

Urban Driving Games with Lexicographic Preferences and Socially Efficient Nash Equilibria

Alessandro Zanardi¹, Enrico Mion², Mattia Bruschetta²,
Saverio Bolognani³, Andrea Censi¹, and Emilio Frazzoli¹

Abstract—We describe Urban Driving Games (UDGs) as a particular class of differential games that model the interactions and incentives of the urban driving task. The drivers possess a “communal” interest, such as not colliding with each other, but are also self-interested in fulfilling traffic rules and personal objectives. Subject to their physical dynamics, the preference of the agents is expressed via a lexicographic relation that put as first priority the shared objective of not colliding. Under mild assumptions, we show that communal UDGs have the structure of a *lexicographic ordinal potential game* which allows us to prove several interesting properties. Namely, socially efficient equilibria can be found by solving a single (lexicographic) optimal control problem and iterated best response schemes have desirable convergence guarantees.

Index Terms—Autonomous Agents; Motion and Path Planning; Optimization and Optimal Control; Game Theory

I. INTRODUCTION

The current progress towards innovative applications of robotics such as drone delivery or autonomous driving has highlighted the importance of interactions among agents. As robots leave the factory floors for a complex world featuring many heterogeneous agents, a rational systematic way of interacting is needed [1]. In this regard, autonomous agents need to step their game up by explicitly taking into account others in their decision making. In autonomous driving for instance, a vehicle first predicts others’ actions and then plans accordingly. Instead, a preferable decision making procedure would directly consider how the others react to the ego planning. To some extent, standard planning pipelines force the agent to become a “passive” road user, making a seamless integration of autonomous agents in society problematic.

An advisable future is one where agents take rational decisions based on others’ reasoning and uncertainties. It can be argued that considering interactions in full generality would require a complete theory of mind; both for robots and humans [2]. While the current state of the art is far from capturing this in its entirety, we feel it is a road that needs to be traveled. A notable effort in this direction has been made in autonomous racing scenarios [3]–[6], where the opponents have the only objective of overcoming their rivals, thus resulting—at

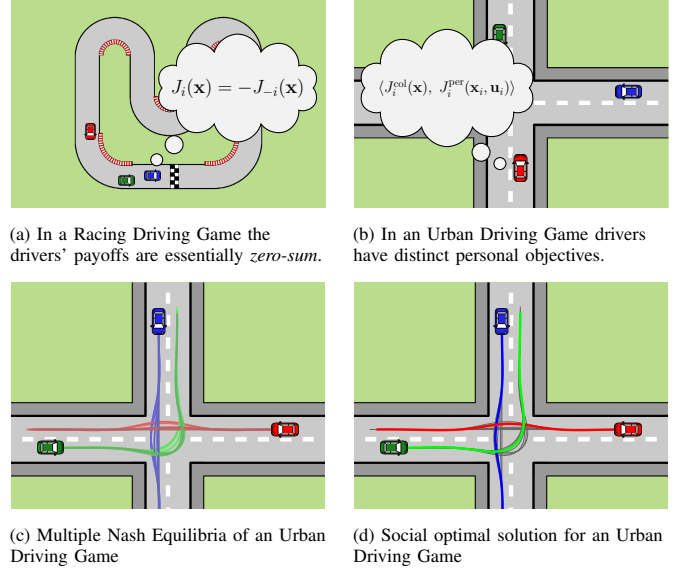


Fig. 1. In contrast to racing scenarios that are inherently adversarial, users of every-day roads are decoupled from each other in terms of personal goals and constraints. The only coupling happens via the shared desire of not colliding. We recognize a particular structure to these interactions that we name Urban Driving Games (UDGs). We show that UDGs with communal collision cost benefit from a *potential* structure that characterizes the socially efficient solutions and provides an optimization-based approach to its computation.

least conceptually—in a zero-sum problem. Instead, we consider the world outside of race tracks, where each agent has their own priorities, driving style, and capabilities. All of these can be considered *personal* features that decouple the agents to their own world of priorities and constraints which are not adversarial. Yet road users are interlaced to each other by the shared wish of not colliding. This speaks to the presence of a distinct structure in urban road interactions, which we explore in this work.

Related work: Motion planning for autonomous robots formalized as differential games can be traced back to the early work of [7]. In the last years, there has been a revival in the field with the rise of applications such as autonomous driving. Most of the recent work investigates racing scenarios, fusing standard control techniques such as receding horizon and game theoretical principles such as best responses to achieve highly interactive behaviors [3]–[6]. Other works include non racing scenarios that showcase a wide range of applications including highway merging [8], [9], truck platooning [10], common intersections [11], interactions with humans [12], and pedestrian walking [13].

“Solving a game” usually means finding the Nash Equilibria (NE), which are in general hard to compute, and significantly

Manuscript received: October, 15, 2020; Revised January, 18, 2021; Accepted February, 13, 2021. This paper was recommended for publication by Editor N. Amato upon evaluation of the Associate Editor and Reviewers’ comments. This work was supported by the Swiss National Science Foundation under NCCR Automation, grant agreement 51NF40_180545.

¹A. Zanardi, A. Censi and E. Frazzoli are with the Institute for Dynamic Systems and Control, ³S. Bolognani is with the Automatic Control Laboratory, ETH Zurich, 8092 Zurich, Switzerland. azanardi@ethz.ch

²E. Mion and M. Bruschetta are with the Department of Information Engineering, University of Padova, Italy.

Digital Object Identifier (DOI): see top of this page.

more complex than single-agent motion planning [14]. To this end, recent works have constructed solvers leveraging different assumptions and approximations to find local equilibria of a Generalized Nash Equilibrium Problem (GNEP) [15]. For instance, [11], [16] solve repeated quadratic games, [9] exploits an augmented Lagrangian approach, and [17] exploits the temporal structure in projected gradient and Douglas-Rachford splitting methods. The peculiarity of urban driving with shared and personal objectives enjoys a more favorable setup than general games (e.g., [18], [19]). These conditions have been partially exploited, in combination with additional assumptions on the dynamics and the payoffs, to obtain scenarios in which it is easier to prove existence of NE and to find them. For example [20] makes convexity assumptions about the best-response sets and how the control signal parametrizes the dynamics, while [21] obtains similar results when paths are fixed and the vehicles control only the acceleration.

We identify two main limitations in these works that motivate our manuscript: 1) In urban scenarios, agents often have so many conflicting goals that the resulting feasible set of a GNEP would be often empty. 2) Local solvers do not explicitly consider the issue of equilibrium selection.

Contribution: We show that a wide range of road interactions get naturally formalized as differential games with decoupling properties and a “shared” objective of not colliding. In contrast to related works, we embed each player’s problem complexity in the (lexicographic) preference of the agent over the outcomes. The only hard constraints are the actuation limits and the dynamics. This allows a more general formulation inspired by the principle of *minimum violation planning* that does not allow infeasibility but only *least worst* solutions [22]–[24].

We then prove that under mild assumptions of communal collision cost (see Def. 3) and personal secondary objectives, Urban Driving Games have an ordinal potential structure (in the lexicographic sense). This structure provides theoretical guarantees for popular game theoretical algorithms such as iterated-best-response (IBR), and suggests that we could compute the *socially efficient* Nash Equilibrium of the game via a single optimization problem.

