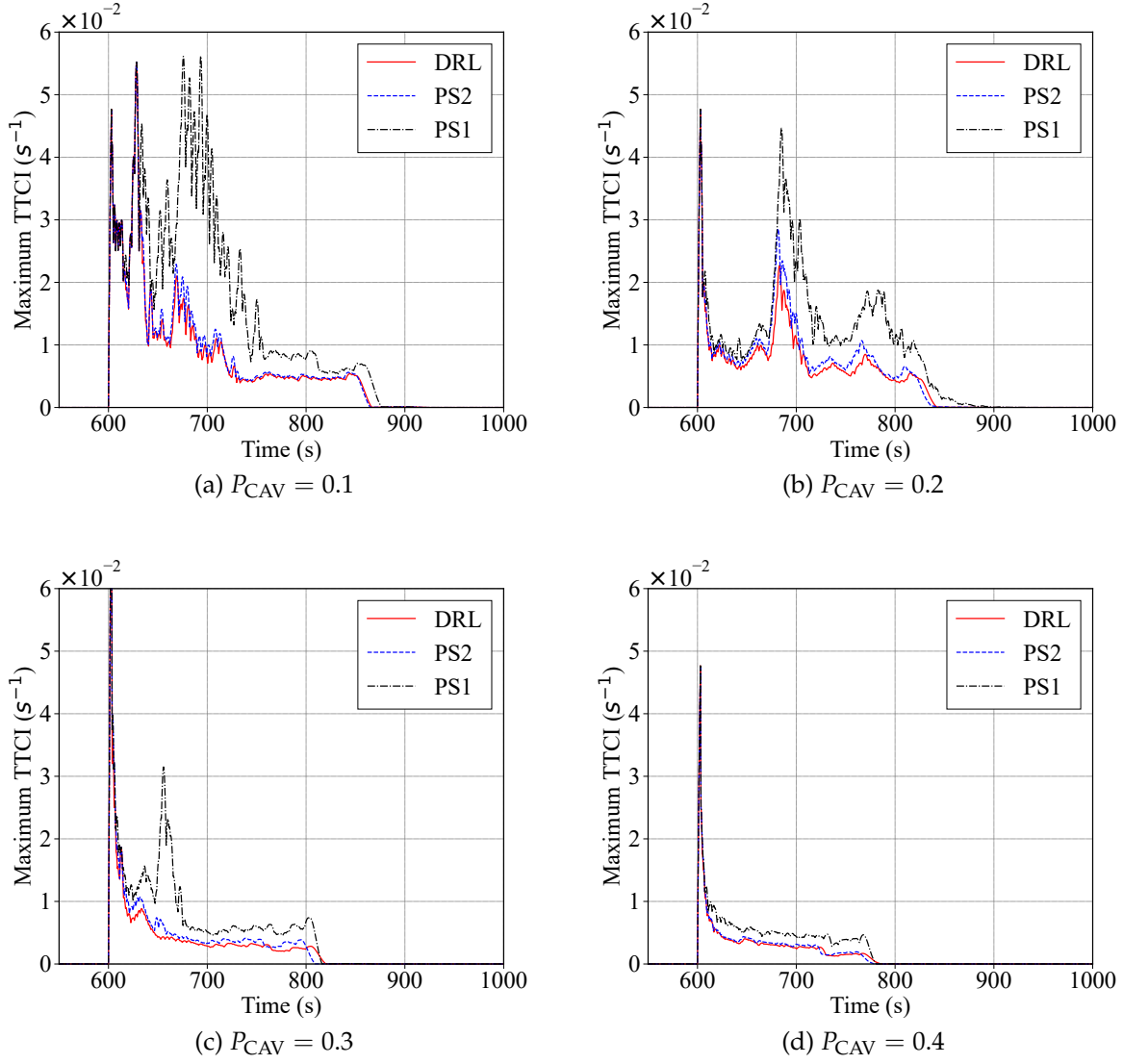


Figure 9: Peak inverse time-to-collision (TTCl) under different penetration rates P_{CAV} .

sponses of average ($\overline{TTCl}$) and maximum ($TTCl_{max}$) with various penetration rates. Experimental data indicate that under high CAV penetration ($P_{CAV} = 0.4$), all three platooning strategies maintain $\overline{TTCl}$ below 0.001 s^{-1} , confirming CAV cooperative control's effectiveness in suppressing risk propagation. Notably, the DRL-based strategy demonstrates significant advantages at low penetration ($P_{CAV} = 0.1$), reducing peak $\overline{TTCl}$ by 52.95% compared to conventional platooning PS1 and by 23.21% versus the hybrid strategy PS2.

Regarding extreme collision risks, Fig. 9(a) reveals violent $TTCl_{max}$ oscillations under conventional platooning PS1 at $P_{CAV} = 0.1$, indicating collision risk accumulation with transient peaks reaching 0.0562 s^{-1} . In contrast, the DRL-based strategy maintains stable extreme risk

values, demonstrating superior control over critical collision scenarios.

