

Potential Game-Based Decision-Making for Autonomous Driving

Mushuang Liu¹, Member, IEEE, Ilya Kolmanovsky², Fellow, IEEE, H. Eric Tseng³, Suzhou Huang⁴, Dimitar Filev⁵, Fellow, IEEE, and Anouck Girard⁶, Senior Member, IEEE

Abstract—Decision-making for autonomous driving is challenging, considering the complex interactions among multiple traffic agents (including autonomous vehicles (AVs), human-driven vehicles, and pedestrians) and the computational load needed to evaluate these interactions. This paper develops two general potential game-based frameworks, namely, finite and continuous potential games, for decision-making in autonomous driving. The two frameworks account for the AVs' two types of action spaces, i.e., finite and continuous action spaces, respectively. The developed frameworks provide theoretical guarantees for the existence of pure-strategy Nash equilibria and for the convergence of the Nash equilibrium (NE) seeking algorithms. The scalability challenge is also addressed. In addition, we provide cost function shaping approaches such that the agents' cost functions not only reflect common driving objectives but also yield potential games. The performance of the developed algorithms is demonstrated in diverse traffic scenarios, including intersection-crossing and lane-changing scenarios. Statistical comparative studies, including 1) finite potential game vs. continuous potential game, 2) best response dynamics vs. potential function optimization, and 3) potential game vs. reinforcement learning (RL) vs. control barrier function (CBF), are conducted to compare the robustness against various surrounding vehicles' strategies and to compare the computational efficiency. It is shown that the developed potential game frameworks have better robustness than RL and than CBF if the surrounding vehicles are not safety-conscious, and are computationally feasible for real-time implementation.

Index Terms—Autonomous driving, game theory, predictive control.

I. INTRODUCTION

DURING the past five years, the autonomous vehicle (AV) market has attracted more than \$50 billion in investment from major carmakers, tech giants and start-ups, and is expected to continue its rapid growth [1]. Although the benefits of AVs are substantial, technical challenges, such as intelligent decision-making in diverse traffic scenarios, still remain to be addressed before the AVs can routinely drive

Manuscript received 14 December 2021; revised 17 October 2022 and 10 February 2023; accepted 28 March 2023. Date of publication 12 April 2023; date of current version 2 August 2023. This work was supported by Ford-University of Michigan Alliance Project under Grant 000362-UM0220. The Associate Editor for this article was D. Wu. (*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).

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

H. Eric Tseng, Suzhou Huang, and Dimitar Filev are with the Ford Research and Innovation Center, Dearborn, MI 48124 USA (e-mail: htseng@ford.com; shuang10@ford.com; dfilev@ford.com).

Digital Object Identifier 10.1109/TITS.2023.3264665

on the roads [2], [3], [4]. Developing practical and reliable decision-making algorithms for AVs is challenging, due to the complexities of the interactions of multiple traffic participants and the requirement that decisions are made in real-time [5], [6], [7].

To generate safe and effective decisions for the ego vehicle, the interactions with its surrounding traffic agents, including AVs, human-driven vehicles, and pedestrians, have to be considered [8], [9], [10], [11]. As each traffic agent has its own objective (also called self-interest), the multi-agent decision-making problem is intrinsically a multi-player game problem [7], [12], [13], [14], [15]. In a game-theoretic setting, each agent aims to optimize its self-interest, which is affected by not only its own actions but also the actions of its surrounding agents. Along these lines, [16] develops a two-player nonzero-sum game to model the interactions between a highway vehicle and a merging vehicle in merging-giveaway scenarios. Papers [12], [13] consider a two-player Stackelberg game and a pairwise normal-form game, respectively, to address the decision-making in lane-changing scenarios. To capture the agents' interactions in intersection-crossing scenarios, a leader-follower game and a normal-form game are introduced, respectively, in [14] and [17]. To solve for Nash equilibrium multi-agent Markov games, a best response dynamics based solution algorithm is proposed in [18], with numerical examples in two-vehicle highway merging scenarios. Although these game-theoretic approaches have shown effectiveness in characterizing agents' interactions and in generating human-like negotiating behaviors, technical challenges, including existence of pure-strategy solutions, convergence of solution seeking algorithms, and computational scalability, still remain to be addressed before the game-theoretic approaches can be widely-accepted as a practical and reliable solution to autonomous driving. This paper aims to address these challenges by developing a novel potential game based framework integrated with receding horizon optimization.

Potential games [19] are a special class of multi-player games where a real-valued global function (called potential function) exists such that a change of an agent's self-interest by its own strategy deviation equals the change of the potential function. Potential games have appealing properties, including existence and attainability of pure-strategy Nash equilibria, and as such, have attracted increasing attention in economics [20], social systems [21], and wireless networks [22]. However, verifying whether a given game is a potential game is not easy [23], limiting broader applications of potential games to engineering problems such as autonomous driving.

This paper proposes a novel formulation of potential games to address the decision-making in autonomous driving. The contributions are fourfold:

- 1) A finite potential game-based framework is proposed, where the AV action space contains a finite number of elements. This framework features an integration of receding horizon optimization and potential games, and a cost function shaping approach such that the designed cost functions not only reflect common driving objectives but also yield potential games. Existence and convergence to pure-strategy NE is guaranteed, and the scalability challenge is also addressed.
- 2) A continuous potential game-based framework is developed, where the AV action space contains an infinite number of elements. Theoretical analysis, cost function shaping approaches, and scalable NE seeking algorithms are provided.
- 3) Practical formulations of the proposed potential games for specific traffic scenarios, including intersection-crossing and lane-changing scenarios, are provided. Numerical studies in these scenarios are conducted.
- 4) Statistical comparisons are performed for a) finite potential game vs. continuous potential game, b) best response dynamics vs. potential function optimization, and c) potential game vs. reinforcement learning vs. control barrier function based approaches. These comparisons highlight the robustness and computational efficiency characteristics of various algorithms.

This paper is organized as follows. Section II formulates the AV decision-making as a multi-player game problem. Section III develops a finite potential game framework for AVs in a setting of finite action spaces. Section IV develops a continuous potential game framework for AVs in a setting of continuous action spaces. Section V conducts numerical studies, and Section VI concludes the paper.

II. PROBLEM FORMULATION

Consider a group of traffic agents $\mathcal{N} = \{1, 2, \dots, N\}$ sharing the road. The dynamics of each agent are described by the following discrete-time equations.

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

where $i \in \mathcal{N}$, $x_i(t) \in \mathcal{X}_i$ and $u_i(t) \in \mathcal{U}_i$ are, respectively, the state and action of agent i at the time step t , \mathcal{X}_i and \mathcal{U}_i are the state space and action space of agent i , respectively, and f_i is the system evolution model of agent i . We denote the dimension of x_i and u_i as n_i and m_i , respectively. In autonomous driving applications, x_i usually contains the vehicle's position and velocity information, and u_i represents maneuvers such as braking and/or lane-changing.

In a driving scenario, every agent has its own objective, e.g., tracking its desired speed while avoiding collisions. We denote the performance index of agent i at time step t as $J_i(x_i(t), u_i(t), x_{-i}(t), u_{-i}(t))$, where x_{-i} and u_{-i} represent the states and actions of all other agents except for agent i , i.e., $x_{-i} = \{x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_N\}$, $u_{-i} = \{u_1, u_2, \dots, u_{i-1}, u_{i+1}, \dots, u_N\}$. Define $x(t)$ as the

global system state, $x(t) = \{x_i(t), x_{-i}(t)\}$, and $u(t)$ as the global action $u(t) = \{u_i(t), u_{-i}(t)\}$. At each time step t , an agent i plans its optimal action sequence (also referred to as strategy) for the next T ($T \geq 1$) time periods, i.e., $\mathbf{u}_i^*(t) = \{u_i^*(t), u_i^*(t+1), \dots, u_i^*(t+T-1)\}$, to minimize the cumulative cost over the prediction horizon $\mathcal{T} = \{t, t+1, \dots, t+T-1\}$, that is,

$$\begin{aligned} \mathbf{u}_i^*(t) &\in \underset{\mathbf{u}_i(t) \in \mathcal{S}_i}{\operatorname{argmin}} V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) \\ &= \underset{\mathbf{u}_i(t) \in \mathcal{S}_i}{\operatorname{argmin}} \sum_{\tau=t}^{t+T-1} J_i(x_i(\tau), u_i(\tau), x_{-i}(\tau), u_{-i}(\tau)), \end{aligned} \quad (2)$$

where \mathcal{S}_i is the strategy space of agent i and is determined by its action space \mathcal{U}_i . Denote $\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2 \times \dots \times \mathcal{S}_N$ and $\mathbf{u}(t) = \{\mathbf{u}_1(t), \mathbf{u}_{-1}(t)\}$, where $\mathbf{u}_{-i} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{i-1}, \mathbf{u}_{i+1}, \dots, \mathbf{u}_N\}$. Here $V_i^t : \mathcal{S} \rightarrow \mathbb{R}$ is the cumulative cost over the prediction horizon. It is assumed that the initial states of all agents are available, i.e., $x(t)$ is known to all $i \in \mathcal{N}$ at time t .

To enable an AV to respond timely to any changes in its environment (e.g., a sudden brake of the front vehicle), we employ the receding-horizon optimization: At time t , an AV plans its optimal action sequence $\mathbf{u}_i^*(t)$ according to (2), implements the first element $u_i^*(t)$ only, and then repeats the process at the next time step $t+1$ with a shifted horizon.

As shown in (2), agent i 's cost is affected by not only its own strategy $\mathbf{u}_i(t)$, but also the strategies of its surrounding agents $\mathbf{u}_{-i}(t)$. As such, the multi-agent optimization problem (2) is intrinsically a multi-player game problem.

We denote the game at t as $\mathcal{G}^t = \{\mathcal{N}, \mathcal{S}, \{V_i^t\}_{i \in \mathcal{N}}\}$, which is the receding horizon multi-player game we aim to solve. The set of all agents' optimal strategies $\{\mathbf{u}_1^*(t), \mathbf{u}_2^*(t), \dots, \mathbf{u}_N^*(t)\}$ that satisfy (2), if nonempty, yields a pure-strategy Nash equilibrium defined as follows.

Definition 1 (Pure-Strategy Nash Equilibrium [24]):
An N -tuple of strategies $\{\mathbf{u}_1^*(t), \mathbf{u}_2^*(t), \dots, \mathbf{u}_N^*(t)\}$ is a pure-strategy Nash equilibrium for the N -player game (2) if and only if

$$V_i^t(\mathbf{u}_i^*(t), \mathbf{u}_{-i}^*(t)) \leq V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}^*(t)), \quad \forall i \in \mathcal{N}, \forall \mathbf{u}_i(t) \in \mathcal{S}_i. \quad (3)$$

Equation (3) implies that if a pure-strategy NE is achieved, then no player would have the incentive to change its strategy unilaterally, and the system is in an equilibrium state. Since we are considering autonomous driving applications and the strategy usually reflects the AV maneuvers, we only focus on pure-strategy NE in this paper, and the word "pure-strategy" may be omitted if no confusion.