Finally, we empirically showcase the derived results in a challenging intersection example where the agents exhibit sensible behaviors naturally deriving from the lexicographic preference. Despite in practice one cannot guarantee global optimality for such non-convex problems, we observe that the potential structure provides socially efficient NE as a natural equilibrium refinement tool.

II. DRIVING GAMES

We are going to define a class of “Driving Games” as differential games that capture the features of a generic driving tasks. Later we will further distinguish Urban and Racing games. In a driving task we recognize these features:

- **Decoupling:** Each agent drives its own vehicle; in this sense the dynamics are decoupled.
- **Collisions:** The drivers share the same physical space and a physical collision ends the game for the players involved.
- **Lexicographic preference:** In any (civilized) driving task, the first concern is not to collide. Therefore, the cost function

will be a *lexicographic order*: first, the agents do not want to collide; second, they want to minimize a personal cost.

In order to provide a formal definition of this class of games, we recall some preliminaries on differential games in state space. Consider a finite set of n players denoted by \mathcal{A} . The state is a vector $\mathbf{x} \in X$ that evolves according to a differential equation $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}_1(t), \dots, \mathbf{u}_n(t), t)$ with boundary conditions given by the initial state $\mathbf{x}(0)$. The control inputs of each agent are $\mathbf{u}_i \in U_i$. The decision space of each agent is a prespecified class Γ_i of mappings $\gamma_i : X \times [0, T] \rightarrow U_i$ such that their control input has the feedback form of $\mathbf{u}_i(t) = \gamma_i(\mathbf{x}(t), t)$. Based on the information structure, one can distinguish between *feedback policies* $(\mathbf{x}(t), t) \mapsto \mathbf{u}(t)$ and *open loop policies* $\mathbf{x}(0) \mapsto \mathbf{u}_i^T$ that depend only the initial state.

Based on these definitions, we introduce the same assumptions as in [25, Theorem 5.1], which guarantee the uniqueness of solutions for the ODE and the well-posedness of a differential game, and we then formalize the class of Driving Games.

Assumption 1 (Well-posedness). *The vector field \mathbf{f} is continuous for $t \in [0, T]$ for every state \mathbf{x} , uniformly Lipschitz in \mathbf{x} , $\mathbf{u}_i \forall i \in \mathcal{A}$. The strategies $\gamma_i \in \Gamma_i (i \in \mathcal{A})$ are continuous in t for each \mathbf{x} and uniformly Lipschitz in \mathbf{x} .*

Definition 1 (Driving Games). *A Driving Game is a deterministic differential game with the following properties:*

- 1) *Each agent $i \in \mathcal{A}$ has a personal state $\mathbf{x}_i(t) \in X_i$ whose dynamics evolve independently of other agents, i.e., $\dot{\mathbf{x}}_i(t) = \mathbf{f}_i(\mathbf{x}_i(t), \mathbf{u}_i(t), t)$ with $\mathbf{x}_i(0) = \mathbf{x}_i^0$.*
- 2) *Each agent $i \in \mathcal{A}$ has a distinct terminal time T_i . The instant T_i is defined as the smallest t for which either the player reaches its goal or it physically collides with another agent.¹ Let $X_i^{\text{goal}} \subset X_i$ be the goal region for agent i , then for each state trajectory let $t_i^{\text{goal}} = \min\{t \in \mathbb{R}_+ \mid \mathbf{x}_i(t) \in X_i^{\text{goal}}\}$. Let $\phi(\mathbf{x})$ be the footprint of an agent in the workspace, then $t_i^{\text{col}} = \min\{t \in \mathbb{R}_+ \mid \phi(\mathbf{x}_i(t)) \cap \phi(\mathbf{x}_j(t)) \neq \emptyset\}$. Thus $T_i = \min\{t_i^{\text{goal}}, t_i^{\text{col}}\}$.*
- 3) *There is a collision cost function that depends on the joint state trajectory of two agents $J_{i,j}^{\text{col}} : X_i^{T_i} \times X_j^{T_j} \rightarrow \mathbb{R}_{\geq 0}$. Therefore the total collision cost for agent i becomes*

$$J_i^{\text{col}}(\gamma) = \sum_{j \in \{-i\}} J_{i,j}^{\text{col}}(\mathbf{x}_i^T, \mathbf{x}_j^T). \quad (1)$$

- 4) *Each agent $i \in \mathcal{A}$ has a personal cost function composed by two functionals $g_i : X \times U_i \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ (the incremental cost) and $s_i : X_i \rightarrow \mathbb{R}$ (the terminal cost):*

$$J_i^{\text{per}}(\gamma) = \int_0^{T_i} g_i(\mathbf{x}(t), \mathbf{u}_i(t), t) dt + s_i(\mathbf{x}(T_i)). \quad (2)$$

- 5) *A strategy profile γ^* is a NE when $\forall i \in \mathcal{A}$*

$$J_i(\gamma^*) \preceq J_i(\gamma_i, \gamma_{-i}^*), \quad \forall \gamma_i \in \Gamma_i, \quad \forall (\mathbf{x}^0, t^0) \in X \times [0, T] \quad (3)$$

where the preference relation \preceq is a lexicographic total order² on the tuple:

$$J_i(\gamma) = \langle J_i^{\text{col}}, J_i^{\text{per}} \rangle. \quad (4)$$

¹This ensures that the dynamics are independent for the time interval that is relevant to each agent’s own cost.

A. On constraints and players' preferences

In this formalization, the only hard constraints are the agents' physical actuation limits, e.g., how much they can steer or accelerate. Agents' decisions are determined by their preference over a multi-objective outcome. This is different, for instance, from the GNEP formulation [18], which considers games with a scalar payoff with physical and soft constraints at the same level.

Our choice is motivated by the paradigm of *minimum violation planning*, where the output of a planning problem cannot be infeasibility—what shall the agent then do?—but rather the trajectory that performs the least worst. The shift in complexity on the agent's preference allows to express a precise hierarchy over conflicting objectives, or soft constraints (e.g. traffic rules violations do not have the same importance as deviating from the desired speed [22]–[24]).

On a technical level, note that this is not a re-formulation of the same. Indeed, lexicographic preferences do not always admit a numerical function representation (a formal proof can be found in the footnotes of [26]), but any scalar weighted sum can be embedded into a lexicographic preference when a trade-off is desired. Indeed lexicographic preferences represent a superset of the scalar case.

The gained expressiveness comes with some limitations. For instance, while it is easy to rank a finite set of trajectories, it becomes cumbersome to optimize over continuous spaces for arbitrary ordinal relations. In Sec. IV we will show how this can be achieved for lexicographic preferences, resulting in non-trivial behaviors.

B. Racing vs Urban Driving Games

A race is intrinsically a zero-sum game: there is a winner and the losers (or multiple tiers of winners). Although the incremental solution of a racing scenarios can also be formulated/approximated as a general-sum game [3], the “outcome of a race” depends on the joint state which defines a winner (+1) and a loser (−1).