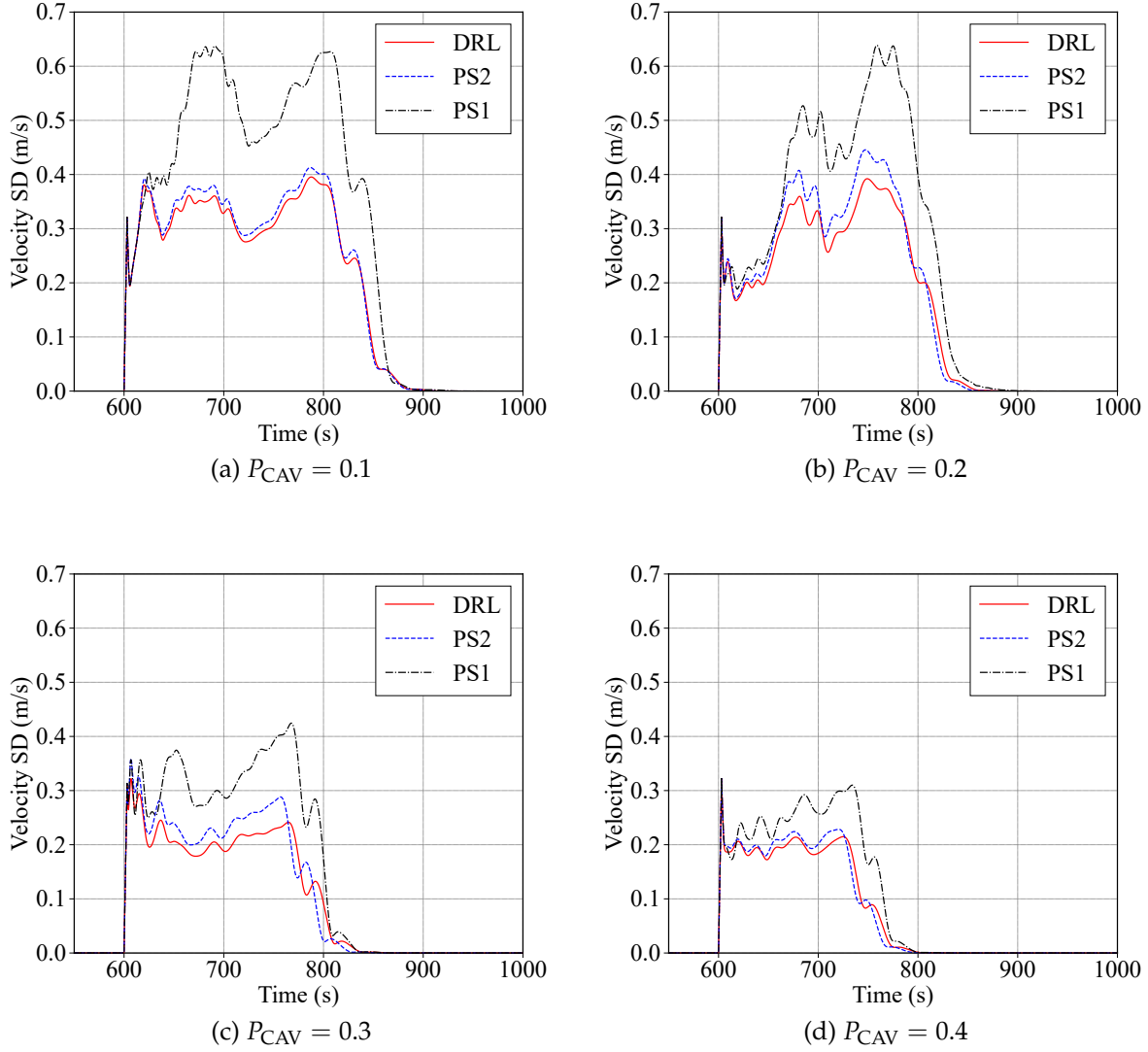


Figure 10: Velocity standard deviation (SD) under different penetration rates P_{CAV} .

Fig. 10 reveals significant stability differences among platooning strategies through velocity SD analysis. The DRL-based scheme reduces peak velocity SD by up to 38.60% compared to conventional platooning and 11.97% versus the hybrid scheme PS2, demonstrating its effectiveness in suppressing velocity oscillation propagation within mixed traffic flows under disturbance scenarios.

The experiments demonstrate platooning strategy effects on traffic safety from various perspectives. The results confirm that the DRL-based strategy, through synergistic platoon structure optimization, effectively mitigates longitudinal collision risks, suppresses velocity oscilla-

tion propagation, and ensures operational safety in mixed traffic environments.

It should be noted that the above safety evaluation is conducted under ideal communication and perception conditions. In practice, communication failures, sensor faults, or observation noise may impair vehicle-state acquisition, especially for HVs/AVs, thereby weakening platoon integrity and safety. Future work will therefore examine fault-induced robustness, fail-safe control strategies, and more comprehensive traffic-risk assessment methods, such as causal graph-based conflict risk prediction (Lyu et al., 2025).

5.4. Energy consumption and emissions

Recent years have witnessed extensive research by scholars and institutions on vehicle fuel consumption and emissions using the Measure of Effectiveness (MOE) framework. As a quantitative metric, MOE quantifies fuel consumption and emissions per unit distance traveled. Studies establish that instantaneous MOE represents a mathematical function of a vehicle's instantaneous velocity and acceleration, yielding unique values when these parameters are specified.

This study employs the VT-Micro model, originally developed by Ahn (1998), and validated through extensive experimental observations, to compute fuel consumption and emissions for vehicles in mixed traffic flows. Within this model, MOE is calculated as:

$$\text{MOE}_e(t) = e^{P(v(t), a(t))} \quad (23)$$

where $\text{MOE}_e(t)$ denotes instantaneous fuel consumption or emission values at time t ; $e \approx 2.71828$ represents the natural logarithm base, and $P(v(t), a(t))$ constitutes a function of instantaneous velocity and acceleration:

$$P(v(t), a(t)) = \sum_{i=0}^3 \sum_{j=0}^3 K_{i,j} v(t)^i a(t)^j \quad (24)$$

where $K_{i,j}$ represents regression coefficients with power exponent i for velocity and j for acceleration.

Fuel consumption and emission test data across diverse vehicle types have revealed distinct energy and emission characteristics during acceleration and deceleration phases (Li et al., 2014). Consequently, three independent sets of regression coefficients $K_{i,j}$ were derived through VT-Micro modeling for fuel consumption, CO_2 , and NO_x emissions, respectively.

