

Safe and Human-Like Autonomous Driving: A Predictor–Corrector Potential Game Approach

Mushuang Liu¹, Member, IEEE, H. Eric Tseng², Dimitar Filev³, Life Fellow, IEEE,
Anouck Girard⁴, Senior Member, IEEE, and Ilya Kolmanovsky¹, Fellow, IEEE

Abstract—This article proposes a novel decision-making framework for autonomous vehicles (AVs), called predictor–corrector potential game (PCPG), composed of a predictor and a corrector. To enable human-like reasoning and characterize agent interactions, a receding-horizon multiplayer game is formulated. To address the challenges caused by the complexity of solving a multiplayer game and by the requirement of real-time operation, a potential game (PG)-based decision-making framework is developed in the PG predictor, where the agents’ cost functions are heuristically predefined. We acknowledge that the behaviors of other traffic agents, for example, human-driven vehicles and pedestrians, may not necessarily be consistent with the predefined cost functions. To address this issue, a best response-based PG corrector is designed. In the corrector, the action deviation between the ego vehicle prediction and the surrounding agents’ actual behaviors are measured and are fed back to the ego vehicle decision-making, to correct the prediction errors caused by the inaccurate predefined cost functions and to improve the ego vehicle strategies. Distinguished from most existing game-theoretic approaches, this PCPG: 1) deals with multiplayer games and guarantees the existence of a pure-strategy Nash equilibrium (PSNE) and the convergence of the PSNE-seeking algorithm; 2) is computationally scalable in a multiagent scenario; 3) guarantees the ego vehicle safety under suitable conditions; and 4) approximates the actual PSNE of the system despite the unknown cost functions of others. Comparative studies between the PG, the PCPG, and the control barrier function (CBF)-based approaches are conducted in diverse traffic scenarios, including oncoming traffic scenarios and multivehicle intersection-crossing scenarios. The results from validation case studies based on a naturalistic dataset are reported.

Index Terms—Autonomous driving, decision-making, game theory, model predictive control, potential games (PGs).

I. INTRODUCTION

FULLY autonomous vehicles (AVs) are expected to improve safety, mobility, and accessibility and reduce energy consumption [1]. However, their broad introduction

Manuscript received 4 August 2022; revised 6 September 2023; accepted 7 November 2023. Date of publication 21 November 2023; date of current version 25 April 2024. This work was supported by Ford Motor Company. Recommended by Associate Editor S. Sacone. (Corresponding author: *Mushuang Liu*.)

Mushuang Liu is with the Department of Mechanical and Aerospace Engineering, University of Missouri, Columbia, MO 65211 USA (e-mail: ml529@missouri.edu).

H. Eric Tseng and Dimitar Filev are with the Ford Research and Innovation Center, Dearborn, MI 48124 USA (e-mail: htsseng@ford.com; dfilev@ford.com).

Anouck Girard and Ilya Kolmanovsky are with the Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: anouck@umich.edu; ilya@umich.edu).

Digital Object Identifier 10.1109/TCST.2023.3332438

still faces significant challenges [2]. One of the key technical challenges lies in the design of AV decision-making algorithms, which aim to generate safe, reliable, and intelligent decisions for AVs in diverse and complex traffic scenarios. An ideal AV decision-making framework is expected to have the following properties.

- 1) *Safety*: Theoretical safety guarantees are desirable if a safe solution exists.
- 2) *Interpretability*: The decision-making process is expected to be interpretable by humans to build trust between humans and AVs.
- 3) *Applicability*: The algorithm should apply to various traffic scenarios and have the capability to handle unfamiliar or even previously unseen situations.
- 4) *Scalability*: The algorithm is computationally scalable to handle a large number of traffic agents.
- 5) *Intelligence*: Human-like negotiating behaviors and reasoning are desirable, especially when AVs interact with human drivers and/or pedestrians.

The existing approaches to AV decision-making can be largely grouped into two categories: model-based and data-driven. Model-based approaches [3], [4], [5], [6], [7], [8], in general, have good explainability and have the potential to provide safety guarantees under certain conditions (this is the case for responsibility-sensitive safety (RSS) [3] and control barrier function (CBF) [4]-based approaches). However, the downside is that they often lead to conservative AV behaviors [9], are highly dependent on assumed model parameters [10], and lack the applicability to diverse traffic scenarios [11]. Data-driven approaches, on the other hand, take advantage of naturalistic traffic datasets [12], [13], [14] and have the capability to generate human-like behaviors [15] in certain scenarios. However, their performance is highly reliant on the quality of the training datasets, lacking assurance in dealing with unfamiliar or unseen situations. Moreover, since deep neural networks (DNNs) are usually employed, especially in end-to-end learning [16], [17], the lack of transparency in the decision-making process becomes a concern in achieving intrinsic interpretability [18] and in building trust and confidence in AVs [19]. In addition, theoretical safety guarantees are generally not available in purely data-driven approaches.

Game-theoretic approaches have the potential to combine the advantages of the model-based and the data-driven approaches [20], [21], [22], [23], [24], [25]. Given the cost

function of each agent, game-theoretic decision-making is consistent with human reasoning: Human behaviors are naturally motivated by pursuing their own interests while accounting for interactions with others [20]. On the other hand, data-driven approaches can facilitate the AV cost function design, leading to human-like behaviors. Examples include supervised learning, reinforcement learning (RL) with value function approximation, and inverse RL [26], [27], [28]. However, conventional game-theoretic approaches often suffer from scalability issues and the lack of knowledge of the surrounding agents' cost functions that reflect the variability in human driving behaviors.

To address the scalability challenge, pairwise games have been widely adopted in the literature [23], [29]. In these kinds of approaches, the ego vehicle is assumed to play multiple two-player games, instead of one multiplayer game, and the most conservative outcomes are selected as the final decision. Such pairwise games may lead to conservative AV behaviors and result in a deadlock even in an ideal environment where all agents employ the pairwise games [29].

To address the challenge of unknown surrounding agents' cost functions, attempts have been made to learn the parameters of the cost function in real-time [25], [30], [31]. These parameters usually characterize driving style [26], aggressiveness [25], or social value orientation (SVO) [32], [33]. However, because the time duration of traffic agents' interaction is usually short, the amount of data may not be sufficient to guarantee the online learning performance. An inaccurate estimate of the cost can mislead the ego vehicle to perform undesirable or even dangerous maneuvers [24].

To address the above limitations and enable AV decision-making to meet the expectations listed in the first paragraph, we propose a novel predictor–corrector potential game (PCPG) framework, composed of two main components: a PG predictor and a PG corrector. The predictor solves a multiplayer game with predefined cost functions to find the optimal strategy for the ego vehicle while taking into consideration the interactions with other traffic agents. The corrector aims to correct the prediction error caused by the inaccurate predefined cost functions of others and to improve the ego vehicle decision-making. With this PCPG:

- 1) the existence of a pure-strategy Nash equilibrium (PSNE) and the convergence of the solution-seeking algorithm is guaranteed;
- 2) the computational scalability challenge is addressed;
- 3) the ego vehicle safety is guaranteed under suitable conditions. These conditions are detailed in Section V;
- 4) the optimal solution for the ego vehicle can be approximated despite the unknown cost functions of others.

To summarize, this PCPG framework inherits the appealing properties of game-theoretic approaches, including human-like reasoning and explicit agent interaction characterization, and addresses the computational scalability and lack of information challenges, improving the practicability and reliability of applying game-theoretic approaches to autonomous driving.

This article is organized as follows. Section II introduces preliminaries to facilitate the analysis in this article. Section III

formulates the AV decision-making problem as a receding-horizon multiplayer game problem. Section IV proposes the PCPG. Section V analyzes the PCPG performance, including safety and optimality. Section VI conducts numerical studies in specific traffic scenarios and reports validations using a naturalistic dataset. Section VII concludes this article.

II. PRELIMINARIES

Let $\mathbb{Z}_+(\mathbb{Z}_{++})$ denote the set of nonnegative (positive) integers and $\mathbb{R}_+(\mathbb{R}_{++})$ denote the set of nonnegative (positive) real numbers.

Definition 1 (Lipschitz Continuous Function [34]): Suppose $(\mathcal{Y}, d_{\mathcal{Y}})$ and $(\mathcal{Z}, d_{\mathcal{Z}})$ are metric spaces and $g : \mathcal{Y} \rightarrow \mathcal{Z}$. If there exists $K \in \mathbb{R}_+$ such that

$$d_{\mathcal{Z}}(g(y_1), g(y_2)) \leq K \cdot d_{\mathcal{Y}}(y_1, y_2) \quad \forall y_1, y_2 \in \mathcal{Y} \quad (1)$$

then g is called a Lipschitz continuous function on \mathcal{Y} with Lipschitz constant K .

To introduce the preliminaries on game theory, let us consider a strategic-form game $\mathcal{G} = \{\mathcal{N}, \mathcal{A}, \{J_i\}_{i \in \mathcal{N}}\}$. Here, $\mathcal{N} = \{1, 2, \dots, N\}$ is the set of players (or agents), $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_N$ with \mathcal{A}_i representing the strategy space of player i , and $J_i : \mathcal{A} \rightarrow \mathbb{R}$ is the cost function of player i . Denote by \mathcal{N}_{-i} the set of all agents except for agent i . We let $\mathbf{a}_i \in \mathcal{A}_i$ represent the strategy of agent i and $\mathbf{a}_{-i} \in \mathcal{A}_{-i}$ represent the set of strategies of all other agents except for agent i , that is, $\mathbf{a}_{-i} = \{\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_N\}$ with \mathcal{A}_{-i} being the domain of \mathbf{a}_{-i} . Denote $\mathbf{a} = \{\mathbf{a}_i, \mathbf{a}_{-i}\} \in \mathcal{A}$. Similarly, let $J_{-i} = \{J_1, \dots, J_{i-1}, J_{i+1}, \dots, J_N\}$ and $J = \{J_i, J_{-i}\}$.

Definition 2 (Best Response [35]): Agent i 's best response to other agents' fixed strategies $\mathbf{a}_{-i} \in \mathcal{A}_{-i}$ is defined as the strategy \mathbf{a}_i^* such that

$$J_i(\mathbf{a}_i^*, \mathbf{a}_{-i}) \leq J_i(\mathbf{a}_i, \mathbf{a}_{-i}) \quad \forall \mathbf{a}_i \in \mathcal{A}_i. \quad (2)$$

Definition 3 (PSNE [35]): An N -tuple of strategies (or strategy profile) $\{\mathbf{a}_1^*, \mathbf{a}_2^*, \dots, \mathbf{a}_N^*\}$ is a PSNE for an N -player game if and only if

$$J_i(\mathbf{a}_i^*, \mathbf{a}_{-i}^*) \leq J_i(\mathbf{a}_i, \mathbf{a}_{-i}^*) \quad \forall \mathbf{a}_i \in \mathcal{A}_i \quad \forall i \in \mathcal{N}. \quad (3)$$

Equation (3) implies that if all agents play their best response, then a PSNE is achieved, and if a PSNE is achieved, then no player would have the incentive to change its strategy.

Next, we define a special class of games, called continuous potential game (PG).

Assumption 1: \mathcal{A}_i is a connected set, $\mathcal{A}_i \neq \emptyset$, $\forall i \in \mathcal{N}$, and \mathcal{A} is compact, that is, closed and bounded.

Definition 4 (Continuous PG [36]): Under Assumption 1, let J_i be everywhere differentiable on an open superset of \mathcal{A} . The game \mathcal{G} is called a continuous PG if and only if there exists a function $F : \mathcal{A} \rightarrow \mathbb{R}$ such that F is everywhere differentiable on an open superset of \mathcal{A} , and

$$\frac{\partial J_i(\mathbf{a}_i, \mathbf{a}_{-i})}{\partial \mathbf{a}_i} = \frac{\partial F(\mathbf{a}_i, \mathbf{a}_{-i})}{\partial \mathbf{a}_i} \quad (4)$$

holds $\forall \mathbf{a}_i \in \mathcal{A}_i$, $\forall \mathbf{a}_{-i} \in \mathcal{A}_{-i}$, and $\forall i \in \mathcal{N}$. The function F is called the potential function.

Throughout this article, the term PG refers to a continuous PG, and the word "continuous" may be omitted.

A PG has appealing properties. We summarize two of them in the following lemmas.

Lemma 1 (Existence of PSNE [36], [37]): Under Assumption 1, if \mathcal{G} is a PG, then it has at least one PSNE. Moreover, if the potential function F is strictly convex, then the PSNE is unique.