This racing scenario does not capture the interaction of every-day urban driving. In this paper we describe the class of Urban Driving Games as those that represent better the real-life interactions; for example, there are no “winners” and “losers”; each agent has its own personal objectives. But there are still collisions, which make everybody a loser. By formalizing this intuition, we show that communal Urban Driving Games enjoy a “potential” structure that makes it possible to find socially efficient NE. First, we ask that the personal cost depends only on the agent's state. We call this specialization an Urban Driving Game.

Definition 2 (Urban Driving Games). *We call Urban Driving Game a driving game with the personal cost depending only on the agent's personal state and actions:*

$$J_i^{\text{per}}(\gamma_i) = \int_{t=0}^{T_i} g_i(\mathbf{x}_i(t), \mathbf{u}_i(t), t) dt + s_i(x_i(T_i)). \quad (5)$$

²Formally, we have that the cost space has the topology given by the order topology and that for $a = \langle a_1, a_2 \rangle \in \mathbb{R}^2$ we have that $(\forall a, a' \in \mathbb{R}^2) \quad a \prec a' \Leftrightarrow a_1 < a'_1 \vee (a_1 = a'_1 \wedge a_2 < a'_2)$.

Without loss of generality, we consider a scalar personal cost. As we will see in Sec. IV, the personal cost can be itself a lexicographic order over personal objectives with different priorities. Second, we introduce another condition on the collision cost. We want the collision cost to be regarded as a communal objective among the agents—nobody wants ever to collide.

Definition 3 (Communal collision cost). *We say that the collision cost is communal iff $\forall i \in \mathcal{A}$ and for every $\gamma_i, \gamma'_i \in \Gamma_i$, $\gamma_{-i} \in \Gamma_{-i}$ it holds:*

$$J_i^{\text{col}}(\gamma') - J_i^{\text{col}}(\gamma) < 0 \Rightarrow \sum_{j \in \{-i\}} [J_{j,i}^{\text{col}}(\gamma') - J_{j,i}^{\text{col}}(\gamma)] \leq 0. \quad (6)$$

In words, (6) requires that if the collision cost improves for a unilateral deviation in strategy of agent i , the collective collision cost of all the other agents interacting with i does not get worse either.

Def. 3 includes a large variety of possible collision costs. For instance, all the *symmetric collision costs*, i.e., functionals for which $J_{i,j}^{\text{col}} = J_{j,i}^{\text{col}} \forall i, j \in \mathcal{A}$. As an example, let m_i be the mass of the agent i and \vec{v}_i its velocity vector, then the cost

$$J_{i,j}^{\text{col}} = m_i m_j |\vec{v}_i - \vec{v}_j|^2 \text{ if } \phi(\mathbf{x}_i(\bar{T})) \cap \phi(\mathbf{x}_j(\bar{T})) \neq \emptyset \quad (7)$$

is a proxy for the kinetic energy transferred and received during the impact. Clearly it is symmetric but different for different agents' pairs. Other examples would include a *minimum safety distance* or a collision indicator. Naturally not all functions satisfy Def. 3, for instance, removing m_j from (7) breaks the potential structure.

III. THE POTENTIAL STRUCTURE OF CUDGS

The importance of *static potential games* is well known since their introduction [27] as a class of games which possess pure-strategy NE, have convergence guarantees for the learning process, and for which efficient solutions are often non-cooperative NE [28]. In this section, we show that Communal Urban Driving Games (CUDGs) have a potential structure. In order to do so, we adapt the standard definition of potential game from [27] to the total lexicographic cost structure of driving games. Because of Assumption 1, each admissible strategy profile corresponds to a unique state trajectory, and in turn to a unique cost. Hence, for compactness, we write the costs as functions of the strategies instead of the state and the control inputs. This is justified from the equivalence among the *loop model*, the *tree model*, and the *normal form* of games [25, Ch 5]. Moreover, we do not distinguish between open and closed loop information structure since the theoretical results hold for both.

Definition 4 (Ordinal potential function for totally ordered set). *Let (X, \preceq) be a totally ordered set. A function $P : \times_i \Gamma_i \rightarrow X$ is said to be an ordinal potential function on X if for every agent $i \in \mathcal{A}$ and every strategy $\gamma_{-i} \in \Gamma_{-i}$ it holds*

$$J_i(\gamma'_i, \gamma_{-i}) \prec J_i(\gamma_i, \gamma_{-i}) \text{ iff } P(\gamma'_i, \gamma_{-i}) \prec P(\gamma_i, \gamma_{-i}) \quad (8)$$

for every $\gamma'_i, \gamma_i \in \Gamma_i$.

Note that for $X = \mathbb{R}$ this is equivalent to the standard notion of potential function from [27]. A game that admits an ordinal potential function is called an *ordinal potential game*. Finally, we present our main theoretical result.

Theorem 1 (CUDGs are socially efficient potential games). *Let \mathcal{G} be a Urban Driving Game with communal collision cost (Communal Urban Driving Game (CUDG)) (Defs. 2 and 3). Then, for compact strategy spaces and lower semi-continuous bounded costs, the minimum of the social cost corresponds to a pure-strategy NE of \mathcal{G} .*

Proof. In turn, we will show that:

- a) \mathcal{G} admits a lexicographic potential function P .
- b) The minimum of P is a pure NE of \mathcal{G} .
- c) The minimum of P achieves the minimum social cost.

Once a) is shown, b) follows from standard game theory results under mild continuity assumptions [27]. On the other hand, c) is a desirable property which is not true in general for potential games. More specifically, we claim that

$$P(\gamma) = \left\langle \frac{1}{2} \sum_{i \in \mathcal{A}} J_i^{\text{col}}(\gamma), \sum_{i \in \mathcal{A}} J_i^{\text{per}}(\gamma_i) \right\rangle \quad (9)$$

is a lexicographic potential for \mathcal{G} .

Recall that $J_i^{\text{col}}(\gamma) = \sum_{j \in \{-i\}} J_{i,j}^{\text{col}}(\gamma_i, \gamma_j)$ and observe that, for a unilateral deviation in strategy of agent i ($\gamma_i \rightarrow \gamma'_i$), hence $\gamma \rightarrow \gamma'$ where $\gamma' = [\gamma_1, \dots, \gamma'_i, \dots, \gamma_n]$, Eq. (9) can be rewritten putting the terms depending on γ'_i in evidence:

$$\begin{aligned} P(\gamma'_i, \gamma_{-i}) &= \left\langle \frac{1}{2} J_i^{\text{col}}(\gamma') + \right. \\ &+ \frac{1}{2} \sum_{j \in \{-i\}} \sum_{k \in \{-i\} \setminus \{j\}} \left[J_{j,k}^{\text{col}}(\gamma_j, \gamma_k) + J_{j,i}^{\text{col}}(\gamma_j, \gamma'_i) \right], \\ &\left. J_i^{\text{per}}(\gamma'_i) + \sum_{j \in \{-i\}} J_j^{\text{per}}(\gamma_j) \right\rangle. \end{aligned} \quad (10)$$