Building upon the disturbance propagation scenario in Section 5.2, Figs. 11-13 delineate the evolution of energy consumption and emission under different platooning strategies. Experi-

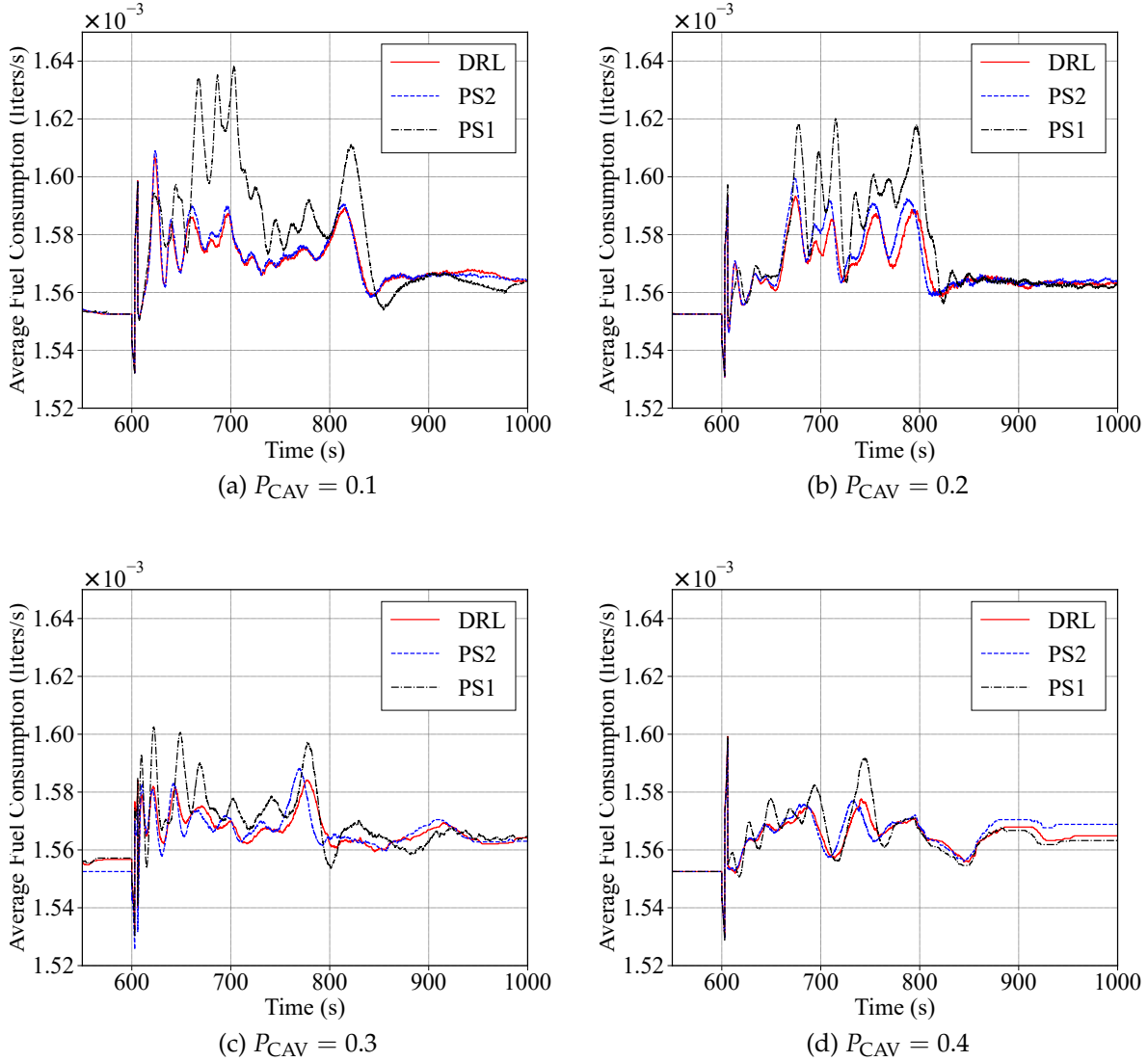


Figure 11: Average fuel consumption under different penetration rates P_{CAV} .

mental results indicate that at low CAV penetration rates ($P_{CAV} \leq 0.3$), the DRL-based strategy reduces peak average energy consumption by up to 1.90% and 0.37% compared to conventional and hybrid platooning schemes (PS1 and PS2) respectively, effectively mitigating incomplete fuel combustion caused by frequent acceleration/deceleration.

Analysis of emission characteristics reveals divergent response mechanisms between CO_2 and NO_x . For CO_2 emissions, the DRL-based platooning strategy demonstrates performance consistent with its energy consumption profile: under low CAV penetration scenarios, it reduces CO_2 emissions by over 1.2% compared to conventional platooning PS1, while achieving up to 7.1% reduction in NO_x emissions. Relative to hybrid platooning PS2, the DRL-based

