Efficient Motion Control for Heterogeneous Autonomous Vehicle Platoon Using Multilayer Predictive Control Framework

Du, Guodong; Zou, Yuan; Zhang, Xudong; Fan, Jie; Sun, Wenjing; Li, Zirui

DOI
10.1109/JIOT.2024.3445460

Publication date
2024

Document Version
Final published version

Published in
IEEE Internet of Things Journal

Efficient Motion Control for Heterogeneous Autonomous Vehicle Platoon Using Multilayer Predictive Control Framework

Guodong Du, Yuan Zou, *Senior Member, IEEE*, Xudong Zhang, *Member, IEEE*, Jie Fan, Wenjing Sun, and Zirui Li

***Abstract*—Autonomous driving technology and platooning driving technology are important directions for the development of intelligent and connected vehicles. Aiming at the motion control problem of autonomous vehicle platoon, this article proposes a multilayer predictive control framework (MPCF) based on heuristic learning agent and improved distributed model. First, the leading autonomous vehicle and following heterogeneous vehicles are modeled, respectively, and the motion control problem of autonomous platoon is described. Then, the multilayer motion control framework is designed, which contains highly automated tracking control optimization for the leading vehicle (LV) and high-precision formation keeping optimization for the following vehicles (FVs). In the upper layer, the heuristic Dyna algorithm-based predictive control (HDY-PC) method is proposed to improve the path tracking performance of the LV. In the lower layer, the improved distributed model-based predictive control (IDM-PC) method is developed to guarantee the motion effectiveness and stability of the vehicle platoon. Besides, the multilayer control framework can handle various communication topologies and dynamic cut-in/cut-out maneuvers. The virtual environment simulation shows that the proposed motion control framework for heterogeneous autonomous vehicle platoon achieves better performance in path tracking and platoon keeping. The adaptability of the framework is also verified using another real-world scene.**

***Index Terms*—Autonomous connected vehicle platoon, heuristic reinforcement learning, improved distributed model, motion control, multilayer predictive control framework (MPCF).**

Manuscript received 19 June 2024; revised 25 July 2024; accepted 14 August 2024. Date of publication 19 August 2024; date of current version 20 November 2024. This work was supported in part by the National Key Research and Development Program of China under Grant 2021YFB2500900; in part by the National Natural Science Foundation of China under Grant 52272410; and in part by the High-End Foreign Experts Recruitment Program under Grant G2023178009L. *(Corresponding author: Yuan Zou.)*

Guodong Du is with the Institute of Dynamic System and Control, ETH Zurich, 8092 Zurich, Switzerland, and also with the National Engineering Laboratory for Electric Vehicles, School of Mechanical Engineering, and the Collaborative Innovation Center of Electric Vehicles in Beijing, Beijing Institute of Technology, Beijing 100081, China (e-mail: guodongdu_robbie@163.com).

Yuan Zou, Xudong Zhang, Jie Fan, and Wenjing Sun are with the National Engineering Laboratory for Electric Vehicles, School of Mechanical Engineering, and the Collaborative Innovation Center of Electric Vehicles in Beijing, Beijing Institute of Technology, Beijing 100081, China (e-mail: zouyuan@bit.edu.cn; xudong.zhang@bit.edu.cn; 384954890@qq.com; sunwenjing@163.com).

Zirui Li is with the Department of Transport and Planning, Delft University of Technology, Delft, 2628 CD, The Netherlands, and also with the School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China (e-mail: 3120195255@bit.edu.cn).

Digital Object Identifier 10.1109/JIOT.2024.3445460

I. Introduction

In THE context of intelligent traffic systems (ITSs), intelligent and connected vehicles (ICVs) and autonomous driving technologies are developing rapidly [1]. Among several important development directions, the platoon of connected vehicles has attracted considerable attention and become typical application in the field of automotive intelligence [2]. Compared to a single ICV, the ICVs platoon has wider application extensibility when performing tasks, benefitting from its larger capacity and higher energy efficiency [3]. To achieve the expected platooning driving effect, the appropriate control scheme is required to optimize the motion performance of both the leading vehicle (LV) and following vehicles (FVs). In basic platooning tasks, the FVs should follow the motion trajectory of the LV while maintaining a preset space between any two adjacent vehicles [4]. In particular, when the autonomy level of connected vehicle platoon is further improved, the platoon system will automatically generate the optimal motion sequence rather than relying on the speed curve reference input from the external environment. Undoubtedly, highly autonomous control framework is a core trend in the future development of vehicle platooning.

A. Literature Review

In early stages of the research on the platoon control, the proposition of advanced platoon control strategies within the multiagent consensus control framework was growing [5], [6]. Considering the range limitation of communication topologies, a distributed receding horizon control approach was designed to guarantee the stability of vehicle platoon [7]. However, these approaches adopted linear dynamics and controllers, which were primarily for the sake of theoretical comprehensiveness while neglecting considerations for inherent model nonlinearities and input constraints.

With the rise of artificial intelligence (AI) algorithms, machine learning has advantages in dealing with complex control models and multidimensional constraints [8], [24]. A deep reinforcement learning (DRL) method was proposed to deal with the uncertainties and control constraints in the platooning behavior, and the theoretical analysis were validated by simulation results [9]. In [10], a communication proximal policy optimization was proposed based on the multiagent reinforcement learning (RL) system to improve the control

efficiency of vehicles platoon. The results showed that the proposed RL method achieved better performance than the traditional platoon control strategy and some existing RL strategies. Moreover, the neural network (NN) was also applied to the cooperative tracking control for platoon of vehicles, and the relevant effectiveness was verified by simulation experiments [11]. However, machine learning models heavily rely on the quality and representativeness of training data, which may lead to lack of interpretability or unintended consequences in different application scenarios. Besides, limited by its solution characteristic of single step decision, the machine learning methods cannot deal with the multiagent control in connected vehicles platoon from global optimization perspective, which results in the suboptimal strategies problem.

In recent years, model predictive control (MPC) method and its variants are widely used in many control problems, which could deal with both constraints and nonlinearities explicitly [12]. Aiming at the robust platoon control, the framework based on tube MPC was proposed [13]. The simulations were provided to evaluate the effectiveness of this control framework. Wang et al. [14] designed a holistic robust motion controller based on MPC algorithm for the ICV platoon. It was validated that this controller realized better suppression on position error propagation than single-structure MPC. In view of the feature of vehicle platoon containing multiple vehicles, several scholars introduced distributed MPC (DMPC) schemes, where each vehicle will be assigned to a locally optimized control problem. Zheng et al. [15] presented a DMPC-based algorithm for heterogeneous vehicle platoon. In the construction of this control system, the actual platoon control problem was simplified, where only the longitudinal motion control was considered, while the lateral control was ignored. The effectiveness of this method was preliminarily verified by using a simple artificially created velocity trajectory in the longitudinal dimension of the LV. Considering the intervehicular spacing constraints, a DMPC-based control scheme was proposed for longitudinal motion control of heterogeneous vehicle platoon [16]. The comparison simulation demonstrated the feasibility of this scheme in several scenarios. Wang et al. [17] proposed a DMPC algorithm for connected vehicle platoon considering predecessor-leader following (PLF) topology, and the consistency of platoon spacing and speed was proved by experiments. Considering more disturbance factors and optimization objectives, other DMPC methods for vehicle platoon control have also been proposed in recent years [18], [19], [20]. With the application of vehicle-to-vehicle (V2V) communication, variable communication topologies are required for the platoon control problem. Different from the above studies which only consider the specific communication topologies, Hu et al. [21] designed a DMPC control scheme for platooning, where several typical topologies (predecessor following (PF), bidirectional, PF leader, two predecessors following) were used.

It is worth noting that the above researches have simplified the motion control problem of vehicle platoon. First, the design of the platoon controller was only based on longitudinal motion, ignoring the lateral motion of different FVs in the platoon. In the actual application scene for vehicle platoon, the driving route is not an ideal straight path, and it is unrealistic to assume that the lateral motion is always consistent and steady [22]. Therefore, designing a motion control method for vehicle platoon to realize longitudinal control and lateral control coordination is meaningful. Besides, in the above researches, the velocity trajectory of the LV that needs to be followed is either a simple artificially created curve or a smoothly changing velocity curve. In practical applications, the velocity trajectory of the LV is often time-varying and fluctuating, which puts forward higher requirements for the designed motion control method to ensure the long-term stability of the entire vehicle platoon.

In [23], a hybrid automaton architecture based on MPC method was constructed for motion planning of connected vehicle platoon, which involved the longitudinal motion and lateral motion. Nonetheless, the scenario in the above paper belongs to a typical three-lane overtaking problem, which is not suitable for platooning control in nonspecific scenarios. Feng et al. [25] proposed a DMPC-based method for vehicle platoon considering longitudinal and lateral coupling. In this hierarchical structure, the lateral motion control was represented by the linear parametric-varying lane-keeping model, and both road radius and the specified look-ahead distance were assumed to be known. The effectiveness of this method was validated by the joint simulation. Different from the lateral control problem of lane keeping, the lateral motion control of platoon needs to be finely modeled in more generalized longitudinal and lateral collaborative application scenarios, such as nonspecific scenarios. Therefore, for a wider range of application requirement, the vehicle platoon control in nonspecific scenarios considering both longitudinal and lateral motions is valuable and necessary. Meanwhile, dynamic cut-in/cut-out maneuvers are another major challenge for vehicle platoon. All of vehicles in the platoon are required to maintain consistency in the lateral direction and a preset spacing in the longitudinal direction after cut-in/cut-out maneuvers.

Furthermore, the development of autonomous driving technology determines the future trend of highly autonomous platoon driving. In [26], the platoon of autonomous vehicles operating was considered for the application in the urban road network, and the traffic congestion phenomena was significantly reduced by the proposed method. The ideal autonomous vehicle platoon is expected to automatically generate the motion states curves instead of receiving the curves reference generated artificially. In this situation, a higher autonomy requirement for the LV control in the platoon is put forward, which needs to drive autonomously based on the external environment and obtained navigation path information. In the current research phase, highly automated motion control optimization for single autonomous vehicle has been carried out [27], [28], [29].

In recent years, the motion control based on MPC has become a popular solution in this field. The MPC method iteratively solves motion control sequences for single autonomous vehicle by using optimization algorithms in the future prediction horizon [30]. An improved kinematic MPC method was proposed for the high-speed motion control of single autonomous vehicle [31]. The results demonstrated

that the developed controller tracked the reference path well. Bao et al. [32] proposed a safety-guarantee MPC method for autonomous vehicle, which dealt with uncertain environments and uncertain models. The results showed that the controlled autonomous vehicle realized collision avoidance for different obstacles, and the proposed method achieved higher computational efficiency.

Furthermore, machine learning algorithms are constantly applied to the highly automated motion control optimization of single autonomous vehicles. Shan et al. [33] proposed a RL-based autonomous motion control strategy, which involved driving smoothness and tracking accuracy. The results proved that this RL-based control was better than traditional MPC-based control and LQR-based control. Another adversarial RL framework was designed for the motion control in the autonomous driving [34]. The effectiveness of this framework was validated by simulation tests. Besides, other variants of RL methods have also been applied to the autonomous motion control research, such as partial RL method [35] and safe RL method [36]. Actually, the RL architecture is suitable for solving dynamic motion control problem of single autonomous vehicle with its powerful self-learning ability. The general RL algorithms belong to Markov decision process (MDP), and generate single-step control action according to the current state. However, the actual global optimal solution of a control sequence is often derived in a long solution time domain. The single-step decision of the existing RL method does not consider the state of the next several steps, and the generated control strategy is only optimal for the current state, which may miss the actual optimal control strategy of the current state from the perspective of global optimization. Therefore, the set of one-step solution of RL method easily falls into the problem of suboptimal. The effects of existing motion control methods for single autonomous vehicle still have the potential of improvement, including tracking accuracy, driving rapidity, and stability.

Facing with the above problem and referring to the advantage of MPC method, the predictive time domain concept of MPC can be introduced into the traditional RL algorithm, and the control optimization based on the future prediction horizon can effectively improve the performance of the global control policy. Therefore, incorporating a forward multistep predictive decision method with RL structure to replace the single-step decision of traditional RL is meaningful, which has the potential to achieve better motion control effect of autonomous vehicle.

Furthermore, the transition from highly automated motion control of a single autonomous vehicle to highly automated motion control of an autonomous platoon is challenging, which involves higher dimensions of collaborative constraints. Especially in nonspecific scenarios, maintaining the safety and stability of the whole vehicle platoon while performing the autonomous driving task without driver input is of great significance.

B. Motivation and Contributions

Inspired by the aforementioned challenges in current motion control problem of autonomous vehicle platoon, the motivation of this article is designing an efficient and high-autonomy control scheme for heterogeneous autonomous vehicle platoon in nonspecific scenarios, which adapts to various communication topologies and dynamic cut-in/cut-out maneuvers.

The innovation points of this research can be summarized as follows: 1) the multilayer predictive control framework (MPCF) is proposed for the motion control of heterogeneous autonomous vehicle platoon, which contains highly automated tracking control optimization for the LV and high-precision formation keeping optimization for the FVs; 2) for the highly automated tracking control optimization of the LV, the heuristic Dyna (HDY) algorithm-based predictive control (HDY-PC) method is proposed to achieve better optimization efficiency and tracking performance; and 3) for the high-precision platoon keeping control optimization of the FVs, the improved distributed model-based predictive control (IDM-PC) method is designed to realize efficient longitudinal and lateral collaborative control.

The contribution details for each innovation point are shown below.

1) The multilayer framework ensures the stability of the longitudinal and lateral movement of the vehicle platoon through the differentiated control logic of the leading and FVs. The two logical layers are coupled and realize real-time transmission of dynamic information based on intelligent connected system. Through the cooperative operation of the two layers, the motion control task of the heterogeneous autonomous vehicle platoon is completed. It can also be applied for different typical communication topologies and dynamic cut-in/cut-out maneuvers in nonspecific scenarios.
2) In the implementation of HDY-PC method, the RL algorithm called HDY is developed for the training of initial motion control agent, which achieves higher training efficiency and optimization effects. The predictive control solution logic is designed, and the prediction horizon is constructed on the trained initial agent to realize the multiple steps decision instead of single step decision, which realizes better tracking performance.
3) In the implementation of IDM-PC method, the generalized longitudinal control and lateral control coordination is used to construct the improved distributed model, which adapts to collaborative application in nonspecific scenarios. The local optimal control solution for each FV is realized by DMPC method, and the update mechanism of the communication topology is also constructed to cope with the cut-in/cut-out maneuvers.