Hence to evaluate whether $P(\gamma'_i, \gamma_{-i}) \prec P(\gamma)$ it is sufficient to consider only the terms depending on the new strategy γ' . Thus, $P(\gamma'_i, \gamma_{-i}) \prec P(\gamma) \Leftrightarrow$

$$\begin{aligned} &\Leftrightarrow \left\langle \frac{1}{2} J_i^{\text{col}}(\gamma') + \frac{1}{2} \sum_{j \in \{-i\}} J_{j,i}^{\text{col}}(\gamma_j, \gamma'_i), J_i^{\text{per}}(\gamma'_i) \right\rangle \prec \\ &\prec \left\langle \frac{1}{2} J_i^{\text{col}}(\gamma) + \frac{1}{2} \sum_{j \in \{-i\}} J_{j,i}^{\text{col}}(\gamma_j, \gamma_i), J_i^{\text{per}}(\gamma_i) \right\rangle. \end{aligned} \quad (11)$$

The second term of the tuples in (11) corresponds exactly to the personal cost deviation of agent i . Instead, the potential deviation of the collision cost sums the collision cost of agent i and the collision cost deviation of all the other agents with respect to agent i . It follows that, under the assumption of communal collision cost (Def. 3), we have a potential game.

For compact strategy sets, and lower semi-continuity of bounded cost functions, a global minimum of P exists. Let $\gamma^* \in \arg \min_{\gamma \in \Gamma} P(\gamma)$. Then γ^* is also a pure NE for \mathcal{G} because of (8) (same reasoning of Proposition 12.2 in [29]).

We now show c). Recall that the *social cost* is defined as $\mathcal{C}(\gamma) = \sum_{i \in \mathcal{A}} J_i(\gamma)$. The claim follows observing that, in our case, $\gamma^* \in \arg \min_{\gamma \in \Gamma} P(\gamma)$ is a NE of \mathcal{G} , but also $\gamma^* \in \arg \min_{\gamma \in \Gamma} \mathcal{C}(\gamma)$.

We underline that the existence of a socially efficient pure NE is not inherent in the definition of potential games. A possible misalignment between the potential function and the social utility is crucially important for instance in the class of *congestion games*, because it gives an explanation to the observed inefficiency of competitive solutions to these problems³. For CUDGs, we now know that it is sufficient to solve a single non-linear optimization problem over the set of pure strategies to find a socially efficient NE.

Moreover, we show in Sec. IV that driving games frequently possess multiple admissible NE, which introduces the issue of equilibrium selection [30]. This cannot be solved without introducing further assumptions (e.g., players agree to refine the set of NE according to the same criteria) or external infrastructure (e.g. additional communication among agents). To this end, the NE obtained via optimization of the potential offer a natural refinement tool for equilibrium selection.

The main drawback of this optimization-based approach is its computational burden, in particular for a large number of agents. This tradeoff is illustrated in the remainder of this section, which presents an easier computational approach to compute NE with no guarantees of social efficiency.

A. Iterated better (and best) response convergence

In Iterated Best Response (IBR) schemes, at each step a single agent updates its strategy based on its *best response* to the others' strategy. Because of its simplicity, IBR has been the predominant algorithm in practical robotic applications that seek NE [4], [5]. Nevertheless, for general games, there are no guarantees of convergence. It turns out that, thanks to the potential structure, these guarantees exist for CUDGs.

Proposition 1 (On IBR-schemes of CUDGs).

- i) Let the assumptions of Thm. 1 hold. Then for an arbitrary small $\varepsilon > 0$, iterated ε -better response (ε -BR) converges to a Nash ε -equilibrium (ε -NE) in a finite number of steps.
- ii) For discrete strategies, IBR-schemes converge to a NE.

Proof. We call a sequence of strategies $\Pi = [\gamma^0, \gamma^1, \dots]$ an ε -improvement path with respect to Γ if

$$\begin{aligned} &\langle J_i^{\text{col}}(\gamma^{k-1}) - \varepsilon, J_i^{\text{per}}(\gamma^{k-1}) \rangle \succ J_i(\gamma^k) \text{ or} \\ &\langle J_i^{\text{col}}(\gamma^{k-1}), J_i^{\text{per}}(\gamma^{k-1}) - \varepsilon \rangle \succ J_i(\gamma^k), \end{aligned}$$

for some $\varepsilon > 0$ and for all $k \geq 1$ where $i \in \mathcal{A}$ is the unique deviator at step k . Adapting the standard notion of ε -NE to the lexicographic structure, we call a strategy profile γ^ε a total order ε -NE iff there exist an $\varepsilon > 0$ and $\forall i \in \mathcal{A}$

$$\begin{aligned} &J_i(\gamma^\varepsilon) \prec \langle J_i^{\text{col}}(\gamma_i, \gamma_{-i}^\varepsilon) + \varepsilon, J_i^{\text{per}}(\gamma_i, \gamma_{-i}^\varepsilon) \rangle \quad \text{and} \\ &J_i(\gamma^\varepsilon) \prec \langle J_i^{\text{col}}(\gamma_i, \gamma_{-i}^\varepsilon), J_i^{\text{per}}(\gamma_i, \gamma_{-i}^\varepsilon) + \varepsilon \rangle, \quad \forall \gamma_i \in \Gamma_i. \end{aligned} \quad (12)$$

Given the above definitions, the proof follows the same reasoning as for finite games (e.g. [27, Lemma 4.1], [29, Proposition 13.1]). If a strategy profile γ^ε cannot find an ε -improvement deviation, then it satisfies (12) and it is a ε -NE. If an ε -improvement path exists, the myopic update of players' strategies induces an $\frac{\varepsilon}{2}$ -improvement path on the potential

³Hence, the famous concepts of *price of anarchy/stability* as measures of the system inefficiency when operating at a NE (see Braess' paradox, etc.).

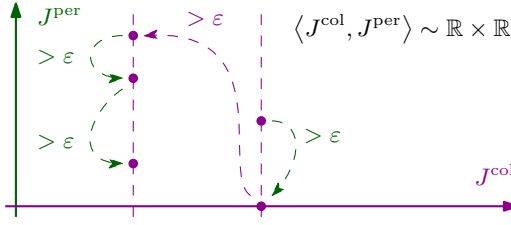


Fig. 2. An example of ε -improvement path. When improving an objective of higher rank, the lower can do arbitrarily worse. As the single objectives are bounded above and below by assumption, and the number of objectives is finite, the length of any ε -improvement path is finite.

function (since (8)). Because P is bounded by assumption, every $\frac{\varepsilon}{2}$ -improvement path of maximal length will converge in a finite number of steps to an ε -NE. Fig. 2 provides an intuitive visualization of the reasoning. Item ii) follows the same logic but with a discrete domain. ■

Remark 1. In general, for compact strategy spaces, we cannot guarantee that a sequence of ε -NE converges to a NE at the limit $\varepsilon \rightarrow 0$ [18], [31]. Nevertheless, convergence to a ε -NE for a predefined and arbitrarily small value of ε is typically satisfactory in practice, when a termination criterion for the iterated refinement of the agents' strategies is anyway necessary.