Although the receding horizon multi-player game (2) can model the AV decision-making nicely, solving such a multi-player game is not straightforward in general. The challenges include 1) Given arbitrary V_i^t , the NE $\mathbf{u}^*(t)$ that satisfies (3) may not always exist [25]; 2) Even if there exists a NE, the NE seeking algorithm, e.g., the best-response dynamics, may not always converge [19]; 3) Even if the algorithm converges, solving for NE is computationally expensive in general, and the computational complexity can exponentially increase

with the number of agents [26]. These challenges limit the practical applicability of multi-player games to autonomous driving. We aim to address these challenges in this paper by formulating the game (2) as a potential game by designing the AV cost functions, V_i^t , appropriately for various driving scenarios so that they can reflect common driving objectives while leading to potential games.

III. FINITE POTENTIAL GAME

In this section, we consider the multi-player game (2) in the setting of finite strategy spaces. Specifically, \mathcal{S}_i has a finite number of elements for all $i \in \mathcal{N}$.

A. Preliminaries

We first recall the definition of finite potential games.

Definition 2 (Finite Exact Potential Game [27]): The game (2) is a finite exact potential game if and only if the strategy space \mathcal{S}_i ($i \in \mathcal{N}$) contains a finite number of elements and a potential function $F^t : \mathcal{S} \rightarrow \mathbb{R}$ exists such that, $\forall i \in \mathcal{N}$,

$$\begin{aligned} V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) - V_i^t(\mathbf{u}'_i(t), \mathbf{u}_{-i}(t)) \\ = F^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) - F^t(\mathbf{u}'_i(t), \mathbf{u}_{-i}(t)), \\ \forall \mathbf{u}_i(t), \mathbf{u}'_i(t) \in \mathcal{S}_i, \text{ and } \forall \mathbf{u}_{-i}(t) \in \mathcal{S}_{-i}. \end{aligned} \quad (4)$$

Equation (4) implies that the change of agent i 's cumulative cost caused by its own strategy deviation leads to exactly the same amount of change in the potential function. Note that the potential function F^t does not have a subscript, implying that it is the same for all agents. This paper only considers exact potential games, and the word “exact” will subsequently be omitted.

Potential games have appealing properties. Two of these properties are summarized by the following two lemmas, which will be used to facilitate the analysis in this paper.

Lemma 1 (Existence of Pure-Strategy Nash Equilibria [19]): If the game (2) is a finite exact potential game, then there exists at least one pure-strategy Nash equilibrium.

Lemma 2 (Equivalence of Nash Equilibrium Sets [27]): If the game (2) is a finite exact potential game with F^t as a potential function, then the set of pure-strategy Nash equilibria of (2) coincides with the set of pure-strategy Nash equilibria of the identical-interest game with all agents' cumulative cost equal to the potential function. That is,

$$NESet(\mathcal{G}^t) = NESet(\tilde{\mathcal{G}}^t), \quad (5)$$

where $NESet$ denotes the set of pure-strategy NE, and $\tilde{\mathcal{G}}^t = (\mathcal{N}, \mathcal{A}, \{F^t\}_{i \in \mathcal{N}})$.

B. Constructing Finite Potential Games for Autonomous Driving

Consider the decision-making in autonomous driving according to (2). In this subsection, we show how to design V_i^t such that (2) is a finite potential game.

Theorem 1 shows that if V_i^t depends only on $\mathbf{u}_i(t)$, then the resulting game is a finite potential game.

Theorem 1: Let the cumulative cost in (2) be of the following form,

$$V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) = V_i^{t, self}(\mathbf{u}_i(t)), \quad (6)$$

where $V_i^{t, self}$ is a function that is determined solely by agent i 's strategy $\mathbf{u}_i(t)$. Then the game (2) is a finite exact potential game with a potential function P^t defined as

$$P^t(\mathbf{u}(t)) = \sum_{i \in \mathcal{N}} V_i^{t, self}(\mathbf{u}_i(t)). \quad (7)$$

Proof: Given the cumulative cost (6), for any unilateral strategy deviation of agent i from \mathbf{u}_i to \mathbf{u}'_i , the following equation holds,

$$\begin{aligned} V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) - V_i^t(\mathbf{u}'_i(t), \mathbf{u}_{-i}(t)) \\ = V_i^{t, self}(\mathbf{u}_i(t)) - V_i^{t, self}(\mathbf{u}'_i(t)). \end{aligned} \quad (8)$$

Consider the function P^t in (7). The strategy deviation from $\mathbf{u}_i(t)$ to $\mathbf{u}'_i(t)$ leads to the change of this function as

$$\begin{aligned} P^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) - P^t(\mathbf{u}'_i(t), \mathbf{u}_{-i}(t)) \\ = V_i^{t, self}(\mathbf{u}_i(t)) + \sum_{j \in \mathcal{N}, j \neq i} V_j^{t, self}(\mathbf{u}_j(t)) \\ - V_i^{t, self}(\mathbf{u}'_i(t)) - \sum_{j \in \mathcal{N}, j \neq i} V_j^{t, self}(\mathbf{u}_j(t)) \\ = V_i^{t, self}(\mathbf{u}_i(t)) - V_i^{t, self}(\mathbf{u}'_i(t)). \end{aligned} \quad (9)$$

Because $V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) - V_i^t(\mathbf{u}'_i(t), \mathbf{u}_{-i}(t)) = P^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) - P^t(\mathbf{u}'_i(t), \mathbf{u}_{-i}(t))$ holds for all $i \in \mathcal{N}$, $\mathbf{u}_i(t), \mathbf{u}'_i(t) \in \mathcal{S}_i$, and $\mathbf{u}_{-i}(t) \in \mathcal{S}_{-i}$, the function P^t is a potential function, and the formulated game is a finite potential game. \square

Remark 1: Theorem 1 states that if the cumulative cost V_i^t satisfies (6), then the game (2) is a finite potential game. Such a cost can characterize self-centered driving objectives, such as ride comfort (often reflected by vehicle jerk [28]), travel efficiency (captured by vehicle speed [29]), fuel efficiency (dependent on speed, acceleration, and jerk [30]), keeping in the center of a lane (reflected by vehicle lateral position), and so on. A specific cost function that motivates the AV to track a desired speed and that satisfies (6) is provided in the simulation section by (30).

Theorem 2 considers the cases where an AV's objective is jointly affected by both $\mathbf{u}_i(t)$ and $\mathbf{u}_{-i}(t)$, in a pairwise symmetric manner.

Theorem 2: Let the cumulative cost in (2) be of the following form,

$$V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) = \sum_{j \in \mathcal{N}, j \neq i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)), \quad (10)$$

where $V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)) = V_{ji}^t(\mathbf{u}_j(t), \mathbf{u}_i(t))$, $\forall i, j \in \mathcal{N}, i \neq j, \forall \mathbf{u}_i(t) \in \mathcal{S}_i, \forall \mathbf{u}_j(t) \in \mathcal{S}_j$. Then the game (2) is a finite exact potential game with the following potential function,

$$G^t(\mathbf{u}(t)) = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}, j < i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)). \quad (11)$$

Proof: The strategy deviation of agent i from \mathbf{u}_i to \mathbf{u}'_i leads to the change in its cumulative cost.

$$\begin{aligned} V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) - V_i^t(\mathbf{u}'_i(t), \mathbf{u}_{-i}(t)) \\ = \sum_{j \in \mathcal{N}, j \neq i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)) - \sum_{j \in \mathcal{N}, j \neq i} V_{ij}^t(\mathbf{u}'_i(t), \mathbf{u}_j(t)). \end{aligned} \quad (12)$$

The deviation from $\mathbf{u}_i(t)$ to $\mathbf{u}'_i(t)$ leads to the change in the function G^t defined in (11) as

$$\begin{aligned} G^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) - G^t(\mathbf{u}'_i(t), \mathbf{u}_{-i}(t)) \\ = \sum_{j \in \mathcal{N}, j \neq i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)) - \sum_{j \in \mathcal{N}, j \neq i} V_{ij}^t(\mathbf{u}'_i(t), \mathbf{u}_j(t)). \end{aligned} \quad (13)$$

Equation (13) holds because $V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)) = V_{ji}^t(\mathbf{u}_j(t), \mathbf{u}_i(t)) \forall i, j \in \mathcal{N}, i \neq j, \forall \mathbf{u}_i(t) \in \mathcal{S}_i, \forall \mathbf{u}_j(t) \in \mathcal{S}_j$. As such, the function G^t is a potential function, and the formulated game is a finite exact potential game. \square

Remark 2: Theorem 2 states that if the cumulative cost V_i^t satisfies (10), then the game (2) is a finite potential game. In a typical driving scenario, vehicles interactions are often brought about by the desire of avoiding collisions. The cost design in (10) can characterize such an interaction, by applying a symmetric collision penalty to both vehicles if a collision between them happens. A specific example is provided in Section V by (32) and (33).

Next theorem shows that if an AV's objective function contains both self-dependent and pairwise components, then the resulted game is still a potential game.

Theorem 3: Let the cumulative cost in (2) be of the following form,

$$\begin{aligned} V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) \\ = \alpha V_i^{t, \text{self}}(\mathbf{u}_i(t)) + \beta \sum_{j \in \mathcal{N}, j \neq i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)), \end{aligned} \quad (14)$$

where $V_i^{t, \text{self}}(\mathbf{u}_i(t))$ and $V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t))$ satisfy (6) and (10), respectively, and $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}$. Then the game (2) is a finite exact potential game with the following potential function,

$$\begin{aligned} F^t(\mathbf{u}(t)) \\ = \alpha P^t(\mathbf{u}(t)) + \beta G^t(\mathbf{u}(t)) \\ = \alpha \sum_{i \in \mathcal{N}} V_i^{t, \text{self}}(\mathbf{u}_i(t)) + \beta \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}, j < i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)). \end{aligned} \quad (15)$$

Proof: The proof is straightforward by combining Theorems 1 and 2. \square

Remark 3: Theorem 3 provides a general template to design AV cost functions that yield finite potential games. The cost (14) can reflect driving objectives of, for example, tracking a desired speed while avoiding collisions with other vehicles. We note that most of the AVs' cost function models in the existing works follow, or can be slightly revised to follow, the form proposed in (14). See [31], [32], [33], and [34] for examples.

Algorithm 1 Best Response Dynamics to Solve (2)

Inputs:

Agent set: \mathcal{N} ;

Global system state: $x(t)$;

Global strategy space: \mathcal{S} ;

Cumulative cost function of each agent: $V_i^t, \forall j \in \mathcal{N}$;

System dynamics of each agent: $f_j, \forall j \in \mathcal{N}$.