C. Organization of Paper

The organizational structure of this article: in Section II, the motion control problem of autonomous vehicle platoon is described, and the leading autonomous vehicle, following heterogeneous vehicles and communication topology are modeled, respectively. Then the multilayer motion control framework based on heuristic learning agent and improved distributed model for heterogeneous autonomous vehicle platoon is proposed in Section III. The virtual environment simulation results and analysis are provided in Section IV, and

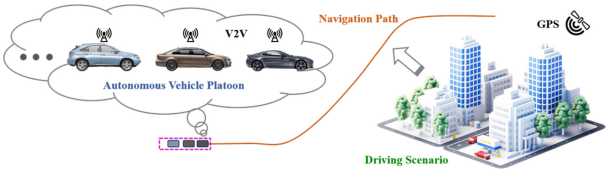


Fig. 1. Visualization of motion control problem for autonomous vehicle platoon.

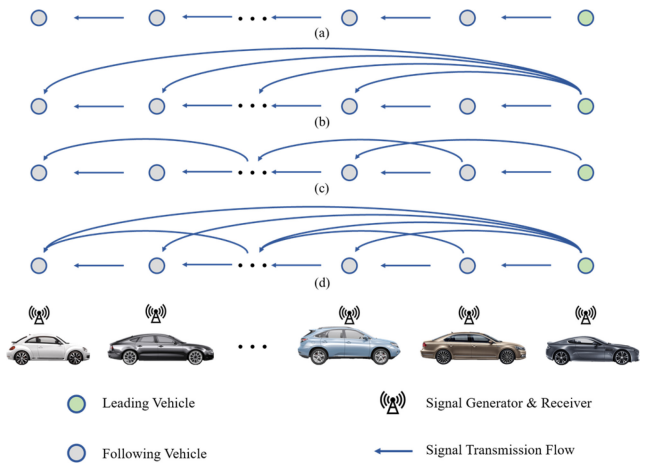


Fig. 2. Variable communication topologies for heterogeneous autonomous vehicle platoon: (a) PF; (b) PLF; (c) two predecessors following (TPF); and (d) two PLF (TPLF).

the adaptability of the proposed framework is also discussed. Finally, Section V concludes this article.

II. Modeling of Motion Control for Heterogeneous Autonomous Vehicle Platoon

A. Description of Problem

The motion control problem for autonomous vehicle platoon is visualized in Fig. 1. Relying on the global positioning system (GPS) and mature navigation software application, the global navigation path can be directly obtained when the starting and target positions in the driving scenario are determined. As shown in the plum red rectangle, the vehicle platoon needs to complete autonomous path tracking task according to the navigation path and external environment information. During the entire task execution process, the tracking accuracy, driving rapidity and smoothness of autonomous platoon should be improved. Meanwhile, the platooning safety and stability should also be guaranteed, which requires vehicles in the platoon to maintain consistency in the lateral direction and a preset spacing in the longitudinal direction.

In this research, variable communication topologies are considered for the heterogeneous platoon, as shown in Fig. 2. These four distinct communication topologies have been widely used in platooning applications [39], [40], [41], [42]. The autonomous platoon contains a LV and several FVs which are indexed from 1 to N. Each vehicle is equipped with communication equipment, and only the vehicles communicating directly with the LV can get the desired set point. The control objective of the LV is to track the navigation path with high autonomy, while the control objective of the FVs is to keep the platoon formation with high precision. The control logic from different parts together generates the complete autonomous platoon driving.

B. Nonlinear Modeling of the Leading Vehicle

In this research, both longitudinal and lateral dynamics of vehicles are considered. In order to ensure the actual execution of the control, nonlinear dynamic modeling for the LV is carried out, including rolling resistance, aerodynamic drag, inertial lag, and drivetrain efficiency. The nonlinear model of LV can be expressed by

$$\begin{cases} xp_0(t+1) = xp_0(t) + v_0(t)\cos\varphi_0(t)\Delta t \\ yp_0(t+1) = yp_0(t) + v_0(t)\sin\varphi_0(t)\Delta t \\ v_0(t+1) = v_0(t) + \frac{1}{m_{\text{leading}}}\left(\frac{\eta_{\text{dri},0}}{R_0}T_0(t) - F_{\text{leading}}(v_0(t))\right)\Delta t \\ T_0(t+1) = T_0(t) - \frac{1}{\tau_0}T_0(t)\Delta t + \frac{1}{\tau_0}a_0(t)\Delta t \\ \varphi_0(t+1) = \varphi_0(t) + \omega_0(t)\Delta t \\ F_{\text{leading}}(v_0(t)) = C_{A,0}{v_0}^2(t) + m_{\text{leading}}gf_{\text{leading}} \end{cases} \tag{1}$$

where $xp_0(t)$ and $yp_0(t)$ represent the position information of LV at x and y coordinates, respectively. Δt is the sampling time. $v_0(t)$ denotes the velocity of the LV, which is decided by vehicle mass m_{leading}, the drivetrain efficiency $\eta_{dri,0}$, the tire radius R_0, the external resistance $F_{\text{leading}}(v_0(t))$ and the integrated torque $T_0(t)$. Considering the inertial lag τ_0, $T_0(t)$ can be obtained by the desired torque command $a_0(t)$. $\varphi_0(t)$ and $\omega_0(t)$ stand for the yaw angle and steering angular speed, respectively. $C_{A,0}$ is the aerodynamic drag coefficient, g is the gravity acceleration, and f_{leading} is the rolling resistance coefficient. For the motion control of LV, the state is represented by $x_0(\text{t}) = [xp_0(\text{t}), yp_0(\text{t}), v_0(t), T_0(t), \varphi_0(t)]^T \in \mathbb{R}^{5\times 1}$, the output is represented by $y_0(t) = [xp_0(t), yp_0(t), v_0(t), \varphi_0(t)]^T \in \mathbb{R}^{4\times 1}$, and the control input can be represented by $u_0(t) = [a_0(t), \omega_0(t)]^T \in \mathbb{R}^{2\times 1}$. Then, the nonlinear model can be rewritten as the following equations:

$$\begin{cases} x_0(t+1) = \phi_0(x_0(t)) + \psi_0 \cdot u_0(t) \\ y_0(t) = \gamma_0 \cdot x_0(t) \end{cases} \tag{2}$$

where the relevant matrices are shown as follows:

$$\phi_0 = \begin{bmatrix} xp_0(t) + v_0(t)\cos\varphi_0(t)\Delta t \\ yp_0(t) + v_0(t)\sin\varphi_0(t)\Delta t \\ v_0(t) + \frac{1}{m_{\text{leading}}}\left(\frac{\eta_{\text{dri},0}}{R_0}T_0(t) - F_{\text{leading}}(v_0(t))\right)\Delta t \\ T_0(t) - \frac{1}{\tau_0}T_0(t)\Delta t \\ \varphi_0(t) \end{bmatrix} \in \mathbb{R}^{5\times 1}$$

$$\psi_0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \frac{\Delta t}{\tau_0} & 0 \\ 0 & \Delta t \end{bmatrix} \in \mathbb{R}^{5\times 2} \quad \gamma_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{4\times 5}. \tag{3}$$

The objective of LV is to track the navigation path with high autonomy while improving the tracking accuracy, smoothness

and rapidity. The navigation path P_{ref} contains a dense sequence of point positions

$$P_{\text{ref}} = \begin{pmatrix} (xp_{\text{ref}}, yp_{\text{ref}})_1 \\ (xp_{\text{ref}}, yp_{\text{ref}})_2 \\ \vdots \\ (xp_{\text{ref}}, yp_{\text{ref}})_{m-1} \\ (xp_{\text{ref}}, yp_{\text{ref}})_m \end{pmatrix} \in \mathbb{R}^{m\times 2}. \tag{4}$$

As the disturbance of autonomous driving control, obstacles in external environment are made up of a large number of pixels. The set of obstacles can be described as follows:

$$\begin{cases} \text{Obs} = \text{obs}_1 \cup \text{obs}_2 \cup \cdots \cup \text{obs}_l \\ \text{obs}_j = \left\{ \left(xp_{\text{obs}}{}^k, yp_{\text{obs}}{}^k \right) | k = 1, 2, \ldots, n; \right. \\ \qquad \left. \left(xp_{\text{obs}}{}^k, yp_{\text{obs}}{}^k \right) \in \mathbb{R}^2 \right\}_j, j = 1, 2, \ldots, l. \end{cases} \tag{5}$$

The updating of tracking control process is realized by the discrete dynamic equation

$$\begin{aligned} x_0(t+1) = & f_{\Delta t}(x_0(t), u_0(t), P_{\text{ref}}, \text{Obs}) \\ & \text{s.t.} \quad A_{\text{det}}(x_0(t+1)) \subseteq A_{\text{pass}}(x_0(t+1), \text{Obs}) \\ & A_{\text{det}}(x_0(t+1)) \cap A_{\text{impass}}(x_0(t+1), \text{Obs}) = \varnothing. \end{aligned} \tag{6}$$

where $f_{\Delta t}$ represents the execution function parameterized by Δt. $A_{\text{det}}(x_0(\text{t+1}))$ is the collision detection region at state $x_0(\text{t+1})$. A_{pass} stands for the passable region, and A_{impass} stands for the impassable region. Then, the optimal motion control strategies of LV u_0^* can be derived by minimizing the cost function $J(x_0(\text{t}), u_0(\text{t}), P_{\text{ref}}, Obs)$

$$\begin{aligned} u^* = & \underset{x_0(i), u_0(j)}{\text{argmin}} \quad J(x_0(i), u_0(j)); \ i = [0:N], j = [0:N-1] \\ & \text{s.t.} \quad \forall k \in \{0, \ldots, N-1\}: \\ & x_0(k+1) = f_{\Delta t}(x_0(k), u_0(k), P_{\text{ref}}, \text{Obs}). \end{aligned} \tag{7}$$

C. Nonlinear Modeling of the Following Vehicles

Similar to the nonlinear dynamics modeling of the LV, the model of the ith FV can be expressed as follows:

$$\begin{cases} x_i(t+1) = \phi_i(x_i(t)) + \psi_i \cdot u_i(t) \\ y_i(t) = \gamma_i \cdot x_i(t) \end{cases}$$

$$\phi_i = \begin{bmatrix} xp_i(t) + v_i(t)\cos\varphi_i(t)\Delta t \\ yp_in + v_i(t)\sin\varphi_i(t)\Delta t \\ v_i(t) + \frac{1}{m_{\text{following},i}}\left(\frac{\eta_{\text{dri},i}}{R_i} T_i(t) - F_{\text{following},i}(v_i(t)) \right) \Delta t \\ T_i(t) - \frac{1}{\tau_i} T_i(t)\Delta t \\ \varphi_i(t) \end{bmatrix} \in \mathbb{R}^{5\times 1}$$

$$\psi_i = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \frac{\Delta t}{\tau_i} & 0 \\ 0 & \Delta t \end{bmatrix} \in \mathbb{R}^{5\times 2} \ \gamma_i = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{4\times 5}. \tag{8}$$

Considering states, outputs and control inputs of all FVs, the corresponding vectors are constructed as follows:

$$\begin{cases} X(t) = \left[x_1{}^T(t), x_2{}^T(t), \ldots, x_N{}^T(t) \right]^T \in \mathbb{R}^{5N\times 1} \\ Y(t) = \left[y_1{}^T(t), y_2{}^T(t), \ldots, y_N{}^T(t) \right]^T \in \mathbb{R}^{4N\times 1} \\ U(t) = \left[u_1{}^T(t), u_2{}^T(t), \ldots, u_N{}^T(t) \right]^T \in \mathbb{R}^{2N\times 1}. \end{cases} \tag{9}$$

Furthermore, the complete dynamics model of all FVs ($1 \sim N$) in the autonomous vehicle platoon can be expressed by:

$$\begin{cases} X(t+1) = \boldsymbol{\Phi}(X(t)) + \boldsymbol{\Psi} \cdot U(t) \\ Y(t+1) = \boldsymbol{\Gamma} \cdot X(t+1) \end{cases}$$

$$\begin{cases} \boldsymbol{\Phi} = \left[\phi_1{}^T(x_1), \phi_2{}^T(x_2), \ldots, \phi_N{}^T(x_N) \right]^T \in \mathbb{R}^{5N\times 1} \\ \boldsymbol{\Psi} = \text{diag}\{\psi_1, \psi_2, \ldots, \psi_N\} \in \mathbb{R}^{5N\times 2N} \\ \boldsymbol{\Gamma} = \text{diag}\{\gamma_1, \gamma_2, \ldots, \gamma_N\} \in \mathbb{R}^{4N\times 5N}. \end{cases} \tag{10}$$

The objective of FVs is to keep the platoon formation with high precision while maintaining a preset spacing in the longitudinal direction and consistency in the lateral direction. The platoon maintenance can be formulated as

$$\begin{cases} \lim_{t\to\infty} \| v_i(t) - v_0(t) \| = 0 \\ \lim_{t\to\infty} \left\| s_i(t) - d_{i,i-1} \right\| = 0 \\ \lim_{t\to\infty} \| \text{lat}_i(t) \| = 0 \\ s_i(t) = \sum_{q=\text{idx}_i}^{\text{idx}_{i-1}-1} \sqrt{ \left(xp_{\text{ref},q+1} - xp_{\text{ref},q} \right)^2 + \left(yp_{\text{ref},q+1} - yp_{\text{ref},q} \right)^2 } \\ \text{lat}_i(t) = \sqrt{ \left(xp_i(t) - xp_{\text{ref},i_{\text{label}}} \right)^2 + \left(yp_i(t) - yp_{\text{ref},i_{\text{label}}} \right)^2 } \\ \text{idx}_i(t) = f_{\text{nearest}}(xp_i(t), yp_i(t), P_{\text{ref}}) \\ \text{idx}_{i-1}(t) = f_{\text{nearest}}(xp_{i-1}(t), yp_{i-1}(t), P_{\text{ref}}) \ \ i \in [1, N] \end{cases} \tag{11}$$

where $d_{i,i-1}$ denotes the preset spacing between the ith FV and the $(i-1)$th FV. $\text{s}_i(\text{t})$ is the distance between projection points of two vehicles on the navigation path. $lat_i(\text{t})$ is the distance from the actual position of the ith FV to its projected point on the navigation path. $f_{\text{nearest}}()$ stands for the calculation function of the projection point index.