Lemma 2 (Equivalence of Nash Equilibria Sets [36], [37]): Under Assumption 1, if \mathcal{G} is a PG, then the set of PSNE coincides with the set of PSNE of the identical-interest game $\mathcal{G}' = \{\mathcal{N}, \mathcal{A}, \{F\}_{i \in \mathcal{N}}\}$, with F being the potential function. That is,

$$\text{NESet}(\mathcal{G}) = \text{NESet}(\mathcal{G}') \quad (5)$$

where NESet denotes the set of PSNE.

III. PROBLEM FORMULATION

Consider a set of traffic agents \mathcal{N} represented by the following discrete-time models:

$$x_i(t+1) = f_i(x_i(t), a_i(t)) \quad (6)$$

where $x_i(t) \in \mathcal{X}_i \subseteq \mathbb{R}^{n_i}$ and $a_i(t) \in \mathcal{U}_i \subseteq \mathbb{R}^{m_i}$ are, respectively, the state and action of agent i at the time step t , f_i is the system evolution model, and $i \in \mathcal{N}$. Denote $x_{-i} = \{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N\}$, $x = \{x_i, x_{-i}\}$, $a_{-i} = \{a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_N\}$, $a = \{a_i, a_{-i}\}$, and $f = \{f_1, f_2, \dots, f_N\}$. Denote the dimension of a_{-i} as m_{-i} , that is, $a_{-i}(t) \in \mathbb{R}^{m_{-i}}$.

In autonomous driving applications, x_i usually contains agent i 's position, velocity, and heading angle information, and a_i is the acceleration and/or the angular velocity. These signals can be measured by/estimated from vehicle-mounted sensors and have been used in driver assistance systems implemented in real vehicles [38]. The system evolution model f_i represents agent dynamics. Examples of f_i include single-mass model [39], [40], unicycle model [41], and bicycle model [42].

Assumption 2: At time t , agent i has access to $x(t)$, $a(t-1)$, and f .

In a driving scenario, every traffic agent has its own driving objective, for example, tracking a desired trajectory without collisions and with ride comfort. We use the cost function J_i to characterize agent i 's objective and formulate the decision-making problem as a receding horizon optimal control problem. That is, at each t , agent i aims to find its optimal action sequence (also called strategy, or behavior) $\mathbf{a}_i^*(t)$ such that

$$\begin{aligned} \mathbf{a}_i^*(t) &\in \arg \min_{\mathbf{a}_i(t) \in \mathcal{A}_i} J_i(\mathbf{a}_i(t), \mathbf{a}_{-i}(t)) \\ &= \arg \min_{\mathbf{a}_i(t) \in \mathcal{A}_i} \sum_{\tau=t}^{t+T-1} \Psi_i(x_i(\tau), x_{-i}(\tau), a_i(\tau), a_{-i}(\tau)) \end{aligned} \quad (7)$$

where $\mathbf{a}_i(t) = \{a_i(t), a_i(t+1), \dots, a_i(t+T-1)\} \in \mathcal{A}_i$, \mathcal{A}_i is determined by \mathcal{U}_i , Ψ_i is the instantaneous cost at one time instant, $T \in \mathbb{Z}_{++}$ is the horizon length, and $J_i : \mathcal{A} \rightarrow \mathbb{R}$ is agent i 's cost function. The expression of $J_i(\mathbf{a}_i(t), \mathbf{a}_{-i}(t))$ can be different at different t due to its state dependence. With a slight abuse of notation, we do not distinguish J_i at different t in this article, as it is clear from the context. After deriving

$\mathbf{a}_i^*(t)$, agent i implements the first element $a_i^*(t)$ and repeats the same procedure at the next time instant $t + 1$.

In (7), J_i depends on both \mathbf{a}_i and \mathbf{a}_{-i} , indicating that agent i 's driving performance is affected by not only its own actions, but also the actions of other traffic agents, characterizing agent interactions. At each t , if every agent i aims to optimize its performance J_i , then the multiagent optimization (7) becomes a multiplayer game, and the set of all agents' optimal strategies $\{\mathbf{a}_1^*(t), \mathbf{a}_2^*(t), \dots, \mathbf{a}_N^*(t)\}$, if nonempty, composes a PSNE according to Definition 3.

Remark 1: The receding horizon game (7) is consistent with human driver decision-making in the following sense.

- 1) Human driver actions are generally motivated by a foreseen gain or loss within a finite prediction horizon. For example, a braking maneuver is usually motivated by a foreseen collision in a few seconds if the current speed is continued, consistent with the predictive control setting of (7).
- 2) Human drivers adjust their strategies frequently according to the latest information. For example, if a human driver is cut off by a car, he/she may have to abort the previous strategy (e.g., keeping speed) and take action to avoid collision. Such real-time adjustment is achieved in (7) since a new strategy is planned at every t .
- 3) Human driver strategies are affected by the surrounding traffic agents. An experienced human driver does not simply react to the actions of others; instead, he/she predicts others' behaviors by considering their driving objectives/interests, to benefit his/her own decision-making. This feature is represented in (7) by the coupled agent cost functions and by the game-theoretic formulation.

Such optimization-based approaches to represent human decision-making during driving have been validated in the literature with naturalistic datasets [33], [43], [44], [45], [46], [47].

The practical use of the formulation (7), although consistent with human reasoning, presents several challenges.

- 1) *Existence of Solution:* Given arbitrary J_i , a strategy profile $\{\mathbf{a}_1^*(t), \mathbf{a}_2^*(t), \dots, \mathbf{a}_N^*(t)\}$ that satisfies (7) for all $i \in \mathcal{N}$ may not always exist.
- 2) *Convergence of Algorithm:* Even if a solution exists, a solution-seeking algorithm, for example, best- and better-response dynamics, may not necessarily converge.
- 3) *Nonuniqueness of PSNE:* The game (7) may have multiple solutions, that is, multiple PSNE. Selecting the one that is preferable to others can also be challenging.
- 4) *Computational Scalability:* Solving (7) generally requires multiple and iterative optimizations, resulting in a high computational burden when N is large.
- 5) *Lack of Information:* To solve (7), the ego vehicle needs to know everyone's cost function, which may not be realistic in a traffic setting.

To differentiate the ego vehicle and its surrounding agents, from now on, we use agent i to represent the ego vehicle, and the set \mathcal{N}_{-i} to represent its surrounding agents, which may include AVs, human-driven vehicles, and pedestrians.

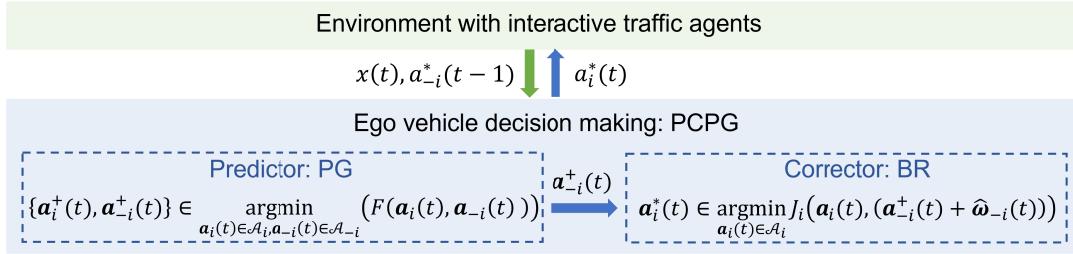


Fig. 1. Diagram of the PCPG at time t .

Remark 2: Although agent i does not know J_{-i} , it may have access to a typical human driver or pedestrian cost function, calibrated using naturalistic traffic datasets [27], [47]. Such a typical cost function corresponds to “common” or “general” human behaviors. However, given a specific human driver, his/her behaviors may not necessarily be consistent with the predetermined typical cost function, as human drivers often have distinct driving habits and styles, and their behaviors can hardly be described by one universal cost function.

Let $\hat{J}_j : \mathcal{A} \rightarrow \mathbb{R}$ be a predefined typical cost function for agent j . \hat{J}_j may not be equal to J_j .

This article aims to develop a decision-making framework to solve (7) and to address the aforementioned challenges.

IV. PREDICTOR–CORRECTOR PG

This section proposes a PCPG approach to solve (7). The predictor–corrector structure is inspired by previous work on CBF-based collision avoidance [48]. Here, we apply this structure in a game-theoretic setting to address the challenges caused by the unknown cost functions of other agents. The predictor solves (7) with the predefined cost function \hat{J}_j , and the corrector corrects the prediction error caused by $\hat{J}_j \neq J_j$. The diagram of the PCPG framework is shown in Fig. 1.

A. Predictor

In the predictor step, the ego vehicle assumes that the surrounding agents’ behaviors are generated from \hat{J}_j . Therefore, it solves the following N -player game at each t :

$$\mathbf{a}_j^+(t) \in \arg \min_{\mathbf{a}_j(t) \in \mathcal{A}_j} \hat{J}_j(\mathbf{a}_j(t), \mathbf{a}_{-j}(t)) \quad (8)$$

where $j \in \mathcal{N}$, $\mathbf{a}_j^+(t)$ is the predicted agent j ’s action sequence at t , and $\hat{J}_i = J_i$.

Assumption 3: \hat{J}_j is everywhere differentiable on an open superset of \mathcal{A} , $\forall j \in \mathcal{N}$.

To ensure a PSNE always exists and is obtainable in real-time, we formulate the game (8) as a PG.

Theorem 1 ([36, Th. 6]): If the cost function \hat{J}_j in (8) satisfies

$$\begin{aligned} \hat{J}_j(\mathbf{a}_j(t), \mathbf{a}_{-j}(t)) \\ = \alpha J_j^{\text{self}}(\mathbf{a}_j(t)) + \beta \sum_{k \in \mathcal{N}, k \neq j} J_{jk}(\mathbf{a}_j(t), \mathbf{a}_k(t)) \end{aligned} \quad (9)$$

where $J_j^{\text{self}} : \mathcal{A}_j \rightarrow \mathbb{R}$ is a function determined solely by agent j ’s action, $J_{jk} : \mathcal{A}_j \times \mathcal{A}_k \rightarrow \mathbb{R}$ satisfies

$$\begin{aligned} J_{jk}(\mathbf{a}_j(t), \mathbf{a}_k(t)) &= J_{kj}(\mathbf{a}_k(t), \mathbf{a}_j(t)) \\ \forall j, k \in \mathcal{N}, j \neq k, \text{ and } \forall \mathbf{a}_j \in \mathcal{A}_j, \mathbf{a}_k \in \mathcal{A}_k \end{aligned} \quad (10)$$

and α and β are two real numbers. Then, the game (8) is a PG with the following potential function:

$$\begin{aligned} F(\mathbf{a}(t)) = \alpha \sum_{j \in \mathcal{N}} J_j^{\text{self}}(\mathbf{a}_j(t)) \\ + \beta \sum_{j \in \mathcal{N}} \sum_{k \in \mathcal{N}, k < j} J_{jk}(\mathbf{a}_j(t), \mathbf{a}_k(t)). \end{aligned} \quad (11)$$

Theorem 1 states that if \hat{J}_j can be represented as a linear combination of two components: J_j^{self} and $\sum_{k \in \mathcal{N}, k \neq j} J_{jk}$, then the resulting game is a PG. The required cost function form, that is, (9), meets the autonomous driving application needs in general. Specifically, the first component J_j^{self} can model self-focused objectives, including tracking a desired speed, minimizing fuel consumption, and maintaining ride comfort. The second component $\sum_{k \in \mathcal{N}, k \neq j} J_{jk}$ can characterize symmetric pairwise agent interaction, such as a pairwise collision penalty. See [47], [49], [50], [51], for example, where the AV cost function design follows, or can be slightly revised to follow, the form (9), and [36], [52], [53], for example, where PGs are employed in the context of driving.

After formulating the game as a PG, according to Lemma 1, a PSNE always exists. According to Lemma 2, a global minimizer to the following optimization problem provides a PSNE to the game (8):

$$\min_{\mathbf{a}(t) \in \mathcal{A}} F(\mathbf{a}(t)) \quad (12)$$

where $F : \mathcal{A} \rightarrow \mathbb{R}$ is the potential function determined by (11). Denote a global minimizer to (12) as $\mathbf{a}^+(t)$, that is, $\mathbf{a}^+(t) \in \arg \min_{\mathbf{a}(t) \in \mathcal{A}} F(\mathbf{a}(t))$.