Iterated best response schemes have been widely adopted because computing the rational reaction set of a single agent is often computationally tractable (e.g. [6]). At least for open loop information structure, this is the case also for Urban Driving Games. More formally, let us denote the rational reaction set of agent i to a strategy profile γ as $\mathcal{R}_i(\gamma) \subseteq \Gamma_i$. Then an agent can compute its best-response solving the following constrained optimization problem:

$$\begin{aligned} \mathcal{R}_i(\gamma_{-i}) = \arg \min_{\gamma_i \in \Gamma_i} \quad & \langle J_i^{\text{col}}(\gamma_i, \gamma_{-i}), J_i^{\text{per}}(\gamma_i) \rangle \\ \text{subject to} \quad & \dot{\mathbf{x}}_i(t) = f_i(\mathbf{x}_i(t), \mathbf{u}_i(t)) \\ & \mathbf{x}_i(0) = \mathbf{x}_i^0. \end{aligned} \quad (13)$$

Algorithm 1: Iterated Best Response Algorithm

```

converged  $\stackrel{?}{=} \gamma_i^k \in \mathcal{R}_i(\gamma_{-i}^k) \quad \forall i \in \mathcal{A}$ 
 $\gamma^0 = [\gamma_1^0, \dots, \gamma_n^0];$ 
 $k = 0;$ 
while  $\neg$ converged do
  for  $i \in \mathcal{A}$  do
     $\gamma_i^{k+1} \in \mathcal{R}_i(\gamma_{-i}^k);$ 
     $\gamma^{k+1} = [\gamma_1^k, \dots, \gamma_i^{k+1}, \dots, \gamma_n^k];$ 
     $k = k + 1;$ 
  end
end

```

Hence, given an initial strategy profile $\gamma^0 \in \Gamma$ and an arbitrary permutation of agents playing IBR (and for a finite stopping criterion $\varepsilon > 0$), there exist a finite \bar{k} such that $\gamma_i^{\bar{k}} \in \mathcal{R}_i(\gamma_{-i}^{\bar{k}}) \forall i \in \mathcal{A}$. That is, $\gamma^{\bar{k}}$ is a ε -NE. Algorithm 1 provides the algorithmic execution. Note that the sequence of agents' updating their strategies is usually considered

fixed between iterations, even though this is not needed for convergence.

Unfortunately, we remark that there are no guarantees on the “quality” of the (ε) -NE found via IBR. Namely:

- $\gamma^{\bar{k}}$ can be a non-admissible NE (i.e., there can be another NE which yields a better outcome for all players).
- $\gamma^{\bar{k}}$ is sensitive to the initial conditions: starting from different γ^0 might lead to different $\gamma^{\bar{k}}$.
- $\gamma^{\bar{k}}$ is sensitive to the order of the sequence of agents updating their best response: a different sequence often yields a different $\gamma^{\bar{k}}$.

Intentionally we did not distinguish between open and feedback strategies as the result applies to both. However, the two diverge significantly in computational complexity. For instance, (13) can be solved by standard optimization tools for open loop strategies, but it becomes significantly cumbersome the case of closed-loop strategies which would require to solve it for every possible state as initial state.

IV. ANALYSIS FOR OPEN-LOOP INFORMATION STRUCTURE

We call a “trajectory game” a CUDG with open-loop information structure. In this case, the set of strategies are mappings from the initial state to a sequence of control commands for the game duration. This corresponds to a scenario in which the drivers commit to an entire trajectory which, in more realistic scenarios with uncertainties and noise, it can be re-computed in a shrinking horizon fashion. Finally, we empirically observe that despite not having numerical guarantees of global optimality (as expected for this kind of non-convex optimization problems), the results are aligned with the theory.

A. Trajectory games at a four-way intersection

We take an uncontrolled four-way intersection with three vehicles as a guiding example. In this scenario, the vehicles are approaching the intersection with the same initial speed, starting from different directions, and each one with a predefined exit. A fourth disabled vehicle is added to the scenario blocking the nominal path of the red car (Fig. 3).

Physical constraints: We use a bicycle model for the dynamics. The vehicle states $\mathbf{x}_i = [x_i, y_i, \phi_i, v_i, \delta_i, a_{x,i}]$ are: the two-dimensional position (x_i, y_i) [m] of the center of mass; the heading angle ϕ_i [rad]; the scalar velocity v_i [m/s], the steering angle δ_i [rad] and the longitudinal acceleration $a_{x,i}$ [m/s²]. The control inputs $\mathbf{u}_i = [\dot{\delta}_i, \dot{a}_{x,i}]$ are the bounded steering angle velocity and the jerk, respectively. In addition, physical constraints on min/max available acceleration and steering, are applied. The nonlinear equations of motion are

$$\dot{x}_i = v_i \cdot \cos(\phi_i), \quad \dot{y}_i = v_i \cdot \sin(\phi_i), \quad \dot{\phi}_i = v_i/l \cdot \tan(\delta_i)$$

where l is the length of the vehicle.

Lexicographic objectives: Each agent has a lexicographic preference with three entries: $\langle J^{\text{col}}, J^{\text{per}(1)}, J^{\text{per}(2)} \rangle$. First, avoid collisions with other vehicles and the environment (J^{col}). Second, respect the traffic rules: the speed limit and the lane boundaries ($J^{\text{per}(1)}$). Finally, favor the personal comfort and the desired cruising speed ($J^{\text{per}(2)}$). More in detail, we have: J^{col} : The collision cost $J_{i,j}^{\text{col}}$ is implemented as the violation of a minimum safety distance and the constraint of staying on

the road. The minimum allowed distance is set to $\underline{d} = 4 [m]$ (the road width is $7 [m]$).

$J^{\text{per}(1)}$: The lane and maximum allowed speed objectives are implemented as soft constraints that minimize the weighted sum of their violation. The lane is defined by an open formulation of a twice differentiable spline curve $\lambda(\eta) : [0, l_\lambda] \rightarrow \mathbb{R}^2$, where $\eta \in [0, l_\lambda]$ represents the vehicle position along the curve λ , closest to the current position of the vehicle $(x_i(t), y_i(t))$, and l_λ is the length of the path. This formulation can be found also in [4]. The maximum tolerated speed is $9 [m/s]$.

$J^{\text{per}(2)}$: A weighted sum of secondary personal objectives:

- *Keep the center of the lane*: Both longitudinal and lateral deviation from the nominal path are penalized quadratically.
- *Maintain the desired speed*: Each vehicle starts with an initial speed of $8.3 [m/s]$ and aims to maintain it throughout the game duration. Deviations from the target velocity are penalized asymmetrically (lower velocities are less penalized) with quadratic terms.
- *Favor comfort*: A quadratic penalty for variations in steering velocity ($\dot{\delta}$) and one for the jerk (\dot{a}_x).

B. Optimization over lexicographic preferences

In urban scenarios oftentimes the high-level planning problem is solved with a discrete strategy space. One selects a trajectory from a finite candidate set [23], [32]. In that case, for each candidate trajectory one computes the associated costs and sorts them according to the lexicographic preference. This is advantageous because most of the computation can be carried out in parallel and the complexity scales linearly with the number of lexicographic levels [23].

