

An Integrated of Decision Making and Motion Planning Framework for Enhanced Oscillation-Free Capability

Zhuoren Li^{ID}, Student Member, IEEE, Jia Hu^{ID}, Member, IEEE, Bo Leng^{ID}, Lu Xiong^{ID}, and Zhiqiang Fu^{ID}

Abstract— Autonomous driving requires efficient and safe decision making and motion planning in dynamic and uncertain environments. Future movement of surrounding vehicles is often difficult to represent. Besides, most existing studies consider decision making and planning/control separately. Both them may lead to the oscillation and unsafe for autonomous driving. This paper proposes an integrated framework of decision making and motion planning with oscillation-free capability. The proposed approach overcomes the shortcomings of autonomous driving for lane change/keeping maneuvers and is able to: i) make oscillation-free behavior decisions given biased prediction; ii) cut through in the traffic efficiently and safely when being in squeezed; iii) accelerate computation efficiency by building a state transfer model based on prediction uncertainty; iv) reduce the dissonance between decision-making and motion planning. A belief decision planner is designed with the uncertainty of the prediction trajectories. Lateral and longitudinal drivable corridors including the reference state and the related boundary constraints are built, which provide better suited information for planning to solve the optimal motion sequence more quickly and stably, and improve its consistency with decision module. Finally, the problem is formulated as an optimal control problem considering the vehicle dynamics and some soft constraints and the motion trajectory is solved by OSQP. Simulation and experimental tests are implemented to evaluate the feasibility and effectiveness of the proposed approach. Test results show that the integrated approach can make proper, safe and continuous decision and planning for autonomous vehicles and the calculation time is very low.

Index Terms— Autonomous driving, decision making, motion planning, uncertainty.

I. INTRODUCTION

EVERY step of the proposal, development and implementation of autonomous vehicle will have a significant impact on the automotive industry, and will also bring about huge changes to the entire automotive industry, even the

Manuscript received 31 March 2023; revised 23 October 2023; accepted 9 November 2023. Date of publication 23 November 2023; date of current version 31 May 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 52372394, in part by the National Key Research and Development Program of China under Grant 2022YFE0117100, and in part by the Science and Technology Commission of Shanghai under Grant 21DZ1203802. The Associate Editor for this article was Y. Zhang. (*Corresponding author: Bo Leng*)

Zhuoren Li, Bo Leng, Lu Xiong, and Zhiqiang Fu are with the School of Automotive Studies, Tongji University, Shanghai 201804, China (e-mail: 1911055@tongji.edu.cn; lengbo@tongji.edu.cn; xiong_lu@tongji.edu.cn; fu0zhiqiang@163.com).

Jia Hu is with the Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 201804, China (e-mail: hujia@tongji.edu.cn).

Digital Object Identifier 10.1109/TITS.2023.3332655

transportation field and the way of travel in society. Autonomous vehicles mainly consist of the environment perception and prediction, decision making, motion planning and vehicle control [1].

Decision making and motion planning, which are regarded as the AV's brain [2], bridge the environment perception and vehicle control. They need to select driving behaviors that meet human driving requirements and habits in a dynamic environment with multiple traffic participants, and generate a feasible trajectory from an initial state to a target state following vehicle constraints and traffic rules [3].

The most challenge for motion planning is that autonomous vehicles tend to be overly aggressive or conservative, or repeatedly jitter in different behaviors in highly dynamic maneuvers such as overtaking in traffic or being cut in. It is difficult to make decision and plan with safety, high-efficiency and speediness in such highly dynamic risk environment [4].

On the one hand, there are a lot of motion uncertainties in the traffic scenarios. The surrounding traffic participants have uncertainty of its behaviors and trajectory poses, and their future behavior is tightly coupled with the future trajectory of the ego vehicle [5]. Describing the optimal solution in such scenarios with a decision policy instead of a trajectory allows for less conservative behavior [6].

Interaction-based decision-making methods, such as Game Theory, Partially Observable Markov Decision Processes (POMDP), et. al., have been extensively studied to solve this problem [7]. They allow to model interactive behavior and uncertain motion with their probabilistic action policy. Based on game theory, a human-like decision-making system is proposed for autonomous vehicles to realize automatic lane change [8], merging [9] and car following [10]. Reference [11] uses game theory to solve the interactive driving strategy, which obtains an action sequences to balance uncertain environmental risks and behavior of ego vehicle in planning horizon. A driving model based on a stackelberg strategy is proposed and tested in highway scenario [12]. Although game theory has been widely used in the studies of interactive decision-making under urban roads, there is currently no unified framework. POMDP considers the observability of state information and adds a belief space to the system to represent the probability distribution of each possible state, thus it has a better ability to deal with uncertainty problems. Reference [13] propose a multi-policy decision-making (MPDM) framework to select the best policy by simultaneously forward-simulating

the state of ego vehicle and other traffic participants through the POMDP model. Reference [14] estimates the distribution of potential policies of surrounding vehicles based on Bayesian change points detection and deduces the interaction between multiple agents. In [15], a longitudinal POMDP decision model that considers the hidden intentions of other vehicles at the intersection without traffic lights is proposed to obtain the optimal acceleration policy for ego vehicle. Reference [16] establish a MDI-POMDP framework integrating behavior decision and motion planning, which consider multi-modal driving risk field based on dynamic collision region. However, since there is a high-dimensional state/action modeling and the solution process considers the motion evolution of others, both POMDP and game-theory are easy to fall into the dimensionality problem, which makes the algorithm difficult to solve.

On the other hand, most existing studies consider decision making and planning/control separately. Although the separating structure has many advantages such as modularization and intuitiveness, however, simple decision results may not be effectively utilized by planning, which tends to make the solution process of motion planning time-consuming, or the planning unable to reach decision expectations, trajectory shaking and even solving failure in dealing with complex scenarios [17].

Many decision methods explicitly define the tasks or functions of planning module, such as avoidance, overtaking, merging, U-turn, accelerating and braking, etc., enabling the planning module to perform specific tasks in specific scenarios [18], [19], [20]. The advantage of such task-based connection is that the tasks for the motion planning are simple and clear, which can effectively provide guidance and constraints. However, most of the work is handled by the decision-making module, which has high requirements for scene modeling and maneuver understanding. Moreover, the decision outputs are discrete semantic actions, which does not inherit enough information and may lead to unreasonable results in planned trajectories. Thus, there are some other decision approach does not explicitly output a task instruction, but implicitly outputs the result according to the method adopted by the motion planning module [21]. Reference [22] proposes a parameter decision-making framework, whose output is not a certain action or a specific task, but a sequence of motion state including position, heading, speed, etc. [23] designs a decision model to build a lane-changing space based on virtual targets, which constructs a series of candidate spaces in the target lane and then selects the target space for trajectory planning. Reference [24] discretizes the time and position coordinates to build a dynamic grid map and calculate whether the lateral space meets the lane-changing requirements, and then provide parameter guidance for the polynomial curve to generate the trajectory. Compared with the task/function output, the method-based output not only provides certain target, but also has more implicit constraints, which can avoid some unforeseen planning results. Some work has attempted to integrate decision-making and planning in a single method through such as Mixed Integer Programming (MIP) [25], [26], but these methods also place enormous challenges to the computation.

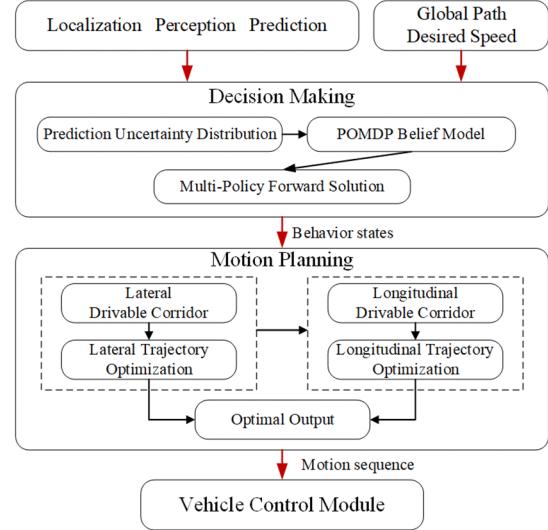


Fig. 1. Integrated framework structure of decision-making and trajectory planning.

In addition, there are reinforcement learning methods can also integrate decision, planning and control [27], some even the perception through End-to-End network [28]. However, due to the inexplicability of learning-based methods, it is difficult to guarantee their security and robustness, which is out of scope of this paper. Therefore, how to effectively combine the decision-making method with the motion planning and ensure the timeliness of the framework is still an issue worth exploring.

To this end, this paper proposes an integrated framework of decision-making and motion planning for autonomous driving focus on the lane change/keeping maneuvers. Firstly, we design a belief POMDP decision planner while building the uncertainty of the prediction trajectories. Through the Multi-policy forward solution, getting the optimal decision action and the corresponding future states with the consideration of the uncertainty risk for surrounding vehicles. Then, based on the decision results of lateral semantic behavior and longitudinal continuous acceleration, we build drivable corridors including the reference state and the related boundary constraints, which provide better suited information for planning to solve the optimal motion sequence more quickly and stably, and improve its consistency with decision module. Finally, we consider the vehicle dynamics and introduce some soft constraints to solve the optimal motion trajectory. The overall structure of the integrated framework is shown in Fig. 1, and the main features of this paper are as follows:

- Able to make oscillation-free behavior decisions given biased prediction.
- Able to cut through in the traffic efficiently and safely when being in squeezed.
- Able to accelerate computation efficiency by building a state transfer model based on prediction uncertainty
- Able to reduce the dissonance between decision-making and motion planning.

The remainder of the paper is organized as follows: In section II, the belief decision-making model considering prediction uncertainty and its solution are introduced. Then, the

motion planner based on optimal approach with lateral and longitudinal drivable corridors are proposed in Section III. In Section IV, simulation and experimental testing and discussion of proposed approach are given. Finally, Section V concludes the paper.

II. BELIEF DECISION-MAKING CONSIDERING PREDICTION UNCERTAINTY

The focus of this section is to solve the best motion state of ego vehicle (EV) when considering the motion uncertainty of others. We construct the problem as a POMDP, using a simplified model of vehicle kinematics to construct the state space and action space. In addition, the multivariate Gaussian distribution is used to describe the uncertainty of the predicted trajectory of other vehicles (OVs), and then the observation space, observation function and belief space of the POMDP model are established.

The solution methods of POMDP are mainly divided into offline methods and online methods. Offline methods mainly include value iteration based on dynamic programming and approximation based on belief points. Online methods mainly include heuristic search algorithm and Monte Carlo tree search algorithm [29], such as POMCP [30], ABT and DESPOT [31]. However, the POMDP model cannot directly observe the intentions of other agents, but samples the state through observation. As the number of samples and the dimension of the system increase, the corresponding computational complexity increases exponentially. The efficiency of many online solvers still suffers from inadequacies, despite various simplifications and discretizations. Drawing on the idea of MDPM [11], the motion state of EV and the surrounding OVs in the future is deduced through forward-simulation. By evaluating the uncertainties of OVs based on the prediction results instead of modeling the travel, the complex POMDP solving problem can be transformed into the optimal strategy selection with the highest reward among a limited number of strategies, which can effectively improve the real-time performance of the operation. The obtained optimal state sequence of EV for a period of time in the future can provide an effective and feasible sub-goal for the motion planning module.

A. Partially Observable Markov Decision Process

A discrete-time POMDP is defined by the tuple $(X, A, T, \Omega, O, \gamma)$. The state space X contains all the states x in the entire environment. The action space A is the set of all possible actions of the agent. The state-transition matrix $T(x'|x, a) = P(x'|x, a)$ is the probability of the agent transitioning to the next state x' when the agent executes action a in the current state x . $R(x, a)$ is the reward for selecting action a in state x , which is used to evaluate the pros and cons of the strategy in the process of model state transfer. Additionally, the discount factor γ is used to proportionally reduce the immediate reward obtained in the future.