D. Modeling of the Communication Topology

The communication topology of the autonomous platoon is modeled using the directed graph $\mathbb{G} = \{\mathbb{V}, \mathbb{E}\}$, where $\mathbb{V} = \{0, 1, 2, \ldots, N\}$ means the set of vehicle nodes and $\mathbb{E} \subseteq \mathbb{V} \times \mathbb{V}$ means the set of connection edges [39]. Then, the directed graph can be disassembled as adjacency matrix A, Laplacian matrix L and pinning matrix P, which are as follows:

$$\begin{aligned} & A = \left[a_{ij} \right] = \begin{cases} a_{ij} = 1, & \{i, j\} \in \mathbb{E} \\ a_{ij} = 0, & \{i, j\} \notin \mathbb{E} \end{cases} \\ & L = D - A = \text{diag}\left\{ \sum_{j=1}^{N} a_{1j}, \sum_{j=1}^{N} a_{2j}, \ldots, \sum_{j=1}^{N} a_{Nj} \right\} - \left[a_{ij} \right] \\ & P = \text{diag}\{p_1, p_2, \ldots, p_N\}, \quad p_i = \begin{cases} 1, & \{i, 0\} \in \mathbb{E} \\ 0, & \{i, 0\} \notin \mathbb{E} \end{cases} \\ & A \in \mathbb{R}^{N\times N}, L \in \mathbb{R}^{N\times N}, P \in \mathbb{R}^{N\times N} \end{aligned} \tag{12}$$

where $\{i, j\} \in \mathbb{E}$ means the ith FV can receive the information from the jth FV or LV (if j =0). Furthermore, we define three sets to describe the communication situations between the ith FV and the other vehicles in the autonomous platoon

$$\begin{aligned} \mathbb{P}_i &= \begin{cases} \{0\}, & p_i = 1 \\ \varnothing, & p_i = 0 \end{cases} \\ \mathbb{N}_i &= \{j | a_{ij} = 1, \ j \in [1, N]\} \\ \mathbb{O}_i &= \{k | a_{ki} = 1, \ k \in [1, N]\} \end{aligned} \tag{13}$$

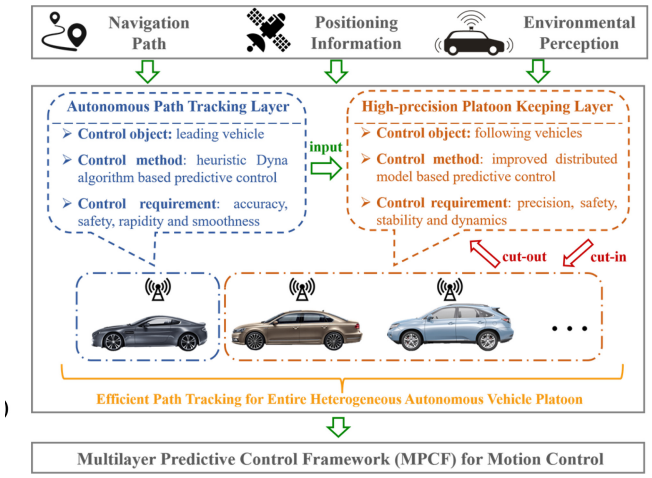


Fig. 3. Implementation structure of the multilayer framework for autonomous vehicle platoon.

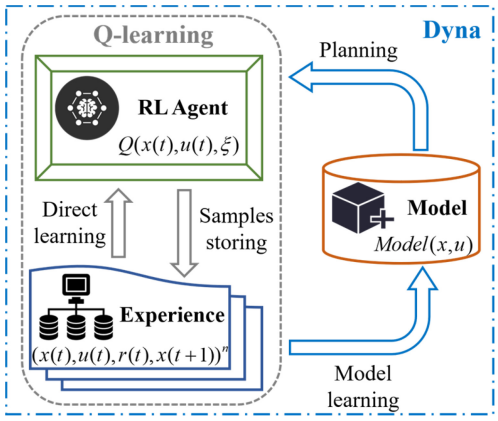


Fig. 4. Training mechanism of the Dyna architecture and the QL architecture.

where $\mathbb{P}_i$ represents the set of information permission from LV. $\mathbb{N}_i$ and $\mathbb{O}_i$ represent the sets of FVs which the ith FV receives information from and sends information to, respectively. Therefore, the set $\mathbb{I}_i = \mathbb{P}_i \cup \mathbb{N}_i$ represents all available information for the local optimal control construction of the ith FV.

III. Multilayer Motion Control Framework for Heterogeneous Autonomous Vehicle Platoon

In this research, the control framework of heterogeneous vehicle platoon has the characteristics of high autonomy, which receives only positioning and perception information, but not reference inputs of control. Therefore, the control framework contains two important logical layers, where the first layer realizes the autonomous path tracking control of the platoon, and the other layer realizes the autonomous formation keeping control of the platoon considering the dynamic cut-in/cut-out maneuvers. In the first layer, the task of efficient path tracking is assigned to the LV. In the second layer, the task of high-precision platoon keeping is assigned to the FVs. On the basis that LV can track the navigation path efficiently, the stable platoon keeping control will ensure that the entire autonomous platoon can achieve the goal of efficient path tracking. Through the cooperative operation of the two logical layers, the motion control task of the heterogeneous autonomous vehicle platoon is completed. The implementation structure of the MPCF is shown in Fig. 3.

A. Heuristic Dyna Algorithm-Based Predictive Control for Leading Vehicle

In the autonomous path tracking layer for LV, the HDY-PC method is proposed to improve the performance in accuracy, safety, rapidity and smoothness. The realization of HDY-PC method contains two stages. The first stage is the initial motion control agent design based on HDY algorithm, and the second stage is the final tracking controller design based on the combination of the initial agent and predictive control.

In the first stage, a RL algorithm with high-training efficiency is developed, named HDY. Different from Q-learning (QL) algorithm and its variants, Dyna algorithm adds steps of model learning and planning in addition to the training mechanism of direct learning [37]. Through the integration of incremental learning and online planning, Dyna architecture tends to perform better than QL architecture. The training mechanism of the Dyna architecture and the QL architecture is shown in Fig. 4.

The tracking control of LV is classified as a sequential decision-making process, and the value of implementing the particular control in the particular state can be assessed. If a control action u_0(t) is performed in state x_0(t) and optimal control strategies are followed thereafter, this control can be evaluated in terms of the expectation of future cumulative rewards. Its optimal value function is expressed by

$$V^*(x_0(t), u_0(t)) = \max_{\pi^*} E\left[\sum_{k=t}^{k=t_{\text{final}}} \gamma^k r_0(x_0(k), u_0(k))|\pi^*\right] \quad (14)$$

where π^* is the subsequent optimal control strategies and $E()$ stands for the expectation function. t_{final} and γ represent the final timestamp and the weight balancing factor for different timestamp rewards, respectively. $r_0(x_0(\text{t}), u_0(\text{t}))$ denotes the reward function in Dyna architecture, which expression is as follows:

$$\begin{aligned} r_0(x_0(t), u_0(t)) = {} & k_{\text{acu}} \cdot f_{\text{tra}}(x_0(t), u_0(t), P_{\text{ref}}) \\ & + k_{\text{saf}} \cdot f_{\text{obs}}(x_0(t), u_0(t), \text{Obs}) \\ & + k_{\text{rap}} \cdot f_{\text{vel}}(x_0(t), u_0(t), v_{\max}) + k_{\text{smo}} \cdot f_{\text{dri}}(u_0(t)) \end{aligned} \quad (15)$$

where k_{acu}, k_{saf}, k_{rap} and k_{com} represent the accuracy weight, safety weight, rapidity weight, and smoothness weight, respectively. Accordingly, f_{tra}, f_{obs}, f_{vel} and f_{dri} stand for tracking error function, obstacles avoidance function, velocity evaluation function and driving smoothness function, respectively.

In the optimization process of Dyna architecture, the Q matrix with strong approximation ability is trained continuously to approximate the optimal value function, which are formulated by

$$\begin{cases} \boldsymbol{Q}(x_0(t), u_0(t)) \\ = r_0(x_0(t), u_0(t)) + \gamma \max\limits_{u_0(t+1)} \boldsymbol{Q}(x_0(t+1), u_0(t+1); \xi) \\ \boldsymbol{Q}^*(x_0(t), u_0(t)) \\ = \max\limits_{u_0(t)} \left(r_0(x_0(t), u_0(t)) + \gamma \max\limits_{u_0(t+1)} \boldsymbol{Q}(x_0(t+1), u_0(t+1); \xi) \right) \end{cases} \quad (16)$$

where ξ_t means the weights set of Q matrix. For the direct learning step, the Q matrix is updated using the experience samples stored

$$\begin{aligned}&\boldsymbol{Q}\big(x_0(t),u_0(t);\xi'\big)\leftarrow\boldsymbol{Q}(x_0(t),u_0(t);\xi)\\&+\alpha\Big[r_0(x_0(t),u_0(t))+\gamma\max_{u_0(t+1)}\boldsymbol{Q}(x_0(t+1),u_0(t+1);\xi)\\&\qquad-\boldsymbol{Q}(x_0(t),u_0(t);\xi)\Big]\end{aligned}\tag{17}$$

where γ and α denote the discount factor and learning rate, respectively. The training of Q matrix is realized by updating the weights set ξ_t.

On the basis of direct learning step, Dyna architecture introduces model learning step and planning step to enhance the effectiveness of RL optimization. The implementation of model learning is as follows:

$$\begin{aligned}&\varepsilon-\text{greedy}(x_0(t),\boldsymbol{Q})\rightarrow u_0(t)\\&\qquad\downarrow\\&f_{\Delta t}(x_0(t),u_0(t),P_{\text{ref}},\text{Obs})\rightarrow x_0(t+1),r_0(t)\\&\qquad\downarrow\\&\textbf{Model}(x_0(t),u_0(t))\leftarrow x_0(t+1),r_0(t).\end{aligned}\tag{18}$$

Then, the model is used for the planning step. In the planning step of traditional Dyna algorithm, the samples information for Q matrix optimization are selected randomly from the model. Nevertheless, different samples from the model have different effects on Q matrix training, and the completely random sampling will restrict the efficiency of planning. Therefore, the heuristic planning concept is developed to improve the efficiency. Through the heuristic evaluation of the training value of different samples, more valuable samples to be used in the planning step are determined, and then the perfection of Q matrix is accelerated. The heuristic evaluation function H is established as follows:

$$\begin{cases}H(x_0(t),u_0(t))=\left|\boldsymbol{Q}^{\text{Target}}(x_0(t),u_0(t))-\boldsymbol{Q}^{\text{Actual}}(x_0(t),u_0(t))\right|\\\boldsymbol{Q}^{\text{Target}}(x_0(t),u_0(t))=r_0(x_0(t),u_0(t))+\gamma\max\limits_{u_0(t+1)}\boldsymbol{Q}(x_0(t+1),u_0(t+1);\xi)\\\boldsymbol{Q}^{\text{Actual}}(x_0(t),u_0(t))=\boldsymbol{Q}(x_0(t),u_0(t);\xi)\end{cases}\tag{19}$$

where $\boldsymbol{Q}^{\text{Target}}$ and $\boldsymbol{Q}^{\text{Actual}}$ represent the target value and actual value of Q matrix at specific state–action pair, respectively. When the heuristic evaluation value of the sample is large, it means that a significant error between the estimation value of this sample by Q matrix and the actual value of this sample exists. Therefore, this sample is more valuable for optimizing the Q matrix, which can accelerate the update of the Q matrix as shown in (17). The implementation of heuristic planning step in HDY can be described by

for $i=1$ **to** k **do**
  get heuristic action $u_0(i)\leftarrow h_a(x_0(i),H)$
  obtain information $x_0(i+1),r_0(i)\leftarrow$ ***Model***
  $\boldsymbol{Q}\big(x_0(i),u_0(i);\xi'\big)\leftarrow\boldsymbol{Q}(x_0(i),u_0(i);\xi)+\alpha[r_0(i)+\gamma\max\limits_{u_0(i+1)}\boldsymbol{Q}(x_0(i+1),u_0(i+1);\xi)-\boldsymbol{Q}(x_0(i),u_0(i);\xi)]$
end for. (20)

Furthermore, the generalized correlation coefficient (GCC) is applied to judge the perfection of Q matrix, which calculates the similarity of Q matrices at two adjacent timestamps. When the value of GCC approaches to 1 infinitely, the training of Q matrix tends to be perfect. Its expression is given by

$$\begin{cases}\text{GCC}(\boldsymbol{Q}(t),\boldsymbol{Q}(t-1))=\frac{\text{tr}[\text{cov}(\boldsymbol{Q}(t),\boldsymbol{Q}(t-1))]}{\{\text{tr}[\text{cov}(\boldsymbol{Q}(t))]\times\text{tr}[\text{cov}(\boldsymbol{Q}(t-1))]\}^{1/2}}\\\text{cov}(\boldsymbol{Q}(t))=\frac{[\boldsymbol{Q}(t)-E(\boldsymbol{Q}(t))]^T\times[\boldsymbol{Q}(t)-E(\boldsymbol{Q}(t))]}{n_{\text{timestamp}}-1}\\\text{cov}(\boldsymbol{Q}(t-1))=\frac{[\boldsymbol{Q}(t-1)-E(\boldsymbol{Q}(t-1))]^T\times[\boldsymbol{Q}(t-1)-E(\boldsymbol{Q}(t-1))]}{n_{\text{timestamp}}-1}\\\text{cov}(\boldsymbol{Q}(t),\boldsymbol{Q}(t-1))=\frac{E\Big\{[\boldsymbol{Q}(t)-E(\boldsymbol{Q}(t))]^T\times[\boldsymbol{Q}(t-1)-E(\boldsymbol{Q}(t-1))]\Big\}}{1}\end{cases}\tag{21}$$

where $tr()$ denotes the trace function and $n_{\text{timestamp}}$ is the number of timestamps. Once the training of optimal Q matrix is completed, it can be used to generate the initial motion control agent with single-step decision-making capability.

In the second stage, the predictive control method based on the trained HDY agent is designed to realize the multiple-step decision instead of single-step decision, named HDY-PC. The prediction horizon is constructed on the HDY decision system. All possible control actions are traversed for the selection of the first control in the prediction horizon, and a series of corresponding next states are generated. Then, the HDY agent is used to iteratively generate subsequent multistep control sequences in the whole forward prediction horizon starting from the next states. After that, the cumulative reward values of different control sequences generated in the prediction horizon are evaluated, and the first control action in the control sequence with the highest cumulative reward is chosen as the optimal tracking control action of current state.