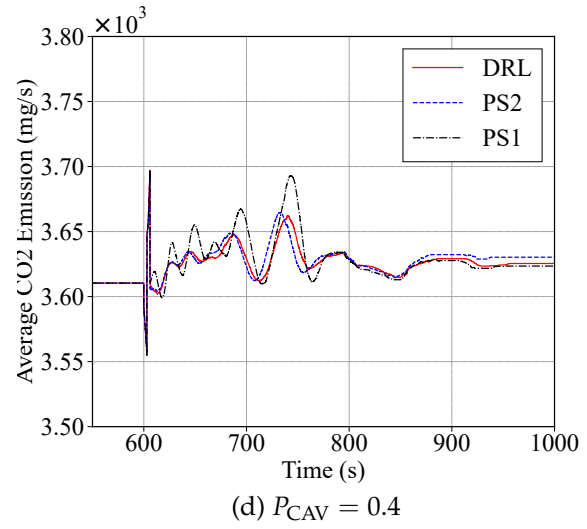
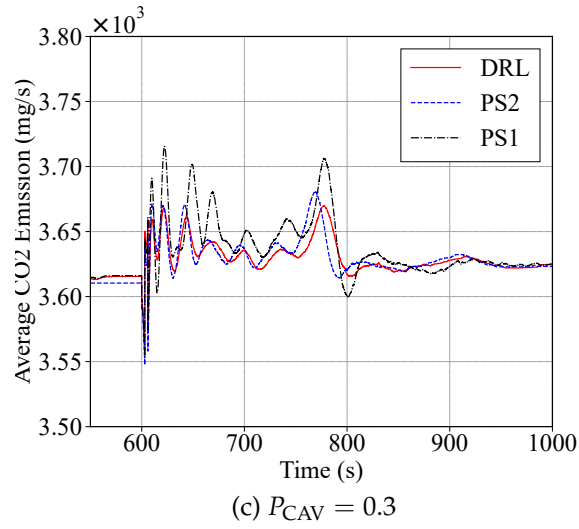
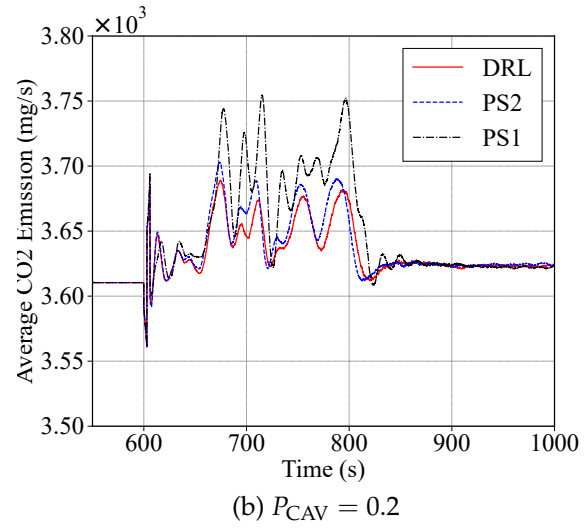
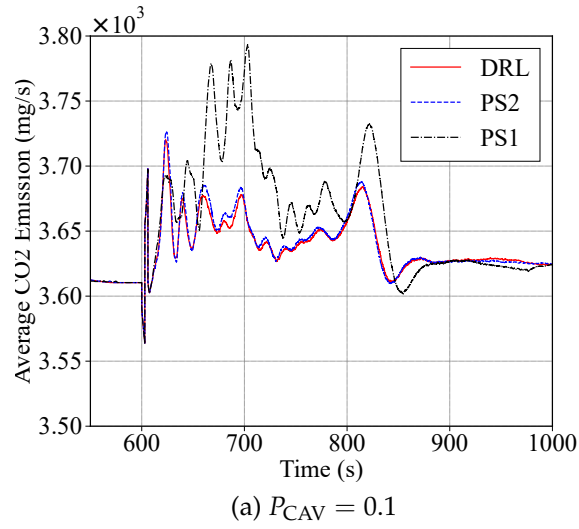


Figure 12: Average CO₂ emissions under different penetration rates P_{CAV} .

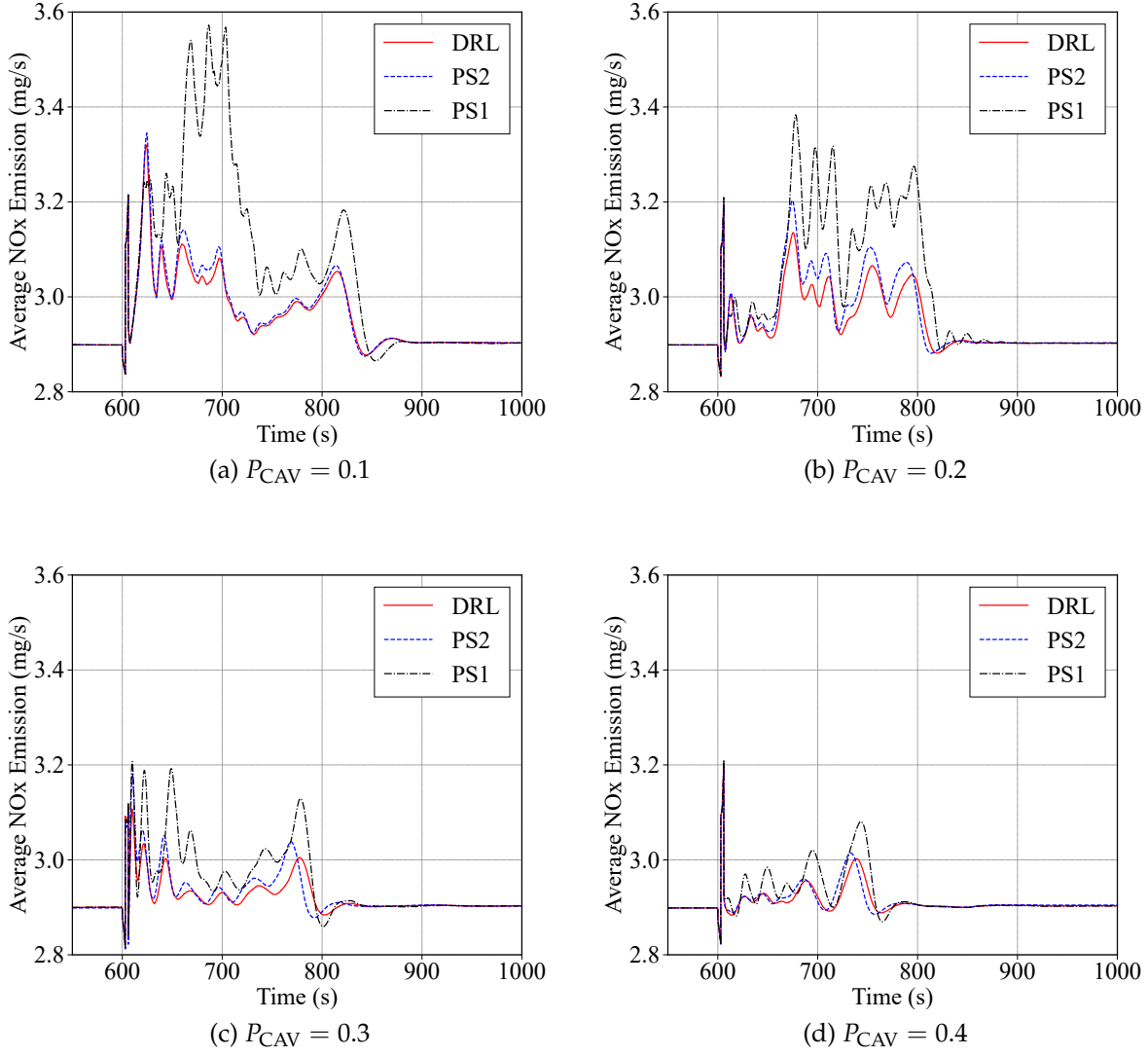


Figure 13: Average NO_x emissions under different penetration rates P_{CAV} .

strategy yields reductions of 0.37% in CO₂ and 2.1% in NO_x emissions. These experimental findings substantiate the energy efficiency and environmental benefits of DRL-based strategy from both fuel consumption and emission perspectives.