Compared with MDP, in order to be able to deal with uncertainties in the environment, observation-related elements are added to express some of the observable properties of the environment, including: (1) observation space O , containing

all possible observations o , (2) observation function $\Omega = P(o|x', a)$, is the probability to observe o in the next state x' after taking action a , (3) belief space B , is the set of belief states $b(x)$, which satisfies as,

$$\begin{aligned} b(x) &= p(x) \\ \sum_{s \in S} b(s) &= 1, b \in B \end{aligned} \quad (1)$$

The update of the belief state can be calculated as (2):

$$b(x') = \eta P(o|x', a) \sum_{x' \in X} T(x'|x, a) b(x) \quad (2)$$

where η is the regularization factor, which guarantees that the probability of the belief state sums to 1. Then the reward function in belief state $b(x_\tau)$ obtained by taking action a_τ is:

$$R(b, a) = \sum_{x \in X} b(x) R(x, a) \quad (3)$$

The solution of the POMDP is the optimal policy, which is the same as the solution of the MDP problem in the belief space, that is, to maximize the cumulative expected return on the future belief space in a limited time,

$$\pi^* = \max_{a \in A} E \left\{ \sum_{x \in X} R(b, a) + \gamma \sum_{o \in O} P(o|a, b) V^*(b') \right\} \quad (4)$$

B. State Space

The state space must provide sufficient information for belief state updates in POMDP, and all vehicles in the scene are represented in the state space as,

$$R(b, a) = \sum_{x \in X} b(x) R(x, a) \quad (5)$$

x_{ev} represents the state of EV and $x_{ov,i}$ ($i \in [1, n]$) the states of the OVs. Our method is applied to structured road scenarios, where the position of the vehicle is represented by Frenet coordinates (s, l) , where s represents the longitudinal distance along the reference path, and l represents the lateral deviation along the tangent to the reference path.

The EV's state is defined as,

$$x_{ev} = \xi_{ev} = [s_{ev} \ l_{ev} \ \theta_{ev} \ v_{ev}]^T \quad (6)$$

where (s_{ev}, l_{ev}) is the EV's position, θ_{ev} is the heading angle in the Frenet coordinates and v_{ev} is the longitudinal velocity of EV. Similarly, the state of the OVs is,

$$x_{ov,i} = [\xi_{ov,i} \ b_{ov,i}]^T = [s_{ov,i} \ l_{ov,i} \ \theta_{ov,i} \ v_{ov,i} \ b_{ov,i}]^T \quad (7)$$

where, $\xi_{ov,i} = [s_{ov,i} \ l_{ov,i} \ \theta_{ov,i} \ v_{ov,i}]$ is the measurable motion state of OVs, and $b_{ov,i}$ is the unmeasurable driving intention such as lane-changing or overtaking. It is a partially observable state and needs to be calculated from the observed data and prediction model.

C. Action Space

In POMDP modeling, we abstract the driving actions of autonomous vehicle into semantic actions. It is divided into lateral discrete actions and longitudinal continuous actions.

The lateral actions include lane keeping (*LK*), lane change left (*LCL*), and lane change right (*LCR*):

$$a_{lat} \in A_{lat} = [LK, LCL, LCR] \quad (8)$$

The longitudinal action is the acceleration within the capability of EV, which is calculated in Sec. III.F.

$$a_{lon} \in A_{lon} = [a_{ev,min}, a_{ev,max}] \quad (9)$$

The final action space is defined as:

$$A = [A_{lon}, A_{lat}] \quad (10)$$

The discrete lateral semantic action cannot be combined well with the motion planning because it provides too little reference information for motion planning. Therefore, we solve the corresponding optimal state in the forward simulation (Sec. II-F), to facilitate the motion planning to better understand the decision results.

D. Transition Model

It is assumed that each vehicle's state transitions in the environment are independent of each other. The state transition model is divided into EV's transition and OVs' transitions. The complete state transition equation can be derived from the probability formula as:

$$T = P(\mathbf{x}'|\mathbf{x}, a) = P(\mathbf{x}'_{ev}|\mathbf{x}_{ev}, a) \prod_{i=1}^n P(\mathbf{x}'_{ov,i}|\mathbf{x}_{ov,i}) \quad (11)$$

where $P(\mathbf{x}'_{ev}|\mathbf{x}_{ev}, a)$ is the state transition probability of EV, and $P(\mathbf{x}'_{tv,i}|\mathbf{x}_{tv,i})$ of the i^{th} OV.

According to the simplified vehicle kinematics model, the state transition model of EV can be obtained as:

$$\begin{cases} s'_{ev} = s_{ev} + v_{ev}\Delta T_d + 0.5a_{ev}\Delta T_d^2 \\ l'_{ev} = l_{ev} + v_{ev}(\theta_{ev} - \theta_r)\Delta T_d \\ \theta'_{ev} = \theta_{ev} + v_{ev}\Delta T_d \tan(\delta_{ev})/L \\ v'_{ev} = v_{ev} + a_{ev}\Delta T_d \end{cases} \quad (12)$$

where ΔT_d is the discrete decision-making time step.

Compared with EV, the state transition probability of the OVs cannot be directly obtained due to the uncertainty of their motion intention. According to the full probability formula, it is decomposed into two parts as:

$$P(\mathbf{x}'_{ov,i}|\mathbf{x}_{ov,i}) = P(\xi'_{ov,i}|\mathbf{x}_{ov,i})P(b'_{ov,i}|\mathbf{x}_{ov,i}) \quad (13)$$

where $P(\xi'_{ov,i}|\mathbf{x}_{ov,i})$ is the motion state transition probability and $P(b'_{ov,i}|\mathbf{x}_{ov,i})$ is the driving intention transition probability. We assume that the intention distribution probability of other cars can be obtained from the multimodal prediction trajectory. The motion state distribution of the OVs needs to be obtained by constructing its uncertainty distribution according to the corresponding predicted trajectory.

For the OVs, the position uncertainty is propagated with time by describing their future motion states through the point mass model [32], the prediction error update equation is,

$$\mathbf{e}' = (\mathbf{A}_{ov} + \mathbf{B}_{ov}\mathbf{K})\mathbf{e} \quad (14)$$

where $\mathbf{e} = [e_s, e_l]$ is the prediction error vector of the lateral and longitudinal position, \mathbf{A}_{ov} and \mathbf{B}_{ov} are the state matrix and control input matrix of the OVs' motion system:

$$\mathbf{A}_{ov} = \begin{bmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{B}_{ov} = \begin{bmatrix} 0.5\Delta t^2 & 0 \\ \Delta t & 0 \\ 0 & 0.5\Delta t^2 \\ 0 & \Delta t \end{bmatrix} \quad (15)$$

\mathbf{K} is the feedback control matrix. It is assumed that the longitudinal and lateral prediction error form a bivariate normal distribution. As we consider Gaussian distributions, the lateral prediction error and the longitudinal are two normally distributed independent random variables. Then a recursive computation of the prediction error covariance matrix Σ_e is,

$$\Sigma'_e = (\mathbf{A}_{ov} + \mathbf{B}_{ov}\mathbf{K})\Sigma_e(\mathbf{A}_{ov} + \mathbf{B}_{ov}\mathbf{K})^T \quad (16)$$

where $\Sigma_e = diag(\sigma_s^2, 0, \sigma_l^2, 0)$, σ_s^2 is the variance of the lateral prediction error and σ_l^2 the longitudinal.

Giving a confidence parameter p , and the sum of the squares of the longitudinal position prediction errors can be formed as a chi-square distribution with 2 degrees of freedom in (17). Its confidence interval quantile is denoted as r_p , which ensures the probability of the true OV's position lying within the isoline is p .

$$\frac{(e_s - \mu_s)^2}{\sigma_s^2} + \frac{(e_l - \mu_l)^2}{\sigma_l^2} \leq r_p \quad (17)$$

The ellipse semi-major and semi-minor axes are then given by

$$a_{e_s} = \sigma_s \sqrt{r_p}, \quad b_{e_l} = \sigma_l \sqrt{r_p} \quad (18)$$

We use $\zeta_{tv,i} = [s_{tv,i}, l_{tv,i}]^T$ to represent the position of OV. Then, the motion state transition probability of the OV is,

$$P(\xi'_{ov,i}|\mathbf{x}_{ov,i}) = P(\zeta'_{ov,i}|\mathbf{x}_{ov,i}) = \frac{\exp\left\{\frac{1}{2}[(\zeta'_{ov,i} - \hat{\zeta}'_{ov,i})^T \Sigma_e (\zeta'_{ov,i} - \hat{\zeta}'_{ov,i})]\right\}}{2\pi p * |\Sigma_e|^{1/2}} \quad (19a)$$

$$\hat{\zeta}'_{ov,i} - \begin{bmatrix} a_{e_s} \\ b_{e_l} \end{bmatrix} \leq \zeta'_{ov,i} \leq \hat{\zeta}'_{ov,i} + \begin{bmatrix} a_{e_s} \\ b_{e_l} \end{bmatrix} \quad (19b)$$

where $\zeta'_{ov,i}$ is the next position state of the OV, $\hat{\zeta}'_{ov,i}$ is the predicted position state of the OV, and $\xi'_{ov,i}$ is the next motion state which means that the OV transfers the motion state along the predicted trajectory according to the predicted error distribution. Fig. 2. shows the uncertainties of OV according to predicted trajectories under different driving intentions $b_{ov,i}$. This transition model that constructed from prediction error uncertainty enables EV to measure the potential risk of OVs in the future. it also avoids aimless state sampling during the decision solving process, thus reducing the computational time-consuming.

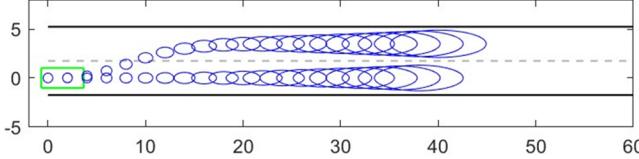


Fig. 2. Trajectory uncertainty under different driving intentions.

E. Observation Space

The observation space O contains the information observed by the environment perception system. It is defined as,

$$O = [o_{ego}, o_1, o_2, \dots, o_i, \dots] \quad (20)$$

where o_{ego} is the observation of EV, which contains the same information as the EV's state $x_{ego,0}$ and $o_i (i \in [1, n])$ is the observation of OVs. The observation model is decomposed for EV and OVs,

$$\Omega = P(o|s', a) = P(o_{ego}|\mathbf{x}'_{ev}) \prod_{i=1}^n P(o_i|\mathbf{x}'_{ov,i}) \quad (21)$$

Since EV is equipped with a high-precision localization system, we assume that the state of EV is completely observable and the observed information is equivalent to the actual state. The observation function of EV is,

$$\Omega_{ego} = P(o_{ego}|\mathbf{x}'_{ev}, a) = \begin{cases} 1 & \text{if } \mathbf{x}'_{ev} \text{ exist} \\ 0 & \text{otherwise} \end{cases} \quad (21a)$$

For OVs, due to the existence of perception errors, there are some differences between the observations and the actual states. We consider the perceived noise of the OVs' longitudinal and lateral positions on the road coordinates, and then define the observation function of the OVs as follows,

$$\Omega_i = P(o_i|\mathbf{x}'_{ov,i}, a) = \{N(\mu_0, \Sigma_0)\} i \in [1, n] \quad (21b)$$

where μ_0 is the observation states of OVs and Σ_0 is the perception error variance matrix.

F. Reward

The reward function is used to judge the value of taking the action. It is one of the cores of POMDP and directly affects the efficiency of the decision-making. In the lane changing scenario, we mainly consider the four aspects of the cost of driving safety, comfort, driving efficiency and continuity of driving actions. The reward function consists of a linear combination of the costs defined above, which is defined as,

$$R = R_{safe} + R_e + R_c + R_a \quad (22)$$

where R_{safe} is the safety reward, R_e is the efficiency reward, R_c is the comfort reward and R_a is the action continue reward.