For continuous search spaces, optimizing over a lexicographic preference relation is a more challenging problem [33]. We consider at least two practical ways.

- 1) Solve sequentially one optimization problem per lexicographic level. The k -th optimization minimizes the k -th objective with the additional constraint of not worsening the previously computed costs (see [34]). For instance, (13) becomes

$$\begin{aligned} (J_i^{\text{col}})^* &= \min_{\substack{\mathbf{u}_i \in U_i \\ \mathbf{x}_i(T_i) \in X_i^{\text{goal}}}} J_i^{\text{col}}(\mathbf{x}_i, \mathbf{x}_{-i}) \\ \text{s. t.} \quad &\dot{\mathbf{x}}_i = f_i(\mathbf{x}_i, \mathbf{u}_i), \mathbf{x}_i^0 \end{aligned} \quad (14)$$

followed by

$$\begin{aligned} \min_{\substack{\mathbf{u}_i \in U_i \\ \mathbf{x}_i(T_i) \in X_i^{\text{goal}}}} J_i^{\text{per}}(\mathbf{x}_i, \mathbf{u}_i) \\ \text{s. t.} \quad &J_i^{\text{col}}(\mathbf{x}_i, \mathbf{x}_{-i}) \leq (J_i^{\text{col}})^* \\ &\dot{\mathbf{x}}_i = f_i(\mathbf{x}_i, \mathbf{u}_i), \mathbf{x}_i^0. \end{aligned} \quad (15)$$

- 2) Solve in parallel (15) for different “bottoms” values of the first objective, for example using $5, 3, 2 [m]$ as minimum safety distance to be respected. The first feasible optimization will provide a solution to the lexicographic problem with a guarantee for that particular level of safety. Clearly this method approximates the original problem discretizing the safety levels. Yet it involves a single stage optimization problem where the solution for different values can be carried out in parallel.

C. IBR and Potential solutions

We explored the set of NE for the trajectory game described in Sec. IV-A with two different methods:

- 1) IBR solutions for different orders of players’ updates;
- 2) NE obtained by minimizing the potential function.

IBR: We initialize the strategy profile solving the *individual* optimal control problems (OCPs) as if no other agents were present. This strategy causes the agents to collide (Fig. 3a). As IBR (Alg. 1) is performed with a round robin of (lexicographic) best-responses, we converge to strategy profiles that are collision-free and minimize agents personal costs: personal traffic rules and personal comfort.

Potential: Finally, the NE of the potential function is obtained minimizing a single optimization problem obtained from the sum of each players’ objectives and union of their physical constraints.

All the optimizations, i.e. the single agent best responses and the potential minimization, are carried out solving three sequential minimization of $\langle J^{\text{col}}, J^{\text{per}(1)}, J^{\text{per}(2)} \rangle$ (the first lexicographic method described in Sec. IV-B) via the commercial solver Forces Pro [35] with primal-dual interior point method.

V. RESULTS AND DISCUSSION

Representative results for IBR and the potential solution are shown in Figs. 3 and 4 while Fig. 5 provides a breakdown of all the resulting personal costs. In these examples, the final collision costs were all negligible, meaning that no significant violation of the minimum safety distance occurs.

Despite the challenging setup, one can appreciate how the lexicographic preference of the agents naturally leads to sensible behaviors respecting the specifications. For instance, the red car is forced out of the lane by a broken down vehicle and trades off its cruising speed tracking objective in the interest of minimally violating the lane keeping objective (see (b),(f) of Figs. 3 and 4). The converse happens in (c), where the red is forced into a more aggressive maneuver with significant steering effort and lateral deviations. This indeed is reflected in higher $J^{\text{per}(1)}$ costs in Fig. 5a.

It stands out that all the NE computed, independently of the method and the sequence, are *admissible* (in the lexicographic sense): i.e., no equilibrium strictly dominates the other. We empirically observed that the IBR schemes converged to an ε -NE in 3 to 8 iterations, sometimes leading to equilibria that are clearly skewed in favor of some agents (see (b),(c) in Figs. 3 and 4 where the green car is forced to slightly violate the speed limit to stay clear of the red car approaching). In other instances, IBR returned equilibria that look very close to the one found minimizing the potential (Pot).

Despite having no guarantees of global optimality (as expected for this kind of non-convex optimization problems), we observed that the results are aligned with Thm. 1. Indeed, by optimizing the potential, we find a NE which is socially more efficient than all the ones obtained otherwise. We speculate that the joint optimization of the costs in the potential symmetrically dispense the amount of inevitable costs. This is hardly possible to guarantee in best response schemes, yet one might envision different IBR-schemes where the iterations happen level-wise