Output:

The ego vehicle optimal strategy (in the sense of NE): $\mathbf{u}_i^*(t)$.

Procedures:

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

$$\mathbf{u}_j^*(t) \in \operatorname{argmin}_{\mathbf{u}_j(t) \in \mathcal{S}_j} V_j^t(\mathbf{u}_j(t), \mathbf{u}_{-j}(t)).$$

- 5: Update $\mathbf{u}_j(t)$ using $\mathbf{u}_j^*(t)$.

End for

- 7: **If** $\mathbf{u}_j^*(t) \in \operatorname{argmin}_{\mathbf{u}_j(t) \in \mathcal{S}_j} V_j^t(\mathbf{u}_j^*(t), \mathbf{u}_{-j}^*(t))$

holds $\forall j \in \mathcal{N}$,

- 9: **Then** Set $NashCondition=True$.

10: **End if**

11: **End while**

C. Solving the Constructed Finite Potential Games

After formulating the game (2) as a finite potential game, according to Lemma 1, a pure-strategy NE always exists. Next we show how to solve this potential game and find the NE.

One of the most commonly-used algorithms to seek NE is the best response dynamics [26], which works as follows: Given an initial guess of all agents' strategies, we find the best response of each agent to the strategies of others, i.e., $\mathbf{u}_j^*(t) \in \operatorname{argmin}_{\mathbf{u}_j(t) \in \mathcal{S}_j} V_j^t(\mathbf{u}_j(t), \mathbf{u}_{-j}(t))$ for all $j \in \mathcal{N}$; After all agents' strategies are updated, we then repeat the process iteratively until no agent has the incentive to change its strategy. The detailed algorithm is shown in Algorithm 1.

Theorem 4: [Convergence of Algorithm 1] Consider the game (2) and the Algorithm 1. If the cumulative cost V_i^t in (2) is designed according to (14), then Algorithm 1 converges within a finite number of iterations.

Proof: According to Theorem (3), if the V_i^t in (2) satisfies (14), then the game (2) is a finite potential game. Because the best response dynamics in Algorithm 1 generates an improvement path, and because every improvement path must terminate within a finite number of iteration steps in finite potential games [19], Algorithm 1 converges to NE within a finite number of iterations. \square

Although the convergence of the best response dynamics is guaranteed, this algorithm is generally computationally expensive, as it requires multiple and iterative optimizations at each t . Next we show that if the game is a potential game, then the pure-strategy NE can be derived by solving only one optimization at each t :

$$\mathbf{u}^*(t) \in \operatorname{argmin}_{\mathbf{u}(t) \in \mathcal{S}} F^t(\mathbf{u}(t)), \quad (16)$$

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

Agent set: \mathcal{N} ;
 Global system state: $x(t)$;
 Global strategy space: \mathcal{S} ;
 Cumulative cost function of each agent: $V_i^t, \forall j \in \mathcal{N}$;
 System dynamics of each agent: $f_j, \forall j \in \mathcal{N}$.

Output:

The ego vehicle optimal strategy (in the sense of NE): $\mathbf{u}_i^*(t)$.

Procedure:

- 1: Find potential function F^t according to (15).
- 2: Find $\mathbf{u}^*(t)$ according to (16).

where F^t is the potential function of the game (2). The detailed algorithm that uses this potential function optimization approach to solve the game (2) is shown in Algorithm 2.

Theorem 5: Consider the game (2) and the Algorithm 2. If the cumulative cost V_i^t in (2) is designed according to (14), then $\mathbf{u}^*(t)$ from (16) is a pure-strategy NE.

Proof: According to Theorem 3, if the V_i^t in (2) satisfies (14), then the game (2) is a finite potential game. According to Lemma 2, if (2) is a finite potential game, then the NE can be derived by solving the identical-interest game (16). \square

Remark 4: In autonomous driving, real-time evaluation and planning are critical for AVs to generate timely response to any changes in the environment. If the game (2) is not a potential game, then solving it generally requires multiple and iterative optimizations, as shown in Algorithm 1; these could be too time-consuming to be employed in autonomous driving.

In addition, the solutions from Algorithms 1 and 2 may be different due to the existence of multiple NE. In this case, the one from Algorithm 2 should be preferable from a “social cost” perspective. It is because the solution from Algorithm 2 is not only a NE but also a global minimizer of the “social cost” characterized by the potential function F^t , which usually takes into consideration all agents’ cost as shown in (15). In an autonomous driving setting, it means that if a maneuver can benefit not only the AV itself (in the sense of individual optimality, $\min V_i^t$), but also the surrounding vehicles and pedestrians (in the sense of social optimality, $\min F^t$), then it should be preferable to the ones that are only individually optimal. Indeed the AV is expected to be considerate to other agents on the road, instead of caring for its self-interest only.

IV. CONTINUOUS POTENTIAL GAMES

In this section, we consider the multi-player game (2) with continuous strategy spaces \mathcal{S}_i . As a continuous action space contains infinitely many elements, the finite potential game results investigated in Section III may not always hold. Therefore, in this section, we develop results for continuous potential games.

A. Preliminaries

Let us first define continuous potential games.

Definition 3 (Continuous Exact Potential Game): The game (2) is a continuous exact potential game if and only if $\mathcal{S}_i \subset \mathbb{R}^{m_i}$ is a connected set and a potential

function $F^t : \mathcal{S} \rightarrow \mathbb{R}$ exists such that, $\forall i \in \mathcal{N}$, $\forall \mathbf{u}_i(t) \in \mathcal{S}_i, \forall \mathbf{u}_{-i}(t) \in \mathcal{S}_{-i}$,

$$\frac{\partial V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t))}{\partial \mathbf{u}_i(t)} = \frac{\partial F^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t))}{\partial \mathbf{u}_i(t)}, \quad (17)$$

where V_i^t and F^t are everywhere differentiable on an open superset of \mathcal{S} , and \mathcal{S}_i is nonempty $\forall i \in \mathcal{N}$.

Note that the above definition is slightly different from the one in [27]. Specifically, \mathbf{u}_i is defined as a scalar in [27], while it is a vector in Definition 3. With this generalization, to ensure that the theoretical analysis still holds, we include the proofs of each lemma in this section in the Appendix.

Proposition 1 shows the equivalence of the finite and continuous potential games under certain conditions.

Proposition 1: Assume that V_i^t and F^t are everywhere differentiable on an open superset of \mathcal{S} , and \mathcal{S}_i is nonempty $\forall i \in \mathcal{N}$. If $\mathcal{S}_i \subset \mathbb{R}^{m_i}$ is a compact (i.e., bounded and closed) and connected set $\forall i \in \mathcal{N}$, then (17) holds if and only if (4) holds.

Proof: See Appendix A. \square

Lemma 3 (Equivalence of Nash Equilibrium Sets): If the game (2) is a continuous exact potential game and F^t is a potential function, then the set of pure-strategy Nash equilibria of (2) coincides with the set of pure-strategy Nash equilibria for the identical-interest game with all agents’ cumulative cost equal to the potential function F^t .

Proof: See Appendix B. \square

Lemma 4 (Existence of Pure-Strategy Nash Equilibria): If the game (2) is a continuous exact potential game with compact \mathcal{S} , then the game has at least one pure-strategy Nash equilibrium. Moreover, if the potential function F^t is strictly convex, then the pure-strategy Nash equilibrium is unique.

Proof: See Appendix C. \square

B. Constructing Continuous Potential Games for Autonomous Driving

In the next theorems, we show that the cost function shaping approaches developed in Section III also work in continuous potential games. Theorem 6 considers self-centered driving objectives, Theorem 7 considers pairwise interactions, and Theorem 8 considers a mix of these two types of costs.

Theorem 6: Assume that the cumulative cost in (2) are everywhere differentiable on an open superset of \mathcal{S} , and

$$V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) = V_i^{t,\text{self}}(\mathbf{u}_i(t)), \quad (18)$$

where $V_i^{t,\text{self}}(\mathbf{u}_i(t))$ is dependent solely on $\mathbf{u}_i(t)$. Then the game (2) is a continuous potential game with the potential function,

$$P^t(\mathbf{u}(t)) = \sum_{i \in \mathcal{N}} V_i^{t,\text{self}}(\mathbf{u}_i(t)). \quad (19)$$

Proof: The following equation holds.

$$\begin{aligned} \frac{\partial P^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t))}{\partial \mathbf{u}_i(t)} &= \frac{\partial \sum_{j \in \mathcal{N}} V_j^{t,\text{self}}(\mathbf{u}_j(t))}{\partial \mathbf{u}_i(t)} \\ &= \frac{\partial V_i^{t,\text{self}}(\mathbf{u}_i(t))}{\partial \mathbf{u}_i(t)} \\ &= \frac{\partial V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t))}{\partial \mathbf{u}_i(t)}. \end{aligned} \quad (20)$$

As such, according to Definition 3, the game is a continuous potential game, and P^t is a potential function. \square

Theorem 7: Assume that the cumulative cost in (2) are everywhere differentiable on an open superset of \mathcal{S} , and

$$V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) = \sum_{j \in \mathcal{N}, j \neq i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)), \quad (21)$$

where $V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)) = V_{ji}^t(\mathbf{u}_j(t), \mathbf{u}_i(t))$, $\forall i, j \in \mathcal{N}, i \neq j, \forall \mathbf{u}_i(t) \in \mathcal{S}_i, \forall \mathbf{u}_j(t) \in \mathcal{S}_j$. Then the game (2) is a continuous potential game with the potential function,

$$G^t(\mathbf{u}(t)) = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}, j < i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)). \quad (22)$$

Proof: Because $V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)) = V_{ji}^t(\mathbf{u}_j(t), \mathbf{u}_i(t))$, $\forall i, j \in \mathcal{N}, i \neq j$, the following equation holds.

$$\begin{aligned} & \frac{\partial G^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t))}{\partial \mathbf{u}_i(t)} \\ &= \frac{\partial \left(\sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}, j < i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)) \right)}{\partial \mathbf{u}_i(t)} \\ &= \frac{\partial \left(\sum_{j \in \mathcal{N}, j \neq i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)) \right)}{\partial \mathbf{u}_i(t)} = \frac{\partial V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t))}{\partial \mathbf{u}_i(t)}. \end{aligned} \quad (23)$$

As such, the game is a continuous potential game with G^t as the potential function. \square

Theorem 8: Assume that the cumulative cost in (2) are everywhere differentiable on an open superset of \mathcal{S} , and

$$\begin{aligned} & V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) \\ &= \alpha V_i^{t, self}(\mathbf{u}_i(t)) + \beta \sum_{j \in \mathcal{N}, j \neq i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)), \end{aligned} \quad (24)$$

where $V_i^{t, self}(\mathbf{u}_i(t))$ and $V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t))$ satisfy the requirements in Theorems 6 and 7, respectively. Then the game (2) is a continuous potential game with the potential function,

$$\begin{aligned} F^t(\mathbf{u}(t)) &= \alpha P^t(\mathbf{u}(t)) + \beta G^t(\mathbf{u}(t)) \\ &= \alpha \sum_{i \in \mathcal{N}} V_i^{t, self}(\mathbf{u}_i(t)) \\ &\quad + \beta \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}, j < i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)). \end{aligned} \quad (25)$$

Proof: This theorem is straightforward by combining Theorems 6 and 7. \square

C. Solving the Constructed Continuous Potential Games

As the action space contains infinitely many elements in continuous potential games, the convergence of the best response dynamics may not be achieved within a finite number of iterations. To address this issue, we define an ϵ -Nash equilibrium.