With (12), the PSNE seeking problem is transferred to a simpler optimization problem. To see the difference clearly, we write down both algorithms here. Algorithm 1 is the best-response dynamics, which is among the most widely used algorithms to solve a multiplayer game. As shown in [54], the number of required optimizations in this algorithm can increase exponentially with the number of game players, thus not scalable. Algorithm 2 is the potential function optimization, applicable if (8) is a PG. This algorithm only requires one optimization, regardless of the number of game players. While this optimization problem is with a larger number of players, its size grows linearly with the number of players. The comparison of the computational times of the two algorithms in the context of five-vehicle intersection-crossing is reported in [36], showing the computational advantages of the PG approach. In addition, Algorithm 1 does not always converge even if a PSNE exists, that is, the NashCondition in

Algorithm 1 Best Response Dynamics to Solve (8)**Inputs:**

Agent set \mathcal{N} ;
 System state $x(t)$;
 System dynamics f ;
 Strategy space \mathcal{A} ;
 Cost functions \hat{J}_j , $j \in \mathcal{N}$;

Output:

PSNE $\mathbf{a}^+(t)$.

Procedures:

- 1: Set $NashCondition=False$
- 2: **While** $NashCondition=False$ **do**
- 3: **For** $j = 1, 2, \dots, N$ **do**
- 4: Find $\mathbf{a}_j^+(t)$ according to

$$\mathbf{a}_j^+(t) \in \operatorname{argmin}_{\mathbf{a}_j(t) \in \mathcal{A}_j} \hat{J}_j(\mathbf{a}_j(t), \mathbf{a}_{-j}(t)).$$
- 5: Update $\mathbf{a}_j(t)$ using $\mathbf{a}_j^+(t)$.
- 6: **End for**
- 7: **If**
- 8: $\mathbf{a}_j^+(t) \in \operatorname{argmin}_{\mathbf{a}_j(t) \in \mathcal{A}_j} \hat{J}_j(\mathbf{a}_j(t), \mathbf{a}_{-j}^+(t))$ holds $\forall j \in \mathcal{N}$,
- 9: **Then**
- 10: Set $NashCondition=True$.
- 11: **End if**
- 12: **End while**

Algorithm 2 Potential Function Optimization to Solve (8)**Input:**

Agent set \mathcal{N} ;
 System state $x(t)$;
 System dynamics f ;
 Strategy space \mathcal{A} ;
 Cost functions \hat{J}_j , $j \in \mathcal{N}$, that satisfy (9);

Output:

PSNE $\mathbf{a}^+(t)$.

Procedures:

- 1: Find $F(\mathbf{a}(t))$ according to (11).
- 2: Find $\mathbf{a}^+(t)$ such that

$$\mathbf{a}^+(t) = \operatorname{argmin}_{\mathbf{a}(t) \in \mathcal{A}} F(\mathbf{a}(t)). \quad (13)$$

Algorithm 1 may never become True. If the game is a PG, then Algorithm 1 is guaranteed to converge [36].

Therefore, by designing (8) as a PG according to Theorem 1, the challenges caused by solution existence, algorithm convergence, and computational scalability are addressed.

Moreover, the game (8) may have multiple PSNE, and Algorithm 1 may converge to any of them, depending on the initial conditions. If using Algorithm 2, then the outcome is not only a PSNE, but also a global minimizer of the potential function F . Since F usually contains everyone's cost, for example, (11), it represents a kind of "social cost." A PSNE that is not only individually optimal for each player, but also socially optimal from a global interest perspective,

is naturally more preferable than others when multiple PSNE exist, addressing the challenge caused by the nonuniqueness of PSNE.

B. Corrector

According to Definition 3, the ego vehicle strategy generated from the predictor, $\mathbf{a}_i^+(t)$, is the best response to $\mathbf{a}_{-i}^+(t)$

$$\mathbf{a}_i^+(t) \in \operatorname{argmin}_{\mathbf{a}_i(t) \in \mathcal{A}_i} J_i(\mathbf{a}_i(t), \mathbf{a}_{-i}^+(t)). \quad (14)$$

In other words, if $\hat{J}_j = J_j \forall j \in \mathcal{N}$, then $\mathbf{a}_i^+(t)$ is the ego vehicle's optimal strategy. However, if $\hat{J}_j \neq J_j$, it is likely that $\mathbf{a}_{-i}^*(t) \neq \mathbf{a}_{-i}^+(t)$, where $\mathbf{a}_{-i}^*(t) = \{a_{-i}^*(t), a_{-i}^*(t+1), \dots, a_{-i}^*(t+T-1)\}$ represents the surrounding agents' actual strategies.

Define action deviation at t as

$$\omega_{-i}(t) = a_{-i}^*(t) - a_{-i}^+(t). \quad (15)$$

In the corrector step, we let the ego vehicle perform the best response to a corrected prediction on the surrounding agents' actions. Specifically, the ego vehicle aims to find $\mathbf{a}_i^*(t)$ such that

$$\mathbf{a}_i^*(t) \in \operatorname{argmin}_{\mathbf{a}_i(t) \in \mathcal{A}_i} J_i(\mathbf{a}_i(t), \hat{\mathbf{a}}_{-i}^*(t)) \quad (16)$$

with

$$\begin{aligned} \hat{\mathbf{a}}_{-i}^*(t) &= \mathbf{a}_{-i}^+(t) + \hat{\boldsymbol{\omega}}_{-i}(t) \\ \hat{\boldsymbol{\omega}}_{-i}(t) &= \mathbf{1}_T \otimes \omega_{-i}(t-1) \\ &= \mathbf{1}_T \otimes (a_{-i}^*(t-1) - a_{-i}^+(t-1)) \end{aligned} \quad (17)$$

where \otimes represents the Kronecker product and $\mathbf{1}_T$ is a vector of ones with T elements. Note that J_i in (16) is the cost function of the ego vehicle and hence can be assumed to be known.

Equation (17) finds the corrected prediction on surrounding agents' actions, that is, $\hat{\mathbf{a}}_{-i}^*(t) = \{\hat{a}_{-i}^*(t), \hat{a}_{-i}^*(t+1), \dots, \hat{a}_{-i}^*(t+T-1)\}$, taking into consideration both the prediction from the PG, that is, $\mathbf{a}_{-i}^+(t)$, and the observed action deviation at $t-1$, that is, $\hat{\boldsymbol{\omega}}_{-i}(t) = \mathbf{1}_T \otimes \omega_{-i}(t-1)$. The ego vehicle best response to this corrected prediction is $\mathbf{a}_i^*(t)$, derived by (16).

At time t , let $\tau \in [t, t+T-1]$, $\tau \in \mathbb{Z}_+$. Define prediction error at τ as

$$e(\tau) = a_{-i}^*(\tau) - \hat{a}_{-i}^*(\tau). \quad (18)$$

The prediction error is of dimension m_{-i} , that is, $e(\tau) \in \mathbb{R}^{m_{-i}}$.

Next theorem shows that with (16) and (17), the prediction error $e(\tau)$ admits a bound.

Theorem 2 (Bounded Prediction Error): Assume that $a_{-i}^* : \mathbb{Z}_+ \rightarrow \mathbb{R}^{m_{-i}}$ and $a_{-i}^+ : \mathbb{Z}_+ \rightarrow \mathbb{R}^{m_{-i}}$ are Lipschitz continuous functions with

$$\|a_{-i}^*(\tau) - a_{-i}^*(t-1)\| \leq K_1 \cdot (\tau - t + 1) \cdot \Delta t \quad (19)$$

$$\|a_{-i}^+(\tau) - a_{-i}^+(t-1)\| \leq K_2 \cdot (\tau - t + 1) \cdot \Delta t \quad (20)$$

where $K_1 \in \mathbb{R}_+$ and $K_2 \in \mathbb{R}_+$ are two constants, $\Delta t \in \mathbb{R}_{++}$ is the sampling time, and $\tau \in [t, t + T - 1]$, $\tau \in \mathbb{Z}_+$. Then, there exists a constant $C \in \mathbb{R}_+$ such that the inequality

$$\|e(\tau)\| \leq C \cdot (\tau - t + 1) \cdot \Delta t \quad (21)$$

holds $\forall \tau \in [t, t + T - 1]$, $\tau \in \mathbb{Z}_+$.

Proof: Substituting (15) and (17) into (18), and since $\hat{\omega}_{-i}(\tau) = \omega_{-i}(t - 1)$, $\forall \tau \in [t, t + T - 1]$, $\tau \in \mathbb{Z}_+$, we have

$$\begin{aligned} e(\tau) &= a_{-i}^*(\tau) - \hat{a}_{-i}^*(\tau) \\ &= (a_{-i}^+(\tau) + \omega_{-i}(\tau)) - (a_{-i}^+(\tau) + \hat{\omega}_{-i}(\tau)) \\ &= \omega_{-i}(\tau) - \hat{\omega}_{-i}(\tau) \\ &= \omega_{-i}(\tau) - \omega_{-i}(t - 1) \\ &= (a_{-i}^*(\tau) - a_{-i}^+(t - 1)) - (a_{-i}^*(t - 1) - a_{-i}^+(t - 1)) \\ &= (a_{-i}^*(\tau) - a_{-i}^*(t - 1)) - (a_{-i}^+(\tau) - a_{-i}^+(t - 1)). \end{aligned} \quad (22)$$

Substituting (19) and (20) into (22), we have

$$\begin{aligned} \|e(\tau)\| &= \|(a_{-i}^*(\tau) - a_{-i}^*(t - 1)) - (a_{-i}^+(\tau) - a_{-i}^+(t - 1))\| \\ &\leq \|a_{-i}^*(\tau) - a_{-i}^*(t - 1)\| + \|a_{-i}^+(\tau) - a_{-i}^+(t - 1)\| \\ &\leq (K_1 + K_2) \cdot (\tau - t + 1) \cdot \Delta t. \end{aligned} \quad (23)$$

By letting $C = K_1 + K_2$, the proof is completed. \square

Theorem 2 shows that the prediction error $e(\tau)$ remains bounded over a finite prediction horizon, and the bound depends on the constants K_1 and K_2 . The value of K_1 can be estimated from the acceleration, jerk, and angular acceleration limits of the individual vehicles. For example, if the actions are the longitudinal acceleration/deceleration, then K_1 can be estimated from the vehicle jerk limit. According to [55], 99.9% vehicle jerk in the studied highway driving is within the range $[-1.47, 1.07] \text{ m/s}^3$. Therefore, one may select $K_1 = 1.47$ for highway scenarios (or select $K_1 = 2.94$ per the six sigma rules [56] to enhance safety). In addition, since it is not likely that the peak jerk lasts for a long period, one may make K_1 smaller when $(\tau - t + 1) \cdot \Delta t \geq 1 \text{ s}$, or limit the maximum acceleration to further narrow the error bounds. The value of K_2 solely depends on the PG design and thus can be manually selected (e.g., by letting $K_2 = K_1$ to make it consistent with real trajectory bound) and be enforced when solving the PG. In addition, it can also be estimated from offline closed-loop simulations of the multiagent system operating according to the PG in Algorithm 2.

Given the bound (21), we denote

$$\mathcal{E}(\tau) = \{\hat{e} \in \mathbb{R}^{m-i} \mid \|\hat{e}\| \leq C \cdot (\tau - t + 1) \cdot \Delta t\} \quad (24)$$

as the set of all possible prediction errors at τ , $\tau \in [t, t + T - 1]$.

The corrector algorithm is summarized in Algorithm 3.

Remark 3: The ego vehicle cost functions in the predictor and the corrector do not need to be identical, although, in the current PCPG, they are designed to be the same to simplify the developments. In the predictor, the ego vehicle cost function needs to follow the form required by Theorem 1 for a PG. However, in the corrector, it is not necessary, and the cost function can be designed more flexibly.

Remark 4: Many of the existing game-theoretic AV decision-making approaches are open-loop [23], [36], [57], [58]. Even if the ego vehicle observes other agents' action

Algorithm 3 PG Corrector

Input:

Output from the Predictor, $\hat{a}_{-i}^+(\tau)$;
Surrounding agents' actions at $t - 1$, $a_{-i}^*(t - 1)$;
Ego vehicle strategy space, \mathcal{A}_i ;
Ego vehicle cost function, J_i ;

Output:

Optimal strategy for the ego vehicle, $a_i^*(t)$.

Procedures:

- 1: Find the corrected prediction $\hat{a}_{-i}^*(\tau)$ according to (17).
 - 2: Find the ego vehicle best response $a_i^*(t)$ according to (16).
-

deviations, it does not adjust its own decision-making. A few works, on the other hand, aim to address this issue by learning driver-specific cost functions in real-time [25], [31], [33]. However, considering the short duration of agent interaction, estimating others' cost functions can be challenging. In contrast, our approach relies on correcting the actions based on other agents' deviations and on exploiting the error bound (21) predetermined offline.