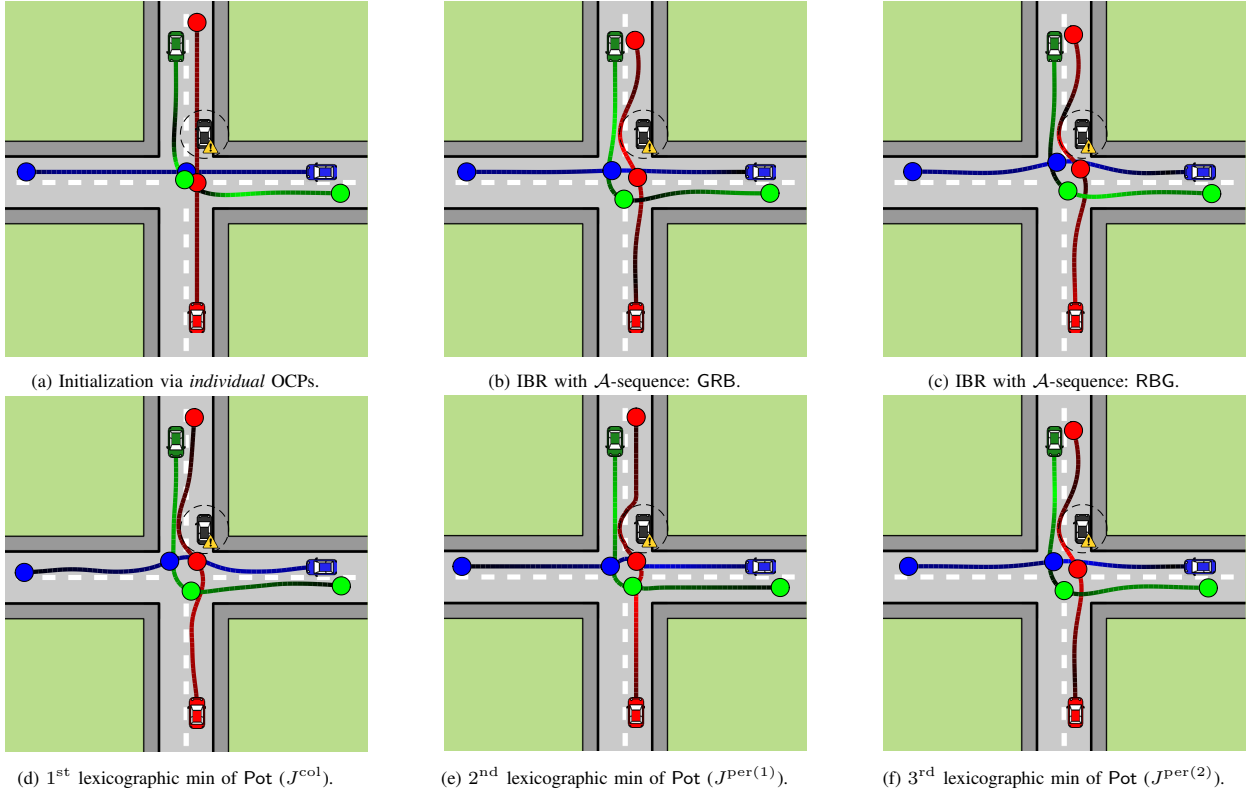


Fig. 3. The initial strategies are obtained solving the individuals optimal control problems (OCPs) which give colliding trajectories (a). Applying IBR, the agents update sequentially till convergence to a NE (b),(c). The agents' update sequence matters, indeed (b) and (c) are both NE but the solutions are asymmetrically skewed in favor of some agents. The asymmetry disappears in (f), that is the result of the lexicographic minimization of the potential costs, for which the intermediate stages are shown in (d) and (e). In (d) the agents care only about not colliding; the solution is refined in (e) by including the traffic rules (lane keeping and speed limit); in the end, (f) includes also the personal objectives. Independently from the optimization scheme, it can be appreciated how the lexicographic preference naturally leads the red car to compromise its desired speed when overtaking the gray vehicle to minimize the lane violation. The shading of trajectories hints the spatial deceleration/acceleration points shown in Fig. 4; darker shades correspond to braking.

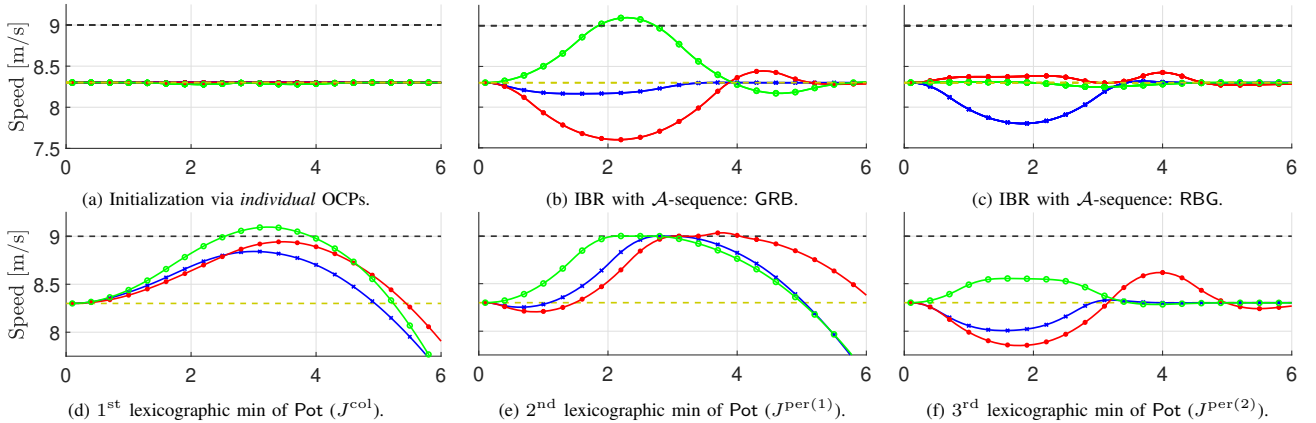


Fig. 4. The unfairness of some NE is more evident when looking at the speed profiles. As it can be seen in (b) an unfortunate sequence of players' updates leads to an equilibrium in which the green car is forced to violate the speed limit to avoid a collision with the incoming red car. The equilibrium obtained minimizing the potential (f) offers better speed profiles and lateral acceleration (not represented, see Fig. 5 for a quantitative comparison).

(in terms of lexicographic importance) and a “low rank” priority of an agent is not traded off with a “high rank” priority of another.

VI. CONCLUSIONS

In this manuscript, we formalized the class of Urban Driving Games as general-sum games with a precise but comprehensive structure which is well-suited to describe everyday vehicle

interactions from a game-theoretical standpoint. By allowing a lexicographic preference of the agents over the outcomes, we introduce the concept of minimum violation in a competitive (game-theoretic) path planning problem.

We showed that CUDGs admit a potential structure which has some convergence guarantees for IBR-schemes. These results hold for lexicographic preference and hence apply also to the standard scalar case. The existence of a socially efficient Nash Equilibrium (and the possibility to compute it via optimization

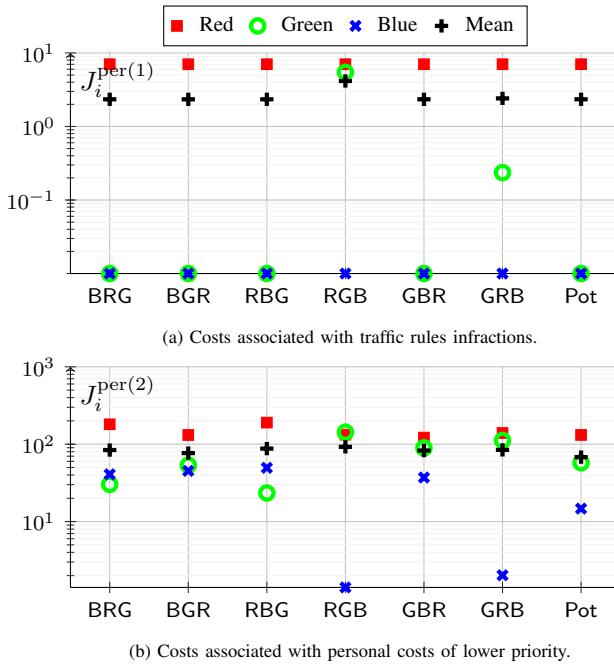


Fig. 5. Both IBR-schemes and the potential led to solutions with negligible collision costs. The only remaining costs are the personal ones. We distinguish between the ones with higher priority (associated to traffic rules) Fig. 5a and the personal desires (mainly comfort) Fig. 5b. One can observe that the average cost of the NE given by the potential (P) provides a lower bound for the other equilibria. Labels on X axis refer to the sequential order of the players (i.e., BRG means Blue first, Red second, Green third), whereas Pot refers to the Potential solution.

tools) demonstrates that efficient sharing of the mobility space may be possible both in theory and in practice.

We overall feel that there is still much to be discovered. Extensions range from feedback policies to computation scalability leveraging the inherent symmetries of the problem. In addition, one could ask for a more “relaxed” preference relations of the agents which might be even *indifferent* to some alternatives. Departing from totally ordered set can be challenging, as not even transitivity can always be preserved (e.g. in lexicographic semi-orders). We imagine that the trip has just begun.