Definition 4 (ϵ -Nash Equilibrium [24]): An N -tuple of strategies $\{\mathbf{u}_1^, \mathbf{u}_2^*, \dots, \mathbf{u}_N^*\}$ is an ϵ -Nash equilibrium for an N -player game if and only if $\exists \epsilon \in \mathbb{R}_+$ such that, $\forall i \in \mathcal{N}, \forall \mathbf{u}_i(t) \in \mathcal{S}_i$,*

$$V_i^t(\mathbf{u}_i^*(t), \mathbf{u}_{-i}^*(t)) \leq V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}^*(t)) + \epsilon. \quad (26)$$

With the ϵ -Nash equilibrium, we can use the best response dynamics, i.e., Algorithm 1, or the potential function optimization, i.e., Algorithm 2, to solve the game, if appropriate changes are made. Specifically, to use Algorithm 1 and to ensure the convergence within a finite number of iteration, the *NashCondition* in Procedure 7 should be replaced by “ $V_j^t(\mathbf{u}_j^*(t), \mathbf{u}_{-j}^*(t)) \leq V_j^t(\mathbf{u}'_j(t), \mathbf{u}_{-j}^*(t)) + \epsilon$ holds $\forall \mathbf{u}'_j \in \mathcal{S}_j, \forall j \in \mathcal{N}$ ”; and the update rule in Procedure 5 should be replaced by “Update $\mathbf{u}_j(t)$ using $\mathbf{u}_j^*(t)$ if $V_j^t(\mathbf{u}_j^*(t), \mathbf{u}_{-j}(t)) \leq V_j^t(\mathbf{u}_j(t), \mathbf{u}_{-j}(t)) - \epsilon$ ”. In Algorithm 2, the potential function F^t should be replaced by (25).

Next lemma shows the guaranteed convergence of the best response algorithm.

Theorem 9 (Convergence to ϵ -Nash Equilibrium):

Consider the game (2) and the Algorithm 1. If the cumulative cost V_i^t in (2) satisfies (24) with \mathcal{S} being a connected and compact set, and if Algorithm 1 is modified according to the ϵ -Nash equilibrium condition as specified above, then Algorithm 1 converges within a finite number of iterations.

Proof: According to Theorem 8, if the cumulative cost V_i^t in (2) satisfies (24), and if \mathcal{S} is a connected set, then the game (2) is a continuous potential game. Because the potential function F^t is continuous in the compact set \mathcal{S} , F^t is bounded. Therefore, the ϵ -improvement path generated by Algorithm 1 must terminate within a finite number of iterations because the ϵ -increment of the potential function is finite. When an ϵ -improvement path comes to an end point, there exists no strategy profile that leads to a decrease of cumulative cost greater than ϵ for any player. As such, this end point is an ϵ -Nash equilibrium. \square

V. NUMERICAL RESULTS

In this section, we apply the developed finite and continuous potential game frameworks to specific traffic scenarios, including intersection-crossing scenarios and lane-changing scenarios.

The vehicles' dynamics are described by a kinematic bicycle model [35], [36]:

$$\begin{aligned} x_i(t+1) &= x_i(t) + v_i(t)\cos(\phi_i(t) + \beta_i(t))\Delta t, \\ y_i(t+1) &= y_i(t) + v_i(t)\sin(\phi_i(t) + \beta_i(t))\Delta t, \\ v_i(t+1) &= v_i(t) + a_i(t)\Delta t, \\ \phi_i(t+1) &= \phi_i(t) + \frac{v_i(t)}{l_r}\sin(\beta_i(t))\Delta t, \\ \beta_i(t+1) &= \beta_i(t) + \tan^{-1}\left(\frac{l_r}{l_r + l_f}\tan(\delta_{i,f}(t))\right)\Delta t, \end{aligned} \quad (27)$$

where $i = 1, \dots, N$ designates the i^{th} agent; x_i and y_i are the longitudinal and lateral position of the center of mass of vehicle i along x and y axes, respectively; v_i and a_i are the velocity and acceleration of the center of mass of vehicle i , respectively; β_i is the angle of the velocity with respect to the longitudinal axis of the vehicle; ϕ_i is the inertial heading; l_f and l_r are the lengths from the center of mass to the front and rear ends of the car, respectively, and are selected as $l_f = l_r = 1.5$ m; $\delta_{i,f}$ is the steering angle of the front wheels, and $|\delta_{i,f}| \leq 20^\circ$. Because for most vehicles, the rear wheels

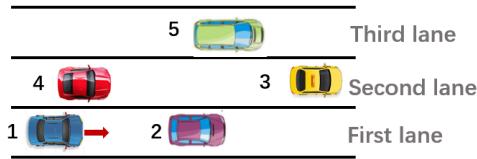


Fig. 1. The lane-changing scenario.

cannot be steered, the rear wheel steering angles are assumed to be zero [35]. The inputs of each vehicle are the acceleration $a_i(t)$ and the steering angle $\delta_{i,f}(t)$.

The following numerical studies are conducted: Section V-A applies the finite potential game framework to lane-changing scenarios. Section V-B applies the continuous potential game framework to intersection-crossing scenarios. Section V-C compares the performance of various algorithms in intersection-crossing scenarios, and Section V-D compares potential games with other decision-making approaches including reinforcement learning and control barrier function based approaches.

A. Finite Potential Game in Lane-Changing Scenarios

Consider the three-lane highway scenario pictured in Figure 1. When driving in highways, a vehicle needs to determine 1) whether to accelerate or decelerate, and 2) which lane to drive in. Therefore, we define the action space of each vehicle as

$$\mathcal{U}_i = \mathcal{A}_i \times \mathcal{L}_i, \quad (28)$$

where

$$\mathcal{A}_i = \{\text{hard brake, brake, mild brake, maintain speed,}\}$$

$$\quad \text{mild accelerate, accelerate, rapid accelerate}\},$$

$$\mathcal{L}_i = \{\text{change to the left lane, maintain the current lane,}\}$$

$$\quad \text{change to the right lane}\}.$$

In (28), both \mathcal{A}_i and \mathcal{L}_i are finite sets, and therefore, \mathcal{U}_i is a finite action space. We let $a_i(t) = -3 \text{ m/s}^2$ (resp. 3 m/s^2) if “hard brake” (resp. “rapid accelerate”) is selected, $a_i(t) = -2 \text{ m/s}^2$ (resp. 2 m/s^2) if “brake” (resp. “accelerate”) is selected, $a_i(t) = -1 \text{ m/s}^2$ (resp. 1 m/s^2) if “mild brake” (resp. “mild accelerate”) is selected, and $a_i(t) = 0 \text{ m/s}^2$ if “maintain speed” is selected. If “change to the left lane” (resp. “change to the right lane”) is selected, then a constant steering angle $\delta_{i,f}(t) = 0.9^\circ$ (resp. -0.9°) is added to the steering angle required to negotiate the road, if any, until the AV is in the center of the target lane. Once the AV reaches the center line, a lane centering controller is activated at 10 Hz sampling rate, applying the steering required to reduce the relative heading between the vehicle and the road. Specifically, the steering angle induces the vehicle side slip angle $\beta_i(t)$ in the amount of $-\frac{1}{2}(\phi_i(t) - \phi_r(t))$, where $\phi_r(t)$ depends on the road orientation. On the other hand, if the decision is “stay in the current lane”, then $\delta_{i,f}(t)$ remains at the steering angle corresponding to the curvature of the road (or that of the turning lanes in an intersection).

In our simulation, we use a simplified strategy space: $\mathbf{u}_i(t) = \{u_i(t), u_i(t+1), \dots, u_i(t+T-1)\}$, where $u_i(t+\tau) =$

$u_i(t) \in \mathcal{U}_i, \forall \tau = 0, 1, \dots, T-1$. This simplified strategy space represents that the AV plans to continue the same action over the horizon T . This assumption is reasonable as it is consistent with a common driving experience. When a human driver plans a maneuver (e.g., lane-changing), he/she usually expects to continue the maneuver for some time to complete it, instead of planning a sequence of different maneuvers for every Δt . Note that although the AV plans one maneuver for the next T periods, it may change its mind and select a different maneuver after Δt , triggered by the change of its environment and by the receding horizon optimization conducted at every t . In the simulation, we select T to correspond to 4 s to cover a lane-changing maneuver duration [37] and select $\Delta t = 0.5 \text{ s}$ to enable the ego vehicle to respond timely to the change of its surrounding vehicles’ states and behaviors [38]. Note that a larger T and a smaller Δt lead to increased computational time and effort for the decision-making algorithm.

In such a multi-vehicle scenario, the driving performance of the ego vehicle (labeled with the number “1” in Figure 1) is correlated with its surrounding vehicles’ states (i.e., positions and velocities) and actions (i.e., lane-changing and/or accelerations). To take into consideration such interactions, at every t , the ego vehicle solves an N -player game as described in (2) to generate its optimal strategy. We set $N = 5$ in our simulations and let vehicles’ initial positions be similar to the ones in Figure 1.