For the solution logic of HDY-PC, the cumulative reward function of prediction horizon can be expressed by

$$\begin{cases}J(x_0(k),u_0(k))_{h_p}=\boldsymbol{Q}(x_0(k),u_0(k))+\sum_{j=k+1}^{k+h_p}\lambda_j\boldsymbol{Q}^*(x_0(j))\\\qquad=\boldsymbol{Q}(x_0(k),u_0(k))+\sum_{j=k+1}^{k+h_p}\lambda_j\max\limits_{u_0(j)}\boldsymbol{Q}(x_0(j),u_0(j))\\x_0(i+1)=f_{\Delta t}(x_0(i),u_0(i),P_{\text{ref}},\text{Obs}),\quad i\in\big[k,k+h_p-1\big]\end{cases}\tag{22}$$

where h_p represents the size of forward prediction horizon and λ_j means the prediction attenuation factor negatively related to the prediction progress. Based on the above cumulative reward function and multiple-step prediction logic, the optimal tracking control strategy at state $x_0(k)$ can be derived by the following equation:

$$\begin{aligned}u_0{}^*(x_0(k))&=\arg\max_{u_0(k)}J(x_0(k),u_0(k))_{h_p}\\&=\arg\max_{u_0(k)}\left[\boldsymbol{Q}(x_0(k),u_0(k))+\sum_{j=k+1}^{k+h_p}\lambda_j\max_{u_0(j)}\boldsymbol{Q}(x_0(j),u_0(j))\right].\end{aligned}\tag{23}$$

In the path tracking layer, the application of predictive control is combined with well-trained HDY control agent which can quickly generate single-step optimal control strategy, rather than with other optimization algorithms and complex dynamic control models in traditional MPC methods. Therefore, the proposed tracking control method has

TABLE I
KEY PARAMETERS OF HDY-PC METHOD

Names	Values
Sampling time (s) Δt	0.1
Learning rate α	0.935
Discount factor γ	0.26
Size of sampling batch n	30
Replay buffer capacity n_{buffer}	6000
Initial exploration factor ε	0.50
Exploration attenuation rate υ	0.98
Length of prediction horizon h_p	6
Initial prediction attenuation factor λ	0.4

fast enough solution speed and real-time application ability. Finally, the tracking controller based on HDY-PC is generated for the LV in autonomous platoon. Besides, the key parameters of HDY-PC are given in Table I.

B. Improved Distributed Model-Based Predictive Control for Following Vehicles

In the high-precision platoon keeping layer for FVs, the IDM-PC method is designed to improve the performance in precision, safety, stability and dynamics. The realization of IDM-PC method contains two stages. The first stage is the improved local open-loop control modeling based on the longitudinal control and lateral control coordination, and the second stage is the platoon keeping controller design based on the DMPC algorithm.

In the first stage, the longitudinal and lateral motion states of each FV are incorporated into the distributed control modeling. From the autonomous path tracking layer, the position and velocity information of the LV is obtained. According to (11), the desired set point of state and control input for the ith FV can be expressed by

$$\begin{cases} x_{\mathrm{des},i}(t) = \left[xp_{\mathrm{des},i}(t), yp_{\mathrm{des},i}(t), v_{\mathrm{des},i}(t), T_{\mathrm{des},i}(t), \varphi_{\mathrm{des},i}(t)\right]^T \\ u_{\mathrm{des},i}(t) = \left[a_i(t), \omega_{\mathrm{des},i}(t)\right]^T \\ \text{s.t.} \\ \left[xp_{\mathrm{des},i}(t), yp_{\mathrm{des},i}(t)\right] = \left[xp_{\mathrm{ref},\mathrm{idx}_{\mathrm{des},i}(t)}, yp_{\mathrm{ref},\mathrm{idx}_{\mathrm{des},i}(t)}\right] \\ \left[v_{\mathrm{des},i}(t), T_{\mathrm{des},i}(t)\right] = \left[v_0(t), a_i(t)\right] \\ \varphi_{\mathrm{des},i}(t) = \frac{yp_{\mathrm{ref},\mathrm{idx}_{\mathrm{des},i}(t)+1} - yp_{\mathrm{ref},\mathrm{idx}_{\mathrm{des},i}(t)-1}}{xp_{\mathrm{ref},\mathrm{idx}_{\mathrm{des},i}(t)+1} - xp_{\mathrm{ref},\mathrm{idx}_{\mathrm{des},i}(t)-1}} \\ \mathrm{idx}_{\mathrm{des},i}(t) = \arg\min_{k_{\mathrm{des}}} \left| \sum_{q=k_{\mathrm{des}}}^{\mathrm{idx}_0-1} \sqrt{\left(xp_{\mathrm{ref},q+1} - xp_{\mathrm{ref},q}\right)^2 + \left(yp_{\mathrm{ref},q+1} - yp_{\mathrm{ref},q}\right)^2} - i \times d_{\mathrm{spc}} \right| \\ \mathrm{idx}_0(t) = f_{\mathrm{nearest}}(xp_0(t), yp_0(t), P_{\mathrm{ref}}) \end{cases} \tag{24}$$

where d_{spc} is the preset spacing constant. Then, the desired output can be obtained by $y_{\mathrm{des},i}(\mathrm{t}) = \gamma_i x_{\mathrm{des},i}(\mathrm{t})$. Different from the constant speed assumption in previous literature, the LV in this research drives with variable speed, which puts forward higher requirements for the stability of platoon keeping. It should also be noted that the motion scale of each FV is 2-D, including both longitudinal and lateral dimensions, rather than the idealized concept of 1-D distance.

For the ith FV, the improved local control problem can be modeled using the information in set $\mathbb{I}_i = \mathbb{P}_i \cup \mathbb{N}_i$. Assuming that the ith FV can receive information from m FVs, the outputs and control inputs of the set $\mathbb{N}_i$ can be defined as the following vectors:

$$\begin{cases} y_{-i}(t) = \left[y_{i,1}^T(t), y_{i,2}^T(t), \ldots, y_{i,m}^T(t)\right]^T \\ u_{-i}(t) = \left[u_{i,1}^T(t), u_{i,2}^T(t), \ldots, u_{i,m}^T(t)\right]^T. \end{cases} \tag{25}$$

For the prediction horizon [t, t+N_p], three trajectories of outputs are introduced: 1) the predicted output trajectory $y_i^p(k|t)$; 2) the optimal output trajectory $y_i^*(k|t)$; and 3) the assumed output trajectory $y_i^a(k|t)$. Analogously, three trajectories of control inputs are introduced: 1) the predicted input trajectory $u_i^p(k|t)$; 2) the optimal input trajectory $u_i^*(k|t)$; and 3) the assumed input trajectory $u_i^a(k|t)$. Then, the local optimal control problem for each FV can be modeled as follows:

$$\begin{aligned} u_i{}^*(x_i(t)) &= \min_{u_i^p(0|t)\sim u_i^p(N_p-1|t)} J_i\left(y_i^p, u_i^p, y_i^a, y_{-i}^a\right) \\ &= \min_{u_i^p(0|t)\sim u_i^p(N_p-1|t)} \left[\sum_{k=0}^{N_p-1} L_i\left(y_i^p(k|t), u_i^p(k|t), y_i^a(k|t), y_{-i}^a(k|t)\right)\right] \\ &\quad \text{s.t.} \quad \forall i \in \{1,2,\ldots,N\}, k = \left[0,1,\ldots,N_p-1\right]: \\ &\qquad x_i^p(k+1|t) = \phi_i\left(x_i^p(k|t)\right) + \psi_i \cdot u_i^p(k|t) \\ &\qquad y_i^p(k|t) = \gamma_i \cdot x_i^p(k|t) \\ &\qquad x_i^p(0|t) = x_i(t) \\ &\qquad u_i^p(k|t) \in \left[\left(a_{i,\min}, a_{i,\max}\right), \left(\omega_{i,\min}, \omega_{i,\max}\right)\right]^T \\ &\qquad y_i^p\left(N_p|t\right) = \frac{1}{|\mathbb{I}_i|} \sum_{j\in\mathbb{I}_i} \left[f_{\mathrm{des}}\left(y_j^a(N_p|t), d_{i,j}\right)\right] \end{aligned} \tag{26}$$

where $u_i^p(0|t) \sim u_i^p(N_p-1|t)$ is the sequence of control inputs which needs to be optimized. $L_i()$ represents the Lyapunov function that characterizes the local cost. $[(a_{i,\min}, a_{i,\max}), (\omega_{i,\min}, \omega_{i,\max})]^T$ means the constraint of the control inputs. $f_{\mathrm{des}}()$ stands for the desired output function of the ith FV at end of prediction horizon, which is calculated based on the terminal assumed output of FVs in set $\mathbb{I}_i$. The calculation logic of $f_{\mathrm{des}}()$ is similar to that in (24). Furthermore, the Lyapunov function $L_i()$ can be expressed by the following equation:

$$\begin{aligned} &L_i\left(y_i^p(k|t), u_i^p(k|t), y_i^a(k|t), y_{-i}^a(k|t)\right) \\ &= \left\|y_i^p(k|t) - y_{\mathrm{des},i}(k|t)\right\|_{\boldsymbol{Q_i}} + \left\|u_i^p(k|t) - u_{\mathrm{des},i}(k|t)\right\|_{\boldsymbol{R_i}} \\ &\quad + \left\|y_i^p(k|t) - y_i^a(k|t)\right\|_{\boldsymbol{F_i}} + \sum_{j\in\mathbb{N}_i} \left\|y_i^p(k|t) - f_{\mathrm{des}}\left(y_j^a(k|t), d_{i,j}\right)\right\|_{\boldsymbol{G_i}} \\ &\quad \text{e.g.} \quad \left\|y_i^p(k|t) - y_{\mathrm{des},i}(k|t)\right\|_{\boldsymbol{Q_i}} \\ &= \left[y_i^p(k|t) - y_{\mathrm{des},i}(k|t)\right]^T \boldsymbol{Q_i} \left[y_i^p(k|t) - y_{\mathrm{des},i}(k|t)\right] \\ &\quad \text{where} \quad \boldsymbol{Q_i} \in \mathbb{R}^{4\times4}, \quad \boldsymbol{R_i} \in \mathbb{R}^{2\times2}, \quad \boldsymbol{F_i} \in \mathbb{R}^{4\times4}, \quad \boldsymbol{G_i} \in \mathbb{R}^{4\times4} \end{aligned} \tag{27}$$

where $\boldsymbol{Q_i}$, $\boldsymbol{R_i}$, $\boldsymbol{F_i}$, and $\boldsymbol{G_i}$ represent the weight matrices which realize regularization. These four weight matrices regularize the penalty of output deviation from desired equilibrium, control input deviation from desired equilibrium, output deviation from assumed value, and output deviation from the assumed

value calculated by neighbor FVs, respectively. If $\mathbb{P}_i$ is empty, the ith FV cannot get information permission from LV, and $\boldsymbol{Q}_i$ is set to 0.

It should be remarked that the above control solution formula of the ith FV only refers to the information of neighbor FVs in set $\mathbb{N}_i$. The desired set point is not needed for the FV which cannot pin to the LV. Therefore, the constructed local optimal control model can handle variable communication topologies shown in Fig. 2.

In the second stage, the DMPC algorithm is developed for the 2-D platoon keeping control. The initialization formulation for the ith FV

$$\begin{aligned}
&\begin{cases} y_i^a(k|0) = y_i^p(k|0) \\ u_i^a(k|0) = \left[\frac{R_i}{\eta_{dri,i}}\left(C_{A,i}{v_i}^2(0) + m_{\text{following},i} g f_{\text{following},i}\right), 0\right]^T \end{cases} \\
&\text{s.t.} \quad k = 0, 1, \ldots, N_p - 1 \\
&\qquad x_i^p(k+1|0) = \phi_i\left(x_i^p(k|0)\right) + \psi_i \cdot u_i^a(k|0) \\
&\qquad y_i^p(k|0) = \gamma_i \cdot x_i^p(k|0) \\
&\qquad x_i^p(0|0) = x_i. \qquad (28)
\end{aligned}$$

Then, the iterative solution process is shown below:

First, deriving the optimal control input sequence $u_i^*(k|t)$ based on current state $x_i(t)$, as shown in (26).

Second, obtaining the optimal state in the prediction horizon executing $u_i^*(k|t)$ as follows:

$$\begin{cases} x_i^*(k+1|t) = \phi_i\left(x_i^*(k|t)\right) + \psi_i \cdot u_i^*(k|t) \\ x_i^*(0|t) = x_i(t). \end{cases} \qquad (29)$$

Third, computing next assumed control input $u_i^a(k|t+1)$ as follows:

$$u_i^a(k|t+1) = u_i^*(k+1|t). \qquad (30)$$

Meanwhile, computing the assumed output as follows:

$$\begin{cases} x_i^a(k+1|t+1) = \phi_i\left(x_i^a(k|t+1)\right) + \psi_i \cdot u_i^a(k|t+1) \\ x_i^a(0|t+1) = x_i^*(1|t) \\ y_i^a(k|t+1) = \gamma_i x_i^a(k|t+1). \end{cases} \qquad (31)$$

Then, sending $y_i^a(k|t+1)$ to FVs in set $\mathbb{O}_i$, while receiving $y_{-i}^a(k|t+1)$ from FVs in set $\mathbb{N}_i$. If $\mathbb{P}_i$ is not empty, computing $y_{\text{des},i}(k|t+1)$ in parallel based on the information of LV.

Finally, implementing the optimal control input from the optimal control sequence: $u_i(x_i(t)) = u_i^*(0|t)$.

By iteratively repeating the above solution process for each FV, the complete platoon keeping strategies can be gradually generated.

Besides, the platoon keeping control based on IDM-PC is required to handle the dynamic cut-in/cut-out maneuvers. Assuming that cut-in/cut-out maneuvers sequentially happen in the autonomous driving process of vehicle platoon, the time stamps of the cut-in/cut-out can be marked by t_{ci} and t_{co}, respectively. Meanwhile, the cut-in maneuver happens before the i_{ci}th FV in autonomous platoon, and the cut-out maneuver of the i_{co}th FV in autonomous platoon happens. Then, the index of FVs will be updated as follows:

$$\forall i(0) \in \{1, 2, \ldots, N\}$$

$$\begin{cases} i(t) = i(0) + 1, & t_{ci} \le t \le t_{co} \ \&\ i(0) \ge i_{ci} \\ i(t) = i(0) - 1, & t > t_{co} \ \&\ i_{co} < i(0) < i_{ci} \\ i(t) = i(0) + 1, & t > t_{co} \ \&\ i_{ci} \le i(0) < i_{co} \\ i(t) = i(0), & \text{else} \end{cases} \qquad (32)$$

where the index of each FV is time varying according to the dynamic maneuvers. The above solution of IDM-PC operates based on the real-time index of FV. Besides, the change in the vehicle platoon caused by the cut-in/cut-out maneuvers will result in the corresponding update of communication topology [38], i.e., $A \in \mathbb{R}^{N\times N} \rightarrow A \in \mathbb{R}^{(N+1)\times(N+1)} \rightarrow A \in \mathbb{R}^{(N+1-1)\times(N+1-1)}$. The relevant information flow in communication topology is also adjusted according to the real-time index of FV.

C. Theoretical Analysis of Improved Distributed Model-Based Predictive Control

According to the modeling section of communication topology, the graph $\mathbb{G}$ belongs to unidirectional topology which includes a spanning tree with the LV as its root. The stability analysis of the proposed IDM-PC method can be implemented by constructing a positive definite Lyapunov function for this motion control problem then proving its monotonically decreasing property.