6. Conclusions and future research

To overcome the limitations of traditional CAV-exclusive platooning in mixed traffic, we propose a flexible platoon formation approach that incorporates HVs/AVs, controlled via a novel DRL framework. We further develop a DRL-based platooning strategy for hybrid platoons in mixed traffic. Our method leverages a POMDP formulation with multi-level state

representations to enable distributed control of heterogeneous platoons. Experimental results show that the proposed strategy improves traffic stability mainly by regulating platoon structures rather than simply maximizing platoon formation. In the training evaluation, the learned policy improves the traffic stability discriminant by 27.9% and reduces the maximum platoon length by 28.3% compared with the conventional hybrid platooning strategy. In the disturbance-propagation evaluation, the trained policy reduces peak average fuel consumption by up to 1.90% and CO₂ emissions by over 1.2% compared with the conventional connected platooning strategy, while reducing NO_x emissions by up to 7.1% under low CAV penetration conditions. The strategy demonstrates robust generalization across varying penetration rates $P_{\text{CAV}} \in (0.1, 0.45)$, effectively balancing capacity-stability tradeoffs while suppressing disturbance propagation throughout the traffic flow. While these findings highlight the effectiveness of the proposed framework, several practical issues remain to be further investigated. The current study adopts a simplified mixed-traffic setting to focus on heterogeneous platoon organization. In real-world traffic, multi-lane interactions, uncertain events, and dynamic platoon splitting and re-forming may jointly affect platoon stability, safety, and efficiency. Future work will extend the framework by incorporating these complex operational conditions and robustness considerations.

References

- Ahn, K., 1998. Microscopic Fuel Consumption and Emission Modeling. Master's thesis. Virginia Polytechnic Institute and State University. Blacksburg, VA, USA.
- Bernstein, D.S., Zilberstein, S., Immerman, N., 2000. The complexity of decentralized control of markov decision processes, in: Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. pp. 32–37.
- Chang, X., Li, H., Rong, J., Zhao, X., et al., 2020. Analysis on traffic stability and capacity for mixed traffic flow with platoons of intelligent connected vehicles. *Physica A: Statistical Mechanics and Its Applications* 557, 124829.
- Du, G., Zou, Y., Zhang, X., Fan, J., Sun, W., Li, Z., 2024. Efficient motion control for heterogeneous autonomous vehicle platoon using multilayer predictive control framework. *IEEE Internet of Things Journal* 11, 38273–38290.
- Feng, S., Song, Z., Li, Z., Zhang, Y., Li, L., 2021. Robust platoon control in mixed traffic flow based on tube model predictive control. *IEEE Transactions on Intelligent Vehicles* 6, 711–722.

- Ghiasi, A., Hussain, O., Qian, Z., Li, X., 2020. Lane management with variable lane width and model calibration for connected automated vehicles. *Journal of Transportation Engineering, Part A: Systems* 146, 04019075.
- Ghiasi, A., Hussain, O., Qian, Z.S., Li, X., 2017. A mixed traffic capacity analysis and lane management model for connected automated vehicles: A markov chain method. *Transportation Research Part B: Methodological* 106, 266–292.
- Guan, H., Wang, H., Meng, Q., Mak, C.L., 2023. Markov chain-based traffic analysis on platooning effect among mixed semi-and fully-autonomous vehicles in a freeway lane. *Transportation Research Part B: Methodological* 173, 176–202.
- Han, Q., Ma, J., Zuo, Z., Wang, X., Guan, X., 2025. Data-driven finite-time platooning control for heterogeneous nonlinear vehicle systems. *IEEE Control Systems Letters* 9, 33–37.
- Hua, M., Chen, D., Jiang, K., Zhang, F., Wang, J., Wang, B., Zhou, Q., Xu, H., 2025. Communication-efficient marl for platoon stability and energy-efficiency co-optimization in cooperative adaptive cruise control of cavs. *IEEE Transactions on Vehicular Technology* 74, 6076–6087.
- Hua, X., Yu, W., Wang, W., Zhao, D., 2023. Impact of multi-class stochastic cyberattacks on vehicle dynamics and rear-end collision risks for heterogeneous traffic. *Physica A: Statistical Mechanics and its Applications* 626, 129095.
- Huang, Z., Chu, D., Wu, C., He, Y., 2018. Path planning and cooperative control for automated vehicle platoon using hybrid automata. *IEEE Transactions on Intelligent Transportation Systems* 20, 959–974.
- Jiang, Y., Zhu, F., Yao, Z., Gu, Q., Ran, B., 2023. Platoon intensity of connected automated vehicles: Definition, formulas, examples, and applications. *Journal of Advanced Transportation* 2023, 3325530.
- Jin, L., Čičić, M., Johansson, K.H., Amin, S., 2021a. Analysis and design of vehicle platooning operations on mixed-traffic highways. *IEEE Transactions on Automatic Control* 66, 4715–4730.
- Jin, S., Sun, D., Liu, Z., 2021b. The impact of spatial distribution of heterogeneous vehicles on performance of mixed platoon: A cyber-physical perspective. *KSCE Journal of Civil Engineering* 25, 303–315.