The cumulative cost of each vehicle is designed as

$$\begin{aligned} V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) \\ = \theta_{i,1} V_i^{t,\text{self},1}(\mathbf{u}_i(t)) + \theta_{i,2} V_i^{t,\text{self},2}(\mathbf{u}_i(t)) \\ + \theta_{i,3} \sum_{j \in \mathcal{N}, j \neq i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)), \end{aligned} \quad (29)$$

where $\theta_{i,1}$, $\theta_{i,2}$ and $\theta_{i,3}$ are constant parameters to tune vehicle i ’s driving styles or aggressiveness. In our simulations, we select $\theta_{i,1} = \theta_{i,2} = 1$ and $\theta_{i,3} = 4$ for all $i \in \mathcal{N}$. Equation (29) facilitates a framework for modeling the surrounding vehicles with a variety of human driver behaviors, where the parameters can be learned or calibrated from naturalistic traffic dataset [31], [39] using inverse optimal control or inverse reinforcement learning [40]. The first term $V_i^{t,\text{self},1}(\mathbf{u}_i(t))$, is to motivate the vehicle to track the desired speed, and is given by

$$V_i^{t,\text{self},1}(\mathbf{u}_i(t)) = \sum_{\tau=t}^{t+T-1} \left(\frac{v_i(\tau) - v_{i,d}}{v_{i,d}} \right)^2, \quad (30)$$

where $v_{i,d}$ is the desired speed of vehicle i . For the ego vehicle, we let $v_{1,d} = 27 \text{ m/s}$ (approximately 60 mph). For the surrounding vehicles, we let $v_{i,d}$ be within the range $v_{i,d} \in [22, 30] \text{ m/s}$ for $i = 2, 3, 4, 5$. (Multiple scenarios with various ego vehicle desired speed in the range of $[10, 30] \text{ m/s}$ are tested, and the ego vehicle performance remains similar in all tested scenarios.) The second term $V_i^{t,\text{self},2}(\mathbf{u}_i(t))$ is to prevent the vehicle from driving outside the road, and is

given by

$$V_i^{t,self,2}(\mathbf{u}_i(t)) = \sum_{\tau=t}^{t+T-1} \gamma \cdot \mathbf{1}_{\{y_i(\tau) < 0 \text{ or } y_i(\tau) > W_d\}} \quad (31)$$

where W_d is the width of the road, $\mathbf{1}_{\{condition\}} = 1$ if *condition* is true and 0 otherwise. The constant γ represents the penalty for driving out of the road and should be a sufficiently large number. We select $\gamma = 1000$ in our simulation while noting that the AV behavior is relatively insensitive to this value as long as it is chosen sufficiently large. The third term $\sum_{j \in \mathcal{N}, j \neq i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t))$ is to avoid collision:

$$V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)) = \sum_{\tau=t}^{t+T-1} J_{ij}(x_i(\tau), y_i(\tau), x_j(\tau), y_j(\tau)), \quad (32)$$

and

$$\begin{aligned} J_{ij}(x_i(\tau), y_i(\tau), x_j(\tau), y_j(\tau)) \\ = \left(\tanh \left(\beta(d_{x,c}^2 - (x_i(\tau) - x_j(\tau))^2) \right) + 1 \right) \\ \cdot \left(\tanh \left(\beta(d_{y,c}^2 - (y_i(\tau) - y_j(\tau))^2) \right) + 1 \right), \end{aligned} \quad (33)$$

where $d_{x,c}$ and $d_{y,c}$ are the longitudinal and lateral collision distances, respectively, and $d_{x,c} = 7 \text{ m}$ and $d_{y,c} = 4.5 \text{ m}$ in the simulation. The parameter β should be sufficiently large such that the \tanh function always takes the two extreme values -1 or 1 , and we set $\beta = 1000$ in the simulation. The cost (33) means that at time τ , if vehicle j is a threat to vehicle i both laterally and longitudinally, then $J_{ij}(x_i(\tau), y_i(\tau), x_j(\tau), y_j(\tau)) \approx 4$; Otherwise, $J_{ij}(x_i(\tau), y_i(\tau), x_j(\tau), y_j(\tau)) \approx 0$. Note that the lane width is 5 m , and the lateral collision distance $d_{y,c}$ is selected to be a bit smaller than the road width to realize the feature that if a vehicle is not in the same lane or in the target lane of the ego vehicle (target lane refers to the lane that the ego vehicle aims to change to), then it should not be considered as a threat to the ego vehicle. The necessity of this feature is illustrated by a comparative study at <https://youtu.be/aWXTfLdovfc>.

According to Theorem 3, the above cost function design makes the N -player game a finite potential game. Therefore, we can employ the potential function optimization algorithm, i.e., Algorithm 2, to solve the game. We let the ego vehicle adopt the generated NE as its strategy, assuming that the surrounding vehicles also behave to optimize their own driving performances characterized by (29). As the actual behavior of the surrounding vehicles may deviate from the assumed one, we perform statistical studies incorporating various surrounding vehicle behaviors in Section V-C. In the specific scenarios 1-4 listed below, the surrounding vehicles use non-intelligent strategies: maintaining a constant speed and heading. These simulation studies illustrate empirically the robustness of potential game approach against various surrounding vehicles' strategies.

Scenario 1: In this scenario, we let vehicle 2 (resp. vehicle 3) move slower (resp. faster) than the ego vehicle. As the front vehicle moves slowly, the ego vehicle decides to perform lane-changing when it is safe. After changing to the second

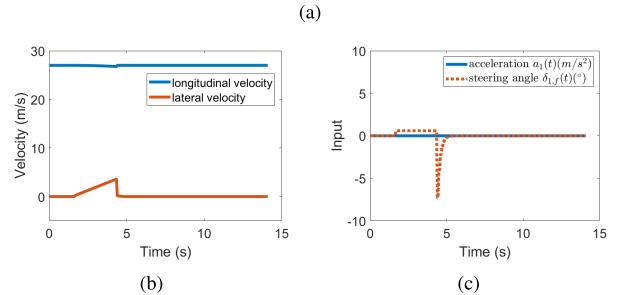
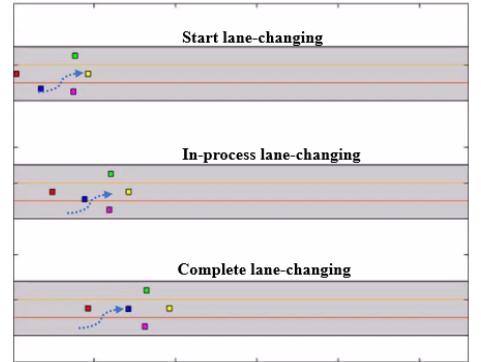


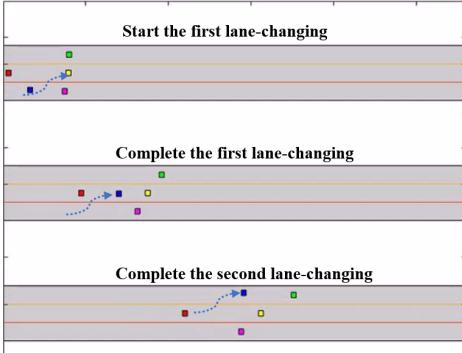
Fig. 2. The ego vehicle behavior in Scenario 1. (a) The ego vehicle trajectory. (b) The ego vehicle velocity. (c) The ego vehicle acceleration and steering angle.

lane, the ego vehicle then keeps its desired speed and drives in the second lane for the rest of the simulation time. The ego vehicle trajectory is shown in Figure 2(a), its velocity and control inputs are shown in Figures 2(b) and 2(c), respectively, and the full animation is provided as “Scenario 1” at <https://youtu.be/nQkpdQRcwEE>.

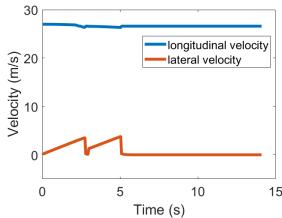
Scenario 2: In this scenario, we let both vehicles 2 and 3 move slower than the ego vehicle, and vehicle 5 moves faster. In this case, the ego vehicle is no longer satisfied driving in the second lane, and as a result, it performs two consecutive lane-changing when it is safe. The ego vehicle behavior is shown in Figure 3, and the full animation is provided as “Scenario 2” at <https://youtu.be/nQkpdQRcwEE>.

Scenario 3: In this scenario, we let all three vehicles in front, i.e., vehicles 2, 3, and 5, move slower than the ego vehicle desired speed. As these three vehicles block all of the three lanes, it is hard for the ego vehicle to keep its desired speed. Based on the game theoretic strategy, the ego vehicle chooses actions that result in both desired speed and safety. It first changes to the second lane and follows vehicle 3. After it overtakes vehicle 2, the ego vehicle then changes back to the first lane to maintain its desired speed. In this scenario, the ego vehicle intelligently prepares itself for the opportunity to overtake vehicle 2, indicating that the ego vehicle has the ability to plan actions over the long run. The ego vehicle trajectory and behavior are shown in Figure 4, and the full animation is shown in “Scenario 3” at <https://youtu.be/nQkpdQRcwEE>.

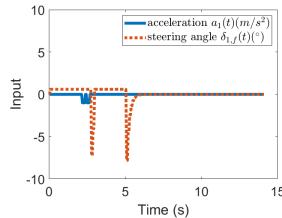
Scenario 4: In this scenario, we test the ego vehicle lane-changing behavior on curved roads. The vehicles' initial positions and velocities are similar to Scenario 1, i.e., vehicle 2 moves slower than the ego vehicle, and vehicle 3 moves faster. To keep the desired speed, the ego vehicle



(a)

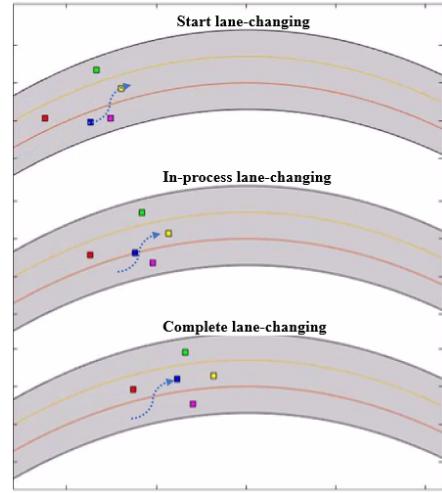


(b)

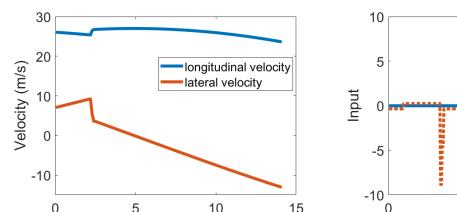


(c)

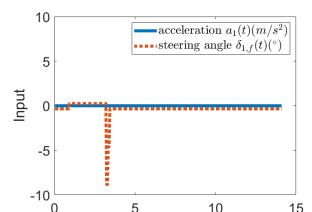
Fig. 3. The ego vehicle behavior in Scenario 2. (a) The ego vehicle trajectory. (b) The ego vehicle velocity. (c) The ego vehicle acceleration and steering angle.



(a)

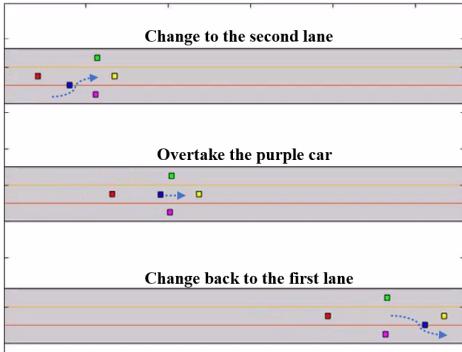


(b)

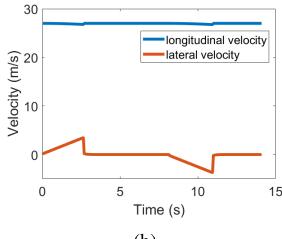


(c)

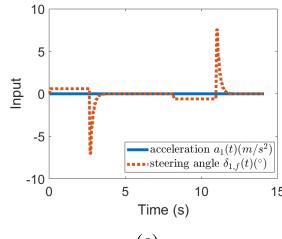
Fig. 3. The ego vehicle behavior in Scenario 2. (a) The ego vehicle trajectory. (b) The ego vehicle velocity. (c) The ego vehicle acceleration and steering angle.