Based on the local optimal control formulation shown in (26), the local optimal cost function of the ith FV at time t can be represented by

$$\begin{aligned}
J_i^*(t) &= J_i^*\left(y_i^*, u_i^*, y_i^a, y_{-i}^a\right) \\
&= \sum_{k=0}^{N_p-1} L_i\left(y_i^*(k|t), u_i^*(k|t), y_i^a(k|t), y_{-i}^a(k|t)\right). \qquad (33)
\end{aligned}$$

In the operation of proposed IDM-PC method, the following equations are satisfied:

$$\begin{cases} y_i^p\left(N_p|t\right) - y_{\text{des},i}\left(N_p|t\right) = 0 \\ u_i^p(:|t+1) - u_i^a(:|t+1) = 0. \end{cases} \qquad (34)$$

Then, the above optimal cost function of the ith FV satisfies

$$\begin{aligned}
&J_i^*(t+1) - J_i^*(t) \\
&\le -L_i\left(y_i^*(0|t), u_i^*(0|t), y_i^a(0|t), y_{-i}^a(0|t)\right) \\
&\quad + \sum_{k=1}^{N_p-1} \begin{bmatrix} \sum_{j\in\mathbb{O}_i} \|{y_i}^*(k|t) - {y_i}^a(k|t)\|_{\boldsymbol{G}_j} \\ -\|{y_i}^*(k|t) - {y_i}^a(k|t)\|_{\boldsymbol{F}_i} \end{bmatrix}. \qquad (35)
\end{aligned}$$

The positive definite Lyapunov function is constructed by using the sum of all local optimal cost functions

$$\begin{aligned}
J_{1:N}^*(t) &= \sum_{i=1}^{N} J_i^*\left(y_i^*, u_i^*, y_i^a, y_{-i}^a\right) \\
&= \sum_{i=1}^{N} \left[\sum_{k=0}^{N_p-1} L_i\left(y_i^*(k|t), u_i^*(k|t), y_i^a(k|t), y_{-i}^a(k|t)\right) \right]. \qquad (36)
\end{aligned}$$

The variation trend of this Lyapunov function can be judged by the following relation:

$$J_{1:N}^*(t+1) - J_{1:N}^*(t)$$

$$\leq -\sum_{i=1}^{N}\left[L_i\left(y_i^*(0|t), u_i^*(0|t), y_i^a(0|t), y_{-i}^a(0|t)\right)\right] + \sum_{i=1}^{N}\left\{\sum_{k=1}^{N_p-1}\left[\begin{array}{r}\sum_{j\in\mathbb{O}_i}\|y_i^*(k|t)-y_i^a(k|t)\|_{\boldsymbol{G_j}} \\ -\|y_i^*(k|t)-y_i^a(k|t)\|_{\boldsymbol{F_i}}\end{array}\right]\right\}. \quad (37)$$

Considering (37), the weight matrices F_i and G_j are set to satisfy the following constraints:

$$\boldsymbol{F_i} \geq \sum_{j\in\mathbb{O}_i}\boldsymbol{G_j}, \qquad i \in \{1, 2, \ldots, N\}. \quad (38)$$

Subsequently, the following situation holds:

$$X^T\left(\sum_{j\in\mathbb{O}_i}\boldsymbol{G_j}-\boldsymbol{F_i}\right)X \leq 0 \quad \forall X \in \mathbb{R}^2. \quad (39)$$

Then, plug the difference between the optimal and assumed outputs into (39)

$$\begin{cases} X = y_i^*(k|t) - y_i^a(k|t) \\ X^T\left(\sum_{j\in\mathbb{O}_i}\boldsymbol{G_j}-\boldsymbol{F_i}\right)X \leq 0 \end{cases} \downarrow \sum_{j\in\mathbb{O}_i}\left\|y_i^*(k|t)-y_i^a(k|t)\right\|_{\boldsymbol{G_j}} - \left\|y_i^*(k|t)-y_i^a(k|t)\right\|_{\boldsymbol{F_i}} \leq 0. \quad (40)$$

Combining (37) and (40), the variation trend of the Lyapunov function can be further derived as

$$\begin{aligned} & J_{1:N}^*(t+1) - J_{1:N}^*(t) \\ & \leq -\sum_{i=1}^{N}\left[L_i\left(y_i^*(0|t), u_i^*(0|t), y_i^a(0|t), y_{-i}^a(0|t)\right)\right] \\ & \quad + \sum_{i=1}^{N}\left\{\sum_{k=1}^{N_p-1}\left[\begin{array}{r}\sum_{j\in\mathbb{O}_i}\|y_i^*(k|t)-y_i^a(k|t)\|_{\boldsymbol{G_j}} \\ -\|y_i^*(k|t)-y_i^a(k|t)\|_{\boldsymbol{F_i}}\end{array}\right]\right\} \\ & \leq -\sum_{i=1}^{N}\left[L_i\left(y_i^*(0|t), u_i^*(0|t), y_i^a(0|t), y_{-i}^a(0|t)\right)\right] \\ & \quad < 0. \end{aligned} \quad (41)$$

Therefore, the constructed positive definite Lyapunov function is monotonically decreasing, then the control stability of the proposed IDM-PC method has been proved.

D. Implementation of Multilayer Predictive Control Framework for Autonomous Vehicle Platoon

Combing the autonomous path tracking layer and high-precision platoon keeping layer, the multilayer control framework for autonomous vehicle platoon is formed. The operation flowchart for the MPCF for heterogeneous autonomous vehicle platoon is visualized in Fig. 5. In the operation of HDY-PC-based layer, the motion control strategy of the LV is generated based on the external environment and navigation path. At the same time, in the operation of IDM-PC-based layer, the motion control strategies of all the FVs are generated based on the dynamic motion information of the LV and navigation path. The information transfer of the entire framework is updated in real time using the intelligent connected system, and the real-time motion information transmitted from the LV to FVs contains the driving velocity and acceleration of the LV. Finally, the autonomous motion control for the complete heterogeneous vehicle platoon is realized, involving the high-accuracy path tracking and high-precision platoon keeping.

TABLE II
PARAMETERS INFORMATION FOR ALL HETEROGENEOUS VEHICLES

Vehicles	Mass (kg)	Tire radius (m)	Inertial lag (s)	Aerodynamic drag coefficient ($N\cdot s^2/m^2$)
LV	1354	0.32	0.70	1.06
FV 1	1129	0.30	0.67	1.03
FV 2	1378	0.33	0.71	1.08
FV 3	1290	0.32	0.70	1.06
FV 4	1576	0.36	0.75	1.12
FV 5	1401	0.34	0.72	1.09
FV 6	1193	0.30	0.68	1.04
FV 7	1362	0.32	0.70	1.06
FV cut-in	1261	0.31	0.69	1.05

IV. RESULTS AND ANALYSIS

To validate the performance of the MPCF for autonomous vehicle platoon, the scene extracted from the real world and virtual driving environment simulation are used. The heterogeneous autonomous platoon contains eight vehicles initially (one LV and seven FVs), and the cut-in of external vehicle will happen between the first and the second FVs, then the fourth FV in platoon will cut out. The different communication topologies during the motion control process of autonomous platoon have been shown in Fig. 2 before. The preset spacing constant between adjacent vehicles is 10 m, and each FV in the autonomous platoon is in the desired position initially. The parameters information for all heterogeneous vehicles are shown in Table II.

A. Evaluation and Validation for Multilayer Predictive Control Framework

The verification scene is generated based on the real world environment, and the navigation path is planned and collected by the real vehicle. For this scene, its satellite map based on GPS, its point cloud map based on simultaneous positioning and mapping (SLAM) and the planned navigation path are shown in Fig. 6. The autonomous platoon is required to efficiently track the navigation path from the start position to the temporary target position, then from the temporary target position to the final position. During the driving process, the platoon should ensure the stability and consistency of the formation.

It should be emphasized that in the initialization setup of this research, the LV is placed in the start position at the beginning of the task, and all FVs are placed behind the LV according to their order in the platoon and preset spacing distance. The vehicle platoon is arranged in the same direction as the

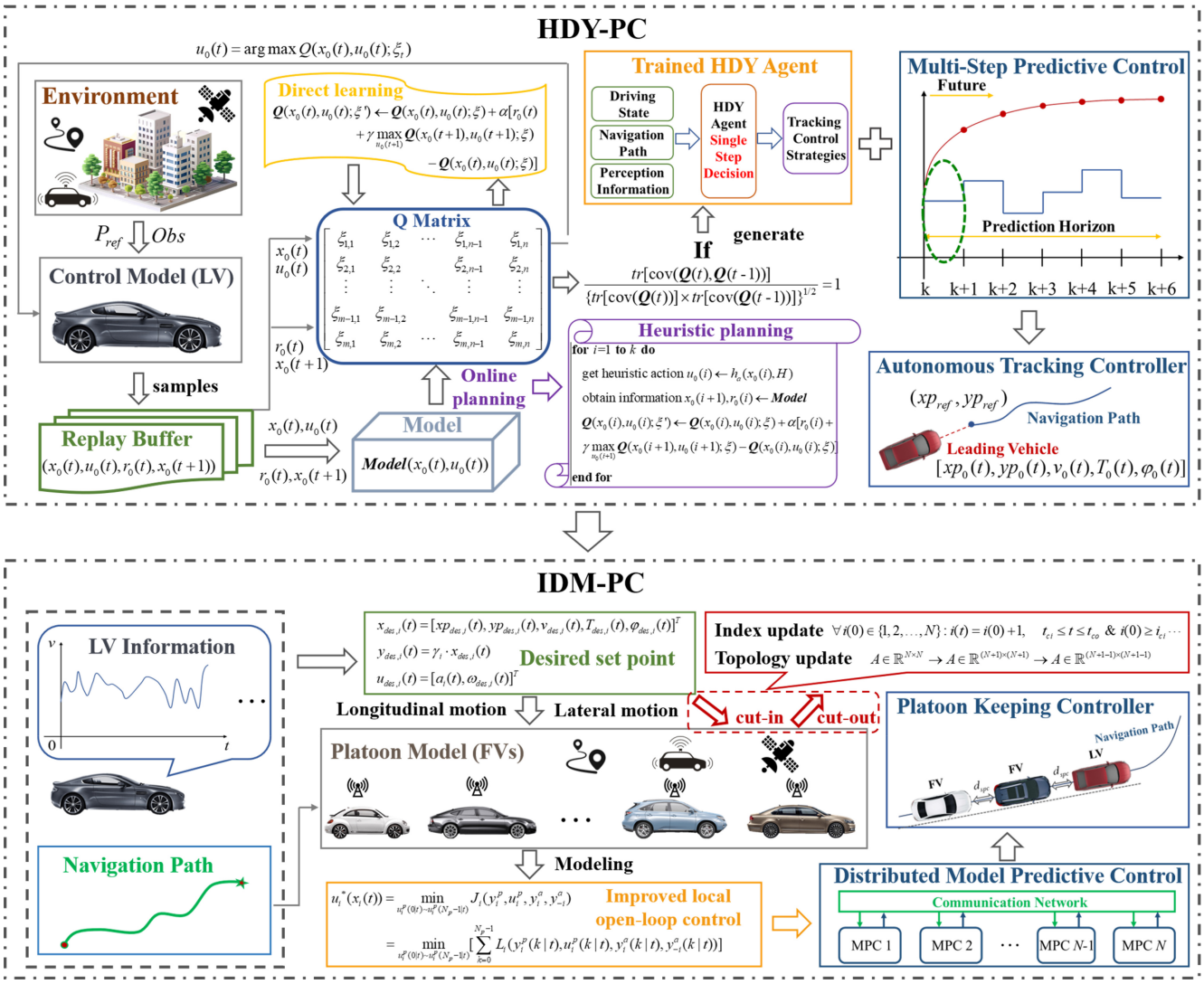


Fig. 5. Flowchart of the MPCF for heterogeneous autonomous vehicle platoon.

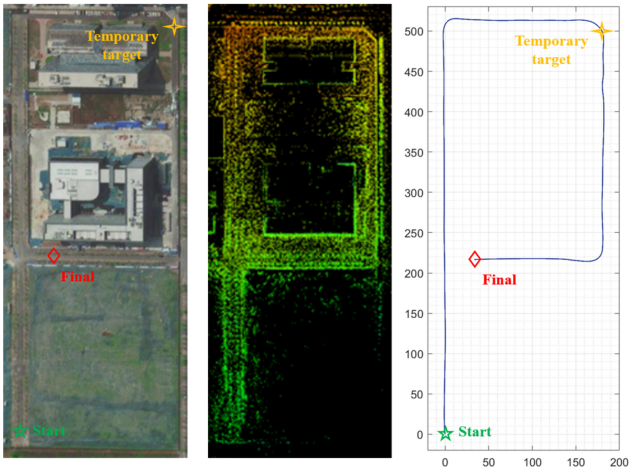


Fig. 6. Driving scene for verification of the MPCF.

initial segment of the navigation path. Before the specific FV passes the start position, its reference path is a straight backward extension of the initial segment of the navigation path. Therefore, the actual reference path of the autonomous platoon includes the navigation path in the global scenario and the virtual path during the initial driving transition of the FVs in the platoon.

To evaluate and analyze the performance of the MPCF, some existing methods are applied for the comparison. In the evaluation of path tracking layer, the proposed HDY-PC method is compared to QL method and Dyna method. In the evaluation of platoon keeping layer, the proposed IDM-PC method is evaluated from both longitudinal and lateral dimensions under various communication topologies. Meanwhile, the dynamic cut-in/cut-out maneuvers will be analyzed in detail.

B. Comparison Results of the Path Tracking Control Based on HDY-PC

The path tracking trajectories of the LV with different methods (HDY-PC, Dyna, QL) are shown in Fig. 7(a). It is easy to find that the LV using proposed HDY-PC method achieves much higher tracking accuracy than the other two

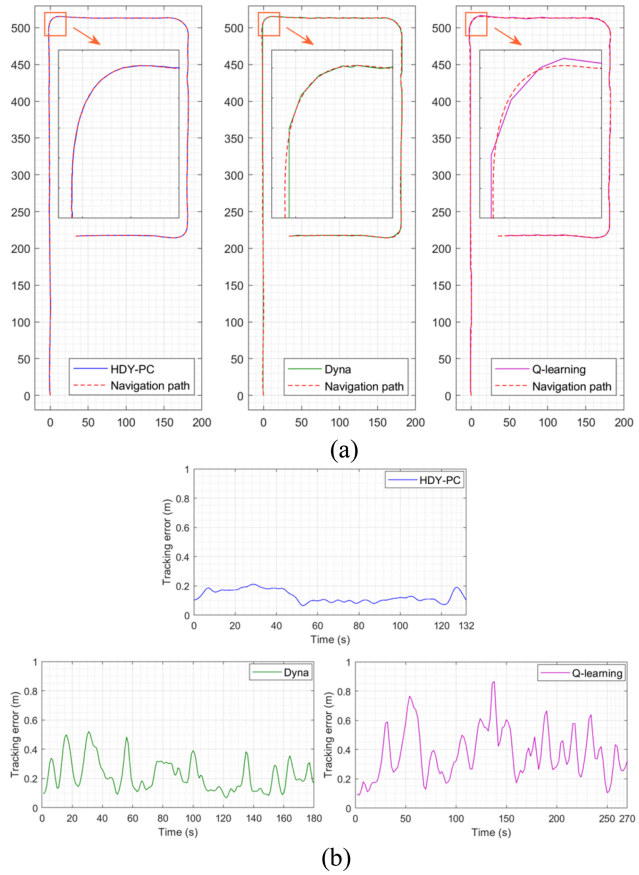


Fig. 7. Results about tracking accuracy of LV with three control methods. (a) Tracking trajectory of LV. (b) Tracking error of LV.