V. SAFETY AND PERFORMANCE ANALYSIS

This section analyzes the PCPG performance, including safety and optimality.

We define a safe set $\mathcal{X}_i^{\text{safe}}(x_{-i}(t))$ as the set of all $x_i(t)$ such that given $x_{-i}(t)$, if the ego vehicle state is within the safe set, that is, $x_i(t) \in \mathcal{X}_i^{\text{safe}}(x_{-i}(t))$, then the ego vehicle is considered safe at t . An example of such a safe set, if x represents vehicle position, is $\mathcal{X}_i^{\text{safe}}(x_{-i}(t)) = \{x_i(t) \mid \|x_i(t) - x_j(t)\| \geq d_{\text{safe}}, \forall j \in \mathcal{N}_{-i}\}$, where $d_{\text{safe}} > 0$ is a predefined safe distance.

At time t , we denote the surrounding agents' future state trajectories generated by their actual action sequence $\mathbf{a}_{-i}^*(t) = \{a_{-i}^*(t), \dots, a_{-i}^*(t + T - 1)\}$ as $\{x_{-i}^*(t + 1), \dots, x_{-i}^*(t + T)\}$, unknown to the ego vehicle. For $\tau \in [t + 1, t + T]$, denote by $\hat{\mathcal{X}}_{-i}(\tau|t)$ the set of $\hat{x}_{-i}(\tau|t)$ generated by the set of action sequences $\{\hat{a}_{-i}^*(t) + \hat{e}(t), \dots, \hat{a}_{-i}^*(\tau - 1) + \hat{e}(\tau - 1)\}$, where $\hat{e}(k) \in \mathcal{E}(k)$, $k \in [t, \tau - 1]$, and $\mathcal{E}(k)$ is defined in (24). With a slight abuse of notation, $\hat{\mathcal{X}}_{-i}(\tau|t)$ can be represented as

$$\begin{aligned} \hat{\mathcal{X}}_{-i}(\tau|t) &= \{\hat{x}_{-i}(\tau|t) \mid \hat{x}_{-i}(\tau|t) \\ &= f_{-i}(x_{-i}(t), \{\hat{a}_{-i}^*(t) + \hat{e}(t), \dots, \\ &\quad \hat{a}_{-i}^*(\tau - 1) + \hat{e}(\tau - 1)\}) \mid \hat{e}(k) \in \mathcal{E}(k), \\ &\quad k \in [t, \tau - 1]\} \end{aligned} \quad (25)$$

where

$$f_{-i}(x_{-i}(t), \{\hat{a}_{-i}^*(t) + \hat{e}(t), \dots, \hat{a}_{-i}^*(\tau - 1) + \hat{e}(\tau - 1)\})$$

represents the surrounding agents' state at τ if the action sequence $\{\hat{a}_{-i}^*(t) + \hat{e}(t), \dots, \hat{a}_{-i}^*(\tau - 1) + \hat{e}(\tau - 1)\}$ is implemented.

Let $\mathcal{A}_i^{\text{safe}}(t, T) \subseteq \mathcal{A}_i$ denote the set of $\mathbf{a}_i(t)$ such that $\forall \tau \in [t + 1, t + T]$

$$\begin{aligned} x_i(\tau) &= f_i(x_i(t), \{a_i(t), \dots, a_i(\tau - 1)\}) \\ &\in \mathcal{X}_i^{\text{safe}}(\hat{x}_{-i}(\tau|t)) \end{aligned} \quad (26)$$

holds $\forall \hat{x}_{-i}(\tau|t) \in \hat{\mathcal{X}}_{-i}(\tau|t)$.

Our next theorem shows that the outcome from the PCPG guarantees the ego vehicle safety, under suitable conditions.

Theorem 3 (Safety): If the ego vehicle cost function J_i is designed such that

$$\arg \min_{\mathbf{a}_i(t) \in \mathcal{A}_i} J_i(\mathbf{a}_i(t), \hat{\mathbf{a}}_{-i}^*(t)) \subseteq \mathcal{A}_i^{\text{safe}}(t, T) \quad (27)$$

and

$$\mathcal{A}_i^{\text{safe}}(t, T) \neq \emptyset \quad (28)$$

then $\mathbf{a}_i^*(t)$ guarantees the ego vehicle safety within the horizon $[t+1, t+T]$, that is, $x_i(\tau) \in \mathcal{X}_i^{\text{safe}}(\hat{x}_{-i}^*(\tau)), \forall \tau \in [t+1, t+T]$.

Proof: According to Theorem 2, $\forall \tau \in [t+1, t+T]$

$$a_{-i}^*(\tau-1) \in \{\hat{a}_{-i}^*(\tau-1) + \hat{e}(\tau-1) \mid \hat{e}(\tau-1) \in \mathcal{E}(\tau-1)\}. \quad (29)$$

Therefore, according to (25)

$$x_{-i}^*(\tau) \in \hat{\mathcal{X}}_{-i}(\tau|t). \quad (30)$$

Combining (26), (27), (28), and (30), we have

$$x_i(\tau) \in \mathcal{X}_i^{\text{safe}}(x_{-i}^*(\tau)). \quad (31)$$

Therefore, the ego vehicle safety is maintained within the horizon $[t+1, t+T]$. \square

Remark 5: Theorem 3 states that the ego vehicle safety is guaranteed with the PCPG if the following two conditions are satisfied: 1) the ego vehicle is safety-conscious, that is, (27) holds, and 2) a safe strategy exists, that is, (28) holds.

A strategy $\mathbf{a}_i(t)$ is said to be safe if it satisfies (26). Intuitively, it means that if a strategy leads to the ego vehicle safety against a set of surrounding agents' strategies, given by the prediction from the PCPG, $\hat{\mathbf{a}}_{-i}^*(t)$, and the bounded prediction error, $e(\tau) \in \mathcal{E}(\tau)$, then it is considered to be safe.

The condition in (27) represents a safety-conscious ego vehicle. That is, if safe strategies exist, then a global minimizer of J_i should be one of them. Such a cost function can be designed by incorporating the safety constraint as a barrier in J_i , using, for example, the interior-point method [59], [60]. A detailed example of how to design such a cost function is provided in Section VI.

The safety guarantee in Theorem 3 does not assume any specific behaviors of the surrounding agents. The surrounding agents can behave freely, be noncooperative, or even be safety-agnostic. In the worst case, where all other agents aim to crash into the ego vehicle, the ego vehicle safety may not be maintained because a safe strategy does not exist. However, as long as a safe strategy exists, that is, $\mathcal{A}_i^{\text{safe}}(t, T) \neq \emptyset$, the PCPG leads to the ego vehicle safety. Here, $\mathcal{A}_i^{\text{safe}}(t, T)$ is defined based on the forward reachable set $\mathcal{X}_i^{\text{safe}}(\hat{x}_{-i}(\tau|t))$ in (26), the computation of which can be realized using, for example, reachability-based trajectory design (RTD) [61].

In a situation where a safe strategy does not exist, that is, $\mathcal{A}_i^{\text{safe}}(t, T) = \emptyset$, the outcome from the PCPG provides a minimum loss strategy for the ego vehicle, in the sense that J_i is minimized subject to $\hat{\mathbf{a}}_{-i}^*(t)$. Note that $\mathcal{A}_i^{\text{safe}}(t, T) = \emptyset$ does not mean that a collision would definitely happen, as whether a collision happens depends on other agents' actions as well. If other agents are also safety-conscious, then

the PCPG enables the ego vehicle to collaborate with others to avoid collision, because the outcome from the PG is not only individually optimal, but also optimizes the "team interest" of all agents, none of which desires a collision.

Remark 6: Theorem 3 does not guarantee recursive feasibility, unless other agents also satisfy certain safety constraints. That is, $\mathcal{A}_i^{\text{safe}}(t+1, T)$ is not guaranteed to be nonempty because we cannot control the surrounding agents' behaviors. As shown in [3], when AVs and human-driven vehicles share the road, it is impossible to guarantee absolute safety for an AV. On the other hand, if other agents' behaviors do satisfy certain safety constraints, for example, the robust CBF (RCBF) constraints (i.e., [48, eq. (2)]), then Theorem 3 can guarantee recursive feasibility, by letting J_i incorporate the RCBF constraints as barriers. However, as we shall see in Section VI if other agents are not safety-conscious, the CBF-based approach can be less preferable than the PCPG.

Our next theorem studies the optimality of $\mathbf{a}_i^*(t)$. According to (16), $\mathbf{a}_i^*(t)$ is optimal if $\hat{a}_{-i}^*(t) = \mathbf{a}_{-i}^*(t)$. With a slight abuse of notation, we denote $(\mathbf{a}_i^*(t), \mathbf{a}_{-i}^*(t))$ (resp., $(\mathbf{a}_i^*(t), \hat{\mathbf{a}}_{-i}^*(t))$) as the strategy profile that the surrounding agents take $\mathbf{a}_{-i}^*(t)$ (resp., $\hat{\mathbf{a}}_{-i}^*(t)$) and the ego vehicle take $\mathbf{a}_i^*(t) \in \arg \min_{\mathbf{a}_i(t) \in \mathcal{A}_i} J_i(\mathbf{a}_i(t), \mathbf{a}_{-i}^*(t))$ (resp., $\mathbf{a}_i^*(t) \in \arg \min_{\mathbf{a}_i(t) \in \mathcal{A}_i} J_i(\mathbf{a}_i(t), \hat{\mathbf{a}}_{-i}^*(t))$).

Theorem 4 (Optimality): Consider the PCPG designed in Section IV and the action deviation in (15). If $\omega_{-i}(t)$ varies slowly with time, that is, $\omega_{-i}(t) - \omega_{-i}(t-1) \rightarrow \mathbf{0}$, $\forall t \in \mathbb{Z}_+$, where $\mathbf{0}$ is a vector of zeros of proper dimensions, then $(\mathbf{a}_i^*(t), \hat{\mathbf{a}}_{-i}^*(t)) \rightarrow (\mathbf{a}_i^*(t), \mathbf{a}_{-i}^*(t))$.

Proof: If $\omega_{-i}(t) - \omega_{-i}(t-1) \rightarrow \mathbf{0}$ and T is finite, then according to (22)

$$\begin{aligned} a_{-i}^*(\tau) - \hat{a}_{-i}^*(\tau) \\ = \omega_{-i}(\tau) - \omega_{-i}(t-1) \\ \rightarrow \mathbf{0} \quad \forall \tau \in [t, t+T-1] \quad \forall t \in \mathbb{Z}_+. \end{aligned} \quad (32)$$

Since from (32), $\hat{a}_{-i}^*(\tau) \rightarrow a_{-i}^*(\tau)$ holds for all $\tau \in [t, t+T-1]$ and $\forall t \in \mathbb{Z}_+$, we have $\hat{a}_{-i}^*(t) \rightarrow \mathbf{a}_{-i}^*(t)$. Therefore, $(\mathbf{a}_i^*(t), \hat{\mathbf{a}}_{-i}^*(t)) \rightarrow (\mathbf{a}_i^*(t), \mathbf{a}_{-i}^*(t))$. \square

Remark 7: Theorem 4 suggests that if the action deviation $\omega(t)$ varies slowly with time, then the outcome from the PCPG accurately approximates the actual PSNE $(\mathbf{a}_i^*(t), \mathbf{a}_{-i}^*(t))$. In practice, a slowly time-varying $\omega(t)$ indicates a consistent driving style. For example, if a driver behaves aggressively at $t-1$, which may be reflected by a larger-than-typical acceleration, that is, $a_j^*(t-1) - a_j^+(t-1) > 0$, it is reasonable to assume that he/she would continue this aggressiveness for a while, that is, $a_j^*(t+k) - a_j^+(t+k) > 0$ for some $k = 0, 1, \dots$

From a reasoning perspective, such a slowly time-varying $\omega(t)$ indicates a consistent reasoning behind agent behaviors. For example, if agent j cares about safety less than a typical driver (i.e., the weight of the collision avoidance term in J_j is smaller than in \hat{J}_j), which may cause more moderate braking than expected, that is, $0 > a_j^*(t-1) > a_j^+(t-1)$, then it is reasonable to assume that this inequality also holds at t , that is, $0 > a_j^*(t) > a_j^+(t)$, since his/her reasoning, characterized by J_j , remains the same at time t .

Note that although a slow time-varying $\omega(t)$ facilitates achieving near optimality, it is not required or necessary in the safety guarantee, that is, Theorem 3.