(a)



(b)



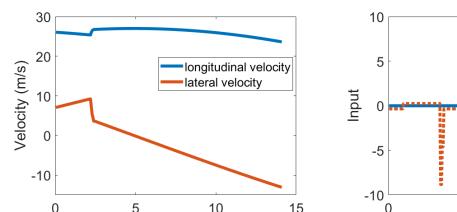
(c)

Fig. 4. The ego vehicle behavior in Scenario 3. (a) The ego vehicle trajectory. (b) The ego vehicle velocity. (c) The ego vehicle acceleration and steering angle.

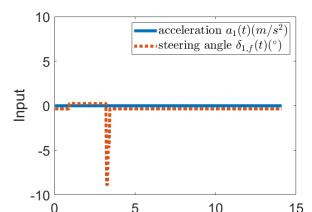
changes to the second lane when it is safe, as shown in Figure 5. The full animation is available as “Scenario 4” at <https://youtu.be/nQkpdQRcwEE>.

B. Continuous Potential Game in Intersection-Crossing Scenarios

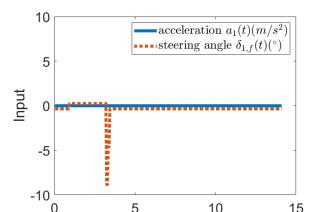
In this section, we apply the continuous potential game framework developed in Section IV to intersection-crossing scenarios.



(a)



(b)



(c)

Fig. 5. The ego vehicle behavior in Scenario 4. (a) The ego vehicle trajectory. (b) The ego vehicle velocity. (c) The ego vehicle acceleration and steering angle.

Consider the scenario pictured in Figure 6. The vehicle labeled with the number “1” is the ego vehicle, aiming to go straight and cross the intersection. The moving directions of its surrounding vehicles (i.e., vehicles 2–5) are assumed to be known to the ego vehicle, and they are marked using grey arrows in Figure 6. We let the action space of each vehicle be $\mathcal{U}_i = [-3, 3] \text{ m/s}^2$, representing that the vehicle acceleration $a_i(t)$ can be any real number from the interval $[-3, 3]$. We also employ the simplified strategy space as illustrated in Section V-A.

The cumulative cost function of each vehicle is chosen as

$$\begin{aligned} V_i^t(\mathbf{u}_i(t), \mathbf{u}_{-i}(t)) \\ = \theta_{i,1} V_i^{t,\text{self},1}(\mathbf{u}_i(t)) + \theta_{i,2} \sum_{j \in \mathcal{N}, j \neq i} V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t)) + \varpi_i, \end{aligned} \quad (34)$$

where ϖ_i is a small positive noise to prevent vehicles from getting stuck in case of identical cost and initial conditions, $V_i^{t,\text{self},1}(\mathbf{u}_i(t))$ takes the form (30) to motivate vehicles to track their desired speeds, and $V_{ij}^t(\mathbf{u}_i(t), \mathbf{u}_j(t))$ takes the form (32) with J_{ij} designed as

$$\begin{aligned} J_{ij}(x_i(\tau), y_i(\tau), x_j(\tau), y_j(\tau)) \\ = \frac{1}{(x_i(\tau) - x_j(\tau))^2 + (y_i(\tau) - y_j(\tau))^2 + \delta}, \end{aligned} \quad (35)$$

if vehicles i and j are in conflict (i.e., their planned trajectories intersect in the intersection region), and $J_{ij}(x_i(\tau), y_i(\tau), x_j(\tau), y_j(\tau)) = 0$ if i and j are not in

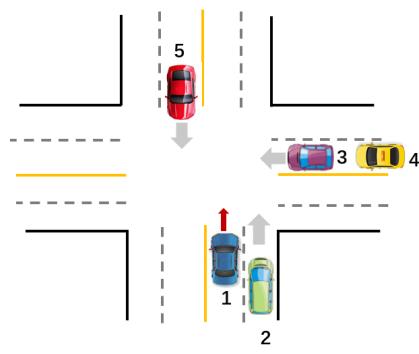


Fig. 6. The lane-changing scenario.

conflict. This inverse-type cost function enables the ego vehicle to adjust its speed smoothly and continuously in the intersection-crossing scenarios. The parameter δ in (35) is used to avoid the denominator being zero during the simulation, and $\delta = 0.01$.

According to Theorem 8, the N -player game is a continuous potential game with the above cost function design.

Figure 7 shows the performance of the continuous potential game in a specific scenario, where all vehicles follow the NE strategies. The speed of each vehicle is shown in the top left corner in Figures 7(a) and 7(b). The ego vehicle first slows down to yield to the purple and yellow vehicles (Figure 7(a)) and then speeds up after these two vehicles cross the intersection (Figure 7(b)). The performance of the best response dynamics and of the potential function optimization algorithm are very close in this scenario. (Comparative animations are available at <https://youtu.be/nQkpQRCwEE>). The iterations of the best response dynamics are shown in Figures 7(c) and 7(d), at the time instance illustrated in Figure 7(a) and that of Figure 7(b), respectively, validating the convergence of the best response dynamics.

We also consider the scenarios where the surrounding vehicles do not take the NE strategies. In these scenarios, the potential function optimization algorithm outperforms the best response dynamics in terms of better safety performance. The specific scenarios are shown in the first part of the intersection-crossing animation at <https://youtu.be/nQkpQRCwEE>, and the statistical comparison is shown in the next subsection.

C. Statistical Comparison of Various Potential Game Algorithms

In this subsection, we conduct two comparative studies in intersection-crossing scenarios: 1) finite potential game vs. continuous potential game, and 2) best response dynamics vs. potential function optimization algorithm. For each study, the comparisons are conducted in terms of a) computational efficiency measured by the average and maximum computational time to complete one decision-making process, b) the ego vehicle safety measured by the collision rate, average and maximum collision speeds, and c) the ego vehicle travel efficiency measured by the ego vehicle average speed during the intersection-crossing. As for the surrounding vehicles, we assign them three different strategies: a) employing the

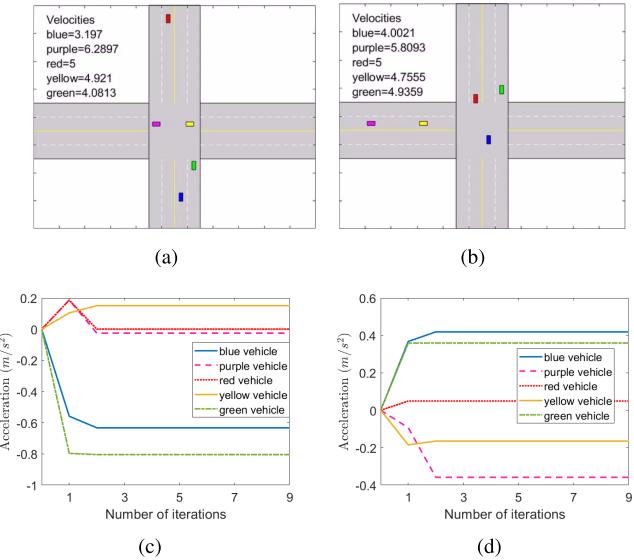


Fig. 7. The ego vehicle behavior in the intersection-crossing scenario, where all vehicles follow the NE strategies. (a) The ego vehicle slows down to avoid crashing into vehicles 3 and 4, which are speeding up to avoid collisions with the ego vehicle. (b) When it is safe, the ego vehicle speeds up to maintain its desired speed. (c) The best response iteration at the time instance of (a). (d) The best response iteration at the time instance of (b).

NE, b) keeping a constant speed, and c) randomly selecting an acceleration from the action space \mathcal{U}_i at each t . The first strategy is rational and intelligent. The second strategy is non-intelligent and safety-agnostic, but may be encountered in real scenarios as human driver may be distracted from driving and therefore cannot respond timely to a collision threat. The third strategy is neither intelligent nor rational but useful for testing robustness of the algorithms. For each algorithm, we test 5000 intersection situations with randomly selected vehicles' initial positions to provide reliable statistical results. In each situation, 5 vehicles get involved (as shown in Figure 6), i.e., the ego vehicle solves a 5-player game for each decision-making.

Table I shows the comparative results between finite potential game and continuous potential game. The running time is collected from MATLAB® on a laptop with an Intel Core i7-10750H processor clocked at 2.60 GHz and 16 GB of RAM. We use the genetic algorithm function 'ga' in Matlab [41], [42] to efficiently solve the non-convex optimization problems. For potential function optimization (Algorithm 2), 10 start points are selected, and for best response dynamics (Algorithm 1), 3 start points are selected for each agent's local optimization. The number of start points is selected based on the criterion that increasing this number does not necessarily improve the ego vehicle performance (in terms of collision rate and average velocity) while decreasing this number clearly leads to worse performance. The collision rate represents the number of situations where a collision happens to the ego vehicle out of the total number of situations (i.e., 5000). Concerning the average speed, note that the ego vehicle desired speed is 5 m/s, and in general, the closer the average speed to 5 m/s, the better the travel efficiency of the ego vehicle. The average and maximum ego vehicle speed at collision indicates the contribution of the ego vehicle to the collision impact,

TABLE I
COMPARATIVE RESULTS: FINITE POTENTIAL GAME VS. CONTINUOUS POTENTIAL GAME

	Finite potential game			Continuous potential game		
	NE	Constant speed	Random acceleration	NE	Constant speed	Random acceleration
Surrounding vehicles' strategies						
Collision rate	0/5000	0/5000	21/5000	0/5000	0/5000	14/5000
Average ego speed (m/s)	3.78	3.05	3.01	3.93	3.21	3.09
Avg/Max relative collision speed (m/s)	N/A	N/A	10.16/13.56	N/A	N/A	10.54/14.68
Avg/Max ego speed at collision (m/s)	N/A	N/A	8.06/9.50	N/A	N/A	8.06/9.61
Avg/Max computational time (s)		0.05/0.10			0.07/0.28	

which relates to the two vehicles' relative speed. The action space of each vehicle is: $\mathcal{U}_i = [-3, 3] \text{ m/s}^2$ for continuous game and $\mathcal{U}_i = \{-3, -2, -1, 0, 1, 2, 3\} \text{ m/s}^2$ for finite game. For both potential games, the NE is determined by the potential function optimization algorithm, i.e., Algorithm 1.