RL methods. It performs well in both straight and turning sections, especially in large angle turn as shown in the local magnification. The optimal performance of the proposed HDY-PC method indicates the effectiveness of heuristic planning and multistep predictive control integrated into the RL system. Besides, the tracking error curves of the LV with three methods are shown in Fig. 7(b). Obviously, the tracking error of the proposed method is less than 0.2 m in most time intervals, which is significantly better than those of other two RL methods. Therefore, the tracking accuracy of the HDY-PC method has been proved preliminarily.

Furthermore, the variations of motion states for the LV equipped with different control methods are illustrated in Fig. 8. Fig. 8(a) depicts the driving velocities of the LV corresponding to three methods. It can be seen that the LV controlled by the HDY-PC method maintains a higher driving velocity in most time intervals. Taking the local velocity curves in the dark blue rectangles as examples, they belong to relatively open road with a basically straight navigation path in scene shown in Fig. 6. Therefore, driving in these sections as fast as possible is reasonable and efficient. The LV controlled by HDY-PC method achieves this goal with the highest speed, and takes only 132 s to complete the entire driving task. Meanwhile, the driving times of LV controlled by Dyna and QL are 180 s and 270 s, respectively. The comparison results validate the effectiveness of establishing heuristic planning and predictive control logic on RL system.

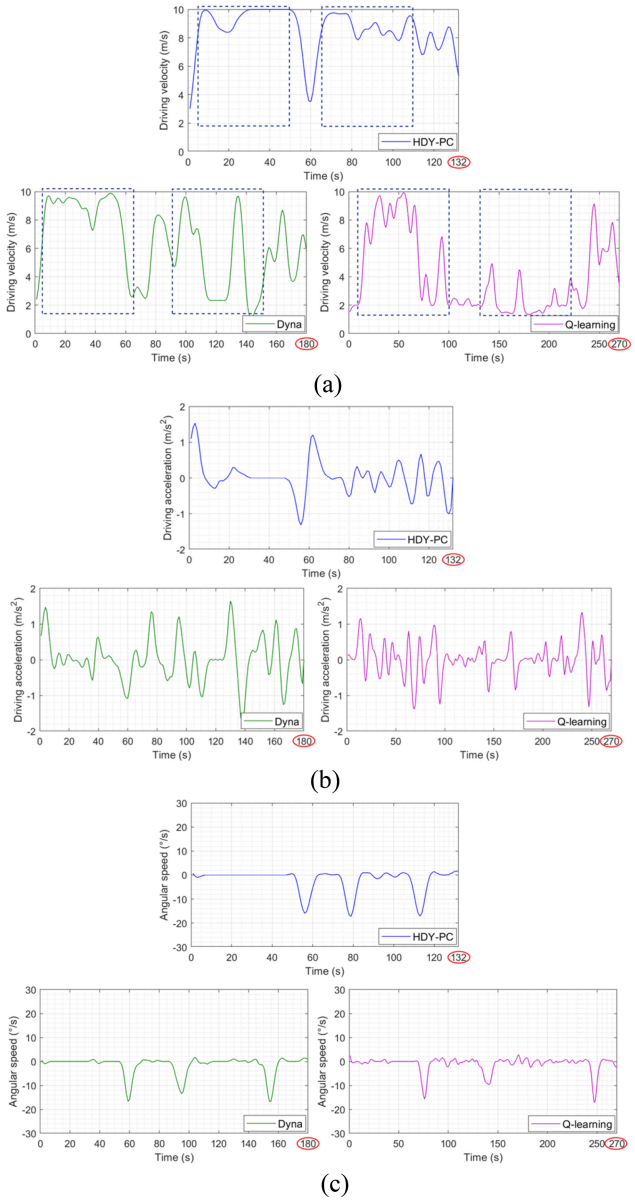


Fig. 8. Variations of motion states for LV equipped with different control methods. (a) Driving velocity. (b) Driving acceleration. (c) Steering angular speed.

Fig. 8(b) depicts the acceleration curves of the LV corresponding to three control methods. As can be seen, the LV controlled by HDY-PC method does not have violent acceleration and deceleration, which is consistent with the result of velocity curve shown in Fig. 8(a). In addition, the steering angular speed curves corresponding to three methods are shown in Fig. 8(c). The proposed method has excellent performance in controlling the steering angular speed of the vehicle, which is stable in $[-20°/s, 10°/s]$ and is basically 0 in some time intervals. Therefore, the driving direction of the LV will not shake violently. Based on the above analysis of motion states, the driving rapidity and smoothness for the HDY-PC method have been demonstrated.

Table III lists the numerical path tracking results of the LV controlled by three methods. Obviously, the proposed HDY-PC method has the smallest average tracking error and the fastest

TABLE III
NUMERICAL PATH TRACKING RESULTS FOR THREE CONTROL METHODS

Methods	Average tracking error (m)	Average driving velocity (m/s)	Computation time (s)
HDY-PC	0.13	8.54	0.02
Dyna	0.32	6.29	0.01
Q-learning	0.39	4.21	0.01

TABLE IV
COMPARISON RESULTS FOR DIFFERENT ALGORITHMS IN ABLATION EXPERIMENTS

Methods	Toal reward value (relative increase)	Training time (s) (relative increase)
QL	132.71 (-)	835.62 (-)
QL+ML	132.71 (+0.0%)	869.16 (+4.0%)
QL+ML+P	157.83 (+18.9%)	752.40 (-10.0%)
QL+ML+HP	166.15 (+25.2%)	605.37 (-27.6%)

average driving velocity, reflecting its advantages in tracking accuracy and driving rapidity. Besides, the computation time of the proposed method is slightly longer than those of other RL methods, which is due to the additional multistep predictive control calculation. 0.02 s is still possible to achieve the real-time control requirements of autonomous vehicle platoon.

Furthermore, to assess and analyze the training effectiveness of HDY algorithm, Dyna algorithm and QL algorithm, a set of ablation experiments was performed. The comparison results for different algorithms are shown in Table IV. QL algorithm is set as the baseline. QL with model learning step (QL+ML) is set as the first comparison object. QL+ML and planning steps (QL+ML+P) is set as the second comparison object. QL+ML and heuristic planning steps (QL+ML+HP) is set as the third comparison object. In these comparison objects, the second object is Dyna, and the third object is HDY.

Obviously, the first comparison object only adds model learning step on the basis of Q-learning structure, and this model is not used for planning RL agent. Therefore, the total reward value does not change compared to the baseline, and the additional model learning step leads to the increase of training time. Then, the second comparison object adds both model learning and planning steps on the basis of QL structure, which realizes integration of incremental learning and online planning for RL agent. As a result, the trained RL agent achieves higher total reward while also reducing training time. Furthermore, the third comparison object combines the heuristic evaluation with the second comparison object. The results show that the third object achieves the highest total reward value and the shortest training time. Finally, the assessments of online planning step and heuristic evaluation step have been completed sequentially.

C. Comparison Results of the Platoon Keeping Control Based on IDM-PC

In this section, for consistency of analysis, the IDM-PC based platoon keeping layer uses the LV controlled by HDY-PC as the motion reference. In particular, the driving velocity of the LV is dynamically time-varying rather than fixed, which actually poses a challenge for high-precision platoon keeping. The longitudinal spacing errors, lateral tracking errors and driving velocities of the autonomous vehicle platoon using different communication topologies are shown in Fig. 9.

It can be found that the autonomous platoon using IDM-PC method remains stable and safe throughout the driving task. In the longitudinal direction, the spacing error for each FV is always kept below 0.1 m without considering the dynamic cut-in/cut-out maneuvers, which indicates that all of FVs keep the preset spacing well during the autonomous platoon driving. In the lateral direction, the tracking error for each FV is always kept below 0.5 m and especially below 0.4 m in nonturning sections, which indicates that all of FVs track the navigation path well while maintaining the stable following of preceding vehicle. Therefore, the platooning safety and stability are guaranteed, where all of vehicles maintain consistency in the lateral direction and a preset spacing in the longitudinal direction. Besides, the driving velocity curves of all FVs are highly coincident with that of the LV, which ensures each FV can keep a precise distance from the preceding vehicle.

Benefitting from the solution logic of the proposed local open-loop control, the desired set point is not needed for the FV which cannot pin to the LV. Hence, under the use of different communication topologies, the IDM-PC based autonomous vehicle platoon can always achieve the expected control effect, as shown in Fig. 9.

For more detailed analysis, the results of the IDM-PC based autonomous vehicle platoon using TPLF topology is taken as an example, and its local magnification is illustrated in Fig. 10. Fig. 10(a) depicts the spacing error variations of different FVs during the cut-in/cut-out maneuvers. When cut-in maneuver happens between the first FV and the second FV, the cut-in vehicle is updated as a unit in the platoon. Meanwhile, the relative spacing error of the second FV suddenly changes, and the relative spacing error of the cut-in FV appears. The sum of these two spacing errors is numerically approximately equal to a preset spacing constant which is 10 m. Then, the relevant longitudinal spacing errors return to 0 within 5 s. When cut-out maneuver of the fourth FV happens, this FV is removed from the platoon. Meanwhile, the preceding vehicle of the fifth FV is updated to the third FV, and the relative spacing error of the fifth FV suddenly increases to 10 m. Then, the relevant longitudinal spacing error returns to 0 within 5 s. Therefore, the IDM-PC method can eliminate the spacing error quickly.

Fig. 10(b) depicts the tracking error variations of different FVs. As can be seen, the lateral error of cut-in vehicle appears when the cut-in maneuver happens, and the lateral errors of the FVs are slightly upward when FVs pass the turning section at higher velocities in order to maintain the preset spacing. Besides, as mentioned above, since the FVs initially track a

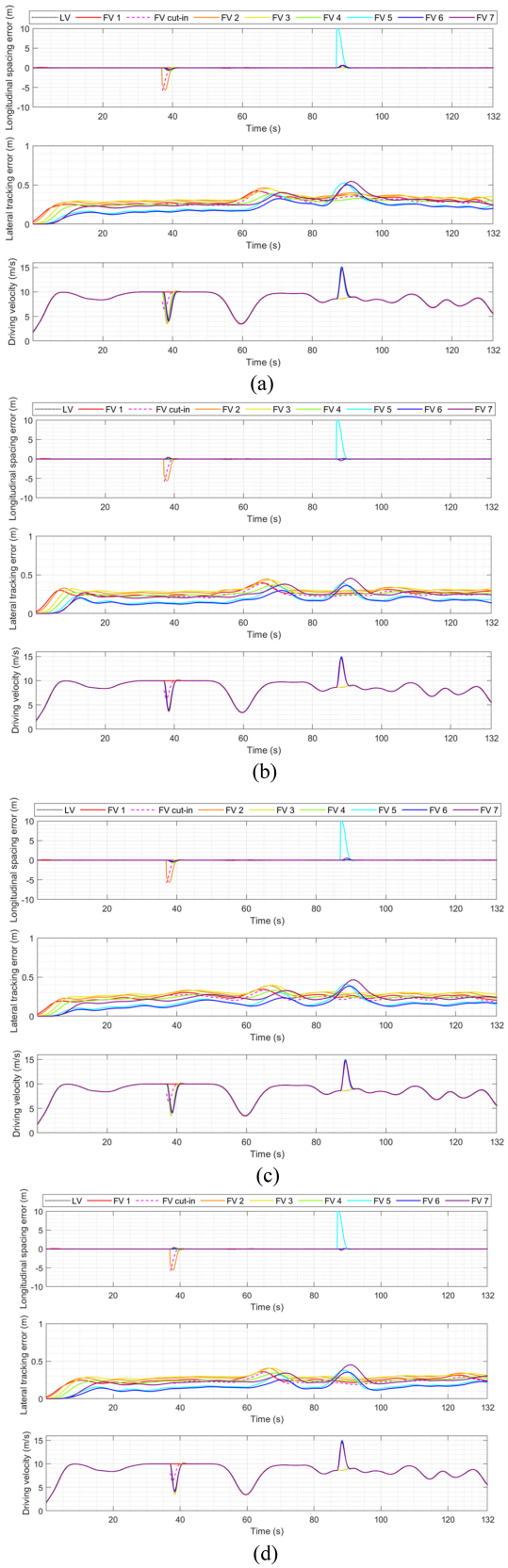


Fig. 9. Control results of the autonomous vehicle platoon using different communication topologies. (a) PF. (b) PLF. (c) TPF. (d) TPLF.

virtual straight path in the same direction as the platoon before passing the start position of the actual navigation path, the lateral errors of FVs are 0 during the initial driving transition.

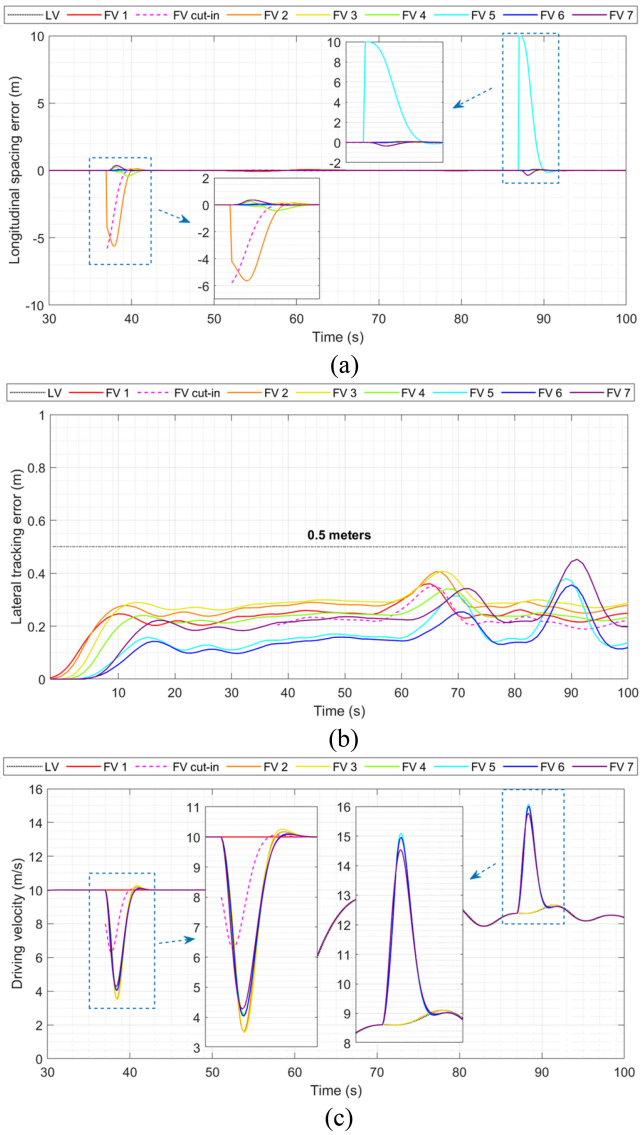


Fig. 10. Local magnification for the control results of the autonomous vehicle platoon using TPLF. (a) Longitudinal spacing error. (b) Lateral tracking error. (c) Driving velocity.