- Lan, J., Zhao, D., Tian, D., 2023. Data-driven robust predictive control for mixed vehicle platoons using noisy measurement. *IEEE Transactions on Intelligent Transportation Systems* 24, 6586–6596.
- Li, K., Wang, J., Zheng, Y., 2022. Cooperative formation of autonomous vehicles in mixed traffic flow: Beyond platooning. *IEEE Transactions on Intelligent Transportation Systems* 23, 15951–15966.
- Li, X., Cui, J., An, S., Parsafard, M., 2014. Stop-and-go traffic analysis: Theoretical properties, environmental impacts and oscillation mitigation. *Transportation Research Part B: Methodological* 70, 319–339.
- Li, Z., Zhou, Y., Zhang, Y., Li, X., 2024. Enhancing vehicular platoon stability in the presence of communication cyberattacks: A reliable longitudinal cooperative control strategy. *Transportation Research Part C: Emerging Technologies* 163, 104660.
- Liang, J., Liu, L., Yu, X., Wang, W., Pan, C., Chen, L., 2025. Cooperative control of mixed cav-hdv platoons with hierarchical optimization for oscillation mitigation. *Physica A: Statistical Mechanics and its Applications* 675, 130844.
- Lin, H., Lyu, C., He, Y., Liu, Y., Gao, K., Qu, X., 2024. Enhancing state representation in multi-agent reinforcement learning for platoon-following models. *IEEE Transactions on Vehicular Technology* 73, 12110–12114.
- Lin, L.J., 1991. Self-improvement based on reinforcement learning, planning and teaching, in: *Proceedings of the Eighth International Conference on Machine Learning*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. pp. 323–327.
- Liu, G., Li, Y., Liu, L., 2026. A Deep Reinforcement Learning Based Cooperative Control for Mixed Platoon with Guaranteed String Stability. *IEEE Transactions on Industrial Electronics* 73, 6324–6334.
- Liu, G., Zheng, N., Wang, H., 2025. Cooperative control method for connected and automated vehicle platoon based on arbitrary time headway switched system. *Transportation Research Part C: Emerging Technologies* 180, 105353.
- Liu, H., Kan, X., Shladover, S.E., Lu, X.Y., Ferlis, R.E., 2018a. Impact of cooperative adaptive cruise control on multilane freeway merge capacity. *Journal of Intelligent Transportation Systems* 22, 263–275.

- Liu, H., Kan, X.D., Shladover, S.E., Lu, X.Y., Ferlis, R.E., 2018b. Modeling impacts of cooperative adaptive cruise control on mixed traffic flow in multi-lane freeway facilities. *Transportation Research Part C: Emerging Technologies* 95, 261–279.
- Luo, D., Wang, J., Wang, Y., Dong, J., 2024. Modeling and analyzing self-resistance of connected automated vehicular platoons under different cyberattack injection modes. *Accident Analysis & Prevention* 198, 107494.
- Lyu, N., Xie, T., Wen, J., Wang, Y., 2025. A real-time conflict risk prediction modeling based on causal inference and graph-model: Applied to multi-configuration expressway diversion zones. *Transportation Research Part C: Emerging Technologies* 178, 105219.
- Ma, Y.S., Che, W.W., Deng, C., Wu, Z.G., 2024. Data-driven distributed vehicle platoon control for heterogeneous nonlinear vehicle systems. *IEEE Transactions on Intelligent Transportation Systems* 25, 2373–2381.
- Maiti, S., Winter, S., Kulik, L., Sarkar, S., 2023. Ad-hoc platoon formation and dissolution strategies for multi-lane highways. *Journal of Intelligent Transportation Systems* 27, 161–173.
- Milanés, V., Shladover, S.E., 2014. Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. *Transportation Research Part C: Emerging Technologies* 48, 285–300.
- Milanés, V., Shladover, S.E., Spring, J., Nowakowski, C., Kawazoe, H., Nakamura, M., 2013. Cooperative adaptive cruise control in real traffic situations. *IEEE Transactions on intelligent transportation systems* 15, 296–305.
- Ngoduy, D., 2013. Analytical studies on the instabilities of heterogeneous intelligent traffic flow. *Communications in Nonlinear Science and Numerical Simulation* 18, 2699–2706.
- Qiang, Z., Dai, L., Chen, B., Xia, Y., 2023. Distributed model predictive control for heterogeneous vehicle platoon with inter-vehicular spacing constraints. *IEEE Transactions on Intelligent Transportation Systems* 24, 3339–3351.
- Qin, Y., Luo, Q., Wang, H., 2023. Stability analysis and connected vehicles management for mixed traffic flow with platoons of connected automated vehicles. *Transportation Research Part C: Emerging Technologies* 157, 104370.

- Sala, M., Soriguera, F., 2020. Macroscopic modeling of connected autonomous vehicle platoons under mixed traffic conditions. *Transportation Research Procedia* 47, 163–170.
- Schulman, J., Moritz, P., Levine, S., Jordan, M., Abbeel, P., 2015. High-dimensional continuous control using generalized advantage estimation .
- Seraj, M., Li, J., Qiu, Z., 2018. Modeling microscopic car-following strategy of mixed traffic to identify optimal platoon configurations for multiobjective decision-making. *Journal of Advanced Transportation* 2018, 1–15.
- Shi, H., Chen, D., Zheng, N., Wang, X., Zhou, Y., Ran, B., 2023. A deep reinforcement learning based distributed control strategy for connected automated vehicles in mixed traffic platoon. *Transportation Research Part C: Emerging Technologies* 148, 104019.
- Shi, Y., Dong, H., He, C.R., Chen, Y., Song, Z., 2025a. Mixed vehicle platoon forming: A multi-agent reinforcement learning approach. *IEEE Internet of Things Journal* 12, 16886–16898.
- Shi, Y., Dong, H., He, C.R., Chen, Y., Song, Z., 2025b. Mixed Vehicle Platoon Forming: A Multiagent Reinforcement Learning Approach. *IEEE Internet of Things Journal* 12, 16886–16898.
- Song, L., Li, S., Chen, G., Zhao, X., Lyu, N., Fan, W.D., 2026. Exploring mechanisms of integrating global perception prediction for connected vehicles with lane-specific reinforcement learning-based variable speed limits. *Expert Systems with Applications* 299, 129958.
- Sun, D., Xie, D.F., Zhao, X.M., Jia, B., 2024. Enhancing stability and flexibility of platoon control with an adaptive spacing policy. *Transportmetrica A Transport Science* .
- Treiber, M., Kesting, A., 2013. Traffic flow dynamics. *Traffic Flow Dynamics: Data, Models and Simulation*, Springer-Verlag Berlin Heidelberg , 983–1000.
- Van Arem, B., Van Driel, C.J., Visser, R., 2006. The impact of cooperative adaptive cruise control on traffic-flow characteristics. *IEEE Transactions on intelligent transportation systems* 7, 429–436.
- Wang, S., Yang, X., Chen, Z., Zhang, Y., 2025. Distributed nonlinear model predictive control of vehicular platoon orienting practical driving conditions. *IEEE Transactions on Transportation Electrification* 11, 2684–2695.