REFERENCES

- [1] B. Jafary, E. Rabiei, M. A. Diaconeasa, H. Masoomi, L. Fiondella, and A. Mosleh, “A survey on autonomous vehicles interactions with human and other vehicles,” in *Probabilistic Safety Assessment and Management PSAM 14*, 2018.
- [2] F. Cuzzolin, A. Morelli, B. Cîrstea, and B. J. Sahakian, “Knowing me, knowing you: theory of mind in AI,” *Psychological Medicine*, vol. 50, no. 7, pp. 1057–1061, May 2020.
- [3] A. Liniger and J. Lygeros, “A Noncooperative Game Approach to Autonomous Racing,” *IEEE Trans. Control Syst. Technol.*, vol. 28, no. 3, May 2020.
- [4] R. Spica, E. Cristofalo, Z. Wang, E. Montijano, and M. Schwager, “A real-time game theoretic planner for autonomous two-player drone racing,” *IEEE Trans. Robot.*, vol. 36, no. 5, pp. 1389–1403, 2020.
- [5] M. Wang, Z. Wang, J. Talbot, J. Christian Gerdes, and M. Schwager, “Game Theoretic Planning for Self-Driving Cars in Competitive Scenarios,” in *Robotics: Science and Systems XV*, Jun. 2019.
- [6] G. Williams, B. Goldfain, P. Drews, J. M. Rehg, and E. A. Theodorou, “Best Response Model Predictive Control for Agile Interactions Between Autonomous Ground Vehicles,” in *IEEE Int. Conf. Robotics and Automation (ICRA)*, May 2018.
- [7] S. M. LaValle, “A game theoretic framework for robot motion planning,” Ph.D. dissertation, University of Illinois, 1995.
- [8] J. H. Yoo and R. Langari, “A Stackelberg game theoretic driver model for merging,” in *ASME Dynamic Syst. and Control Conf.*, 2013.
- [9] S. Le Cleac’h, M. Schwager, and Z. Manchester, “ALGAMES: A Fast Solver for Constrained Dynamic Games,” in *Robotics: Science and Systems XVI*, Jul. 2020.
- [10] E. Stefansson, J. F. Fisac, D. Sadigh, S. S. Sastry, and K. H. Johansson, “Human-robot interaction for truck platooning using hierarchical dynamic games,” in *18th Eur. Control Conf. (ECC)*, Jun. 2019.
- [11] D. Fridovich-Keil, E. Ratner, L. Peters, A. D. Dragan, and C. J. Tomlin, “Efficient Iterative Linear-Quadratic Approximations for Nonlinear Multi-Player General-Sum Differential Games,” in *IEEE Int. Conf. Robotics and Automation (ICRA)*, 2019.
- [12] D. Sadigh, N. Landolfi, S. S. Sastry, S. A. Seshia, and A. D. Dragan, “Planning for cars that coordinate with people: leveraging effects on human actions for planning and active information gathering over human internal state,” *Autonomous Robots*, vol. 42, no. 7, pp. 1405–1426, 2018.
- [13] A. Turnwald and D. Wollherr, “Human-Like Motion Planning Based on Game Theoretic Decision Making,” *Int. J. Social Robotics*, vol. 11, no. 1, pp. 151–170, Jan. 2019.
- [14] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou, “The Complexity of Computing a Nash Equilibrium,” *SIAM J. Computing*, vol. 39, no. 1, pp. 195–259, 2009.
- [15] F. Facchinei and C. Kanzow, “Generalized Nash equilibrium problems,” *4OR*, vol. 5, pp. 173–210, Sep. 2007.
- [16] D. Fridovich-Keil, V. Rubies-Royo, and C. J. Tomlin, “An Iterative Quadratic Method for General-Sum Differential Games with Feedback Linearizable Dynamics,” in *IEEE Int. Conf. Robotics and Automation (ICRA)*, 2020.
- [17] B. Di and A. Lamperski, “Local First-Order Algorithms for Constrained Nonlinear Dynamic Games,” in *Amer. Control Conf. (ACC)*, 2020.
- [18] F. Facchinei, V. Piccialli, and M. Sciandrone, “Decomposition algorithms for generalized potential games,” *Computational Optimization and Applications*, vol. 50, no. 2, pp. 237–262, 2011.
- [19] J.-S. Pang and G. Scutari, “Nonconvex Games with Side Constraints,” *SIAM J. Optimization*, vol. 21, no. 4, pp. 1491–1522, Oct. 2011.
- [20] A. Dreves and M. Gerdts, “A generalized Nash equilibrium approach for optimal control problems of autonomous cars,” *Optimal Control Applications and Methods*, vol. 39, no. 1, pp. 326–342, Jan. 2018.
- [21] A. Britzelmeier and A. Dreves, “A decomposition algorithm for Nash equilibria in intersection management,” *Optimization*, Jun. 2020.
- [22] J. Tumova, L. I. R. Castro, S. Karaman, E. Frazzoli, and D. Rus, “Minimum-violation LTL planning with conflicting specifications,” in *Amer. Control Conf. (ACC)*, 2013.
- [23] A. Censi, K. Slutsky, T. Wongpiromsarn, D. Yershov, S. Pendleton, J. Fu, and E. Frazzoli, “Liability, Ethics, and Culture-Aware Behavior Specification using Rulebooks,” in *Int. Conf. Robotics and Automation (ICRA)*, May 2019.
- [24] T. Wongpiromsarn, K. Slutsky, E. Frazzoli, and U. Topcu, “Minimum-Violation Planning for Autonomous Systems: Theoretical and Practical Considerations,” *arXiv:2009.11954 [cs.RO]*, Sep. 2020.
- [25] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, 2nd ed. SIAM, 1998.
- [26] G. Debreu, “Representation of a preference ordering by a numerical function,” *Decision processes*, vol. 3, pp. 159–165, 1954.
- [27] D. Monderer and L. S. Shapley, “Potential Games,” *Games and Economic Behavior*, vol. 14, no. 1, pp. 124–143, May 1996.
- [28] D. González-Sánchez and O. Hernández-Lerma, “A survey of static and dynamic potential games,” *Science China Mathematics*, vol. 59, no. 11, pp. 2075–2102, Nov. 2016.
- [29] J. P. Hespanha, *Noncooperative game theory: an introduction for engineers and computer scientists*. Princeton University Press, 2017.
- [30] J. Harsanyi and Selten R., *A general theory of equilibrium selection in games*. MIT press, 1988.
- [31] M. J. Powell, “On search directions for minimization algorithms,” *Mathematical Programming*, vol. 4, no. 1, pp. 193–201, 1973.
- [32] T. Phan-Minh, E. C. Grigore, F. A. Boulton, O. Beijbom, and E. M. Wolff, “CoverNet: Multimodal Behavior Prediction Using Trajectory Sets,” in *IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, 2020.
- [33] V. Fedorov, “Optimization problems with an ordered set of constraints,” *USSR Computational Mathematics and Mathematical Physics*, vol. 15, no. 5, pp. 32–45, 1975.
- [34] K.-H. Chang, “Multiobjective Optimization and Advanced Topics,” in *e-Design*. Elsevier, 2015, pp. 1105–1173.
- [35] A. Zanelli, A. Domahidi, J. Jerez, and M. Morari, “FORCES NLP: an efficient implementation of interior-point methods for multistage nonlinear nonconvex programs,” *Int. J. Control*, vol. 93, no. 1, pp. 13–29, 2020.