Fig. 10(c) depicts the driving velocity variations of different FVs during the cut-in/cut-out maneuvers. When cut-in maneuver happens, the preceding spacing for the cut-in FV and the second FV is much lower than 10 m. Therefore, both of them decelerate to create longer spacing distance. Consequently, the other FVs behind second FV also decelerate to keep the preset platoon spacing. When the preceding spacing returns to 10 m, all FVs will once again drive at the same velocity as the LV. Similarly, when cut-out maneuver happens, the fifth FV accelerates to catch up with the third FV, and the sixth and seventh FVs also accelerate to keep the preset platoon spacing. The analysis of Fig. 10(c) is consistent with that of Fig. 10(a). Above all, the proposed IDM-PC method can handle various communication topologies and dynamic cut-in/cut-out maneuvers.

Above all, the MPCF for heterogeneous autonomous vehicle platoon has been validated to achieve expected performance in path tracking and platoon keeping.

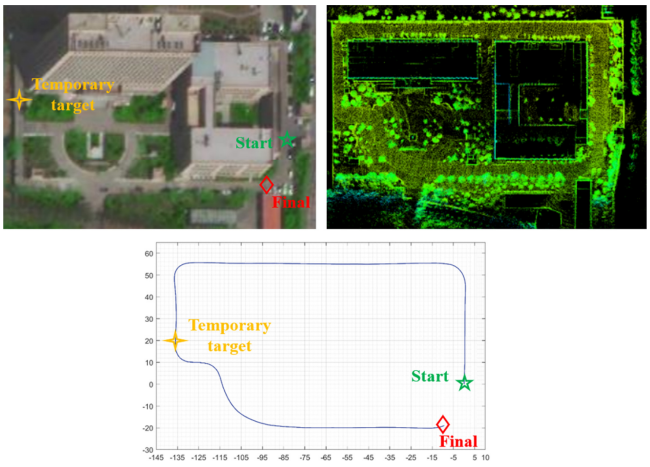


Fig. 11. Driving scene for adaptability verification of the MPCF.

TABLE V
NUMERICAL PATH TRACKING RESULTS FOR ADAPTABILITY VERIFICATION

Methods	Average tracking error (m)	Average driving velocity (m/s)	Computation time (s)
HDY-PC	0.09	4.02	0.02
Dyna	0.17	3.09	0.01
Q-learning	0.19	1.86	0.01

D. Adaptability Verification for the Multilayer Predictive Control Framework

To further validate the adaptability performance of the MPCF for heterogeneous autonomous vehicle platoon, another scene extracted from the real world is applied. The relevant satellite map based on GPS, the point cloud map based on SLAM and the planned navigation path are shown in Fig. 11. The preset spacing constant between adjacent vehicles is 5 m, and each FV in the autonomous platoon is in the desired position initially.

The path tracking results of the LV with different methods are shown in Fig. 12. It is evident that the LV controlled by HDY-PC of MPCF still achieves higher tracking accuracy than the other two RL methods in the second driving scene. Especially in the large angle turn shown in Fig. 12(a), the actual motion trajectory of the LV with HDY-PC is highly consistent with the navigation path. Moreover, the tracking error of the LV with proposed method is always less than 0.2 m, which is still significantly better than those of other two RL methods. Therefore, the tracking accuracy of the MPCF in the second driving scene has been demonstrated.

Table V lists the numerical path tracking results of the LV controlled by three methods. It is obvious that the LV with HDY-PC of MPCF still has the smallest average tracking error and the fastest average driving velocity. Besides, the computation time still meets the real-time control requirements of autonomous vehicle platoon. Therefore, the adaptability of the HDY-PC-based path tracking layer in MPCF has been proved.

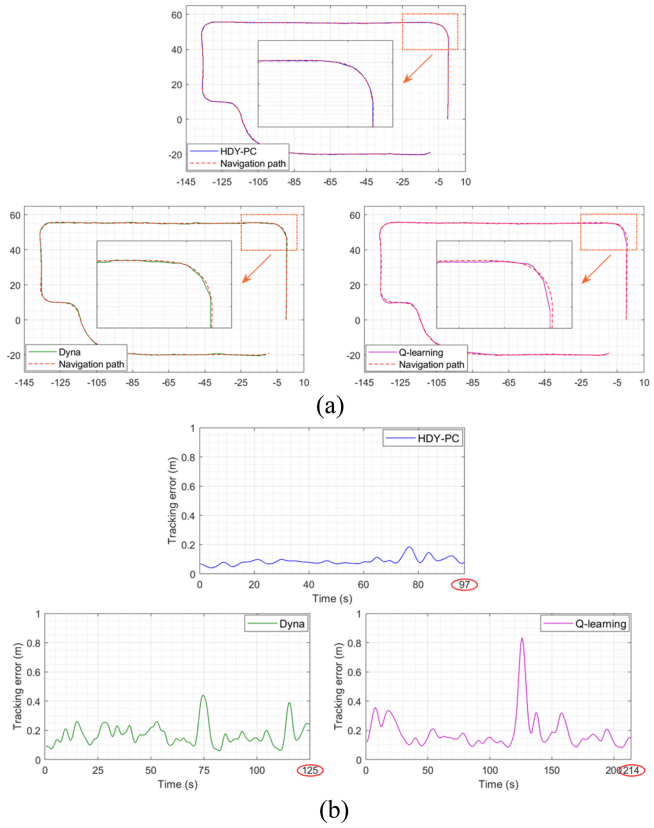


Fig. 12. Path tracking results of the LV with different methods for adaptability verification. (a) Tracking trajectory. (b) Tracking error.

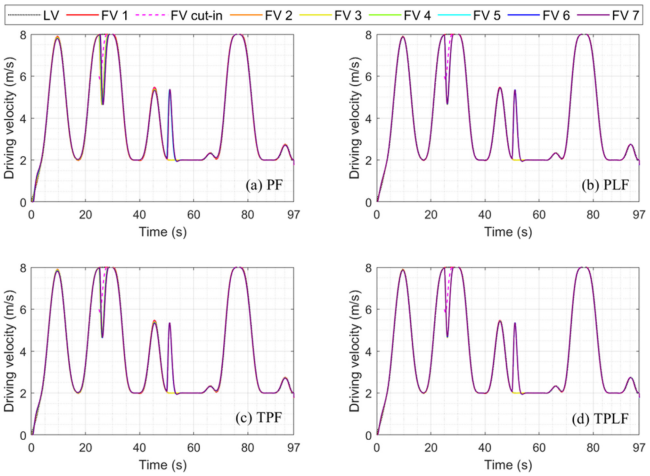


Fig. 13. Driving velocities of the autonomous vehicle platoon for adaptability verification. (a) PF. (b) PLF. (c) TPF. (d) TPLF.

Furthermore, the driving velocities of the autonomous vehicle platoon using different communication topologies are illustrated in Fig. 13. Obviously, the driving velocity curves of all FVs are still highly coincident with that of the LV in the second scene, which ensures each FV can keep a precise distance from the preceding vehicle. When cut-in maneuver happens between the first FV and the second FV, the cut-in FV and the second FV decelerate to create longer spacing distance. Meanwhile, the other FVs behind second FV also decelerate to keep the preset platoon spacing. Similarly, when cut-out

TABLE VI
NUMERICAL PLATOON KEEPING RESULTS FOR ADAPTABILITY VERIFICATION

Topologies	Vehicles	Maximum spacing errors (m)		Maximum tracking errors (m)	
PF	FV 1; FV 2;	0.08;	0.06;	0.37;	0.39;
	FV 3; FV 4;	0.07;	0.06;	0.38;	0.33;
	FV 5; FV 6;	0.07;	0.09;	0.32;	0.40;
	FV 7; FV cut-in	0.10;	0.06	0.40;	0.36
PLF	FV 1; FV 2;	0.07;	0.04;	0.36;	0.39;
	FV 3; FV 4;	0.06;	0.05;	0.38;	0.32;
	FV 5; FV 6;	0.07;	0.08;	0.33;	0.32;
	FV 7; FV cut-in	0.09;	0.06	0.39;	0.34
TPF	FV 1; FV 2;	0.07;	0.04;	0.30;	0.37;
	FV 3; FV 4;	0.06;	0.05;	0.37;	0.28;
	FV 5; FV 6;	0.07;	0.08;	0.38;	0.35;
	FV 7; FV cut-in	0.09;	0.07	0.39;	0.29
TPLF	FV 1; FV 2;	0.06;	0.04;	0.29;	0.31;
	FV 3; FV 4;	0.05;	0.05;	0.32;	0.27;
	FV 5; FV 6;	0.06;	0.03;	0.34;	0.31;
	FV 7; FV cut-in	0.06;	0.04	0.37;	0.28

maneuver of the fourth FV happens, the fifth FV accelerates to catch up with the third FV, and the sixth and seventh FVs also accelerate to keep the preset platoon spacing. When the preceding spacing of each FV returns to preset constant, all FVs will once again drive at the same velocity as the LV. Therefore, the proposed control framework can handle various communication topologies and dynamic cut-in/cut-out maneuvers in the second driving scene.

Table VI shows the numerical platoon keeping results under different communication topologies without considering sudden errors caused by the dynamic cut-in/cut-out maneuvers. It can be found that the spacing error for each FV is always kept below 0.1 m. Meanwhile, the tracking error for each FV is always kept below 0.4 m. Therefore, the platooning stability and robustness in the second driving scene are demonstrated, where all of vehicles maintain a preset spacing in the longitudinal direction and consistency in the lateral direction.

Above all, in the second driving scene, the MPCF for heterogeneous autonomous vehicle platoon performs well in both path tracking and platoon keeping. The adaptability of the framework has been validated.

E. Further Validation for the Multilayer Predictive Control Framework

To further evaluate the progressiveness of the MPCF, six recently proposed methods are introduced into the comparative experiment. The path tracking performance of the proposed framework (MPCF) is compared with the adaptive MPC (AMPC) proposed in [43], the robust MPC (RMPC) proposed in [44] and the forward predictive control method based on the QL algorithm (QL-FPC) proposed in [8]. Meanwhile, the platoon keeping performance of the proposed framework is compared with the objective-based decoupled controller (ODC) proposed in [45], the integrated motion control framework (IMCF) proposed in [46] and the holistic robust motion controller framework (RMCF) proposed in [14]. The validation scene used for comparative experiment is shown in Fig. 6, and PLF is selected as the communication topology of the autonomous platoon. For the consistency of the experimental comparison, the cut-in/cut-out maneuvers will not be included in the comparative experiment.

TABLE VII
NUMERICAL PATH TRACKING RESULTS OF DIFFERENT METHODS IN COMPARATIVE EXPERIMENT

Methods	Average tracking error (m)	Average driving velocity (m/s)
MPCF	0.13	8.54
AMPC	0.17	7.23
RMPC	0.21	6.56
QL-FPC	0.18	7.62

TABLE VIII
NUMERICAL PLATOON KEEPING RESULTS OF DIFFERENT METHODS IN COMPARATIVE EXPERIMENT

Method	Vehicles	Average spacing errors (m)		Average tracking errors (m)	
MPCF	FV 1; FV 2;	0.03;	0.02;	0.22;	0.21;
	FV 3; FV 4;	0.02;	0.02;	0.23;	0.20;
	FV 5; FV 6;	0.02;	0.02;	0.21;	0.23;
	FV 7;	0.02;		0.19;	
ODC	FV 1; FV 2;	0.05;	0.04;	0.34;	0.37;
	FV 3; FV 4;	0.04;	0.03;	0.31;	0.29;
	FV 5; FV 6;	0.03;	0.04;	0.35;	0.33;
	FV 7;	0.03;		0.31;	
IMCF	FV 1; FV 2;	0.06;	0.06;	0.35;	0.32;
	FV 3; FV 4;	0.06;	0.05;	0.34;	0.39;
	FV 5; FV 6;	0.04;	0.05;	0.40;	0.39;
	FV 7;	0.05;		0.33;	
RMCF	FV 1; FV 2;	0.04;	0.04;	0.30;	0.28;
	FV 3; FV 4;	0.03;	0.03;	0.32;	0.31;
	FV 5; FV 6;	0.02;	0.03;	0.34;	0.30;
	FV 7;	0.03;		0.30;	

The path tracking results of different methods in comparative experiment are shown in Table VII. Among these control methods, it is obvious that the proposed MPCF still achieves the smallest average tracking error and the fastest average driving velocity, reflecting its superiority in tracking accuracy and driving rapidity. Furthermore, the platoon keeping results of different methods in comparative experiment are shown in Table VIII. It can be found that FVs in autonomous platoon controlled by proposed MPCF also have the smallest spacing errors and tracking errors compared to FVs controlled by other methods. Therefore, the platoon keeping precision of the proposed framework is the highest in both longitudinal and lateral directions.

To sum up, the MPCF performs better than these six comparison methods, and the progressiveness of this research has been validated.

V. Conclusion

This article proposes a MPCF based on heuristic learning agent and improved distributed model for the heterogeneous autonomous vehicle platoon in nonspecific scenarios. The multilayer framework ensures the stability of the longitudinal and lateral movement of the autonomous vehicle platoon through the differentiated control logic of the leading and FVs. In the upper layer, the HDY-PC method is proposed to improve the path tracking performance of the LV. In the lower layer, the IDM-PC method is developed to guarantee the high-precision formation and stability of the vehicle platoon. The virtual environment simulation and real-world scene validate the superior performance of the MPCF. This framework has also been proved to handle various communication topologies and dynamic cut-in/cut-out maneuvers.

Further research is applying the MPCF to the real vehicle platoon test with disturbances and uncertainties. The time delay and packet loss will also be considered in the next research step.

Acknowledgment

The authors would like to thank them for their support and help.