1) For the safety reward, it is necessary to evaluate the probability of collision between the future state of EV and OVs, and give the corresponding penalties. The security reward function is defined as,

$$R_{safe} = R_{collision} + R_{dis,lon} + R_{dis,lat} \quad (23a)$$

where $R_{collision}$ is the collision reward, and $R_{dis,lon}$ and $R_{dis,lat}$ are the longitudinal and lateral state safety rewards respectively.

In this paper, the method of Monte Carlo sampling is used to approximate the collision probability. Based on the uncertainty distribution of the predicted trajectory of the OV in Sec. III.D, the position of the OV is sampled in this area. U_{bool} is defined to judge whether the future position of EV intersects with the OVs. $U_{bool} = 1$, if the collision happened, otherwise $U_{bool} = 0$. The collision probability of EV can be written as,

$$P_{collision} = \frac{1}{n} \sum_{i=1}^n U_{bool} \in [0, 1] \quad (23b)$$

The collision reward function can be obtained as,

$$R_{collision} = \begin{cases} (1 - P_{collision}) * w_{collision} & \text{if } P_{collision} < P_{safe} \\ -10^6 & \text{other} \end{cases} \quad (23c)$$

where $w_{collision}$ is the corresponding weight and P_{safe} is the safe threshold probability.

The longitudinal safe following distance is calculated by the time to collision (Time To Collision) as,

$$d_{safe,lon} = TTC * (v_{lon,ego} - v_{lead}) + d_{default}, v_{lon,ego} > v_{head} \quad (23d)$$

where v_{lead} is the speed of the front vehicle, $d_{default}$ is a safety threshold.

The longitudinal state safety reward is defined as,

$$R_{dis,lon} = \begin{cases} -1000 & \text{if } d_{lon} < d_{safe,lon} \\ 0 & \text{other} \end{cases} \quad (23e)$$

where d_{lon} is the distance between EV and the front vehicle at the current moment.

Finally, the lateral state safety reward function is defined as,

$$R_{dis,lat} = \begin{cases} w_{dis} (\min(d_{lat}, d_{max}) - d_{min}) & \text{if } d_{lat} > d_{min} \\ -1000 & \text{other} \end{cases} \quad (23f)$$

where d_{min} and d_{max} are the minimum and maximum lateral distances respectively, w_{dis} is the lateral distance weight.

2) For the efficiency reward, it mainly includes three parts: firstly, the driving efficiency of EV during the free lane change process, that is, the speed difference between EV and the speed limit of the target lane $|v_{ego} - v_{target}|$; secondly, the driving efficiency of the target lane, which means the speed difference between the preceding vehicle and the speed limit of the target lane $|v_{lead} - v_{target}|$; finally, the lane-changing revenue, which is the lateral distance deviation between the target lane and the current lane $|l_{ref} - l_i|$.

The efficiency reward function is defined as,

$$R_e = w_e^t |v_{ego} - v_{target}| + w_e^l |v_{lead} - v_{target}| + w_e^r |l_{ref} - l_i| \quad (24)$$

where w_e^t , w_e^l and w_e^r are the corresponding weights.

3) For the comfort reward, the main factor affecting the driving comfort is the change of the speed and acceleration of EV, and it is defined as,

$$R_c = -(\Delta v_{lon,ego}^2 + \Delta a_{lon,ego}^2) * w_c \quad (25)$$

where $\Delta v_{lon,ego} = v'_{lon,ego} - v_{lon,ego}$ and $\Delta a_{lon,ego} = a'_{lon,ego} - a_{lon,ego}$ represent the speed difference and acceleration difference between the next state and the current state. w_c is the corresponding weight

4) In order to improve the consistency of decision results, the action consistency reward is defined as:

$$R_a = \begin{cases} w_a & \text{if } a' = a \\ 0 & \text{otherwise} \end{cases} \quad (26)$$

G. Multi-Policy Forward Simulation Solution

For the traditional solution of POMDP, it starts from the initial belief state $b(s_0)$ and expands discretely in the action space A and observation space O until it exceeds the current decision time domain. Then, the belief tree is constructed in a continuous loop on each belief node, and the Bellman equation is applied to iteratively solve the optimal policy. Accurately solving a POMDP problem is often considered to be computationally infeasible.

To solve this problem, we use the cyclic forward simulation to obtain the future state of EV. For different lateral semantic actions, EV may generate different states. The forward simulation updates the state of EV by solving the control variables, which are divided into longitudinal and lateral control solutions. Among them, the Intelligent Driver Model (IDM) is used to solve the future speed state in longitudinal, and the Model Predictive Control (MPC) based on the simple vehicle kinematics is used to solve the steering angle in the lateral.

After calculating the control input command, the state transfer of EV can be updated in the corresponding belief state under current semantic decision action. The corresponding reward function R is calculated, and then the system solves the next actions a_{lat} and a_{lon} in decision-making step until the decision time domain. After evaluating the all state-action sequences of EV under different policy in the future, the action set with the largest reward is selected as the optimal policy π^* (4)). The pseudo-code of the multi-policy loop forward simulation solution is in follow,

According to the optimal policy π^* , the optimal state sequence x_{ev}^* in the future decision time domain can be obtained.

III. MOTION PLANNING BASED ON MULTI-OBJECTIVE OPTIMIZATION

At the motion planning module, the planning problem is decoupled into lateral and longitudinal multi-objective optimization problems along the reference path. In Section III, the decision results obtained by solving the belief model is the state point set of EV in the Frenét coordinate system. On this basis, the drivable corridors are established to determine the solution space of the motion planning module, which makes

Algorithm 1 Multi-Policy Loop Forward Simulation Solution

Input: EV's current state x_0 , EV's semantic action space A , action of the last moment, Uncertainty distribution of surrounding vehicles' trajectory;

```

1   foreach  $a_{lat,i} \in A_{lat}$  do
2     foreach  $x_{tv,n} \in x_{tv}$  do
3       foreach  $k \in [0, T_d]$  do
4         sample OV's state  $x_{ov,n,k}$  by predicted trajectory  $traj_n$ 
5         compute EV's action  $\{a_{lat,k} a_{lon,k}\}$ 
6         update EV's state  $x_{ev,k}$ 
7         compute reward  $R_k$ 
8         select the best policy  $\pi^* = a_{lat,k} \times a_{lon,k}, [0, T_d]$ 
```

Ouput: action policy π^* and the state sequence x_{ev}^*

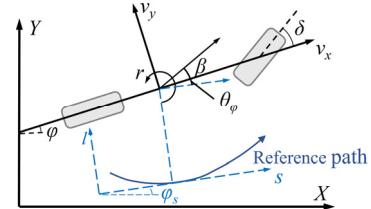


Fig. 3. Vehicle single-track dynamics model.

the decision results and motion planning results have better consistency. Then, the lateral and longitudinal optimization planning problems are constructed respectively based on the vehicle dynamics model. The optimal path control sequence and the optimal speed control sequence are solved respectively, which are merged into the final optimal motion sequence.

A. Lateral Optimization Planner

Lateral planning problems can also be called path planning problems. We construct a convex space for the lateral optimization problem near the initial solution of the POMDP decision and consider the steering characteristics of EV and multiple optimization objectives. The future lateral motion sequence of EV along the reference line is solved and the corresponding state sequence is the lateral path.

1) *Vehicle Dynamics Model in Frenét System:* Considering the two freedom degrees in lateral and yaw motion of the vehicle, the simplified single-track dynamics model of the vehicle is shown in Fig. 3. Here, β is the sideslip angle, r is the yaw rate, δ is the front wheel steering angle, φ is the global heading angle of EV, φ_s is the heading angle of the reference point and their deviation is θ_φ . v_x is the velocity in the heading direction. When the vehicle is running, θ_φ keeps in a small range, so it can be assumed that $\cos(\theta_\varphi) \approx 1$ and $\sin(\theta_\varphi) \approx \theta_\varphi$. Then the vehicle dynamics model can be derived as follows:

$$\begin{aligned} \dot{\theta}_\varphi &= r - v_x k(s) \\ \dot{\beta} &= \frac{-2(C_f + C_r)}{mv_x} \beta + \left(\frac{-2(C_f l_f - C_r l_r)}{mv_x^2} - 1 \right) r + \frac{2C_f}{mv_x} \delta \\ \dot{r} &= \frac{-2(C_f l_f - C_r l_r)}{I_z} \beta + \left(\frac{-2(C_f l_f^2 + C_r l_r^2)}{I_z v_x} \right) r + \frac{2C_f l_f}{I_z} \delta \\ \dot{\delta} &= \tau \end{aligned} \quad (27)$$

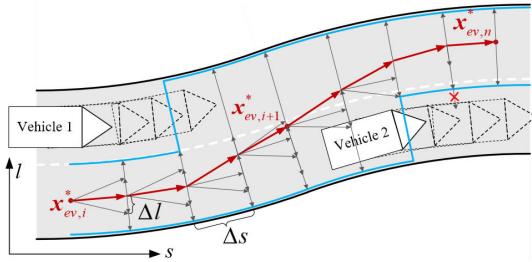


Fig. 4. Sampling process of the drivable corridor.

where m is the mass of the vehicle, C_f and C_r are the cornering stiffness of front and rear tires, I_z is the yaw moment of inertia. The state vector $\mathbf{x}_{lat,k} = [l_k, \theta_{\phi,k}, \beta_k, \gamma_k, \delta_k]^T$ and the control input $\mathbf{u}_{lat,k} = \tau_k$ is the derivative of the steering angle. The system discrete model can be modified as:

$$\begin{aligned} \mathbf{x}_{lat,k+1} &= (\mathbf{I} + \mathbf{A}_{lat,k}\Delta T/2)(\mathbf{I} - \mathbf{A}_{lat,k}\Delta T/2)^{-1}\mathbf{x}_{lat,k} \\ &\quad + \mathbf{B}_{lat,k}\Delta T\mathbf{u}_{lat,k} + \mathbf{C}_{lat,k}\Delta T \\ &= f^{lat}(\mathbf{x}_{lat,k}, \mathbf{u}_{lat,k}, \Delta T) \end{aligned} \quad (28)$$

where

$$\mathbf{A}_{lat,k} = \begin{bmatrix} 0 & v_x & v_x & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{-2(C_f+C_r)}{mv_x} & \frac{-2(C_f l_f - C_r l_r)}{mv_x^2} - 1 & 0 \\ 0 & 0 & \frac{-2(C_f l_f - C_r l_r)}{I_z} & \frac{-2(C_f l_f^2 + C_r l_r^2)}{I_z v_x} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{B}_{lat,k} = [0 \ 0 \ 0 \ 0 \ 1], \mathbf{C}_{lat,k} = [0 \ -v_x k(s) \ 0 \ 0 \ 0]$$

2) *Lateral Drivable Corridor*: After the POMDP decision, a lateral drivable corridor is generated by combining the search and sampling method. Taking the future state points of the POMDP decision result as the target, space sampling is performed between two adjacent state points. The lateral sampling distance is Δl , and the longitudinal sampling distance is Δs . As shown in Fig. 4, node expansion sampling is performed from the previous decision state point $\mathbf{x}_{ev,i}^*$ to the next decision state point $\mathbf{x}_{ev,i+1}^*$. The expansion of the current node must satisfy lateral constraints to avoid unreasonable steering. The lateral constraints between the front and rear nodes are defined as:

$$\theta_k - \Delta\theta \leq \frac{l_{k+1} - l_k}{\Delta s} \leq \theta_k + \Delta\theta \quad (29)$$

where θ_k is the heading angle of the current node.