- Wijnbergen, P., Jeeninga, M., de Haan, R., Lefeber, E., 2025. Longitudinal and lateral control of vehicle platoons: A unifying framework to prevent corner cutting. *Automatica* 177, 112340.
- Wu, R., Jia, H., Huang, Q., Tian, J., Gao, H., Wang, G., 2024. Multi-lane unsignalized intersection cooperation strategy considering platoons formation in a mixed connected automated vehicles and connected human-driven vehicles environment. *IEEE Transactions on Intelligent Transportation Systems* 25, 1569–1585.
- Wu, Y., Zuo, Z., Wang, Y., Han, Q., Hu, C., 2023. Distributed data-driven model predictive control for heterogeneous vehicular platoon with uncertain dynamics. *IEEE Transactions on Vehicular Technology* 72, 9969–9983.
- Xie, D.F., Zhao, X.M., He, Z., 2019. Heterogeneous traffic mixing regular and connected vehicles: Modeling and stabilization. *IEEE Transactions on Intelligent Transportation Systems* 20, 2060–2071.
- Xin, Q., Fu, R., Ukkusuri, S.V., Yu, S., Jiang, R., 2021. Modeling and impact analysis of connected vehicle merging accounting for mainline random length tight-platoon. *Physica A: Statistical Mechanics and its Applications* 563, 125452.
- Yang, J., Chu, D., Yin, J., Pi, D., Wang, J., Lu, L., 2023. Distributed model predictive control for heterogeneous platoon with leading human-driven vehicle acceleration prediction. *IEEE Transactions on Intelligent Transportation Systems* , 1–16.
- Yang, Z., Zheng, Z., Kim, J., Rakha, H., 2025. Eco-cooperative adaptive cruise control for platoons in mixed traffic using single-agent and multi-agent reinforcement learning. *Transportation Research Part D: Transport and Environment* 142, 104658.
- Yao, Z., Gu, Q., Jiang, Y., Ran, B., 2022. Fundamental diagram and stability of mixed traffic flow considering platoon size and intensity of connected automated vehicles. *Physica A: Statistical Mechanics and its Applications* 604, 127857.
- Yao, Z., Ma, Y., Ren, T., Jiang, Y., 2024. Impact of the heterogeneity and platoon size of connected vehicles on the capacity of mixed traffic flow. *Applied Mathematical Modelling* 125, 367–389.
- Yao, Z., Wu, Y., Wang, Y., Zhao, B., Jiang, Y., 2023. Analysis of the impact of maximum platoon size of cavs on mixed traffic flow: An analytical and simulation method. *Transportation Research Part C: Emerging Technologies* 147, 103989.

- Zhang, M., Wang, C., Zhao, W., Liu, J., Zhang, Z., 2025. A multi-vehicle self-organized cooperative control strategy for platoon formation in connected environment. *IEEE Transactions on Intelligent Transportation Systems* 26, 4002–4018.
- Zhou, J., Zhu, F., 2021. Analytical analysis of the effect of maximum platoon size of connected and automated vehicles. *Transportation Research Part C: Emerging Technologies* 122, 102882.
- Zhou, L., Ruan, T., Ma, K., Dong, C., Wang, H., 2021. Impact of cav platoon management on traffic flow considering degradation of control mode. *Physica A: Statistical Mechanics and its Applications* 581, 126193.
- Zong, F., Yue, S., Zeng, M., He, Z., Ngoduy, D., 2025. Platoon or individual: An adaptive car-following control of connected and automated vehicles. *Chaos, Solitons and Fractals* 191, 115850.

Appendix A. Definitions and Mappings of Time Headway Patterns

Applying the CAV demotion criterion and CHV demotion criterion established, these 32 car-following types can be consolidated into 6 distinct time headway patterns. In a car-following pair written as $X - Y$ (or $\Delta T_{X,Y}$), X denotes the leading vehicle and Y denotes the following vehicle; the time headway pattern is primarily determined by the following vehicle's capability and whether cooperation is available.

Appendix A.1. Fundamental autonomous headway (ΔT_A)

Applicable when the following vehicle is an AV or demoted CAV. The corresponding time headway is expressed as:

$$\begin{aligned}
 \Delta T_A &= \Delta T_{AV,CAV} = \Delta T_{HV,CAV} = \Delta T_{CAV,AV} \\
 &= \Delta T_{CAV_{pla},AV_{pla}} = \Delta T_{CHV,AV} = \Delta T_{CHV_{pla},AV_{pla}} \\
 &= \Delta T_{AV,AV} = \Delta T_{AV_{pla},AV_{pla}} = \Delta T_{HV,AV} \\
 &= \Delta T_{HV_{pla},AV_{pla}}
 \end{aligned} \tag{A.1}$$

Appendix A.2. Manual driving headway (ΔT_H)

Corresponds to cases where the following vehicle is an HV or demoted CHV. The corresponding time headway is expressed as:

$$\begin{aligned}
 \Delta T_H &= \Delta T_{AV,CHV} = \Delta T_{HV,CHV} = \Delta T_{CAV,HV} \\
 &= \Delta T_{CAV_{pla},HV_{pla}} = \Delta T_{CHV,HV} = \Delta T_{CHV_{pla},HV_{pla}} \\
 &= \Delta T_{AV,HV} = \Delta T_{AV_{pla},HV_{pla}} = \Delta T_{HV,HV} \\
 &= \Delta T_{HV_{pla},HV_{pla}}
 \end{aligned} \tag{A.2}$$

Appendix A.3. Connected autonomous headway outside platoons (ΔT_{CA})