Table I leads to the following observations:

- 1) Both potential game frameworks are effective in ensuring the ego vehicle safety: No collision happens in any of the 5000 scenarios if the surrounding vehicles follow the NE strategies or maintain constant speeds. Although absolute safety is not guaranteed if the surrounding vehicles randomly accelerates at each t (which is too irrational to happen in real life), the collision rate is still low (less than 1%).
- 2) In both games, the ego vehicle shows better travel efficiency if the surrounding vehicles follow the NE strategies, compared to the one of constant speeds. It implies that the more intelligent the surrounding vehicles are, the better the ego vehicle travel efficiency is. Meanwhile, even if the surrounding vehicles are non-intelligent (maintaining constant speeds), the ego vehicle travel efficiency remains satisfactory (3.21 m/s in continuous game and 3.05 m/s in finite game), while ensuring safety.
- 3) Computational cost is affordable in both potential game frameworks (less than 0.1 s on average).
- 4) In general, the continuous potential game outperforms the finite potential game, as it leads to better travel efficiency with similar safety and computation performance.

Table II compares the results from the two solution algorithms (best response dynamics and potential function optimization) in the continuous potential game framework. The simulation settings are selected to be the same as the ones for Table I.

Table II leads to the following observations:

- 1) Both solution algorithms are effective in ensuring the ego vehicle safety when the surrounding vehicles follow the NE strategies: No collision happens in any of the 5000 scenarios. If the surrounding vehicles are non-intelligent (i.e., maintaining constant speeds or using random accelerations), the potential function optimization algorithm performs notably better than the best response dynamics. This is because the potential function optimization finds the global optimal NE, while the best response dynamics may converge to any NE.
- 2) With regard to the ego vehicle travel efficiency, although the best response dynamics lead to higher average speeds, we cannot conclude that it has better travel

efficiency. This is because the two solution algorithms have different safety levels, and although the best response dynamics result in higher speeds, it puts the ego vehicle into more dangerous situations.

- 3) Computational cost is affordable in both solution algorithms (0.07 s in potential function optimization and 0.24 s in best response dynamics, on average). Compared to the best response dynamics, the potential function optimization has better computational efficiency.
- 4) In general, the potential function optimization algorithm outperforms the best response dynamics, as it requires less computational time and leads to better safety for the ego vehicle.

D. Statistical Comparison of Various AV Decision-Making Algorithms

In this subsection, we compare potential games with other existing decision-making approaches, including reinforcement learning and control barrier function. The simulation scenarios and settings are the same as the ones in Section V-C, and the agents' action spaces are finite.

For potential games, we employ the potential function optimization algorithm.

For reinforcement learning, we employ a double deep Q-network (DDQN) algorithm [43], [44]. Similar to the potential game, the ego vehicle's cost function in RL is also composed of two parts: One is to track the desired speed as in (30), and the other is to avoid collision (specifically, if a collision happens, then a large penalty is applied). During the training stage, the surrounding vehicle behaviors are governed by a simplified intelligent driver model (IDM) [45]: If vehicle i senses a collision threat with vehicle j (i.e., inter-vehicle distance is less than a pre-defined safe distance), then vehicle i brakes if $T_{ij}^c \geq T_j^c$ and accelerates otherwise, with constant acceleration/deceleration. Here T_{ij}^c represents time-to-collision for vehicle i with vehicle j (see [14], [44] for the mathematical expression). We train the RL agent for 100000 episodes, where each episode contains 24 samples (corresponds to the 12 s duration of the intersection crossing with 0.5 s sampling time) or less (if collisions occur). For the execution stage, we let the surrounding vehicles follow one of two strategies, NE or constant speed, to test the robustness of the performance. In this intersection-crossing scenario, the vehicle behaviors from the NE and from the IDM are very close, as they both motivate the vehicles to cooperatively avoid collisions by accelerating (resp. decelerating) the vehicle that is closer (resp. farther) from the conflict point.

TABLE II
COMPARATIVE RESULTS: BEST RESPONSE DYNAMICS VS. POTENTIAL FUNCTION OPTIMIZATION

Surrounding vehicles' strategies	Best response dynamics			Potential function optimization		
	NE	Constant speed	Random acceleration	NE	Constant speed	Random acceleration
Collision rate	0/5000	61/5000	106/5000	0/5000	0/5000	14/5000
Average ego speed (m/s)	4.18	3.70	3.58	3.93	3.21	3.09
Avg/Max relative collision speed (m/s)	N/A	8.84/9.62	9.62/12.88	N/A	N/A	10.54/14.68
Avg/Max ego speed at collision (m/s)	N/A	7.19/8.36	6.67/8.70	N/A	N/A	8.06/9.61
Ave/Max computational time (s)	0.24/0.93			0.07/0.28		

TABLE III
COMPARATIVE RESULTS: POTENTIAL GAME VS. REINFORCEMENT LEARNING VS. CONTROL BARRIER FUNCTION

Surrounding vehicles' strategies	Potential game		Control barrier function		Reinforcement learning	
	NE	Constant speed	NE	Constant speed	NE	Constant speed
Collision rate	0/5000	0/5000	0/5000	2229/5000	0/5000	2403/5000
Average ego speed (m/s)	3.78	3.05	5.00	4.19	4.57	4.89
Avg/Max relative collision speed (m/s)	N/A	N/A	N/A	5.23/7.07	N/A	6.82/7.07
Avg/Max ego speed at collision (m/s)	N/A	N/A	N/A	0.72/5.00	N/A	4.63/5.00
Ave/Max computational time (s)	0.05/0.10		< 0.01 / < 0.01		< 0.01 / < 0.01	

For the CBF-based approach, because the ego vehicle cannot control the surrounding vehicles' behaviors, we use a variant of the centralized CBF – the predictor-corrector collision avoidance (PCCA) of [46]. Note that differently from [46], the ego vehicle's action space is bounded in the driving scenario. If the generated action from the CBF exceeds the bound, then the bound value is applied in our implementation.

The statistical comparison of the above approaches is shown in Table III. (Performance in specific scenarios is shown in the last part at <https://youtu.be/nQkpdQRcwEE>.) Table III leads to the following observations:

- 1) The reinforcement learning approach performs well when the surrounding vehicles follow the NE strategies, while its performance significantly degrades when other vehicles follow the constant speed strategy. It is because the performance of RL is highly dependent on the training dataset. Since the deep Q-network is trained by IDM-governed surrounding vehicles, if a vehicle does not follow the IDM and is safety-agnostic, it can easily lead to a collision with the ego vehicle. In real world, it is extremely hard, if not impossible, to reflect all situations and all vehicle behaviors in the training dataset, and therefore, the robustness of the RL-based algorithm can be a major concern in autonomous driving.
- 2) The control barrier function based approach is not able to ensure safety either, when the surrounding vehicles follow the constant speed strategy. It is because to guarantee the forward invariance of a pre-defined safety set, the CBF based approach generally requires all agents' adherence to the CBF constraints [47]. In a decentralized setting where other agents may not adhere to the CBF constraints, to ensure recursive feasibility, an unbounded action space of the ego vehicle is usually required [46], which is unrealistic in driving applications.
- 3) In contrast to RL and CBF, the potential game approach demonstrates significantly better robustness: It

successfully avoids all collisions even if the surrounding vehicles follow the constant speed strategy.

- 4) All three approaches are computationally feasible, as the running time is always less than the sampling time in all tested scenarios.

VI. CONCLUSION

This paper developed two potential game based frameworks to formulate the process of decision-making in autonomous driving when multiple traffic agents are involved. Theoretical guarantees for the existence of NE and convergence of NE seeking algorithms were presented. Scalability challenge was addressed. Approaches to constructing potential games for autonomous driving applications were provided. By comprehensive numerical studies, we showed that our developed potential game frameworks can generate safe and effective decisions for the ego vehicle in diverse traffic scenarios including lane-changing and intersection-crossing. Comparative studies with the RL and the CBF based approaches showed that the developed potential game approach leads to the best robustness against safety-agnostic surrounding vehicles. Future work will include verification and validation of the proposed potential game approach in various traffic situations, and development of prediction and learning modules to account for uncertain intentions and/or unknown cost functions of other vehicles.

APPENDIX

A. Proof of Proposition 1

Necessity. We first prove that if (4) holds, then (17) holds. Let \mathbf{r} be a unit vector in \mathbb{R}^{m_i} . If (4) holds, then the following equation holds,

$$\begin{aligned} & \lim_{h \rightarrow 0} \frac{V_i^t(\mathbf{u}_i + h\mathbf{r}, \mathbf{u}_{-i}) - V_i^t(\mathbf{u}_i, \mathbf{u}_{-i})}{h} \\ &= \lim_{h \rightarrow 0} \frac{F^t(\mathbf{u}_i + h\mathbf{r}, \mathbf{u}_{-i}) - F^t(\mathbf{u}_i, \mathbf{u}_{-i})}{h}, \quad \forall \mathbf{r}. \end{aligned} \quad (36)$$

Because $\lim_{h \rightarrow 0} \frac{V_i^t(\mathbf{u}_i + h\mathbf{r}, \mathbf{u}_{-i}) - V_i^t(\mathbf{u}_i, \mathbf{u}_{-i})}{h} = \frac{\partial V_i^t(\mathbf{u}_i, \mathbf{u}_{-i})}{\partial \mathbf{u}_i} \cdot \mathbf{r}$ and $\lim_{h \rightarrow 0} \frac{F^t(\mathbf{u}_i + h\mathbf{r}, \mathbf{u}_{-i}) - F^t(\mathbf{u}_i, \mathbf{u}_{-i})}{h} = \frac{\partial F^t(\mathbf{u}_i, \mathbf{u}_{-i})}{\partial \mathbf{u}_i} \cdot \mathbf{r}$, Equation (36) leads to (17).

Sufficiency. Now we prove that if (17) holds, then (4) holds. Because \mathcal{S}_i is a connected set, for any given $\mathbf{u}_i \in \mathcal{S}_i$ and $\mathbf{u}'_i \in \mathcal{S}_i$, there must exist a continuous curve γ that starts at \mathbf{u}_i and ends at \mathbf{u}'_i . As such, (17) leads to

$$\int_{\gamma} \frac{\partial V_i^t(\tilde{\mathbf{u}}_i, \mathbf{u}_{-i})}{\partial \tilde{\mathbf{u}}_i} d\tilde{\mathbf{u}}_i = \int_{\gamma} \frac{\partial F^t(\tilde{\mathbf{u}}_i, \mathbf{u}_{-i})}{\partial \tilde{\mathbf{u}}_i} d\tilde{\mathbf{u}}_i, \quad (37)$$

where \int_{γ} represents the line integral along the curve γ . According to the gradient theorem for line integrals [48], $\int_{\gamma} \frac{\partial V_i^t(\tilde{\mathbf{u}}_i, \mathbf{u}_{-i})}{\partial \tilde{\mathbf{u}}_i} d\tilde{\mathbf{u}}_i = V_i^t(\mathbf{u}'_i, \mathbf{u}_{-i}) - V_i^t(\mathbf{u}_i, \mathbf{u}_{-i})$, and $\int_{\gamma} \frac{\partial F^t(\tilde{\mathbf{u}}_i, \mathbf{u}_{-i})}{\partial \tilde{\mathbf{u}}_i} d\tilde{\mathbf{u}}_i = F^t(\mathbf{u}'_i, \mathbf{u}_{-i}) - F^t(\mathbf{u}_i, \mathbf{u}_{-i})$. As such, Equation (37) leads to (4).