References

[1] B. Deng, J. Nan, W. Cao, and W. Wang, "A survey on integration of network communication into vehicle real-time motion control," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 4, pp. 2755–2790, 4th Quart., 2023.

[2] Y.-S. Ma, W.-W. Che, C. Deng, and Z.-G. Wu, "Data-driven distributed vehicle platoon control for heterogeneous nonlinear vehicle systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 3, pp. 2373–2381, Mar. 2024, doi: 10.1109/TITS.2023.3321750.

[3] H. Zhang, J. Peng, H. Dong, F. Ding, and H. Tan, "Integrated velocity optimization and energy management strategy for hybrid electric vehicle platoon: A multiagent reinforcement learning approach," *IEEE Trans. Transp. Electrif.*, vol. 10, no. 2, pp. 2547–2561, Jun. 2024, doi: 10.1109/TTE.2023.3298365.

[4] Y. Liu, D. Yao, L. Wang, and S. Lu, "Distributed adaptive fixed-time robust platoon control for fully heterogeneous vehicles," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 1, pp. 264–274, Jan. 2023.

[5] P. Barooah, P. Mehta, and J. Hespanha, "Mistuning-based control design to improve closed-loop stability margin of vehicular platoons," *IEEE Trans. Autom. Control*, vol. 54, no. 9, pp. 2100–2113, Sep. 2009.

[6] Y. Zheng, S. Li, K. Li, and L. Wang, "Stability margin improvement of vehicular platoon considering undirected topology and asymmetric control," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 4, pp. 1253–1265, Jul. 2016.

[7] W. Dunbar and D. Caveney, "Distributed receding horizon control of vehicle platoons: Stability and string stability," *IEEE Trans. Autom. Control*, vol. 57, no. 3, pp. 620–633, Mar. 2012.

[8] G. Du, Y. Zou, X. Zhang, Z. Li, and Q. Liu, "Hierarchical motion planning and tracking for autonomous vehicles using global heuristic based potential field and reinforcement learning based predictive control," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 8, pp. 8304–8323, Apr. 2023.

[9] L. Lei, T. Liu, K. Zheng, and L. Hanzo, "Deep reinforcement learning aided platoon control relying on V2X information," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 5811–5826, Jun. 2022.

[10] M. Li, Z. Cao, and Z. Li, "A reinforcement learning-based vehicle platoon control strategy for reducing energy consumption in traffic oscillations," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 12, pp. 5309–5322, Dec. 2021.

[11] Y. Liu, D. Yao, H. Li, and R. Lu, "Distributed cooperative compound tracking control for a platoon of vehicles with adaptive NN," *IEEE Trans. Cybern.*, vol. 52, no. 7, pp. 7039–7048, Jul. 2022.

[12] S. Feng, Y. Zhang, S. Li, Z. Cao, X. Liu, and L. Li, "String stability for vehicular platoon control: Definitions and analysis methods," *Annu. Rev. Control*, vol. 47, Jun. 2019, pp. 81–97.

[13] S. Feng, Z. Song, Z. Li, Y. Zhang, and L. Li, "Robust platoon control in mixed traffic flow based on tube model predictive control," *IEEE Trans. Intell. Veh.*, vol. 6, no. 4, pp. 711–722, Dec. 2021.

[14] H. Wang et al., "A holistic robust motion control framework for autonomous platooning," *IEEE Trans. Veh. Technol.*, vol. 72, no. 12, pp. 15213–15226, Dec. 2023.

[15] Y. Zheng, S. Li, K. Li, F. Borrelli, and J. Hedrick, "Distributed model predictive control for heterogeneous vehicle platoons under unidirectional topologies," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 3, pp. 899–910, May 2017.

[16] Z. Qiang, L. Dai, B. Chen, and Y. Xia, "Distributed model predictive control for heterogeneous vehicle platoon with inter-vehicular spacing constraints," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 3, pp. 3339–3351, Mar. 2023.

[17] P. Wang, H. Deng, J. Zhang, L. Wang, M. Zhang, and Y. Li, "Model predictive control for connected vehicle platoon under switching communication topology," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 7817–7830, Jul. 2022.

[18] A. Gratzer, S. Thormann, A. Schirrer, and S. Jakubek, "String stable and collision-safe model predictive platoon control," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 19358–19373, Oct. 2022.

[19] D. Pi, P. Xue, B. Xie, H. Wang, X. Tang, and X. Hu, "A platoon control method based on DMPC for connected energy-saving electric vehicles," *IEEE Trans. Transp. Electrif.*, vol. 8, no. 3, pp. 3219–3235, Sep. 2022.

[20] M. Xu, Y. Luo, W. Kong, and K. Li, "A distributed model predictive control method combined with delay compensator for multiple vehicle platoons," *IET Intell. Transp. Syst.*, vol. 17, no. 2, pp. 357–372, Sep. 2022.

[21] X. Hu, L. Xie, L. Xie, S. Lu, W. Xu, and H. Su, "Distributed model predictive control for vehicle platoon with mixed disturbances and model uncertainties," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 17354–17365, Oct. 2022.

[22] M. Qiu, D. Liu, S. Baldi, G. Yin, W. Yu, and M. Cao, "Vehicular platooning on curved paths considering collision avoidance and string stability," *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 2, pp. 1896–1908, Mar./Apr. 2024, doi: 10.1109/TNSE.2023.3333027.

[23] S. Wang, Z. Li, B. Wang, and M. Li, "Collision avoidance motion planning for connected and automated vehicle platoon merging and splitting with a hybrid automaton architecture," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 2, pp. 1445–1464, Feb. 2024, doi: 10.1109/TITS.2023.3315063.

[24] T. Liu, L. Lei, K. Zheng, and K. Zhang, "Autonomous platoon control with integrated deep reinforcement learning and dynamic programming," *IEEE Internet Things J.*, vol. 10, no. 6, pp. 5476–5489, Mar. 2023.

[25] Y. Feng et al., "Distributed MPC of vehicle platoons considering longitudinal and lateral coupling," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 3, pp. 2293–2310, Mar. 2024.

[26] L. D'Alfonso, F. Giannini, G. Franzè, G. Fedele, F. Pupo, and G. Fortino, "Autonomous vehicle platoons in urban road networks: A joint distributed reinforcement learning and model predictive control approach," *IEEE/CAA J. Automatica Sinica*, vol. 11, no. 1, pp. 141–156, Jan. 2024.

[27] S. Aradi, "Survey of deep reinforcement learning for motion planning of autonomous vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 740–759, Feb. 2022.

[28] S. Cheng, L. Li, X. Chen, J. Wu, and H. Wang, "Model-predictive-control-based path tracking controller of autonomous vehicle considering parametric uncertainties and velocity-varying," *IEEE Trans. Ind. Electron.*, vol. 68, no. 9, pp. 8698–8707, Sep. 2021.

[29] S. Teng et al., "Motion planning for autonomous driving: The state of the art and future perspectives," *IEEE Trans. Intell. Veh.*, vol. 8, no. 6, pp. 3692–3711, Jun. 2023.

[30] H. Dong, and J. Xi, "Model predictive longitudinal motion control for the unmanned ground vehicle with a trajectory tracking model," *IEEE Trans. Veh. Technol.*, vol. 71, no. 2, pp. 1397–1410, Feb. 2022.

[31] L. Tang, F. Yan, B. Zou, K. Wang, and C. Lv, "An improved kinematic model predictive control for high-speed path tracking of autonomous vehicles," *IEEE Access*, vol. 8, pp. 51400–51413, Mar. 2020.

[32] H. Bao et al., "Moment-based model predictive control of autonomous systems," *IEEE Trans. Intell. Veh.*, vol. 8, no. 4, pp. 2939–2953, Jan. 2023.

[33] Y. Shan, B. Zheng, L. Chen, L. Chen, and D. Chen, "A reinforcement learning-based adaptive path tracking approach for autonomous driving," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 10581–10595, Oct. 2020.
[34] V. Behzadan, and A. Munir, "Adversarial reinforcement learning framework for benchmarking collision avoidance mechanisms in autonomous vehicles," *IEEE Intell. Transp. Syst. Mag.*, vol. 13, no. 2, pp. 236–241, Jun. 2021.
[35] L. Ding, S. Li, H. Gao, C. Chen, and Z. Deng, "Adaptive partial reinforcement learning neural network-based tracking control for wheeled mobile robotic systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 7, pp. 2512–2523, Jul. 2020.
[36] L. Zhang, R. Zhang, T. Wu, R. Weng, M. Han, and Y. Zhao, "Safe reinforcement learning with stability guarantee for motion planning of autonomous vehicles," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 12, pp. 5435–5444, Dec. 2021.
[37] G. Du, Y. Zou, X. Zhang, T. Liu, J. Wu, and D. He, "Deep reinforcement learning based energy management for a hybrid electric vehicle," *Energy*, vol. 201, Jun. 2020, Art. no. 117591.
[38] M. Basiri, B. Ghojogh, N. Azad, S. Fischmeister, F. Karray, and M. Crowley, "Distributed nonlinear model predictive control and metric learning for heterogeneous vehicle platooning with cut-in/cut-out maneuvers," in *Proc. 59th IEEE Conf. Decis. Control (CDC)*, 2020, pp. 2849–2856.
[39] Y. Zheng, S. Eben Li, J. Wang, D. Cao, and K. Li, "Stability and scalability of homogeneous vehicular platoon: Study on the influence of information flow topologies," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 1, pp. 14–26, Jan. 2016.
[40] Y. Yan, H. Du, D. He, and W. Li, "Pareto optimal information flow topology for control of connected autonomous vehicles," *IEEE Trans. Intell. Veh.*, vol. 8, no. 1, pp. 330–343, Jan. 2023.
[41] S. E. Li, Y. Zheng, K. Li, and J. Wang, "An overview of vehicular platoon control under the four-component framework," in *Proc. IEEE Intell. Veh. Symp. (IV)*, 2015, pp. 286–291.
[42] Y. Zhang, Z. Xu, Z. Wang, X. Yao, and Z. Xu, "Impacts of communication delay on vehicle platoon string stability and its compensation strategy: A review," *J. Traffic Transp. Eng.*, vol. 10, no. 4, pp. 508–529, Aug. 2023.
[43] Z. Zhang, L. Zheng, Y. Li, S. Li, and Y. Liang, "Cooperative strategy of trajectory tracking and stability control for 4WID autonomous vehicles under extreme conditions," *IEEE Trans. Veh. Technol.*, vol. 72, no. 3, pp. 3105–3118, Mar. 2023.
[44] J. Liang, Q. Tian, J. Feng, D. Pi, and G. Yin, "A polytopic model-based robust predictive control scheme for path tracking of autonomous vehicles," *IEEE Trans. Intell. Veh.*, vol. 9, no. 2, pp. 3928–3939, Feb. 2024.
[45] N. Zhang, X. Li, J. Chen, H. Li, Y. Nie, and H. Zhang, "Robust lateral and longitudinal control for vehicle platoons with unknown interaction topology subject to multiple communication delays," *IEEE Trans. Intell. Veh.*, early access, Feb. 13, 2024, doi: 10.1109/TIV.2024.3365666.
[46] L. Song, J. Li, Z. Wei, K. Yang, E. Hashemi, and H. Wang, "Longitudinal and lateral control methods from single vehicle to autonomous platoon," *Green Energy Intell. Transp.*, vol. 2, no. 2, Apr. 2023, Art. no. 100066.

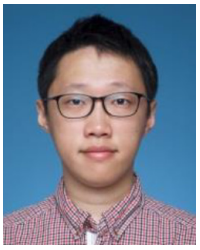
Guodong Du received the B.S. degree in mechanical engineering from Beijing Institute of Technology, Beijing, China, in 2019, where he is currently pursuing the Ph.D. degree in automobile engineering.

He also serves as an Academic Guest with ETH Zürich, Zürich, Switzerland. His research interests include autonomous driving, motion planning and control, reinforcement learning algorithm, vehicle dynamics control, and energy management of hybrid electric vehicles.

Dr. Du won the "Best Student Paper Award" in the 35th IEEE Intelligent Vehicles Symposium (IV 2024).

Yuan Zou (Senior Member, IEEE) received the Ph.D. from Beijing Institute of Technology, Beijing, China, in 2005.

He is currently a Professor with Beijing Collaborative and Innovative Center for Electric Vehicles and School of Mechanical Engineering, Beijing Institute of Technology. He also serves as a Co-Director with ETHZ-BIT Joint Research Center for New Energy Vehicle Dynamic System and Control. He conducted research about ground vehicle propulsion modeling and optimal control with The University of Michigan at Ann Arbor, Ann Arbor, MI, USA, and ETH Zürich, Zürich, Switzerland. His research interests include modeling and control for electrified vehicle and transportation system.



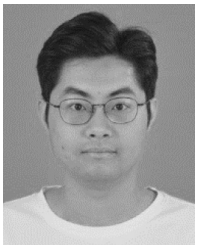
Xudong Zhang (Member, IEEE) received the B.S. and M.S. degrees in mechanical engineering from Beijing Institute of Technology, Beijing, China, in 2011 and 2014, respectively, and the Ph.D. degree (with Distinction) in mechanical engineering from the Technical University of Berlin, Berlin, Germany, in 2017.

Since 2017, he has been an Associate Professor with Beijing Institute of Technology. His research interests include vehicle dynamics control, autonomous driving, and energy management for electric vehicles.

Dr. Zhang was the recipient of the Young Elite Scientists Sponsorship Program supported by CAST (China Association for Science and Technology) and CSAE (China Society of Automotive Engineers).

Jie Fan received the M.S. degree from Beijing Institute of Technology, Beijing, China, in 2019, where he is currently pursuing the Ph.D. degree with the National Engineering Research Center for Electric Vehicles.

His current research interests include the state estimation of power battery, intelligent energy management of hybrid electric vehicle, and path planning for autonomous vehicles.



Wenjing Sun received the B.S. degree from Beijing Institute of Technology, Beijing, China, in 2019, where he is currently pursuing the Ph.D. degree with the National Engineering Laboratory for Electric Vehicles.

He is also an Academic Guest with The City university of Hong Kong, Hong Kong. His current research interests include time-sensitive networking and vehicle electrical/electronic architecture.

Zirui Li received the B.S. degree from Beijing Institute of Technology, Beijing, China, in 2019, where he is currently pursuing the Ph.D. degree in mechanical engineering under the supervision of Prof. Jianwei Gong.

From June 2021 to July 2022, he was a Visiting Researcher with Delft University of Technology, Delft, Netherlands, with CSC Funding from China. Since August 2022, he has been the Visiting Researcher with the Chair of Traffic Process Automation, Faculty of Transportation and Traffic Sciences "Friedrich List" of the TU Dresden. His research focuses on interactive behavior modeling, risk assessment, and motion planning of automated vehicles.