To find all optimal nodes, the cost of candidate nodes is calculated separately, including priority cost and heuristic cost. The priority cost function consists of four terms, including reference cost $c_{lat,ref}$, obstacle distance cost $c_{lat,obs}$, consistency cost c_c and lateral acceleration cost c_a ,

$$\begin{aligned} c_{lat} &= w_{lat,ref}c_{lat,ref} + w_{lat,obs}c_{lat,obs} \\ &\quad + w_{lat,c}c_{lat,c} + w_{lat,a}c_{lat,a} \end{aligned} \quad (30)$$

where $w_{lat,ref}$, $w_{lat,obs}$, $w_{lat,c}$ and $w_{lat,a}$ are the corresponding weights. To ensure that EV is not far from the original reference line when avoiding obstacles, the reference cost is defined as follows:

$$c_{lat,ref} = (l_i - l_{ref})^2 \quad (30a)$$

TABLE I
CONSTRAINTS IN LATERAL OPTIMIZATION

constraint	application	class
safety constraints	prevent collision and ensure safety.	hard constraints
model constraints	ensure the feasibility and smoothness of the path curve	hard constraints
dynamics constraints	ensure the stability of vehicle	hard constraints
smoothness constraints	ensure the smoothness and comfort, and prevent the optimization solution failure.	soft constraints

We project the predicted trajectory of the OVs on the space map step by step. The distance of each node to the obstacle is calculated, and unsafe candidate nodes are removed. The obstacle collision cost function is designed as follows:

$$c_{lat,obs} = \begin{cases} (d_{obs,max} - d_{obs})^2, & d_{obs} \leq d_{obs,max} \\ 0, & \text{otherwise} \end{cases} \quad (30b)$$

where d_{obs} is the distance from the node to the nearest obstacle, and $d_{obs,max}$ is the safety distance. To ensure the continuity and comfort between two nodes, the consistency cost and lateral acceleration cost are designed as follows:

$$c_{lat,c} = \left(\frac{l_i - l_{i-1}}{\Delta s} \right)^2, c_{lat,a} = \left(\frac{l'_i - l'_{i-1}}{\Delta s} \right)^2 \quad (30c)$$

3) *Constraints*: For lateral optimization, the constraints mainly come from the vehicle motion constraints inside and the environment constraints outside. It mainly includes lateral safety boundary constraints, vehicle stability constraints, kinematics constraints and smoothness constraints. Although the optimization problem can consider these hard constraints, in some extreme cases the hard constraints may cause the optimization problem fail to solve. Hence, a soft constraint on steering speed is introduced in this paper to improve the robustness of the lateral optimization algorithm. All constraints are summarized in the Table I.

According to the searched optimal nodes $(s_{r,k}, l_{r,k}, \theta_{r,k})$ (see the green lines in Fig. 3), the safe path boundaries are sequentially determined from both sides of the optimal node along the reference road, which are shown as blue lines in Fig. 3. The lateral boundary constraint of the k^{th} step is described as follows:

$$\max(l_k - n_{r,k}\Delta l, l_{r,b}) \leq l_k \leq \min(l_k + n_{l,k}\Delta l, l_{l,b}) \quad (31a)$$

The safety boundary constraints are the drivable corridor boundary (31a) and the model constraints are according to (28). In terms of vehicle dynamics stability constraint, the constraint of vehicle stability envelope is added. The envelope of stable vehicle driving is defined by the limits of yaw rate r and sideslip angle β , as shown in Fig. 5. While the vehicle state remains inside this envelope, the vehicle can maintain steady driving.

It is assumed that the vehicle does not produce longitudinal slip during driving, that is, the longitudinal force on the tire is zero. Then the yaw rate limits indicated by ① and ③ in Fig. 4

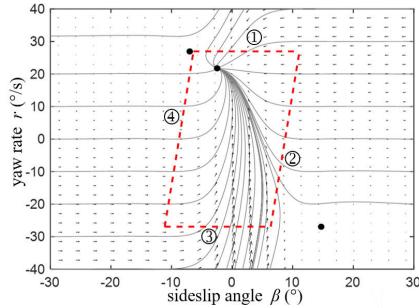


Fig. 5. Schematic diagram of vehicle stability envelope.

can be obtained as:

$$|r_{ss,\text{lim}}| = \frac{\mu g}{v_x} \quad (31\text{b})$$

where μ represents the adhesion coefficient of road surface and g is the gravitational acceleration. According to the rear wheel slip angle of the maximum lateral force in the tire Pacejka model [33], the limit of the sideslip angle β_{ss} shown on ② and ④ in Fig. 4. can be obtained,

$$\beta_{ss} = |\alpha_{r,\text{peak}}| + \frac{l_r r}{v_x}, \alpha_{r,\text{peak}} = \tan^{-1}\left(3 \frac{\mu m g l_f}{C_r (l_f + l_r)}\right) \quad (31\text{c})$$

The other constraints include heading angle constraints, front wheel steering angle and its change rate constraints. They are defined as follows:

$$\begin{aligned} \theta_{\varphi,\text{min}} &\leq \theta_{\varphi,k} \leq \theta_{\varphi,\text{max}} \\ \delta_{\text{min}} &\leq \delta_k \leq \delta_{\text{max}} \\ \tau_{\text{min}} &\leq \tau_k \leq \tau_{\text{max}} \end{aligned} \quad (31\text{d})$$

However, in some emergency situations, the hard constraints on the change rate of the front wheel steering angle can cause the optimization problem to fail. Therefore, this paper introduces a relaxation factor ρ , and then constructs the corresponding soft constraint as follows.:

$$\tau_{\text{min}} - \rho_{\text{lat}} \leq \tau_k \leq \tau_{\text{max}} + \rho_{\text{lat}}, 0 \leq \rho_{\text{lat}} \quad (31\text{e})$$

4) Optimization Objective Function and Solution: The lateral multi-objective optimization is responsible for finding a safe, efficient and smooth path. We integrate multiple objectives of lateral optimization in a single objective function [34] by means of a weight matrix. It can be written as:

$$\begin{aligned} J_{\text{lat}} = \sum_{k=1}^N & (\mathbf{x}_{\text{lat},k} - \mathbf{x}_{\text{lat},r,k})^T \mathbf{Q}_{\text{lat}} (\mathbf{x}_{\text{lat},k} - \mathbf{x}_{\text{lat},r,k}) \\ & + \sum_{k=1}^{N-1} \mathbf{u}_{\text{lat},k}^T \mathbf{R}_{\text{lat}} \mathbf{u}_{\text{lat},k} + \sum_{k=1}^{N-1} \boldsymbol{\rho}_{\text{lat},k}^T \boldsymbol{\lambda}_{\text{lat}} \boldsymbol{\rho}_{\text{lat},k} \end{aligned} \quad (32)$$

$\mathbf{x}_{\text{lat},r,k} = [l_{r,k}, \theta_{\varphi,r,k}, 0, 0, 0]^T$ is the reference state and $l_{r,k}, \theta_{\varphi,r,k}$ is from the optimal node coordinates solved above. $\mathbf{Q}_{\text{lat}} = \text{diag}\{q_L, q_\theta, q_r, q_\beta, q_\delta\}$ and $\mathbf{R}_{\text{lat}} = [q_\tau]$ are the weight matrices of system states and control inputs, respectively. $\boldsymbol{\lambda}_{\text{lat}} = \{\lambda_{\text{lat},k}\}$ is the weight of the corresponding soft constraints.

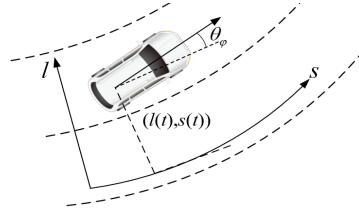


Fig. 6. Longitudinal motion model in Frenet system.

With the definition of the objective function and the constraints, min the final problem formulation is given as,

$$\begin{aligned} \min J_{\text{lat}} &= \min(32) \\ \text{s.t. } \mathbf{x}_{\text{lat},k+1} &= \mathbf{f}^{\text{lat}}(\mathbf{x}_{\text{lat},k}, \mathbf{u}_{\text{lat},k}, \Delta T) \\ \mathbf{x}_{\text{lat},k,\text{min}} &\leq \mathbf{x}_{\text{lat},k} \leq \mathbf{x}_{\text{lat},k,\text{max}} \\ \mathbf{u}_{\text{lat},k,\text{min}} - \boldsymbol{\rho}_{\text{lat},k} &\leq \mathbf{u}_{\text{lat},k} \leq \mathbf{u}_{\text{lat},k,\text{max}} + \boldsymbol{\rho}_{\text{lat},k} \\ 0 &\leq \boldsymbol{\rho}_{\text{lat},k} \\ k &\in [1, N] \end{aligned} \quad (33)$$

with the linear function \mathbf{f}^{lat} is according to (28). $\mathbf{x}_{\text{lat},k,\text{min}}$ and $\mathbf{x}_{\text{lat},k,\text{max}}$ are the limits of the states, $\mathbf{u}_{\text{lat},k,\text{min}}$ and $\mathbf{u}_{\text{lat},k,\text{max}}$ are the limits of the control input. The problem (34) can be rearranged as a QP form to solve.

B. Longitudinal Optimization Planner

The longitudinal speed planning is to generate the corresponding speed curve based on the path planning. In this part, we pay more attention to the influence of the surrounding OVs. The speed optimization problem is converted into a convex optimization problem according to the state results of the POMDP decision, which can generate a speed curve that satisfies the driving safety, comfort and efficiency for autonomous vehicle.

1) Longitudinal Motion Model in Frenet System: In the Frenet coordinate system, the longitudinal motion of the EV is along the reference path. When the road curve has some curvature, the direction of the EV's heading and speed are not the same, and the influence of road curvature needs to be considered. Hence, as shown in Fig.6, the longitudinal motion model is written as:

$$\begin{aligned} \dot{s}(t) &= \frac{v_x(t) \cos \theta_\varphi}{1 - \kappa(s) l(t)} \\ \dot{v}_x(t) &= a_x(t) \\ \dot{a}_x(t) &= j(t) \end{aligned} \quad (34)$$

Defining the longitudinal states $\mathbf{x}_{\text{lon},k} = [s_k, v_{x,k}, a_{x,k}]^T$ and the jerk as control input $\mathbf{u}_{\text{lon},k} = [j_k]$, the discrete state equation of the model (34) under the sampling of time interval ΔT is,

$$\begin{aligned} \mathbf{x}_{\text{lon},k+1} &= \mathbf{A}_{\text{lon},k} \mathbf{x}_{\text{lon},k} + \mathbf{B}_{\text{lon},k} \mathbf{u}_{\text{lon},k} \\ &= \mathbf{f}^{\text{lon}}(\mathbf{x}_{\text{lon},k}, \mathbf{u}_{\text{lon},k}, \Delta T) \end{aligned} \quad (35)$$

where

$$\mathbf{A}_{\text{lon},k} = \begin{bmatrix} 1 & \frac{v_x(t) \cos \theta_\varphi \Delta T}{1 - \kappa_k l} & 0 \\ 0 & 1 & \Delta T \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{B}_{\text{lon},k} = [0 \ 0 \ \Delta T]$$

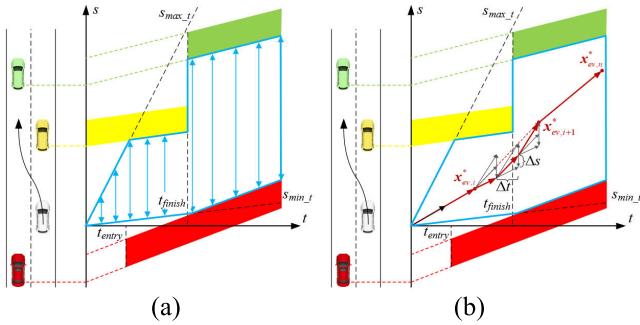


Fig. 7. Longitudinal motion model in Frenet system.