Applicable when CHV/CAV engage in cooperative control without platoon coordination. This pattern comprises the following configurations: CAV-CAV and CHV-CAV. The corresponding time headway is expressed as:

$$\Delta T_{CA} = \Delta T_{CAV,CAV} = \Delta T_{CHV,CAV} \tag{A.3}$$

Appendix A.4. Connected manual headway outside platoons (ΔT_{CH})

Applies to manual control scenarios involving CHVs/CAVs. This pattern comprises the configurations as CAV-CHV and CHV-CHV. The corresponding time headway is expressed as:

$$\Delta T_{CH} = \Delta T_{CAV,CHV} = \Delta T_{CHV,CHV} \quad (A.4)$$

Appendix A.5. Intra-platoon cooperative autonomous headway (ΔT_{PA})

Applies to autonomous cooperative driving scenarios within a platoon. This pattern comprises the following configurations: CAVpla-CAVpla, CHVpla-CAVpla, AVpla-CAVpla, and HVpla-CAVpla. The corresponding time headway is expressed as:

$$\begin{aligned} \Delta T_{PA} &= \Delta T_{CAV_{pla},CAV_{pla}} = \Delta T_{CHV_{pla},CAV_{pla}} \\ &= \Delta T_{AV_{pla},CAV_{pla}} = \Delta T_{HV_{pla},CAV_{pla}} \end{aligned} \quad (A.5)$$

Appendix A.6. Intra-platoon cooperative manual headway (ΔT_{PH})

Applies to manual following scenarios within a platoon. This pattern comprises the following configurations: CAVpla-CHVpla, CHVpla-CHVpla, AVpla-CHVpla, and HVpla-CHVpla. The corresponding time headway is expressed as:

$$\begin{aligned} \Delta T_{PH} &= \Delta T_{CAV_{pla},CHV_{pla}} = \Delta T_{CHV_{pla},CHV_{pla}} \\ &= \Delta T_{AV_{pla},CHV_{pla}} = \Delta T_{HV_{pla},CHV_{pla}} \end{aligned} \quad (A.6)$$

Appendix B. Car-Following Model Formulations and Parameters

To accurately characterize heterogeneous driving behaviors in mixed traffic flows, this study employs three car-following models for distinct vehicle types.

Appendix B.1. The intelligent driver model for human-driven vehicles

The intelligent driver model (IDM) (Treiber and Kesting, 2013), featuring minimal parameters with clear physical interpretations, is employed to simulate human-driven vehicle. It is formulated as:

$$\begin{aligned} \dot{v}_n(t) = & A \left(1 - \left(\frac{v_n(t)}{v_f} \right)^4 \right. \\ & \left. - \left(\frac{s_{\min} + T_{\text{des}} v_n(t) - \frac{v_n(t) \Delta v_n(t)}{2\sqrt{AB}}}{s_n(t)} \right)^2 \right) \end{aligned} \quad (B.1)$$

where A and B are the maximum acceleration and desired deceleration, respectively; $s_{\min}$ represents the jam gap, that is the minimum spacing between two adjacent vehicles; T_{des} denotes the safe time headway; and v_f is the free-flow velocity.

Appendix B.2. The cooperative adaptive cruise control model for CAVs

When CAVs operate within platoons, they move in cooperative adaptive cruise control (CACC) mode. This study employs the CACC model calibrated by PATH Laboratory through field tests (Milanés et al., 2013; Milanés and Shladover, 2014), mathematically expressed as follows:

$$v_n(t + \Delta t) = v_n(t) + k_p e_n(t) + k_d \dot{e}_n(t) \quad (\text{B.2})$$

$$e_n(t) = s_n(t) - T_{\text{des}} v_n(t) \quad (\text{B.3})$$

where Δt denotes length for each time step; k_p and k_d represent control gains; and $e_n(t)$ quantifies the deviation between actual and desired spacing for vehicle n at time t .

To ensure consistency with the general car-following model form in Equation (10), the CACC model is approximated through first-order Taylor expansion:

$$\dot{v}_n(t) = \frac{k_p(s_n(t) - T_{\text{des}} v_n(t)) + k_d \Delta v_n(t)}{k_d T_{\text{des}} + \Delta t} \quad (\text{B.4})$$

Appendix B.3. The adaptive cruise control model for AVs

For AVs and degraded CAVs without connection, this study employs the adaptive cruise control (ACC) model developed by PATH Laboratory (Milanés et al., 2013; Milanés and Shladover, 2014) to characterize their car-following behavior. The model is mathematical formulated as:

$$\dot{v}_n(t) = k_1(s_n(t) - T_{\text{des}} v_n(t)) + k_2 \Delta v_n(t) \quad (\text{B.5})$$

where k_1 denotes the spacing error control gain, and k_2 represents the velocity difference control gain.

Based on the three car-following models, specific behaviors are defined for each scenario in Appendix A: Car-following patterns (1)-(2) occur in non-communication environments where following behavior depends solely on the follower type: IDM for human-driven vehicles (HVs), ACC for AVs. Car-following patterns (3)-(6) feature V2V communication enabling vehicle state sharing: CACC for autonomous-capable followers, IDM for human-driven followers. Time headway parameters and corresponding model coefficients are listed in Table B.3 (Yao et al., 2023; Qin et al., 2023; Jiang et al., 2023; Yao et al., 2024).

Table B.3: Parameters of car-following models.

Parameter	Value	Unit
Free-flow velocity, v_f	33.3	m/s
Minimum spacing, $s_{\min}$	2.0	m
Maximum acceleration, A	1.0	m/s ²
Desired deceleration, B	2.0	m/s ²
Gain, k_p	0.45	s ⁻¹
Gain, k_d	0.25	-
Simulation time step, Δt	0.01	s
Spacing error control gain, k_1	0.23	s ⁻¹
Velocity difference control gain, k_2	0.07	s ⁻¹
ΔT_H	1.8	s
ΔT_{CH}	1.2	s
ΔT_{PH}	0.8	s
ΔT_{CA}	1.0	s
ΔT_A	1.4	s
ΔT_{PA}	0.6	s