B. Proof of Lemma 3

Let $(\mathbf{u}_i^*, \mathbf{u}_{-i}^*)$ be a NE of the game (2), i.e., $\mathbf{u}_i^* \in \operatorname{argmin} V_i^t(\mathbf{u}_i, \mathbf{u}_{-i}^*) \forall i \in \mathcal{N}$. From (17), it follows that $V_i^t(\mathbf{u}_i, \mathbf{u}_{-i}) - F^t(\mathbf{u}_i, \mathbf{u}_{-i}) = C(\mathbf{u}_{-i})$ for some function $C(\mathbf{u}_{-i})$, and hence $\mathbf{u}_i^* \in \operatorname{argmin} F^t(\mathbf{u}_i, \mathbf{u}_{-i}^*) \forall i \in \mathcal{N}$, i.e., $(\mathbf{u}_i^*, \mathbf{u}_{-i}^*)$ is also NE for the identical-interest game with all agents' cumulative costs equal to F^t . On the other hand, if $(\mathbf{u}_i^*, \mathbf{u}_{-i}^*)$ is a NE for the identical-interest game, then similar arguments imply that it is also a NE of the game (2).

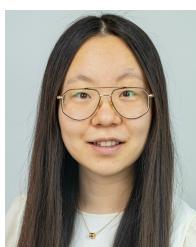
C. Proof of Lemma 4

For continuous potential games with compact \mathcal{S} and continuous F^t , a minimum of F^t as a function of \mathbf{u} always exists, according to Weierstrass theorem [48]. Hence, a NE exists according to Lemma 3. If F^t is strictly convex, then it has a unique minimum, and as such, the game has a unique NE.

REFERENCES

- [1] BuyShares. (2021). *Global Autonomous Car Market to Grow by 36% and Hit a \$37b Value By 2023*. [Online]. Available: <https://buyshares.co.uk/global-autonomous-car-market-to-grow-by-36-and-hit-a-37b-value-by-2023/>
- [2] M. Campbell, M. Egerstedt, J. P. How, and R. Murray, "Autonomous driving in urban environments: Approaches, lessons and challenges," *Philos. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 368, no. 1928, pp. 4649–4672, Oct. 2010.
- [3] 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, Nov. 2016.
- [4] 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, Feb. 2021.
- [5] N. Samyeul, "Decision-making framework for autonomous driving at road intersections: Safeguarding against collision, overly conservative behavior, and violation vehicles," *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 3275–3286, Apr. 2019.
- [6] 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.
- [7] 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.
- [8] A. Rasouli and J. K. Tsotsos, "Autonomous vehicles that interact with pedestrians: A survey of theory and practice," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 900–918, Mar. 2019.
- [9] L. Hou, L. Xin, S. E. Li, B. Cheng, and W. Wang, "Interactive trajectory prediction of surrounding road users for autonomous driving using structural-LSTM network," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 11, pp. 4615–4625, Nov. 2020.
- [10] C. Xu, W. Zhao, L. Li, Q. Chen, D. Kuang, and J. Zhou, "A Nash Q-learning based motion decision algorithm with considering interaction to traffic participants," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 12621–12634, Nov. 2020.
- [11] 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.
- [12] Q. Zhang, R. Langari, H. E. Tseng, D. Filev, S. Szabolowski, 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, Nov. 2019.
- [13] V. Lopez, F. 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.
- [14] 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.
- [15] F. Camara et al., "Pedestrian models for autonomous driving Part II: high-level models of human behavior," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 9, pp. 5453–5472, Sep. 2021.
- [16] H. Kita, "A merging-giveaway interaction model of cars in a merging section: A game theoretic analysis," *Transp. Res. A, Policy Pract.*, vol. 33, nos. 3–4, pp. 305–312, Apr. 1999.
- [17] N. Li, Y. Yao, I. Kolmanovsky, E. Atkins, and A. R. Girard, "Game-theoretic modeling of multi-vehicle interactions at uncontrolled intersections," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 1428–1442, Feb. 2020.
- [18] 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.
- [19] D. Monderer and L. S. Shapley, "Potential games," *Games Econ. Behavior*, vol. 14, no. 1, pp. 124–143, 1996.
- [20] D. Cai, S. Bose, and A. Wierman, "On the role of a market maker in networked Cournot competition," *Math. Oper. Res.*, vol. 44, no. 3, pp. 1122–1144, Aug. 2019.
- [21] J. Wu and D. Zusai, "A potential game approach to modelling evolution in a connected society," *Nature Human Behaviour*, vol. 3, no. 6, pp. 604–610, Apr. 2019.
- [22] K. Yamamoto, "A comprehensive survey of potential game approaches to wireless networks," *IEICE Trans. Commun.*, vol. 98, no. 9, pp. 1804–1823, 2015.
- [23] Y. Hino, "An improved algorithm for detecting potential games," *Int. J. Game Theory*, vol. 40, no. 1, pp. 199–205, Feb. 2011.
- [24] Y. Shoham and K. Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
- [25] H. Lu, "On the existence of pure-strategy Nash equilibrium," *Econ. Lett.*, vol. 94, no. 3, pp. 459–462, Mar. 2007.
- [26] S. Durand and B. Gaujal, "Complexity and optimality of the best response algorithm in random potential games," in *Proc. Int. Symp. Algorithmic Game Theory*, 2016, pp. 40–51.
- [27] Q. D. Lã, Y. H. Chew, and B.-H. Soong, *Potential Game Theory*. Switzerland: Springer, 2016.
- [28] M. Elbanhawi, M. Simic, and R. Jazar, "In the passenger seat: Investigating ride comfort measures in autonomous cars," *IEEE Intell. Transp. Syst. Mag.*, vol. 7, no. 3, pp. 4–17, Jul. 2015.
- [29] M. Zhou, Y. Yu, and X. Qu, "Development of an efficient driving strategy for connected and automated vehicles at signalized intersections: A reinforcement learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 1, pp. 433–443, Jan. 2020.
- [30] L. Zhang, K. Peng, X. Zhao, and A. J. Khattak, "New fuel consumption model considering vehicular speed, acceleration, and jerk," *J. Intell. Transp. Syst.*, vol. 27, pp. 1–13, Oct. 2021.

- [31] Q. Dai, D. Shen, J. Wang, S. Huang, and D. Filev, "Calibration of human driving behavior and preference using naturalistic traffic data," 2021, *arXiv:2105.01820*.
- [32] 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.
- [33] 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.
- [34] 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. 2020.
- [35] 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.
- [36] P. Polack, F. Altché, B. d'Andrea-Novel, and A. De La Fortelle, "The kinematic bicycle model: A consistent model for planning feasible trajectories for autonomous vehicles?" in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2017, pp. 812–818.
- [37] 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.
- [38] X. Cao, W. Young, and M. Sarvi, "Exploring duration of lane change execution," in *Proc. Australas. Transp. Res. Forum*, Brisbane, QLD, Australia, Oct. 2013, pp. 1–17.
- [39] J. Xiao et al., "Vehicle trajectory interpolation based on ensemble transfer regression," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 7680–7691, Jul. 2021.
- [40] 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.
- [41] MathWorks. *Find Minimum of Function Using Genetic Algorithm*. [Online]. Available: <https://www.mathworks.com/help/gads/ga.html>
- [42] D. I. Arkhipov, D. Wu, T. Wu, and A. C. Regan, "A parallel genetic algorithm framework for transportation planning and logistics management," *IEEE Access*, vol. 8, pp. 106506–106515, 2020.
- [43] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in *Proc. AAAI Conf. Artif. Intell.*, vol. 30, no. 1, 2016, pp. 1–7.
- [44] 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.
- [45] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 62, no. 2, p. 1805, Aug. 2000.
- [46] M. Santillo and M. Jankovic, "Collision free navigation with interacting, non-communicating obstacles," in *Proc. Amer. Control Conf. (ACC)*, May 2021, pp. 1637–1643.
- [47] 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.
- [48] R. Williamson and H. Trotter, *Multivariable Mathematics*. London, U.K.: Pearson Education, 2004.



Moshuang Liu (Member, IEEE) received the B.S. degree from the University of Electronic Science and Technology of China in 2016 and the Ph.D. degree from The University of Texas at Arlington in 2020. She was a Post-Doctoral Researcher with the Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI, USA, from 2021 to 2022. 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 and learning for multi-agent systems.



Ilya Kolmanovsky (Fellow, IEEE) received the Ph.D. degree in aerospace engineering from the University of Michigan, Ann Arbor, MI, USA, in 1995. Prior to joining the University of Michigan as a Faculty Member in 2010, he was with Ford Research and Advanced Engineering, Dearborn, MI, for close to 15 years. He is currently a Professor with the Department of Aerospace Engineering, University of Michigan. His research interests include control theory for systems with state and control constraints, and in 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*.



H. Eric Tseng received the B.S. degree from the 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, in 1991 and 1994, respectively. In 1994, he joined Ford Motor Company. At Ford, 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 vehicles implementation. 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 Control Engineering Practice Award in 2013. He has over 100 U.S. patents and over 120 publications. He is an NAE Member.



Suzhou Huang received the Ph.D. degree in theoretical physics from MIT. He is currently an Executive Technical Leader in research and advanced engineering with Ford Motor Company. Prior to his current position, he was the Director of Analytics Research and Development in Global Data Insights and Analytics with Ford Motor and Ford Credit, developing advanced analytics models in a variety of areas, such as production planning, inventory management, pricing, marketing program evaluation, auto financing, risk management, capital allocation, and regulation compliance. Since he joined Ford in 1998, he transformed himself into an applied micro-economist/econometrician. His research interests include autonomous and connected vehicle technologies, human driving behavior modeling, and smart transportation systems using methodologies of ML/AI and game theory.



Dimitar Filev (Fellow, IEEE) is currently a Senior Henry Ford Technical Fellow of Control and AI with Research & Advanced Engineering Ford Motor Company. His research interests include computational intelligence, AI and intelligent control, and their applications to autonomous driving, vehicle systems, and automotive engineering. He holds over 100 granted U.S. patents and has been awarded with the IEEE SMCS 2008 Norbert Wiener Award and the 2015 Computational Intelligence Pioneer's Award. He is a member of the National Academy of Engineering. He was the President of the IEEE Systems, Man, & Cybernetics Society (2016–2017).



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 in 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. She 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.