2) *Longitudinal Drivable Corridor*: According to the state results of the POMDP decision, in the ST-graph, a longitudinal drivable corridor is formed by expanding up and down for collision detection in each time interval ΔT . As shown in Fig. 7(a), there is an overlap between the driving path of EV and OVs. While, the collision can be avoided as long as their future spatial positions do not overlap at the same instant.

The OVs are classified as a following target or an overtaking target. For the following target, the longitudinal position of EV should be behind it when the paths overlap, and the corresponding safety boundary constraint is:

$$s_{ev}(t) + \frac{v_{ev}(t)^2}{2a_{ev}^{dec}} + s_{safe} - s_{ov,i}(t) - \frac{v_{ov,i}(t)^2}{2a_{ov,i}^{dec}} \leq 0$$

$$t_{u,s} \leq t \leq t_{u,e} \quad (36a)$$

where $t_{u,s}$ is the starting time when the OV and the EV are in the same lane, and $t_{u,e}$ is the ending time. For the overtaking target, the longitudinal position of EV should be in front of it when the paths overlap, and the corresponding safety boundary is:

$$s_{ev}(t) + \frac{v_{ev}(t)^2}{2a_{ev}^{dec}} - s_{safe} - s_{ov,i}(t) - \frac{v_{ov,i}(t)^2}{2a_{ov,i}^{dec}} \geq 0$$

$$t_{l,s} \leq t \leq t_{l,e} \quad (36b)$$

In the decision results of POMDP, when the lateral position exceeds the current lane, it is considered that EV starts to change lane. According to that, the entry time t_{entry} of lane change and the finish time t_{finish} can be calculated. Like the lateral drivable corridor, the state points $x_{ev,i}^*$ of the POMDP decision are projected into the ST-graph. When the time is in $[0, t_{entry}]$, EV is in the current lane, and the yellow OV in front is the following target. In $[t_{entry}, t_{finish}]$, EV is in the process of lane changing while it is necessary to consider both yellow OV (following target) and the red OV (overtaking target). The EV has completely entered the target lane after t_{finish} , and it needs to avoid the green OV (following target) in front.

In addition, the maximum and minimum speed limits should also be taken into account as,

$$s_{min_t} = v_{min}t \leq s(t) \leq v_{max}t = s_{max_t} \quad (36c)$$

The longitudinal nodes are sampled and expanded between adjacent state nodes until the above boundaries (36) are exceeded, so as to obtain the longitudinal position constraints as:

$$s_{lb,k} \leq s_k \leq s_{ub,k} \quad (37)$$

TABLE II
CONSTRAINTS IN LONGITUDINAL OPTIMIZATION

constraint	application	class
safety constraints	prevent collision and satisfy the road speed limits	hard constraints
model constraint	ensure the feasibility of the speed curve	hard constraints
dynamics constraints	ensure the stability of vehicle	hard constraints
smoothness constraints	ensure the smoothness and comfort, and prevent the optimization solution failure.	soft constraints

Then, as shown in Fig. 7(b), the optimal nodes $(s_{r,k}, t_k)$ are searched with the reference cost $c_{lon,ref}$, obstacle collision cost $c_{lon,obs}$ and the comfort cost $c_{lon,v}$. where $w_{lon,ref}$, $w_{lon,obs}$, and $w_{lon,v}$ are the corresponding weights,

$$c_{lon} = w_{lon,ref} c_{lon,ref} + w_{lon,obs} C_{lon,obs} + w_{lon,v} C_{lon,v} \quad (38a)$$

Unlike the reference cost in lateral drivable corridor, $c_{lon,ref}$ represents the relationship between position and time. The connection line between two adjacent state nodes directly represents the search tendency, and the distance difference between the nodes in the longitudinal direction is used as the reference cost:

$$c_{lon,ref} = (s_i - s_{ref})^2 = \Delta s_{ref}^2 \quad (38b)$$

the obstacle collision cost $c_{lon,obs}$ and the comfort cost $c_{lon,v}$ are given as:

$$c_{lon,obs} = \begin{cases} (s_{obs} - s_{obs,max})^2, & s_{obs} \leq s_{obs,max} \\ 0, & \text{otherwise} \end{cases} \quad (38c)$$

$$c_{lon,v} = \left(\frac{s_i - s_{i+1}}{\Delta T} - \frac{s_{i+1} - s_{i+2}}{\Delta T} \right)^2 \quad (38d)$$

where s_{obs} represents the distance between the node and the nearest obstacle along the s direction, and the $s_{obs,max}$ is the threshold distance. The longitudinal reference state sequence $x_{lon,r,k} = (s_{r,k}, v_{r,k}, a_{r,k})$ can be obtained by differential calculation according to the searched optimal nodes result.

3) *Constraints*: The longitudinal non-convex optimization problem is transformed into a convex optimization problem, which limits the solution space and improves the optimization solution efficiency. All constraints are shown in the Table. II

The safety constraints are the boundary of the longitudinal drivable corridor (42). As for the dynamics constraints, the longitudinal motion planning needs to consider the limit of the friction ellipse. The corresponding acceleration constraints are:

$$\left(\frac{a_k}{m} \right)^2 + \left(\frac{v_k^2 \tan \delta_k}{nL} \right)^2 \leq 1, \quad \frac{v_k^2 \tan \delta_k}{L} = a_{y,k} \in [-0.4, 0.4] \quad (39)$$

The smoothness constraints are constructed with jerk. However, in some emergency situations, comfort may not be guaranteed while timely braking is required to ensure driving

safety. The relaxation factor $\rho_{lon,k}$ is introduced and the longitudinal soft constraints are constructed as follows:

$$j_{\min} - \rho_{lon,k} \leq j_k \leq j_{\max} + \rho_{lon,k}, 0 \leq \rho_{lon,k} \quad (40)$$

4) *Objective Function and Problem Solution:* The objective function of the longitudinal planning is,

$$\begin{aligned} J_{lon} = & \sum_{k=1}^N (\mathbf{x}_{lon,k} - \mathbf{x}_{lon,r,k})^T \mathbf{Q}_{lon} (\mathbf{x}_{lon,k} - \mathbf{x}_{lon,r,k}) \\ & + \sum_{k=1}^{N-1} \boldsymbol{\rho}_{lon,k}^T \boldsymbol{\lambda}_{lon} \boldsymbol{\rho}_{lon,k} + \sum_{k=1}^{N-1} \mathbf{u}_{lon,k}^T \mathbf{R}_{lon} \mathbf{u}_{lon,k} \end{aligned} \quad (41)$$

where $\mathbf{Q} = \text{diag}\{q_s, q_v, q_a\}$ and $\mathbf{R} = \{q_j\}$ are the weight matrices of system states and control inputs, respectively. $\lambda_{lon} = \{\lambda_{lon,k}\}$ is the weight of the corresponding soft constraints. With the model constraint is given according to f^{lon} (39), the final longitudinal problem is given as,

$$\begin{aligned} \min J_{lon} = & \min(41) \\ \text{s.t. } & \mathbf{x}_{lon,k+1} = f^{lon}(\mathbf{x}_{lon,k}, \mathbf{u}_{lon,k}, \Delta T) \\ & \mathbf{x}_{lon,k,\min} \leq \mathbf{x}_{lon,k} \leq \mathbf{x}_{lon,k,\max} \\ & \mathbf{u}_{lon,k,\min} - \rho_{lon,k} \leq \mathbf{u}_{lon,k} \leq \mathbf{u}_{lon,k,\max} + \rho_{lon,k} \\ & 0 \leq \rho_{lon} \\ & k \in [1, N] \end{aligned} \quad (42)$$

Like the lateral planning optimization problem, (42) can be also transformed into a QP form to solve by OSQP [35].

C. Optimal Motion Sequence Output

After solving the horizontal planning problem and the vertical planning problem, the optimal state sequence is taken respectively, and a safe, comfortable and stable optimal motion sequence in the Frenét coordinate system is synthesized as the final output trajectory.

IV. SIMULATION AND EXPERIMENT RESULTS

To evaluate the proposed integrated algorithm, we designed simulations and real vehicle tests. We configure the same prediction module and motion control module to verify the effectiveness of our algorithm. Table. III shows the main algorithm parameters.

A. Simulation Analysis

The simulation platform is the Virtual Test Drive (VTD), which is a complete modular simulation tool with vehicle dynamics for autonomous driving and could provide road environment and dynamic scenarios. Simulation tests were conducted in the C++ 11/Ubuntu 18.04 with ZCM messaging mechanism. There are two different simulation scenarios: 1) Cut in scenario: EV is located on the right lane, and OV cut in from the left lane at different speeds. 2) Overtaking scenario: there are three OV's driving around the EV, EV should change lane and overtakes at the right time. To better evaluate the performance of our method, we selected similar prior work using POMDP and optimal approach [36] as a baseline for comparison.

TABLE III
GENERAL ALGORITHM PARAMETERS

scalars		matrices
$l_f = 1.4$	$l_r = 1.3$	$w_a = 5$
$C_f = 133.8 \times 10^3$	$w_{lat,ref} = 35$	$\mathbf{K} = \begin{bmatrix} 0.13 & -0.08 & 0 & 0 \\ 0 & 0 & 0.07 & -0.03 \end{bmatrix}$
$C_r = 125.4 \times 10^3$	$w_{lat,obs} = 45$	$\Sigma_0 = \text{diag}(0.44, 0, 0.09, 0)$
$I_z = 3716$	$w_{lat,c} = 20$	$\mathbf{Q}_{lat} = \text{diag}(10, 0, 10, 25, 10)$
$r_p = 0.9$	$w_{lat,d} = 10$	$\mathbf{R}_{lat} = \{2\}$
$w_{collision} = 50$	$w_{lon,obs} = 40$	$\mathbf{Q}_{lon} = \text{diag}(1, 20, 5)$
$w_{dis} = 20$	$w_{lon,obs} = 40$	$\mathbf{R}_{lon} = \{1\}$
$w_e^t = 40$	$w_{lon,obs} = 20$	$\mathbf{u}_{lat,max} = \{0.5\}$
$w_e^l = 30$	$\Delta T_d = 5$	$\mathbf{u}_{lat,min} = \{-0.5\}$
$w_e^r = 10$	$T_d = 0.5$	$\mathbf{u}_{lon,max} = \{5\}$
$w_c = 5$	$\Delta T = 0.1$	$\mathbf{u}_{lon,min} = \{-10\}$

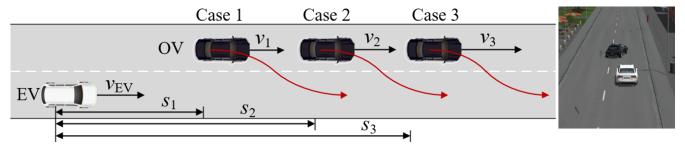


Fig. 8. Cut-in scenario with three initial state of OV ($s_1 = 30m$, $v_1 = 45km/h$, $s_2 = 40m$, $v_2 = 40km/h$, $s_3 = 50m$, $v_3 = 35km/h$).

1) *Cut-in Scenario:* In this scenario, EV is driving in the current lane in 50km/h, OV cuts in after 2 seconds and the lane-changing time is 3s. We set different initial position and speed of OV to verify our approach. We set OV's different longitudinal distances $s_{1\sim 3}$ from EV and their initial velocities $v_{1\sim 3}$ in the 3cases (in Fig. 8). In all cases, the OV cuts into the current lane from the left lane within 2-5 seconds. Thus, the longitudinal relative distance between EV and OV decreases rapidly, and EV decelerates quickly and smoothly to maintain a safe distance. After that, OV ends the cut-in behavior and drive at a constant speed in the current lane, therefore EV increases the speed appropriately to follow OV.