To summarize, with the PCPG, the ego vehicle plans its strategy according to \hat{J}_j (predictor) and gets the strategy improved by incorporating others' action deviations into its decision-making (corrector). Intuitively, although others' actual cost functions are unknown, the ego vehicle can infer how different they are from the assumed ones, for example, whether the aggressiveness is underestimated or overestimated, by observing $a_j^*(t-1) - a_j^+(t-1)$. With this information, the prediction error on others' actions can be effectively compensated without requiring accurate driver-specific cost functions.

Remark 8: Note that the agent set \mathcal{N} can be a subset of the traffic agents, depending on the application needs and the road structure. For example, in the three-lane highway scenario studied in [12], up to six surrounding vehicles are considered interactive agents. In the merging scenario studied in [62], three surrounding vehicles are selected according to a headway-based rule. These agent selection rules are also applicable in the proposed PCPG setting. The performance of the PCPG with a subset of traffic agents is reported in Section VI-C.

VI. NUMERICAL STUDIES

This section applies the PCPG framework to specific traffic scenarios. To illustrate the scalability of the PCPG to varied environments/traffic scenarios, we test the PCPG performance in three different scenarios: a two-vehicle oncoming traffic scenario, a five-vehicle intersection-crossing scenario, and a multivehicle highway merging scenario. In addition, validation results using a naturalistic driving dataset are also reported in highway merging scenarios.

The vehicles' dynamics are described by the following equations [51]:

$$\begin{aligned} X_i(t+1) &= X_i(t) + v_i(t)\Delta t \\ v_i(t+1) &= v_i(t) + a_i(t)\Delta t \end{aligned} \quad (33)$$

where $i = 1, 2, \dots, N$ are the traffic agents, $X_i = [x_i, y_i]^T$ represents agent i 's position, $v_i = [v_{x,i}, v_{y,i}]^T$ is the speed, and $a_i = [a_{x,i}, a_{y,i}]^T$ is the acceleration, and here, agent i 's action. Let the sampling time Δt be 0.5 s in the simulation.

Agents' cost functions are designed as

$$J_i(\mathbf{a}_i(t), \mathbf{a}_{-i}(t)) = \theta_i J_i^{\text{self}}(\mathbf{a}_i(t)) + \sum_{j \in \mathcal{N}, j \neq i} J_{ij}(\mathbf{a}_i(t), \mathbf{a}_j(t)) \quad (34)$$

where $\mathbf{a}_i(t) = \{a_i(t), a_i(t+1), \dots, a_i(t+T-1)\}$ is agent i 's strategy generated at t over the prediction horizon of length T , and θ_i is a constant that characterizes agent i 's aggressiveness.

The first term in (34) is designed to track the desired position and speed. Specifically

$$\begin{aligned} J_i^{\text{self}}(\mathbf{a}_i(t)) \\ = \sum_{\tau=t}^{t+T-1} (\Delta X_i(\tau)^T Q_i \Delta X_i(\tau) + \Delta v_i(\tau)^T R_i \Delta v_i(\tau)) \end{aligned} \quad (35)$$

where $\Delta X_i(\tau) = X_i(\tau) - X_i^d$, $\Delta v_i(\tau) = v_i(\tau) - v_i^d$, X_i^d , and v_i^d are the desired position and speed, respectively. The desired position usually represents, for example, staying in the center of a lane. Here, Q_i and R_i are weighting matrices, Q_i is positive semidefinite, and R_i is positive definite.

The second term in (34) is to avoid collision and is designed to be

$$J_{ij}(\mathbf{a}_i(t), \mathbf{a}_j(t)) = \sum_{\tau=t}^{t+T-1} \frac{d_d^2}{d_{ij}^2(\tau) + \delta} \quad (36)$$

where d_d is the comfortable (or desired) intervehicle distance and $\delta > 0$ is a positive small number to avoid the denominator being 0. The cost (36) discourages the actions that lead to small intervehicle distance $d_{ij}(\tau)$

$$d_{ij}(\tau) = \sqrt{(x_i(\tau) - x_j(\tau))^2 + (y_i(\tau) - y_j(\tau))^2}. \quad (37)$$

As shown from (34)–(37), agent i 's cost is affected by not only its own states and actions, but also the states and actions of its surrounding agents. In light traffic, where d_{ij} is sufficiently large for all j , agent i may be able to track its desired speed v_i^d , since its behavior is primarily governed by (35) in this case. However, in dense traffic, agent i may not be able to maintain its desired speed, since the safety-related cost (36) and (37) would increase dramatically with the decrease of intervehicle distances.

In our simulation, the strategy space \mathcal{A}_i is selected such that a constant action is planned over the prediction horizon, that is, $a_i(\tau) = a_i(t)$, $\forall \tau \in [t, t+T-1]$. Note that although the AV plans one maneuver over the horizon, it may change its mind and select another maneuver after Δt , triggered by the receding horizon control. This setting is consistent with a common driving experience: When a human driver plans a maneuver, for example, steering to change lane, he/she usually expects to continue this maneuver for some time, for example, 4 s for lane-changing [63].

A. Oncoming Traffic Scenario

This section considers the oncoming traffic scenario, where two vehicles encounter each other on a narrow road, as shown in Fig. 2. In such a scenario, both vehicles desire to keep their lateral positions, maintain the lateral and longitudinal velocities, and avoid collisions. To safely pass each other, vehicles have to compromise and deviate their trajectories and velocities from the individually desired ones when they cross each other. How much a vehicle would compromise depends on its aggressiveness, that is, θ_i . This scenario is challenging to handle even for human drivers, as it requires careful interaction with the other driver. Wrong prediction on the other's intention/strategy can easily lead to a collision. We consider limited control authority for both vehicles, that is, $a_{x,i}(t) \in [-3, 3] \text{ m/s}^2$ and $a_{y,i}(t) \in [-3, 3] \text{ m/s}^2$, $i = 1, 2$.

Study 1 (PG Alone): We first test the performance of the predictor alone, that is, the PG with predefined cost functions. The game is solved using the potential function optimization, that is, Algorithm 2.

We first let $\hat{J}_2 = J_2$, that is, vehicle 2's cost function is known to the ego vehicle. In this situation, the two



Fig. 2. Oncoming traffic scenario: two vehicles encounter each other on a narrow road.

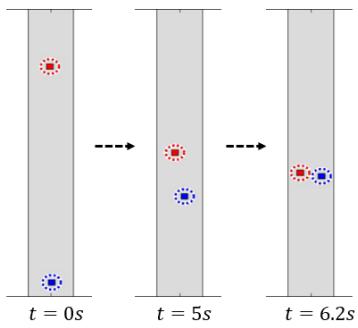


Fig. 3. Oncoming traffic scenario with PG alone. $\hat{J}_2 = J_2$. Two vehicles collaboratively bypass each other. (Video is available at <https://www.youtube.com/watch?v=UOefR0Dhhqk>.)

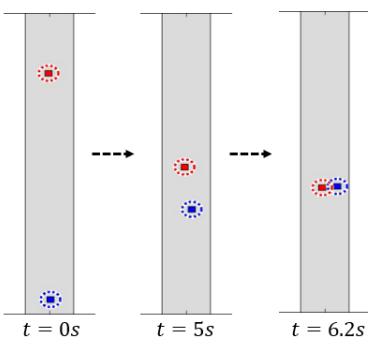


Fig. 4. Oncoming traffic scenario with PG alone. $\hat{J}_2 \neq J_2$. A collision happens at $t = 6.2$ s.

vehicles successfully bypass each other without collision, as shown in Fig. 3. (The animation is available on Youtube at <https://www.youtube.com/watch?v=UOefR0Dhhqk>.) The dotted circles represent the safe distance and have a radius of 2 m for each vehicle. If the circles cross, a collision is considered to happen.

We then consider the case $\hat{J}_2 \neq J_2$. Specifically, the ego vehicle assumes $\theta_2 = 1$ (corresponding to \hat{J}_2), however, $\theta_2 = 10$ (corresponding to J_2), representing a more aggressive vehicle 2 than the ego vehicle expects. In this situation, the ego vehicle expects cautious behavior from vehicle 2, that is, deviating its desired trajectory and speed sufficiently to maintain safety. However, vehicle 2 does not compromise so much, leading to the collision at $t = 6.2$ s, as shown in Fig. 4.

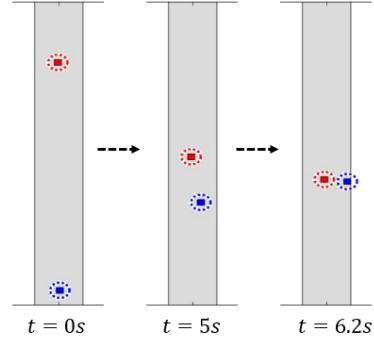


Fig. 5. Oncoming traffic scenario with the PCPG. Collision is successfully avoided despite $\hat{J}_2 \neq J_2$.

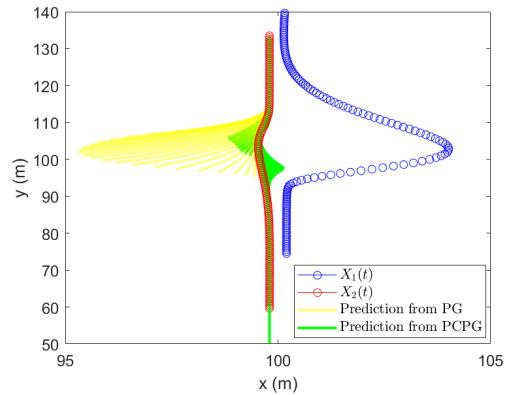


Fig. 6. Trajectories of vehicles 1 and 2. Blue circles: $X_1(t)$. Red circles: $X_2(t)$. Yellow lines: predicted $X_2(\tau)$ according to \mathbf{a}_2^+ (PG). Green lines: predicted $X_2(\tau)$ according to $\hat{\mathbf{a}}_2^*$ (PCPG), where $\tau \in [t, t + 2](s)$.

Study 2 (PCPG): With the same setting and initial conditions as in Fig. 4, we then test the performance of the PCPG. The result is shown in Fig. 5. Thanks to the corrector, the ego vehicle notices that vehicle 2 does not behave as expected, and therefore, it adjusts its decision-making by feeding back vehicle 2's actual behaviors according to Algorithm 3. The collision is now successfully avoided, despite the misinformation $\hat{J}_2 \neq J_2$.

To better visualize the two vehicles' behaviors and the ego vehicle's expectations, we plot their trajectories in Fig. 6. The blue and red circles represent the trajectories of vehicles 1 and 2, that is, $X_1(t)$ and $X_2(t)$, respectively. The ego vehicle trajectory is from the PCPG. The yellow lines represent the ego vehicle predicted vehicle 2 trajectory in the PG, at each t with a horizon $T = 2$ s. The green lines represent the corrected prediction in the PCPG. It is clear that the prediction error is significantly reduced in the PCPG compared to the PG, as the green lines, compared to the yellow ones, are much closer to the red. With the corrected prediction error, the ego vehicle can better respond to the approaching of vehicle 2.

Study 3 (PCPG versus CBF): In addition, we also compare the PCPG with a CBF-based quadratic program (QP) approach. As vehicles are controlled in a decentralized manner, the centralized CBF approach of [4] is not applicable. As such, we employ a variant of the centralized CBF—the predictor-corrector collision avoidance (PCCA) of [48]. In the PCCA,

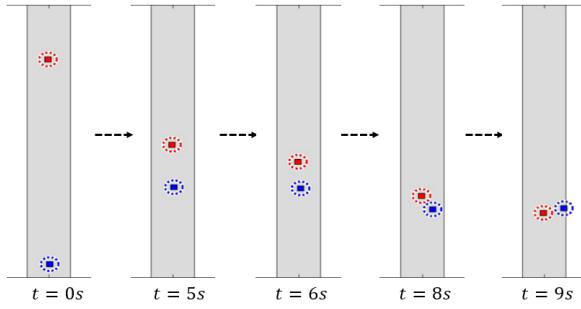


Fig. 7. Oncoming traffic scenario with the PCCA.

the ego vehicle solves the following constrained optimization at each t :

$$\begin{aligned} \min_{a_1(t), a_2(t)} & (\|a_1(t) - a_{10}(t)\|^2 + \|a_2(t)\|^2) \\ \text{s.t. } & b(t) + c(t)a_1(t) - c(t)a_2(t) - c(t)\tilde{\omega}_2(t) \geq 0 \end{aligned} \quad (38)$$

where $\tilde{\omega}_2(t) = a_2^*(t-1) - \tilde{a}_2(t-1)$, $\tilde{a}_2(t-1)$ is vehicle 2 action from the PCCA at $t-1$, and $a_{10}(t)$ is a baseline controller and is from a linear quadratic regulator (LQR). The constraint in (38) is from the RCBF condition, with a relative-degree-two CBF. Specifically

$$\begin{aligned} b(t) &= 2v_{12}^T(t)v_{12}(t) + 2l_1 X_{12}(t)^T v_{12}(t) \\ &\quad + l_0(X_{12}^T(t)X_{12}(t) - d_{\text{safe}}^2) \\ c(t) &= 2X_{12}^T(t) \end{aligned} \quad (39)$$

where $v_{12}(t) = v_1(t) - v_2(t)$, $X_{12}(t) = X_1(t) - X_2(t)$, and l_1 and l_0 are two constants and are selected to be the same as in [48]. Vehicle 2 behaviors remain the same as in Study 2.

As the original PCCA does not consider the control saturation, we first test the PCCA with the unlimited control authority, that is, unbounded action space $a_1(t) \in \mathcal{U}_1 = \mathbb{R}^2$. All other settings and initial conditions are selected to be the same as in Study 2. The result is shown in Fig. 7. As expected, the PCCA successfully avoids collision. Notably, the ego vehicle performs different behaviors with the PCPG and with the PCCA, by comparing Figs. 5 and 7. Specifically, in the PCCA, the ego vehicle responds at a much later time compared to in the PCPG (e.g., at $t = 5$ s, the ego vehicle is well-prepared for the bypassing in the PCPG, while it does not respond yet in the PCCA, since the constraint in (38) is not active yet). Because of this later response, a much larger $\|a_1(t)\|$ is triggered after the constraint in (38) becomes active (e.g., the ego vehicle in Fig. 7 moves backward during 6–8 s to avoid collision). To quantify this observation and the AV driving performance in addition to safety, let us consider the following two performance metrics.

- 1) *Deviation From the Desired Speed:* We calculate both the average and the maximum longitudinal speed deviations, which are related to travel efficiency, according to

$$\begin{aligned} \Delta\bar{v}_{y,1} &= \frac{\sum_{t=0}^{t_f} \Delta v_{y,1}(t)}{t_f} \\ \Delta v_{y,1}^{\max} &= \max_{t \in [0, t_f]} \Delta v_{y,1}(t) \end{aligned} \quad (40)$$

where

$$\Delta v_{y,1}(t) = (v_{y,1}(t) - v_{y,1}^d) \quad (41)$$

$t_f = (15/0.5 \text{ s}) = 30$ is the total time steps of the simulation, and $v_{y,1}^d = 5 \text{ m/s}$ is the ego vehicle desired longitudinal speed. Smaller $\Delta\bar{v}_{y,1}$ and $\Delta v_{y,1}^{\max}$ indicate better adherence to the desired speed, and thus, are preferred.

- 2) *Deviation From the Desired Heading:* The average and the maximum heading angle deviations, which are related to ride comfort [64], are calculated by

$$\begin{aligned} \Delta\bar{\phi}_{y,1} &= \frac{\sum_{t=0}^{t_f} \Delta\phi_1(t)}{t_f} \\ \Delta\phi_1^{\max} &= \max_{t \in [0, t_f]} \Delta\phi_1(t) \end{aligned} \quad (42)$$

where

$$\Delta\phi_1(t) = \phi_1(t) - \phi_1^d = \arctan \frac{v_{y,1}(t)}{v_{x,1}(t)} - \phi_1^d \quad (43)$$

$\phi_1^d = 90^\circ$ is the ego vehicle desired heading angle. Small $\Delta\bar{\phi}_{y,1}$ and $\Delta\phi_1^{\max}$ are preferred.

With the above metrics, we compare the performance of the PCPG and the PCCA. The PCPG leads to

$$\begin{aligned} \Delta\bar{v}_{y,1} &= 0.22 \text{ m/s}, \quad v_{y,1}^{\max} = 1.29 \text{ m/s} \\ \Delta\bar{\phi}_{y,1} &= 6.48^\circ, \quad \Delta\phi_1^{\max} = 36^\circ. \end{aligned}$$

The PCCA leads to

$$\begin{aligned} \Delta\bar{v}_{y,1} &= 1.82 \text{ m/s}, \quad v_{y,1}^{\max} = 8.63 \text{ m/s} \\ \Delta\bar{\phi}_{y,1} &= 36.68^\circ, \quad \Delta\phi_1^{\max} = 170.77^\circ. \end{aligned}$$

Thus, the PCPG results in better adherence to the desired trajectory, thanks to its early response realized by the ego vehicle predictive capability on vehicle 2 strategies, enabling the ego vehicle to plan ahead and to take necessary and moderate maneuvers in advance. In contrast, in the PCCA, the ego vehicle simply reacts to the instantaneous states and actions of vehicle 2 and cannot predict vehicle 2 future behaviors or plan ahead.

Next, let us consider limited control authority for the ego vehicle, that is, $a_{x,1}(t) \in [-3, 3] \text{ m/s}^2$ and $a_{y,1}(t) \in [-3, 3] \text{ m/s}^2$. Specifically, if $a_{x,1}(t)$ or $a_{y,1}(t)$ from (38) is larger than 3 m/s^2 (resp., less than -3 m/s^2), we make the ego vehicle implement 3 m/s^2 (resp., -3 m/s^2). In this case, due to the late response and the bounded action space, a collision happens if the ego vehicle employs the PCCA. The performance is similar to Fig. 4.

Study 4 (Statistical Studies): To make the comparative results more convincing, we conduct statistical studies. 500 scenarios are tested with randomly selected initial lateral positions of the two vehicles, that is, $x_2(0) - x_1(0)$ is uniformly distributed in $[0.2, 2] \text{ m}$, and vehicle 2's aggressiveness, that is, θ_2 is uniformly distributed in $[1, 10]$. The ego vehicle always assumes $\theta_2 = 1$ in all scenarios. Both vehicles have limited control authorities. The statistical results are shown in Table I, where “collision rate” represents the number of scenarios where collision happens divided by the total tested

TABLE I
STATISTICAL COMPARISON

	Two-vehicle oncoming traffic			Five-vehicle intersection-crossing		
Ego vehicle controller	PG	PCPG	PCCA	PG	PCPG	PCCA
Collision rate	264/500	0/500	37/500	28/500	0/500	78/500
Ave/Max longitudinal speed deviation (m/s)	0.11/1.14	0.20/3.63	0.90/8.30	1.00/5.90	1.56/8.20	0.74/7.50
Ave/Max heading angle deviation	3.75°/38.26°	6.50°/69.70°	14.85°/179.8°	0°/0°	0°/0°	0°/0°
Ave/Max computational time (s)	0.06/0.25	0.10/0.33	< 0.01 / < 0.01	0.08/0.30	0.11/0.34	< 0.01 / < 0.01

scenarios, “speed deviation” and “heading angle deviation” are calculated according to (40)–(43), and “computational time” represents the running time for each decision-making collected from MATLAB¹ on a laptop with an Intel Core i7-10750H processor clocked at 2.60 GHz and 16 GB of RAM. The optimization in the PG and the PCPG approaches is performed using the MATLAB genetic algorithm function “*ga*” [65], and in the PCCA approach using the quadratic programming function “*quadprog*” [66].

As we can observe from Table I, among the three approaches, PG, PCPG, and PCCA, the PCPG is the only one that ensures the ego vehicle safety in all tested scenarios. Meanwhile, the PCPG also leads to reasonably small speed deviations and heading angle deviations. Moreover, all three approaches are computationally practical, as all their running time is significantly less than the sampling time of 0.5 s. The PCCA has the lightest computational load and its running time is always less than 0.01 s. While the PCPG has the best performance in both collision avoidance and mobility, between PCCA and PG, the PCCA is more effective in avoiding collisions in the tested scenarios. It is because safety is a hard constraint in the PCCA, while a soft constraint in the PG, and the effectiveness of the PG is heavily dependent on reliable prediction of the surrounding agents’ behaviors.

B. Multivehicle Intersection Crossing

This section considers a multivehicle intersection crossing scenario, as shown in Fig. 8. The vehicle labeled with the number “1” is the ego vehicle, aiming to go straight to cross the intersection. The heading directions of other vehicles are marked using gray arrows and are known to the ego vehicle. Each vehicle controls its own longitudinal acceleration, and the lateral acceleration is zero, that is, no lane-changing while crossing an intersection. The action space for each vehicle is $\mathcal{U}_i = [-3, 3] \text{ m/s}^2$.

The cost function is designed according to (34)–(36), where $\theta_1 = 1$ and $\theta_j \in [1, 100]$ for $j = 2, 3, 4, 5$, unknown to the ego vehicle. Note that when $\theta_j = 100$, the collision avoidance term in (34) is very lightly weighted, indicating a safety-agnostic vehicle. To make the scenario more challenging for the ego vehicle, we also make the surrounding vehicle’s desired speed v_j^d unknown to the ego vehicle. The ego vehicle assumes that $v_j^d = 5 \text{ m/s}$, $\forall j$, but $v_j^d \in [5, 15] \text{ m/s}$ for $j = 2, 3, 4, 5$.

With the above setting, we test and compare the performance of the PG, PCPG, and PCCA.

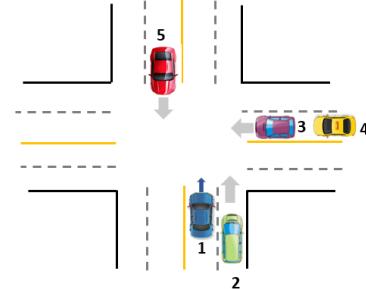


Fig. 8. Multivehicle intersection crossing scenario.

In the PG, the ego vehicle solves a five-player game with the assumed cost functions \hat{J}_j , $j = 2, 3, 4, 5$. When \hat{J}_j and J_j are close, the ego vehicle performs reasonably well in terms of both safety and mobility. However, if J_j deviates significantly from \hat{J}_j , a collision may happen. Fig. 9 shows one such scenario. Vehicles’ aggressiveness, instantaneous, and desired speeds are all labeled in the figure. In this pictured scenario, the purple and the yellow vehicles are both safety-agnostic ($\theta_j = 100$) and have much larger desired speed than the ego vehicle expects ($v_3^d = 13 \text{ m/s}$ and $v_4^d = 15 \text{ m/s}$). At around 4 s, a collision with the yellow vehicle happens, as the ego vehicle expects the yellow to slow down, while the yellow speeds up to track its desired speed.

With the same setting as in Fig. 9, we test the PCPG. As shown in Fig. 10, the collision is now successfully avoided. The ego vehicle notices the action deviation of the purple and the yellow vehicles and improves its strategies accordingly. It can be seen by comparing $v_1(t)$ at $t = 2.5 \text{ s}$ in the PCPG (Fig. 10) and in the PG (Fig. 9).

For the PCCA, similar to the oncoming traffic scenario, if the ego vehicle has unlimited control authority, then collisions can always be avoided. If the control authority is limited, then collisions may happen. Due to the page limit, we do not show the detailed scenario pictures, instead, we present the statistical results in Table I.

From Table I, the PCPG performs the best among the three approaches in terms of safety, as it is the only one that ensures no collision. Different from the two-vehicle scenarios, the PG in the five-vehicle scenarios performs better than the PCCA in terms of safety. It is because the PG encodes the global agent interactions in its decision-making, while the PCCA encodes agent interactions in a pairwise manner (i.e., the CBF constraints are for each vehicle pair separately [48]). It is possible that a strategy to solve the conflict with one vehicle worsens the situation with another. Since the CBF

¹Registered trademark.

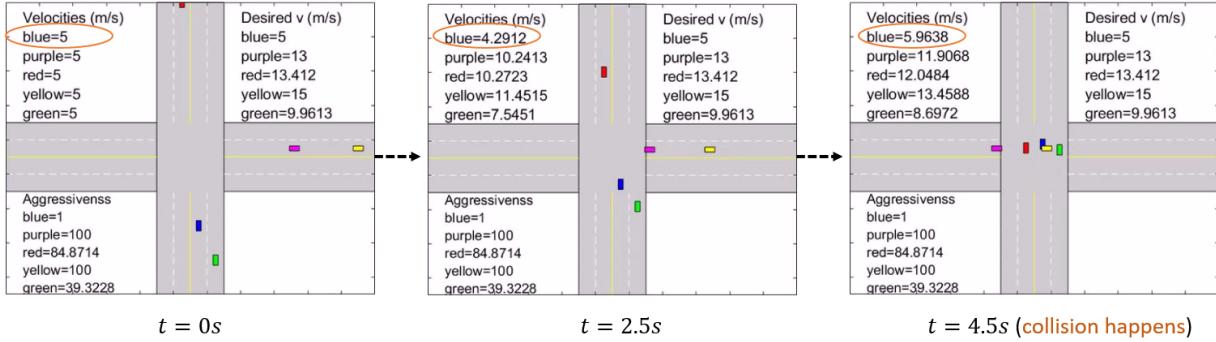


Fig. 9. Multivehicle intersection crossing with PG alone. The blue is the ego vehicle. A collision happens due to the ego vehicle misinformation.