Fig. 9 record the EV's driving results of two approaches. The overall speed and acceleration of our approach are flat, and the trajectory with no oscillation. Compared to baseline, the longitudinal relative distance between EV and OV using our approach has smaller oscillations and it can reduce the occurrence of large braking when OV invades the current lane. In contrast, the baseline produced larger velocity and acceleration oscillations at closer longitudinal distances to the OV. Besides, in the three cases, the minimum TTC in the process of our approach is 11.07s, 9.64s and 7.52s respectively. More statistics results are listed in TABLE IV. As depicted in the table, our proposed approach can improve the minimum relative distance to the OV by more than 30% compared to baseline in all cases. Meanwhile, there is no decrease in the speed of the vehicle, which improves the minimum speed and reduces the maximum deceleration. In particular, we analyze the standard deviation of longitudinal relative distance $\sigma(d)$, velocity $\sigma(v)$ and acceleration $\sigma(a)$ to illustrate that our method enhances the oscillation-free capability of the generated trajectories. As compared to baseline in cases 1 and 2,

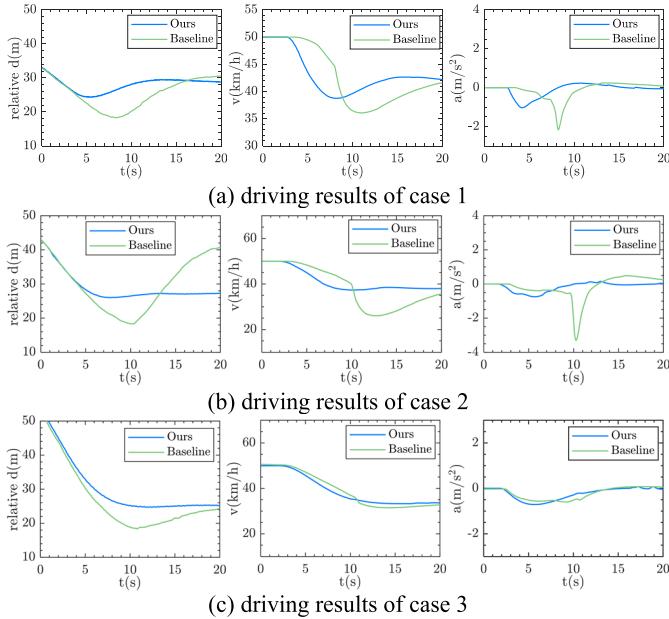


Fig. 9. EV's driving results of all cases under cut-in scenario: In each subfigure, the results from left to right are the longitudinal relative distance between EV and the OV in front, the velocity of EV, and the acceleration of EV.

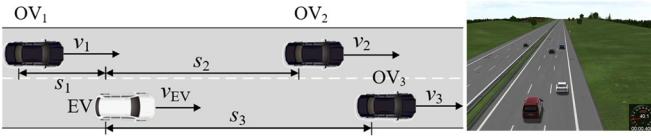


Fig. 10. The overtaking scenario with 3 OVs in different initial states.

where the OV cut in closer to the EV, the trajectories traveled by our method reduce the $\sigma(a)$ by 27.4% and 62.1%, and the $\sigma(v)$ by 32.6% and 48.2%. These results above show that our approach could make quick decision reaction to smoothly decelerate to ensure safety.

2) *Overtaking Scenario*: We designed two overtaking case with three OV₁~₃ around EV as shown in Fig. 10. OV₁ and OV₂ are in the left lane, and OV₃ is in the front of current lane. In case 1, the initial speed of EV is 30km/h while the target speed 40km/h. The initial state of OV₁ is ($s_1 = 5$ m, $v_1 = 15$ km/h) and it will accelerate to 25km/h. The initial state of OV₂ is ($s_2 = 30$ m, $v_2 = 25$ km/h) and it will accelerate to 35km/h. The initial state of OV₃ is ($s_3 = 40$ m, $v_3 = 40$ km/h) and it will decelerate to 20km/h.

The calculation results of our approach at different moments are illustrated in Fig. 11. In Fig. 11(a), the OV₃ is slowing down and the front vehicle in left lane (OV₂) has a higher speed, EV will try to change lanes left to get a higher driving efficiency, while overtake the vehicle behind the left lane (OV₁) to ensure safety. The lane change maneuver is finished at the moment in Fig. 11 (b), and then EV will follow the OV₂ and overtake the OV₃. In Fig. 11 (c), EV find there is enough driving space in the right lane and perform a lane change action again and finally complete the whole process of overtaking safely.

The state change during the whole process is shown in Fig. 12.

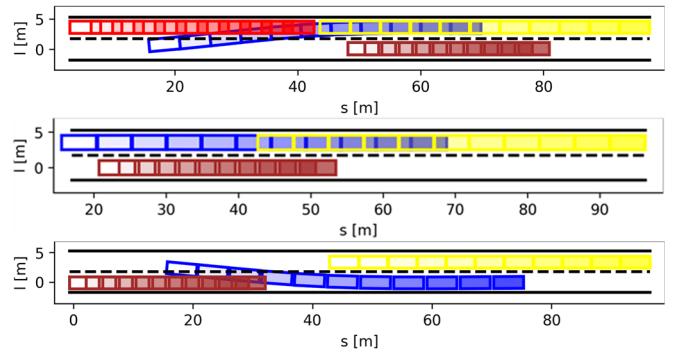


Fig. 11. The blue squares are the results of the planned trajectory of EV, and the rest indicate the predicted trajectories of the OVs (OV₁ is red, OV₂ is yellow, and OV₃ is brown). The time interval of each square is 0.5s. (a), (b) and (c) are correspond to the moments at 3.8s, 9.2s, 14.5s, respectively.

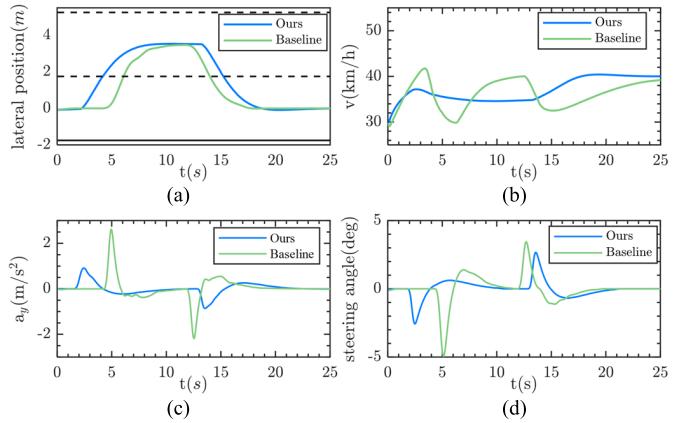


Fig. 12. Driving results of overtaking case 1: (a) lateral position (b) longitudinal speed, (c) lateral acceleration, (d)steering angle.

In Fig. 12(a), the lateral position w.r.t. the reference path of the whole overtaking process of both approach change with no over-shooting, and the average lane change time is around 5s. Other indicators including velocity, lateral acceleration, and steering angle are shown in the (b)-(d). The baseline performs more conservatively, making overtakes only after pulling away enough from OV₁, and its lane-changing driving maneuvers have a shorter time duration, which create larger state oscillations. As a comparison, our approach makes lane-change overtaking behavior earlier and the changes in speed, acceleration, and steering angle are smoother. It completes the overtaking target smoothly with less driving jitter. The absolute value of the maximum lateral acceleration is 0.86m/s², and the maximum steering angle is 2.59deg. All of the indications are within a reasonable range, which illustrates the strength of the integrated decision making-motion planning framework.

In order to further verify the impact of the OV behind during the lane-change overtaking process, the target speed of OV₁ is set to 35km/h in case 2. At the moment in Fig. 13(a), EV does not have enough safe distance for lane change because of the OV₁ is accelerating, and then EV decides to slow down to follow the vehicle in front (OV₃). At the moment in Fig. 13(b), the OV₁ has overtaken EV and its speed is greater than OV₃. In order to obtain a higher driving efficiency, EV decides to switch lanes to the left at this time. Finally, in Fig. 13(c), EV returns to the right lane to continue driving when there is ample safety space.

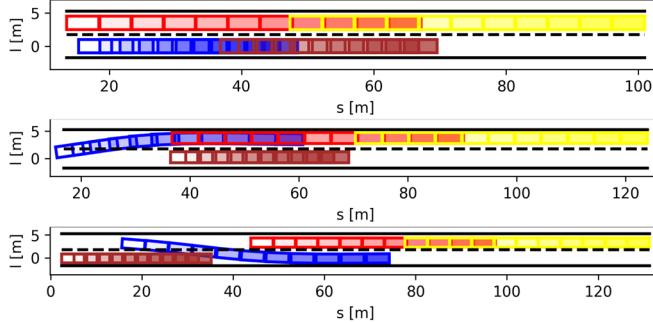


Fig. 13. The planned trajectory of EV, and the predicted trajectories of the OVs. (a), (b) and (c) are correspond to the moments 7.2s, 11.9s, and 22.5s, respectively.

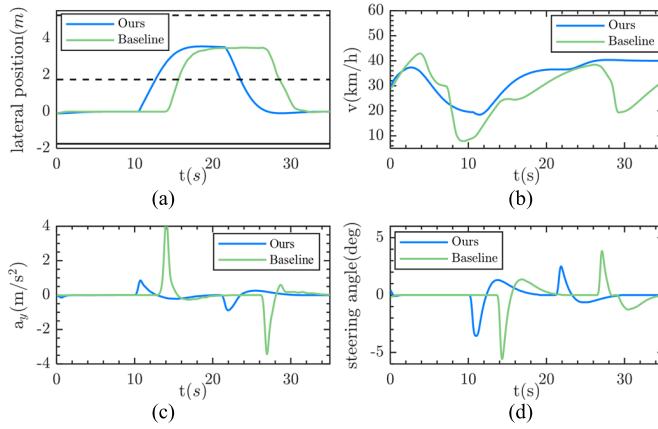


Fig. 14. Driving results of over-taking case 1: (a) lateral position (b) longitudinal speed, (c) lateral acceleration, (d)steering angle.

The state change during the process is shown in Fig. 14. In Fig. 14(a), our approach starts to change lanes at about 10 seconds because it has been following the slowing vehicle ahead (OV₃) for some time. The longitudinal speed of EV reduces firstly to maintain the safe distance, and then increases to reach the target speed in Fig. 14(b). Similar to case1, baseline requires a more conservative driving space to generate lane-change overtaking decisions. From Fig. 14(b)-(d), it can be seen that the longitudinal speed and lateral steering are not stabilized when interacting with other OVs during the overtaking process. More statistics results are list in TABLE V. It mainly represents the performance of our method and baseline in terms of average speed \bar{v} , lateral acceleration a_y and steering angle δ . It can be clearly found that our approach demonstrates excellent smoothness during dynamic overtaking, which is significantly improved compared to baseline in these 3 indicators. More importantly, the standard deviations of mean speed, lateral acceleration and steering angle are greatly reduced. It indicates that our method makes EVs faster to overtake while the driving maneuver is more stable, which is also in keeping with our expected results.

B. Real Vehicle Experiments

The test vehicle platform, as shown as Fig. 15, is a modified electric vehicle from BYD Qin-Pro, equipped with a high precision positioning system included GPS and INS. The real time update rate of GPS/INS system is 100 Hz and

TABLE IV
COMPARISON RESULTS OF ALL CASES IN CUT-IN SCENARIO

		$d_{\min}(\text{m})$	$\sigma(d)$	$v_{\min} (\text{km/h})$	$\sigma(v)$	$a_{\max} (\text{m/s}^2)$	$\sigma(a)$
Case1	our	24.30	2.01	38.76	3.60	-1.03	0.33
	baseline	18.31	4.35	36.08	5.34	-2.16	0.46
Case2	our	26.02	4.19	37.34	4.69	-0.75	0.26
	baseline	18.31	7.63	26.05	9.05	-3.30	0.69
Case3	our	24.70	8.10	33.27	6.52	-0.60	0.26
	baseline	18.42	9.31	31.48	7.50	-0.71	0.26

TABLE V
COMPARISON RESULTS OF OVERTAKING SCENARIO

		\bar{v} (km/h)	$\sigma(\bar{v})$	$a_{\max} (\text{m/s}^2)$	$\sigma(a_{\max})$	δ (deg)	$\sigma(\delta)$
Case1	our	36.99	2.49	0.91	0.26	2.66	0.66
	baseline	36.09	3.12	2.62	0.50	4.95	1.01
Case2	our	33.16	7.15	0.89	0.22	3.56	0.77
	baseline	27.93	8.98	4.51	0.69	5.57	0.98

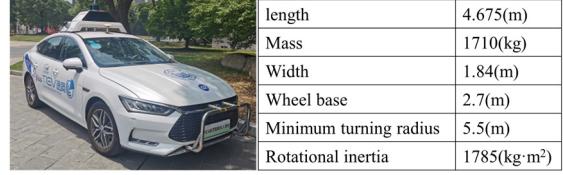


Fig. 15. The test vehicle platform and the major parameters.

the accuracy of position is on the centimeter scale. The surrounding obstacle perception system is generated by the fusion of multiple LiDARs (including one Robosense128-beam LiDAR and four Short-range blind spot LiDAR), which could also give the multimodal prediction results by the algorithm of our research group [37], [38]. Our algorithm runs on the industrial computer (i.e. Intel(R) Core(TM) i9-9900 CPU @ 3.10GHz). The proposed framework makes decisions and planning receiving the localization, perception, prediction and the vehicle state messages and then sending the generated real-time trajectory to the control module. The motion control module (including our previous NMPC trajectory tracking approach [39]) manipulates the vehicle by outputting acceleration and steering wheel angle to the chassis. The different software modules also use the ZCM messaging mechanism. The control module and chassis communicate with each other using the CANbus. The major parameters of the test vehicle are shown in TABLE IV.

As shown in Fig. 16(a), we constructed an overtaking experimental scenario with double OVs. The blue box indicates EV, the red box indicates OV₁ behind in the left lane with 10km/h and the brown box indicates OV₂ in the front with 5km/h speed. EV accelerates from 0km/h to desired speed 22km/h and overtakes two OVs, and the whole overtaking process is shown in Fig. 16. and Fig. 17. are the travel trajectories of EV and OVs during overtaking process in UTM coordinate system. The EV with blue trajectory first accelerates and opens the longitudinal gap with OV₁ while approaching OV₂, and then crosses through the gap between two OVs (brown trajectory

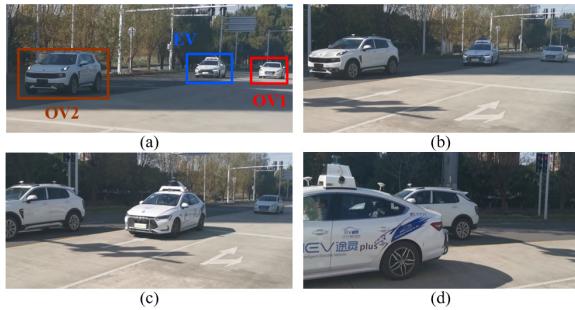


Fig. 16. The real vehicle overtaking test process.

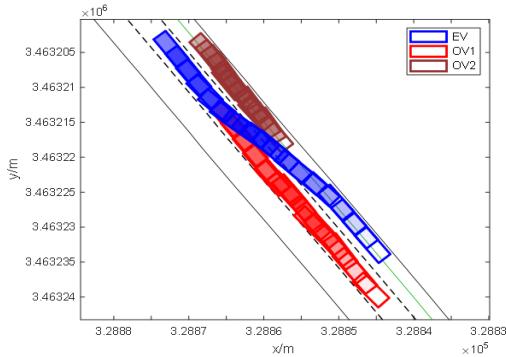


Fig. 17. Travel trajectories of EV and OVs during overtaking process in UTM coordinate system.

TABLE VI
REAL DRIVING RESULTS OF EXPERIMENTAL SCENARIO

$\min l_1(\text{m})$	$\bar{v} (\text{km/h})$	$\sigma(\bar{v})$	$a_{\text{max}} (\text{m/s}^2)$	$\sigma(a_{\text{max}})$
3.15	20.59	4.62	1.49	0.50
$\min l_2(\text{m})$	$a_{\text{max}} (\text{m/s}^2)$	$\sigma(a_{\text{max}})$	$\delta (\text{deg})$	$\sigma(\delta)$
1.01	1.12	0.33	3.87	1.12

of the front OV₁ and red trajectory of the left rear OV₂) at the right time, and finally, completes lane change overtaking.

The states change of EV are shown in Fig. 18. The planning results of EV are continuous and smooth during driving process. The EV accelerates to 22km/h gently, the maximum absolute value of the longitudinal acceleration is less than 2m/s² and that of the lateral acceleration is 1m/s². The maximum transverse angular velocity is 0.02rad/s. Also, the wheel angle changes with a maximum of 5 degrees. There are some statistics data for the entire trajectory in Table VI. According to the TABLE VI, the l_2 (minimum longitudinal distance from OV₂) is 1.01 meters, and l_1 (the longitudinal distance from OV₁) is 3.15 meters. This reflects that EV fully considers the different speeds of the two OVs, and finds the correct time to overtake while ensuring safety. In addition, the EV travels with small oscillations in all indicators. All of them are within a reasonable range, which make the overtaking process comfortable.

C. Calculation Time

More importantly, the proposed approach has nice real-time performance. We take real vehicle test as an example, the calculation time of each module is shown in Fig. 19. It can be

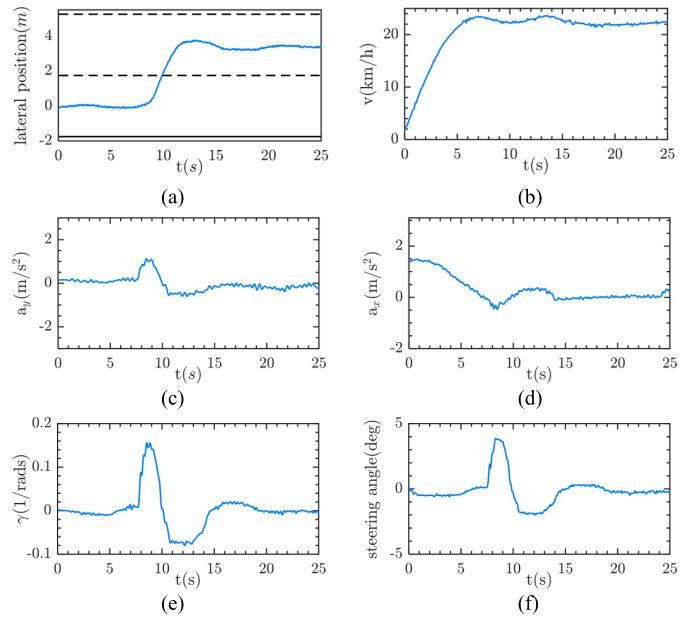


Fig. 18. Results of over-taking real test: (a) lateral position, (b) longitudinal speed, (c) lateral acceleration, (d) longitudinal acceleration, (e) yaw rate, (f) steering angle.

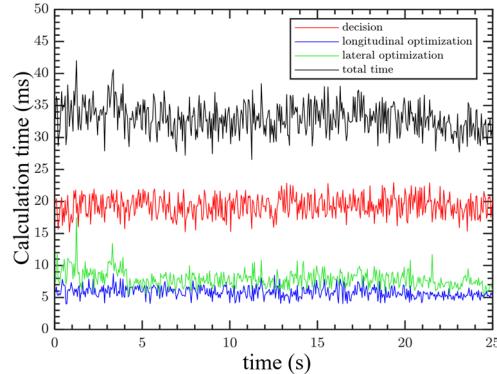


Fig. 19. Calculation time of each algorithm module: decision making, motion planning, and total calculation time.

seen that the computational time is mainly consumed in the belief decision module, but the average solution time is also only 19.31ms (red curve), and the average solution time of motion planning in lateral and longitudinal are 7.89ms (green curve) and 5.81ms (blue curve). The whole algorithm takes less than 45ms with the average of 33.01ms (black curve). Contrastingly, most approaches for lane-changing maneuvers that use POMDP to account for uncertainty (e.g. baseline) is typically 100ms-level or more, which is why these methods are rarely used in the real world at present.

D. Results Discussion

Through results analysis, it can be concluded that our approach can automatically generate proper, continuous and smooth maneuvers in a dynamic environment. The baseline method uses discrete lateral and longitudinal speeds as decision actions. Such reference action has some discrete jumps for motion planning, and due to the lack of consideration of vehicle dynamics, it generates more jittery driving behaviors in overtaking scenarios, including more drastically varying speeds and steering angle. Instead of common decision

methods like baseline, we do not directly use the discrete semantic behavior action, but use IDM and kinematic MPC to simulate and sample the behavioral actions to solve the continuous future reference position state and acceleration. On this basis, we carry out the construction of lateral and longitudinal drivable corridors, and further decide some boundary constraints required for planning. By this means, it provides better suited information for planning and reduces the dissonance between the decision-making and planning modules, which makes it possible to solve the motion sequence more quickly and stably, as can be seen from the simulation results in both two scenarios. Specially, in the cut-in scenario, this difference has a more significant impact on vehicle driving. Discontinuous decision results within a short time window also leads to larger speed and acceleration oscillations for baseline method.

It is also worth mentioning that in this paper, by constructing an iterative model of prediction error in the decision model, we avoid a large number of state sampling solutions while measuring the environmental uncertainty. Meanwhile, the drivable corridor clarifies the reference and boundary of motion planning, and thus the optimization problem is solved very fast even if vehicle dynamics are considered. It infers that the proposed integrated framework cannot only enhance the driving maneuver, but also has strong real-time performance, which can be satisfied with the current computation time requirements of decision making and motion planning in the industrial field.

V. CONCLUSION

This research presents an integrated framework of decision making and motion planning with oscillation-free capability. The proposed approach overcomes the shortcomings of autonomous driving for lane change/keeping maneuvers and is able to: i) make oscillation-free behavior decisions given biased prediction; ii) cut through in the traffic efficiently and safely when being in squeezed; iii) accelerate computation efficiency by building a state transfer model based on prediction uncertainty; iv) reduce the dissonance between decision-making and motion planning. The belief decision model is designed with the uncertainty of the predicted trajectories. It not only outputs continuous decision states, but also avoids unneeded extensive sampling. Based on the decision results, the drivable corridors are constructed to provide the reference state and the related boundary for planning. It leads to a better understanding of the decision information and thus to a faster and more stable solution of optimal trajectory. Comprehensive simulation analysis and experiments were conducted to evaluate the performance of the proposed framework. The results demonstrated that our approach can effectively measure the uncertain risk of the environment and output continuous and stable driving trajectory. It avoids some unnecessary oscillations in autonomous driving. In summary, the proposed framework can safely guide the vehicle at complex driving conditions while driving at its desired velocity as much as possible. In addition, it has also low computation time. This also demonstrates the potential of our approach for practical applications in decision planning modules for autonomous driving.

It should be noted that the proposed method has some requirements on the performance of the prediction system. And in fact, our approach ignores some interaction characteristics between vehicles. Future work will consider the impact of social preference of surrounding vehicles. More importantly, some conditions like crossroads will be investigated to expand the universality of the proposed approach.

REFERENCES

- [1] D. González, J. Pérez, V. Milanés, and F. Nashashibi, "A review of motion planning techniques for automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 4, pp. 1135–1145, Apr. 2016.
- [2] P. Hang, C. Lv, Y. Xing, C. Huang, and Z. Hu, "Human-like decision making for autonomous driving: A noncooperative game theoretic approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 4, pp. 2076–2087, Apr. 2021.
- [3] L. Claussmann, M. Revilloud, D. Gruyer, and S. Glaser, "A review of motion planning for highway autonomous driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 5, pp. 1826–1848, May 2020.
- [4] G. Li et al., "Risk assessment based collision avoidance decision-making for autonomous vehicles in multi-scenarios," *Transp. Res. C, Emerg. Technol.*, vol. 122, Jan. 2021, Art. no. 102820.
- [5] C. Hubmann, J. Schulz, M. Becker, D. Althoff, and C. Stiller, "Automated driving in uncertain environments: Planning with interaction and uncertain maneuver prediction," *IEEE Trans. Intell. Vehicles*, vol. 3, no. 1, pp. 5–17, Mar. 2018.
- [6] C. Hubmann, N. Quetschlich, J. Schulz, J. Bernhard, D. Althoff, and C. Stiller, "A POMDP maneuver planner for occlusions in urban scenarios," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 2172–2179.
- [7] Y. Rahmati, M. K. Hosseini, and A. Talebpour, "Helping automated vehicles with left-turn maneuvers: A game theory-based decision framework for conflicting maneuvers at intersections," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 11877–11890, Aug. 2022.
- [8] H. Yu, H. E. Tseng, and R. Langari, "A human-like game theory-based controller for automatic lane changing," *Transp. Res. C, Emerg. Technol.*, vol. 88, pp. 140–158, Mar. 2018.
- [9] C. Wei, Y. He, H. Tian, and Y. Lv, "Game theoretic merging behavior control for autonomous vehicle at highway on-ramp," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 21127–21136, Nov. 2022.
- [10] M. Wang, S. P. Hoogendoorn, W. Daamen, B. van Arem, and R. Happee, "Game theoretic approach for predictive lane-changing and car-following control," *Transp. Res. C, Emerg. Technol.*, vol. 58, pp. 73–92, Sep. 2015.
- [11] M. Bahram, A. Lawitzky, J. Friedrichs, M. Aeberhard, and D. Wollherr, "A game-theoretic approach to replanning-aware interactive scene prediction and planning," *IEEE Trans. Veh. Technol.*, vol. 65, no. 6, pp. 3981–3992, Jun. 2016.
- [12] N. Li, D. W. Oyler, M. Zhang, Y. Yildiz, I. Kolmanovsky, and A. R. Girard, "Game theoretic modeling of driver and vehicle interactions for verification and validation of autonomous vehicle control systems," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 5, pp. 1782–1797, Sep. 2018.
- [13] A. G. Cunningham, E. Galceran, R. M. Eustice, and E. Olson, "MPDM: Multipolicy decision-making in dynamic, uncertain environments for autonomous driving," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2015, pp. 1670–1677.
- [14] E. Galceran, A. G. Cunningham, R. M. Eustice, and E. Olson, "Multipolicy decision-making for autonomous driving via changepoint-based behavior prediction: Theory and experiment," *Auto. Robots*, vol. 41, no. 6, pp. 1367–1382, Aug. 2017.
- [15] C. Hubmann, M. Becker, D. Althoff, D. Lenz, and C. Stiller, "Decision making for autonomous driving considering interaction and uncertain prediction of surrounding vehicles," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2017, pp. 1671–1678.
- [16] L. Li, W. Zhao, and C. Wang, "POMDP motion planning algorithm based on multi-modal driving intention," *IEEE Trans. Intell. Vehicles*, vol. 8, no. 2, pp. 1777–1786, Feb. 2023.
- [17] Y. Liang, Y. Li, A. Khajepour, Y. Huang, Y. Qin, and L. Zheng, "A novel combined decision and control scheme for autonomous vehicle in structured road based on adaptive model predictive control," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 16083–16097, Sep. 2022.

- [18] F. Damerow and J. Eggert, "Risk-aversive behavior planning under multiple situations with uncertainty," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Sep. 2015, pp. 656–663.
- [19] G. Xiong, Z. Kang, H. Li, W. Song, Y. Jin, and J. Gong, "Decision-making of lane change behavior based on RCS for automated vehicles in the real environment," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2018, pp. 1400–1405.
- [20] P. Hang, C. Lv, C. Huang, J. Cai, Z. Hu, and Y. Xing, "An integrated framework of decision making and motion planning for autonomous vehicles considering social behaviors," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 14458–14469, Dec. 2020.
- [21] T. Zhang, W. Song, M. Fu, Y. Yang, X. Tian, and M. Wang, "A unified framework integrating decision making and trajectory planning based on spatio-temporal voxels for highway autonomous driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 10365–10379, Aug. 2022.
- [22] Y. Zhang, B. Gao, L. Guo, H. Guo, and M. Cui, "A novel trajectory planning method for automated vehicles under parameter decision framework," *IEEE Access*, vol. 7, pp. 88264–88274, 2019.
- [23] H. Chae and K. Yi, "Virtual target-based overtaking decision, motion planning, and control of autonomous vehicles," *IEEE Access*, vol. 8, pp. 51363–51376, 2020.
- [24] Z. Cui, J. Hu, and H. Guan, "A lane-changing trajectory planning and assistant decision-making method for autonomous vehicle," in *Proc. CICTP*, Jul. 2018, pp. 87–101.
- [25] F. Fabiani and S. Grammatico, "Multi-vehicle automated driving as a generalized mixed-integer potential game," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 1064–1073, Mar. 2020.
- [26] R. A. Dollar and A. Vahidi, "Predictively coordinated vehicle acceleration and lane selection using mixed integer programming," in *Proc. ASME Dyn. Syst. Control Conf. (DSCC)*, vol. 1, 2018, p. V001T09A006.
- [27] X. Hou, J. Zhang, C. He, C. Li, Y. Ji, and J. Han, "Crash mitigation controller for unavoidable T-bone collisions using reinforcement learning," *ISA Trans.*, vol. 130, pp. 629–654, Nov. 2022.
- [28] B. Peng et al., "End-to-end autonomous driving through dueling double deep Q-Network," *Automot. Innov.*, vol. 4, no. 3, pp. 328–337, Aug. 2021.
- [29] H. Bey, M. Tratz, M. Sackmann, A. Lange, and J. Thielecke, "Tutorial on sampling-based POMDP-planning for automated driving," in *Proc. 6th Int. Conf. Vehicle Technol. Intell. Transp. Syst.*, 2020, pp. 312–321.
- [30] Z. Sunberg and M. J. Kochenderfer, "Improving automated driving through POMDP planning with human internal states," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 20073–20083, Nov. 2022.
- [31] C. Xia, M. Xing, and S. He, "Interactive planning for autonomous driving in intersection scenarios without traffic signs," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 12, pp. 24818–24828, Dec. 2022.
- [32] T. Brüdigam, M. Olbrich, D. Wollherr, and M. Leibold, "Stochastic model predictive control with a safety guarantee for automated driving," *IEEE Trans. Intell. Vehicles*, vol. 8, no. 1, pp. 22–36, Jan. 2023.
- [33] J. H. Ryu, D. Ogay, S. Bulavintsev, H. Kim, and J. S. Park, "Development and experiences of an autonomous vehicle for high-speed navigation and obstacle avoidance," in *Frontiers of Intelligent Autonomous Systems*, J.-H. Ryu, D. Ogay, and S. Bulavintsev, Eds. Cham, Switzerland: Springer, 2013, pp. 105–116.
- [34] K. Deb and K. Deb, "Multi-objective optimization," in *Search Methodologies*. Cham, Switzerland: Springer, 2014, pp. 403–499.
- [35] B. Stellato, G. Banjac, P. Goulart, A. Bemporad, and S. Boyd, "OSQP: An operator splitting solver for quadratic programs," 2017, *arXiv:1711.08013*.
- [36] C. Tang, Y. Liu, H. Xiao, and L. Xiong, "Integrated decision making and planning framework for autonomous vehicle considering uncertain prediction of surrounding vehicles," in *Proc. IEEE 25th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2022, pp. 3867–3872.
- [37] X. Tian, Z. Zhu, J. Zhao, G. Tian, and C. Ye, "DL-SLOT: Dynamic LiDAR SLAM and object tracking based on collaborative graph optimization," 2022, *arXiv:2212.02077*.
- [38] Z. Wang, S. Zhou, Y. Huang, and W. Tian, "DsMCL: Dual-level stochastic multiple choice learning for multi-modal trajectory prediction," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, Sep. 2020, pp. 1–6.
- [39] L. Xiong, M. Liu, X. Yang, and B. Leng, "Integrated path tracking for autonomous vehicle collision avoidance based on model predictive control with multi-constraints," in *Proc. IEEE 25th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2022, pp. 554–561.



Zhuoren Li (Student Member, IEEE) received the B.S. degree from the School of Aerospace Engineering and Applied Mechanics, Tongji University, Shanghai, China, in 2019. He is currently pursuing the Ph.D. degree with the School of Automotive Studies, Tongji University. His current research interests include interaction decision-making, motion planning, and the safe reinforcement learning of autonomous vehicles.

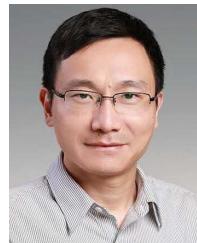


Jia Hu (Member, IEEE) is currently the ZhongTe Distinguished Chair of cooperative automation with the College of Transportation Engineering, Tongji University. Before joining Tongji University, he was a Research Associate with the Federal Highway Administration, USA (FHWA). He is a member of the TRB (a Division of the National Academies) Vehicle Highway Automation Committee, the Freeway Operation Committee, the Simulation Subcommittee of the Traffic Signal Systems Committee, the CAV Impact Committee, and the Artificial Intelligence Committee of ASCE Transportation and Development Institute. He is an Associate Editor of IEEE TRANSACTION ON INTELLIGENT TRANSPORTATION SYSTEMS, IEEE TRANSACTION ON INTELLIGENT VEHICLE, *Journal of Transportation Engineering* (American Society of Civil Engineers), and IEEE OPEN JOURNAL IN INTELLIGENT TRANSPORTATION SYSTEMS; an Assistant Editor of *Journal of Intelligent Transportation Systems*; and an Advisory Editorial Board Member for *Transportation Research—C: Emerging Technologies*. He has been an Associate Editor of the IEEE Intelligent Vehicles Symposium since 2018 and the IEEE Intelligent Transportation Systems Conference since 2019.



selected into the Young Association for Science and Technology in 2022.

Bo Leng received the Ph.D. degree in vehicle engineering from Tongji University, Shanghai, China. He is currently an Associate Professor with the School of Automotive Studies, Tongji University. His current research interests include the dynamic control of distributed drive electric vehicles and motion planning and control of intelligent vehicles. He has won the First Prize in the China Automobile Industry Technology Invention Award and the First Prize in the Shanghai Science and Technology Progress Awards in 2020 and 2022. He has been Elite Scientists Sponsorship Program of the China Association for Science and Technology in 2022.



Lu Xiong received the Ph.D. degree in vehicle engineering from Tongji University, Shanghai, China, in 2005. He is currently the Vice President and a Professor with the School of Automotive Studies, Tongji University. His current research interests include the dynamic control of distributed drive electric vehicles, motion planning and control of intelligent vehicles, and all-terrain vehicles. He won the First Prize in the Shanghai Science and Technology Progress Awards in 2013, 2020, and 2022. He was a recipient of the National Science Fund for Distinguished Young Scholars.



Zhiqiang Fu received the master's degree from Chinesisch-Deutsches Hochschulkolleg, Tongji University, Shanghai, China, in 2018, and the Ph.D. degree from the School of Automotive Studies, Tongji University, in 2022. His current research interests include motion controlling, trajectory planning, and decision-making for autonomous vehicles.