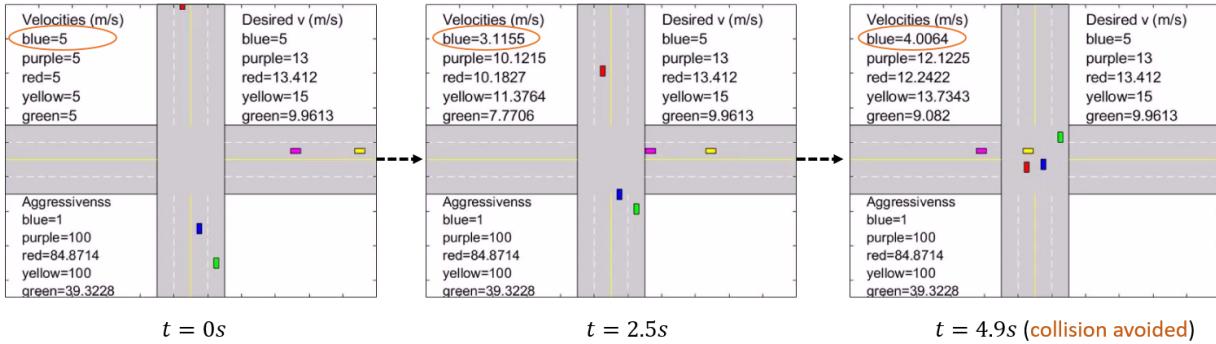


Fig. 10. Multivehicle intersection crossing with PCPG. The collision is successfully avoided despite $\hat{J}_j \neq J_j$, $j \in \mathcal{N}_{-i}$.

constraints are not always active for all vehicles, it is possible that when the constraint becomes active with one vehicle, the ego vehicle is handling the conflict with another, and their handling strategies contradict each other, leading to a collision with at least one of the two vehicles. Therefore, the CBF-based approach can be “short-sighted” in multivehicle scenarios, due to the inability to capture global agent interactions.

From the computational perspective, all three approaches are practical, since their running time is always less than the sampling time. Comparing the running time in the two-vehicle and the five-vehicle scenarios, we can see that the time cost does not increase much with the increase in the number of agents. Specifically, with the PCPG, the average computational time is 0.10 s in the two-vehicle scenario and 0.11 s in the five-vehicle scenario, indicating good computational scalability.

C. Validation With Naturalistic Traffic Data

This section validates the PCPG performance using the naturalistic traffic dataset provided by the Federal Highway Administration’s (FHWA’s) Next Generation Simulation (NGSIM) program [67]. The data was collected on a segment of Interstate 80 (I-80) in Emeryville (San Francisco), California. A snapshot of the I-80 freeway and the traffic situation is shown in Fig. 11. We focus on the merging area (the same road segment and data extraction as in [68]) and aim to validate whether the PCPG enables safe and efficient merging in dense traffic. To this end, we test the PCPG performance on ten merging vehicles and compare it with the original human-driven vehicle trajectories. All vehicles’ initial conditions and



Fig. 11. Snapshot of I-80 traffic video [67].

the surrounding vehicles’ movements are from the I-80 dataset. The merging vehicles’ movements are controlled by the PCPG (purple vehicle in Fig. 12) and by the human drivers as recorded (yellow vehicle in Fig. 12), respectively. In the PCPG, a limited number of surrounding vehicles within the field of view are selected as game players, according to their time-to-collision (TTC) with the ego vehicle [24]. Specifically, a vehicle j is considered a game player if the difference between its TTC (denoted as T_{j1}^c) and the ego vehicle TTC (T_{1j}^c) is within a threshold, that is, $|T_{j1}^c - T_{1j}^c| \leq \Delta T^c$. This criterion selects the most “dangerous” vehicles for the ego vehicle.

The merging vehicle performance is characterized by two metrics: The minimum longitudinal distance with the

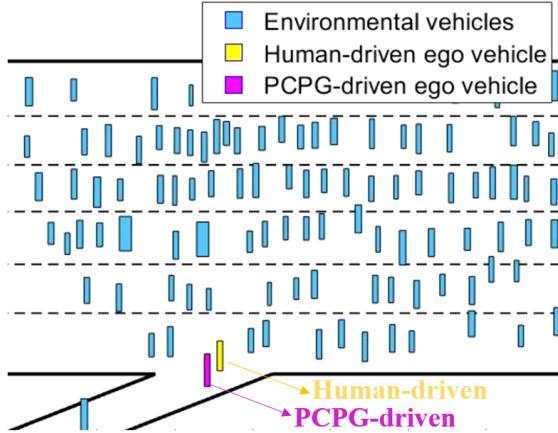


Fig. 12. Snapshot of the validation simulation video, where the merging vehicle trajectories are from the PCPG (purple) and the original human driver (yellow).

TABLE II
STATISTICS OF MERGING BEHAVIORS

Vehicle	Min. inter-vehicle distance		Ave. ego vehicle speed	
	Human driver	PCPG	Human driver	PCPG
1	12m	16m	13m/s	13m/s
2	21m	29m	13m/s	13m/s
3	85m	79m	20m/s	19m/s
4	12m	19m	13m/s	13m/s
5	35m	37m	13m/s	16m/s
6	12m	21m	13m/s	13m/s
7	25m	20m	14m/s	14m/s
8	12m	16m	12m/s	12m/s
9	8m	10m	12m/s	12m/s
10	12m	14m	7m/s	7m/s

surrounding vehicles and the average ego vehicle speed during merging. The first metric indicates how safe this merge is, as longer intervehicle distances are safer. The second metric suggests how efficient the merge is, as higher speed corresponds to higher efficiency. The validation results are reported in Table II, which leads to the following observations.

- 1) In all tested situations, both the PCPG-driven and the human-driven vehicles can successfully merge into the highway, validating the effectiveness of the PCPG.
- 2) In safety-critical situations, where the minimum intervehicle distance is less than 40 m, the PCPG is safer than the corresponding human drivers, as reflected by longer intervehicle distances.
- 3) The average speeds of the human-driven and the PCPG-driven vehicles are almost the same, validating the travel efficiency of the PCPG.

VII. CONCLUSION

In this article, a PCPG framework has been proposed to address the AV decision-making problem. A receding horizon multiplayer game was formulated to mimic human driver reasoning and to characterize agent interactions. To ensure solution existence, algorithm convergence, and computational scalability, a PG-based predictor is developed. To account for inaccurate cost functions of surrounding agents, a best-response-based corrector is introduced. By feeding back the

surrounding agents' action deviations to the ego vehicle decision-making, the prediction error on other agents' future trajectories is corrected, leading to improved ego vehicle strategies. This framework guarantees the ego vehicle safety under suitable conditions and approximates the optimal solution despite the lack of information. Comparative studies between PG, PCPG, and CBF-based approaches show that the PCPG ensures the ego vehicle safety, while the CBF-based approach is not always able to do so if the ego vehicle has limited control authority. It is because the PCPG enables the ego vehicle to always keep in mind the global situation, while the CBF constraints become active only when a collision threat is clear, that is, passively reacting to the environment instead of proactively predicting and planning. With this framework, the long-standing challenges in AV decision-making, including safety, interpretability, computational scalability, applicability to diverse scenarios, and human-like intelligence, are all appropriately addressed. In future work, we will further improve the practicability of the PCPG framework by considering more advanced corrector designs, for example, filtered versions of previous action deviations, to account for possible sensing noises.

REFERENCES

- [1] IGNSS. (2022). *What is the Future of Autonomous Vehicles?* [Online]. Available: <https://insidegnss.com/q-what-is-the-future-of-autonomous-vehicles/>
- [2] CNBC. (2022). *Where the Billions Spent on Autonomous Vehicles by US and Chinese Giants is Heading.* [Online]. Available: <https://www.cnbc.com/2022/05/21/why-the-first-autonomous-vehicles-winners-wont-be-in-your-driveway.html>
- [3] S. Shalev-Shwartz, S. Shammah, and A. Shashua, "On a formal model of safe and scalable self-driving cars," 2017, *arXiv:1708.06374*.
- [4] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," *IEEE Trans. Autom. Control*, vol. 62, no. 8, pp. 3861–3876, Aug. 2017.
- [5] D. D. Salvucci, E. R. Boer, and A. Liu, "Toward an integrated model of driver behavior in cognitive architecture," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1779, no. 1, pp. 9–16, Jan. 2001.
- [6] P. Hidas, "Modelling lane changing and merging in microscopic traffic simulation," *Transp. Res. C, Emerg. Technol.*, vol. 10, nos. 5–6, pp. 351–371, Oct. 2002.
- [7] S. Brechtel, T. Gindele, and R. Dillmann, "Probabilistic decision-making under uncertainty for autonomous driving using continuous POMDPs," in *Proc. 17th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2014, pp. 392–399.
- [8] J. Pérez, V. Milanés, E. Onieva, J. Godoy, and J. Alonso, "Longitudinal fuzzy control for autonomous overtaking," in *Proc. IEEE Int. Conf. Mechatronics*, Apr. 2011, pp. 188–193.
- [9] S. Liu et al., "Calibration and evaluation of responsibility-sensitive safety (RSS) in automated vehicle performance during cut-in scenarios," *Transp. Res. C, Emerg. Technol.*, vol. 125, Apr. 2021, Art. no. 103037.
- [10] H. Königshof, F. Oboril, K.-U. Scholl, and C. Stiller, "A parameter analysis on RSS in overtaking situations on German highways," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2022, pp. 1081–1086.
- [11] S. Kuutti, R. Bowden, Y. Jin, P. Barber, and S. Fallah, "A survey of deep learning applications to autonomous vehicle control," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 2, pp. 712–733, Feb. 2021.
- [12] S. Nageshrao, H. E. Tseng, and D. Filev, "Autonomous highway driving using deep reinforcement learning," in *Proc. IEEE Int. Conf. Syst., Man Cybern. (SMC)*, Oct. 2019, pp. 2326–2331.
- [13] X. Xu, L. Zuo, X. Li, L. Qian, J. Ren, and Z. Sun, "A reinforcement learning approach to autonomous decision making of intelligent vehicles on highways," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 10, pp. 3884–3897, Oct. 2020.
- [14] D. C. K. Ngai and N. H. C. Yung, "A multiple-goal reinforcement learning method for complex vehicle overtaking maneuvers," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 509–522, Jun. 2011.

- [15] L. Li, K. Ota, and M. Dong, "Humanlike driving: Empirical decision-making system for autonomous vehicles," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 6814–6823, Aug. 2018.
- [16] H. Xu, Y. Gao, F. Yu, and T. Darrell, "End-to-end learning of driving models from large-scale video datasets," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3530–3538.
- [17] Y. Xiao, F. Codevilla, A. Gurram, O. Urfalioglu, and A. M. López, "Multimodal end-to-end autonomous driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 1, pp. 537–547, Jan. 2022.
- [18] M. Du, N. Liu, and X. Hu, "Techniques for interpretable machine learning," *Commun. ACM*, vol. 63, no. 1, pp. 68–77, Dec. 2019.
- [19] L. Wells and T. Bednarz, "Explainable AI and reinforcement learning—A systematic review of current approaches and trends," *Frontiers Artif. Intell.*, vol. 4, May 2021, Art. no. 550030.
- [20] 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.
- [21] M. Liu, Y. Wan, F. L. Lewis, and V. G. Lopez, "Adaptive optimal control for stochastic multiplayer differential games using on-policy and off-policy reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 12, pp. 5522–5533, Dec. 2020.
- [22] M. Liu, Y. Wan, V. G. Lopez, F. L. Lewis, G. A. Hewer, and K. Estabridis, "Differential graphical game with distributed global Nash solution," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 3, pp. 1371–1382, Sep. 2021.
- [23] V. G. Lopez, F. L. Lewis, M. Liu, Y. Wan, S. Nageshrao, and D. Filev, "Game-theoretic lane-changing decision making and payoff learning for autonomous vehicles," *IEEE Trans. Veh. Technol.*, vol. 71, no. 4, pp. 3609–3620, Apr. 2022.
- [24] M. Liu, Y. Wan, F. L. Lewis, S. Nageshrao, and D. Filev, "A three-level game-theoretic decision-making framework for autonomous vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 20298–20308, Nov. 2022.
- [25] Q. Zhang, R. Langari, H. E. Tseng, D. Filev, S. Sztabowski, and S. Coskun, "A game theoretic model predictive controller with aggressiveness estimation for mandatory lane change," *IEEE Trans. Intell. Vehicles*, vol. 5, no. 1, pp. 75–89, Mar. 2020.
- [26] D. Silver, J. A. Bagnell, and A. Stentz, "Learning autonomous driving styles and maneuvers from expert demonstration," in *Experimental Robotics*. Berlin, Germany: Springer, 2013, pp. 371–386.
- [27] J. Liu, L. N. Boyle, and A. G. Banerjee, "An inverse reinforcement learning approach for customizing automated lane change systems," *IEEE Trans. Veh. Technol.*, vol. 71, no. 9, pp. 9261–9271, Sep. 2022.
- [28] B. R. Kiran et al., "Deep reinforcement learning for autonomous driving: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 6, pp. 4909–4926, Jun. 2022.
- [29] R. Mandiau, A. Champion, J.-M. Auberlet, S. Espié, and C. Kolski, "Behaviour based on decision matrices for a coordination between agents in a urban traffic simulation," *Int. J. Speech Technol.*, vol. 28, no. 2, pp. 121–138, Apr. 2008.
- [30] S. Hong, J. Lu, and D. P. Filev, "Driving behavior evaluation for future mobility: Application of online transition probability estimation," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 2, pp. 782–791, Feb. 2021.
- [31] J. Lu, D. Filev, and F. Tseng, "Real-time determination of driver's driving behavior during car following," *SAE Int. J. Passenger Cars-Electron. Electr. Syst.*, vol. 8, no. 2, pp. 371–378, Apr. 2015.
- [32] V.-A. Le and A. A. Malikopoulos, "A cooperative optimal control framework for connected and automated vehicles in mixed traffic using social value orientation," 2022, *arXiv:2203.17106*.
- [33] W. Schwarting, A. Pierson, J. Alonso-Mora, S. Karaman, and D. Rus, "Social behavior for autonomous vehicles," *Proc. Nat. Acad. Sci. USA*, vol. 116, no. 50, pp. 24972–24978, 2019.
- [34] M. O'Searcoid, *Metric Spaces*. Berlin, Germany: Springer, 2006.
- [35] Y. Shoham and K. Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
- [36] M. Liu, I. Kolmanovsky, H. E. Tseng, S. Huang, D. Filev, and A. Girard, "Potential game-based decision-making for autonomous driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 8, pp. 8014–8027, Aug. 2023.
- [37] Q. D. Lă, Y. H. Chew, and B.-H. Soong, *Potential Game Theory*. Berlin, Germany: Springer, 2016.
- [38] Tesla. (2022). *Tesla AI Day*. [Online]. Available: https://www.youtube.com/watch?v=ODSJsviD_SU
- [39] C.-J. Hoel, K. Driggs-Campbell, K. Wolff, L. Laine, and M. J. Kochenderfer, "Combining planning and deep reinforcement learning in tactical decision making for autonomous driving," *IEEE Trans. Intell. Vehicles*, vol. 5, no. 2, pp. 294–305, Jun. 2020.
- [40] N. Li, H. Chen, I. Kolmanovsky, and A. Girard, "An explicit decision tree approach for automated driving," in *Proc. Dyn. Syst. Control Conf.*, vol. 58271, 2017, Art. no. V001T45A003.
- [41] S. Kousik, S. Vaskov, M. Johnson-Roberson, and R. Vasudevan, "Safe trajectory synthesis for autonomous driving in unforeseen environments," in *Proc. Dyn. Syst. Control Conf.*, vol. 58271, 2017, Art. no. V001T44A005.
- [42] J. Kong, M. Pfeiffer, G. Schildbach, and F. Borrelli, "Kinematic and dynamic vehicle models for autonomous driving control design," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2015, pp. 1094–1099.
- [43] N. Ratliff et al., "Inverse optimal heuristic control for imitation learning," in *Proc. 12th Int. Conf. Artif. Intell. Statist.*, 2009, pp. 424–431.
- [44] B. D. Ziebart, A. L. Maas, J. A. Bagnell, and A. K. Dey, "Maximum entropy inverse reinforcement learning," in *Proc. AAAI*, vol. 8, 2008, pp. 1433–1438.
- [45] B. D. Ziebart, A. L. Maas, A. K. Dey, and J. A. Bagnell, "Navigate like a cabbie: Probabilistic reasoning from observed context-aware behavior," in *Proc. 10th Int. Conf. Ubiquitous Comput.*, Sep. 2008, pp. 322–331.
- [46] O. Siebinga, A. Zgonnikov, and D. Abbink, "A human factors approach to validating driver models for interaction-aware automated vehicles," *ACM Trans. Hum.-Robot Interact.*, vol. 11, no. 4, pp. 1–21, Dec. 2022.
- [47] Q. Dai, D. Shen, J. Wang, S. Huang, and D. Filev, "Calibration of human driving behavior and preference using vehicle trajectory data," *Transp. Res. C, Emerg. Technol.*, vol. 145, Dec. 2022, Art. no. 103916.
- [48] M. Santillo and M. Jankovic, "Collision free navigation with interacting, non-communicating obstacles," in *Proc. Amer. Control Conf. (ACC)*, May 2021, pp. 1637–1643.
- [49] 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.
- [50] J. F. Fisac, E. Bronstein, E. Stefansson, D. Sadigh, S. S. Sastry, and A. D. Dragan, "Hierarchical game-theoretic planning for autonomous vehicles," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 9590–9596.
- [51] R. Tian, N. Li, I. Kolmanovsky, Y. Yildiz, and A. R. Girard, "Game-theoretic modeling of traffic in unsignalized intersection network for autonomous vehicle control verification and validation," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 2211–2226, Mar. 2022.
- [52] C. Cenedese, M. Cucuzzella, J. M. A. Scherpen, S. Grammatico, and M. Cao, "Highway traffic control via smart e-mobility—Part I: Theory," 2021, *arXiv:2102.09354*.
- [53] 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.
- [54] S. Durand and B. Gaujal, "Complexity and optimality of the best response algorithm in random potential games," in *Proc. Int. Symp. Algorithmic Game Theory*. Cham, Switzerland: Springer, 2016, pp. 40–51.
- [55] F. Feng, S. Bao, J. R. Sayer, C. Flannagan, M. Manser, and R. Wunderlich, "Can vehicle longitudinal jerk be used to identify aggressive drivers? An examination using naturalistic driving data," *Accident Anal. Prevention*, vol. 104, pp. 125–136, Jul. 2017.
- [56] D. C. Montgomery and W. H. Woodall, "An overview of six sigma," *Int. Stat. Rev./Revue Internationale de Statistique*, vol. 76, pp. 329–346, Dec. 2008.
- [57] Q. Dai, X. Xu, W. Guo, S. Huang, and D. Filev, "Towards a systematic computational framework for modeling multi-agent decision-making at micro level for smart vehicles in a smart world," *Robot. Auto. Syst.*, vol. 144, Oct. 2021, Art. no. 103859.
- [58] S. Coskun, Q. Zhang, and R. Langari, "Receding horizon Markov game autonomous driving strategy," in *Proc. Amer. Control Conf. (ACC)*, Jul. 2019, pp. 1367–1374.
- [59] F. A. Potra and S. J. Wright, "Interior-point methods," *J. Comput. Appl. Math.*, vol. 124, nos. 1–2, pp. 281–302, Jan. 2000.
- [60] M. Wright, "The interior-point revolution in optimization: History, recent developments, and lasting consequences," *Bull. Amer. Math. Soc.*, vol. 42, no. 1, pp. 39–56, Jan. 2005.
- [61] S. Kousik, S. Vaskov, F. Bu, M. Johnson-Roberson, and R. Vasudevan, "Bridging the gap between safety and real-time performance in receding-horizon trajectory design for mobile robots," *Int. J. Robot. Res.*, vol. 39, no. 12, pp. 1419–1469, Oct. 2020.

- [62] K. Liu, N. Li, H. E. Tseng, I. Kolmanovsky, and A. Girard, "Interaction-aware trajectory prediction and planning for autonomous vehicles in forced merge scenarios," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 1, pp. 474–488, Jan. 2023.
- [63] L. Yang, X. Li, W. Guan, H. M. Zhang, and L. Fan, "Effect of traffic density on drivers' lane change and overtaking maneuvers in freeway situation—A driving simulator-based study," *Traffic Injury Prevention*, vol. 19, no. 6, pp. 594–600, Aug. 2018.
- [64] S. Dominguez, A. Ali, G. Garcia, and P. Martinet, "Comparison of lateral controllers for autonomous vehicle: Experimental results," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2016, pp. 1418–1423.
- [65] MathWorks. *Find Minimum of Function Using Genetic Algorithm*. Accessed: Nov. 17, 2023. [Online]. Available: <https://www.mathworks.com/help/gads/ga.html>
- [66] MathWorks. *Quadratic Programming*. Accessed: Nov. 17, 2023. [Online]. Available: <https://www.mathworks.com/help/optim/ug/quadprog.html>
- [67] *Next Generation Simulation (NGSIM) Program I-80 Videos*, UD Transportation Federal Highway Administration, Washington, DC, USA, 2023, doi: [10.21949/1504477](https://doi.org/10.21949/1504477).
- [68] M. R. Hajidavalloo, Z. Li, D. Chen, A. Louati, S. Feng, and W. B. Qin, "Mechanical system inspired microscopic traffic model: Modeling, analysis, and validation," *IEEE Trans. Intell. Vehicles*, vol. 8, no. 1, pp. 301–312, Jan. 2023.



Mushuang Liu (Member, IEEE) received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2016, and the Ph.D. degree from the University of Texas at Arlington, Arlington, TX, USA, in 2020.

She was a Post-Doctoral Researcher with the Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI, USA. She is currently an Assistant Professor with the Department of Mechanical and Aerospace Engineering, University of Missouri, Columbia, MO, USA. Her research interests include control, learning, and games for multiagent systems.



H. Eric Tseng received the B.S. degree from National Taiwan University, Taipei, Taiwan, in 1986, and the M.S. and Ph.D. degrees in mechanical engineering from the University of California at Berkeley, Berkeley, CA, USA, in 1991 and 1994, respectively.

In 1994, he joined Ford Motor Company, Dearborn, MI, USA, where he is currently a Senior Technical Leader of controls and automated systems in research and advanced engineering. Many of his contributed technologies led to production vehicle implementation. He has over 100 U.S. patents and over 120 publications.

Dr. Tseng is an NAE Member. His technical achievements have been recognized internally seven times with Ford's Highest Technical Award—the Henry Ford Technology Award and externally by the American Automatic Control Council with the Control Engineering Practice Award in 2013.



Dimitar Filev (Life Fellow, IEEE) is currently a Senior Henry Ford Technical Fellow in control and AI with research and advanced engineering with Ford Motor Company, Dearborn, MI, USA. He holds over 100 granted U.S. patents. His research interests include computational intelligence, AI, and intelligent control, and their applications to autonomous driving, vehicle systems, and automotive engineering.

Dr. Filev is a member of the National Academy of Engineering. He received the IEEE SMCS 2008 and the 2015 Computational Intelligence Pioneer's Award. He was the President of the IEEE Systems, Man, and Cybernetics Society from 2016 to 2017.



Anouck Girard (Senior Member, IEEE) received the Ph.D. degree in ocean engineering from the University of California at Berkeley, Berkeley, CA, USA, in 2002.

She has been with the University of Michigan, Ann Arbor, MI, USA, since 2006, where she is currently a Professor of aerospace engineering. She has coauthored the book *Fundamentals of Aerospace Navigation and Guidance* (Cambridge University Press, 2014). Her current research interests include vehicle dynamics and control systems.

Dr. Girard was a recipient of the Silver Shaft Teaching Award from the University of Michigan and the Best Student Paper Award from the American Society of Mechanical Engineers.



Ilya Kolmanovsky (Fellow, IEEE) received the Ph.D. degree in aerospace engineering from the University of Michigan, Ann Arbor, MI, USA, in 1995.

Before joining the University of Michigan as a Faculty Member in 2010, he was with Ford Research and Advanced Engineering, Dearborn, MI, USA, for close to 15 years. He is currently a Pierre T. Kabamba Collegiate Professor with the Department of Aerospace Engineering, University of Michigan. His research interests include control theory for systems with state and control constraints, and control applications to aerospace and automotive systems.

He is a fellow of IFAC and NAI and a Senior Editor of IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY.