

# Adaptive Robust Game-Theoretic Decision Making for Autonomous Vehicles

Gokul S. Sankar, Kyoungseok Han and Ilya Kolmanovsky

**Abstract**—In a typical traffic scenario, autonomous vehicles are required to share the road with other road participants, e.g., human driven vehicles, pedestrians, etc. To successfully navigate the traffic, a cognitive hierarchy theory such as level- $k$  framework, can be used by the autonomous agents to categorize the agents based on their depth of strategic thought and act accordingly. However, mismatch between the vehicle dynamics and its predictions, and improper classification of the agents can lead to undesirable behavior or collision. Robust approaches can handle the uncertainties, however, might result in a conservative behavior of the autonomous vehicle. This paper proposes an adaptive robust decision making strategy for autonomous vehicles to handle model mismatches in the prediction model while utilizing the confidence of the driver behavior to obtain less conservative actions. The effectiveness of the proposed approach is validated for a lane change maneuver in a highway driving scenario.

## I. INTRODUCTION

Despite many recent advances in connected and automated vehicles technologies, the full automation vehicle in all conditions that provides better or at least as good as driving proficiency compared to the human drivers are still inherently flawed to be deployed in the market [1]. One of the most significant challenges is to plan the motion at the mixed-traffic conditions where the automated car coexists with the other road participants, e.g., human-driven vehicles (HV) and pedestrians [2]. In particular, describing the human decision-making process is the most difficult since the humans do not always exhibit the optimized behavior due to limited rationality [3].

To keep the robustness against all possible traffic situations, a conservative driving policy of the autonomous vehicle (AV) has been suggested [4], [5], but it may cause the disharmony with the other road participants. In other words, too conservative behavior of the AV sometimes causes the adverse effect on the overall traffics, e.g., road congestion, car accidents. However, the aggressive driving policy with very little collision potential is not acceptable even if it performs well for the most cases.

If the HV's response according to traffics, i.e., behavior of the surrounding vehicles, can be predicted, far less conservative motion planning of the AV can be designed [6]. As mentioned, however, modeling of the interaction between the vehicles may not be very accurate. Moreover, since the

Gokul S. Sankar, Kyoungseok Han and Ilya Kolmanovsky are with the Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI, USA. Emails: ggowrisa@umich.edu (G. Sankar), kyoungsh@umich.edu (K. Han) and ilya@umich.edu (I. Kolmanovsky)

communication time with the surrounding vehicles in reality is not sufficient to build the deterministic human driver models.

To overcome such shortcomings, we exploit the “level- $k$  game theory” framework where the human thought processes in strategic games, i.e., driving at the mixed-traffic, are modeled in hierarchical structure [7]. The game theoretic approach has already proved its effectiveness when describing the interactions between the vehicles [8], [9], [10]. Although these approaches effectively describe rational decision-making for the HVs, the model uncertainties according to the accuracy of the identified opponent vehicle's level that influence to the vehicle position are not considered. That is, constant safety constraint of the opponent vehicle is assumed and the AV is controlled to avoid the overlap with this constraint.

In this paper, the AV adaptively adjusts the uncertainty size of the interactive vehicles according to how much the AV believes the estimated interactive vehicle's degree of aggressiveness. With this particular strategy, the proposed approach can plan the balanced behavior of AV that is between the aggressive and conservative motion plan, describing more reasonable behavior. Specifically, the degree of interactive vehicle's aggressiveness is estimated in real-time and it is assumed to be proportional to the size of position uncertainty. With this information in mind, AV's future trajectory is planned in such a way as to maximize the cumulative reward. In [11], the constant confidence level is introduced, which describes the degree of aggressiveness of the HV. However, a *priori* assumption on the confidence level cannot reflect the varying confidence level of interactive vehicles, which can happen. To resolve this issue, we incorporate the level- $k$  game theory with the confidence level so that adaptive confidence level is available in our approach, which enables adjustable motion planning of AV depending on the degree of interactive vehicle's aggressiveness.

The contributions of this paper are summarized as follows. Firstly, we modeled the human-driven vehicle's decision-making processes at the high way using game theory, representing the a degree of the aggressiveness. Secondly, the uncertainties that is imposed in vehicle position is represented by the identified drivers aggressiveness and these are exploited to make the robust decision-making strategy for the autonomous vehicles. Lastly, we verify the benefit of the proposed approach through simulations by comparing with the results that hold their belief of interactive vehicle's degree of aggressiveness.

The rest of the paper is organized as follows. In Section II, the problem treated in this paper is stated, and Section III introduces the vehicle dynamic model with the bounded uncertainty. We next present the game-theoretic decision making

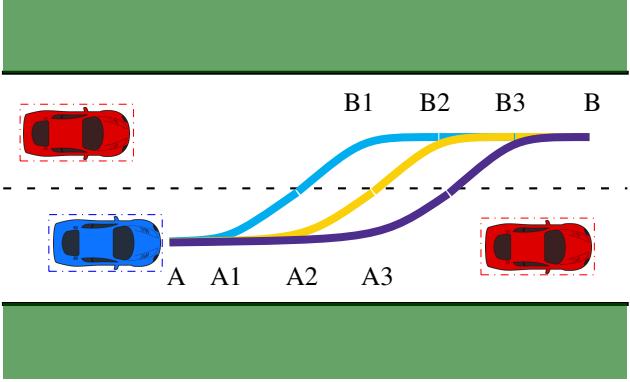


Fig. 1. Possible lane changing sequences that can be chosen by the autonomous vehicle (blue) based on the motion of other non-autonomous vehicles (red) are  $\{(A, A_1), (A_1, B_1), (B_1, B)\}$ ,  $\{(A, A_2), (A_2, B_2), (B_2, B)\}$  and  $\{(A, A_3), (A_3, B_3), (B_3, B)\}$ .

algorithm and motion planning procedure in Section IV, the effectiveness of the proposed approach is described in Section V, and the paper is concluded in Section VI.

[12] [13] [14] [15] [16]

#### A. Notation

The symbol  $\mathbb{Z}_{[a, b]}$  denotes a set of consecutive integers from  $a$  to  $b$  and  $2\mathbb{Z}^+$  denotes set of positive even integers;  $\emptyset$  denotes an empty set. For a vector  $x$ ,  $x > 0$  denotes element-wise inequality. The operator  $\oplus$  denotes the Minkowski addition, defined for the sets  $\mathcal{A}$  and  $\mathcal{B}$  as  $\mathcal{A} \oplus \mathcal{B} := \{a + b | a \in \mathcal{A}, \forall b \in \mathcal{B}\}$ .

## II. PROBLEM STATEMENT

As illustrated in Fig. 1, depending on the motions of interactive vehicles (red), the AV (blue) can take the different motion plans. If the interactive vehicles behave aggressively, the AV should plan the motion conservatively, and vice versa. In addition, to avoid the too conservative or aggressive driving policy, the balanced motion planning is preferred in the actual application.

The problem we treat is to model the intent of the interactive vehicles using game theoretical approach, then AV exploits the estimated interactive vehicle's driving intention when making the decision. The confidence level of the estimated intention is represented by dashed-dotted boxes in Fig. 1 and their sizes are adjusted adaptively according to the estimated confidence levels.

## III. VEHICLE DYNAMIC MODEL

The vehicle dynamics are represented by the following discrete model [17]

$$\begin{aligned} x(t+1) &= x(t) + v(t) \cos(\psi(t) + \beta(t)) \Delta t + w_x(t) \\ \end{aligned} \quad (1a)$$

$$y(t+1) = y(t) + v(t) \sin(\psi(t) + \beta(t)) \Delta t + w_y(t) \quad (1b)$$

$$\psi(t+1) = \psi(t) + \frac{v(t)}{l_r} \sin(\beta(t)) \Delta t \quad (1c)$$

$$v(t+1) = v(t) + a(t) \Delta t \quad (1d)$$

$$\beta(t) = \arctan\left(\frac{l_r}{l_r + l_f} \tan(\delta_f(t))\right), \quad (1e)$$

where  $t$  denotes the discrete time instant; the pair  $(x(t), y(t))$  represent the global position of the center of mass of the vehicle; the vehicle's speed is denoted by  $v(t)$ ;  $\beta(t)$  is the angle of  $v(t)$  with respect to the longitudinal axis of the vehicle;  $\psi(t)$  denotes the vehicle's yaw angle (the angle between the vehicle's heading direction and the global x-direction);  $a(t)$  denotes the vehicle's acceleration at time  $t$ ;  $\Delta t$  denotes the time step size;  $\delta_f(t)$  represents the front steering angle; and  $l_f$  and  $l_r$  are the distance of the center of the mass of the vehicle to the front and rear axles, respectively;  $w_x(t)$  and  $w_y(t)$  denote the uncertainty in the position of the center of mass, respectively. It is assumed the uncertainties originate from a closed and compact disturbance set defined as

$$\mathcal{W} := \{w = (w_x, w_y) | \zeta w \leq \theta, \zeta \in \mathbb{R}^{a \times 2}, \theta \in \mathbb{R}^b\}, \quad (2)$$

where  $b \in 2\mathbb{Z}^+$ . The disturbance set is assumed to contain the origin. Furthermore, it is assumed that the rear wheels cannot be steered. Therefore, the control input to the model (1), represented by  $\gamma(t) = (a(t), \delta_f(t))$ , is the acceleration and front steering angle pair.

## IV. ROBUST GAME-THEORETIC DECISION MAKING

At each time instant, each vehicle selects an input pair from the finite action set,  $\Gamma = \{(0, 0), (0, \delta_{f, \text{nom}}), (0, -\delta_{f, \text{nom}}), (a_{\text{nom}}, 0), (-a_{\text{nom}}, 0), (a_{\max}, 0), (-a_{\max}, 0), (a_{\text{nom}}, \delta_{f, \text{max}}), (a_{\text{nom}}, -\delta_{f, \text{max}})\}$ , where  $a_{\text{nom}}$ ,  $\delta_{f, \text{nom}}$  and  $a_{\max}$ ,  $\delta_{f, \text{max}}$  are the nominal and maximum acceleration, front steer angle, respectively. The inputs pairs in  $\Gamma$  correspond to the actions, {“maintain”, “turn slightly left”, “turn slightly right”, “accelerate”, “decelerate”, “maximum acceleration”, “maximum deceleration”, “turn left and accelerate”, “turn right and accelerate”}, respectively. The input pair to be applied at every time step is decided based on optimizing a reward function.

#### A. Reward function

The decision making process of the vehicle in choosing the optimal input pair follows a receding horizon strategy. The

nominal model used within the prediction horizon is given by

$$x_{t+j+1} = x_{t+j} + v_{t+j} \cos(\psi_{t+j} + \beta_{t+j}) \Delta t \quad (3a)$$

$$y_{t+j+1} = y_{t+j} + v_{t+j} \sin(\psi_{t+j} + \beta_{t+j}) \Delta t \quad (3b)$$

$$\psi_{t+j+1} = \psi_{t+j} + \frac{v_{t+j}}{l_r} \sin(\beta_{t+j}) \Delta t \quad (3c)$$

$$v_{t+j+1} = v_{t+j} + a_{t+j} \Delta t \quad (3d)$$

$$\beta_{t+j} = \arctan\left(\frac{l_r}{l_r + l_f} \tan(\delta_{f,t+j})\right), \quad (3e)$$

where  $j \in \mathbb{Z}_{[0, N-1]}$  represents the prediction step, and  $\gamma_{t+j} = (a_{t+j}, \delta_{f,t+j})$  denotes the input pair applied to (3) at a prediction step  $j$ . A sequence of actions,  $\gamma_t = \{\gamma_t, \gamma_{t+1}, \dots, \gamma_{t+N-1}\}$ , is chosen that maximizes a cumulative reward given by

$$\mathcal{R}(\gamma_t) = \sum_{j=0}^{N-1} \lambda^j R_{t+j}, \quad (4)$$

where  $R_{t+j}$  is the stage reward at a prediction step  $j$  determined at time step  $t$  for an input,  $\gamma_{t+j} \in \Gamma$ ;  $\lambda \in [0, 1]$  is the discount factor. By the receding horizon strategy, the input applied to (1),  $\gamma(t)$ , is the first element of  $\gamma_t^* = \{\gamma_t^*, \gamma_{t+1}^*, \dots, \gamma_{t+N-1}^*\}$  is applied at each time instant  $t$ . The stage reward at a prediction step  $j$ ,  $R_{t+j}$ , is defined as

$$R_{t+j} = \alpha^T \phi_{t+j}, \quad (5)$$

where  $\phi_{t+j} = \{\phi_{1,t+j}, \phi_{2,t+j}, \dots, \phi_{m,t+j}\}$  is the feature vector at step  $j$  and the weights for these features are in  $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_m\}$ , in which  $\alpha_i > 0, \forall i \in \mathbb{Z}_{[0:m]}$ . For the lane changing scenario in Fig. 1, the features considered are described below.

Rectangular outer approximation of the geometric contour of each vehicle is considered as shown by the dash-dotted boxes in Fig. 1. This outer approximation is referred as the collision avoidance zone (c). The features,  $\phi_{1,t}$ ,  $\phi_{2,t}$  and  $\phi_{3,t}$ , are indicator functions based on the collision avoidance zone of the vehicles that respectively characterize:

- Collision status - The intersection of the collision avoidance zone of the ego vehicle with that of any other vehicle indicates a collision or a danger of collision. If an overlap is detected then  $\phi_{1,t}$  is assigned a value  $-1$ ; and  $0$ , otherwise.
- On-road status - The intersection of the collision avoidance zone of the ego vehicle with that of green regions shown in Fig. 1 indicates that the ego vehicle is outside the road boundaries. The feature  $\phi_{2,t} = -1$  if an overlap is detected;  $\phi_{2,t} = 0$ , otherwise.
- Safe zone violation status - A safe zone (s) of a vehicle is a rectangular area that subsumes the collision avoidance zone of the vehicles with a safety margin. The safety margin is chosen based on the minimum distance to be maintained from the surrounding vehicles. If an overlap of the safe zone of the ego vehicle with that of another vehicle is detected then  $\phi_{3,t}$  is assigned a value  $-1$ ; and  $0$ , otherwise.

The other features considered in this work characterize:

- Distance to objective - In order to encourage the ego vehicle to change lane and reach a reference point in the new lane,  $(x^{\text{ref}}, y^{\text{ref}})$ , the feature  $\phi_{4,t}$  is defined as

$$\phi_{4,t} = -(|x_t - x^{\text{ref}}| + |y_t - y^{\text{ref}}|). \quad (6)$$

- Distance to lane center - The feature,  $\phi_{5,t}$ , defined as

$$\phi_{5,t} = -|y_t - y^{lc}|, \quad (7)$$

where  $y^{lc}$  is the y-coordinate of the center of the current lane that is included to encourage the ego vehicle to be at the middle of the current lane.

- Velocity error - The deviation of the velocity of the ego vehicle from a reference velocity,  $v^{\text{ref}}$ , is described by the feature  $\phi_{6,t}$  as

$$\phi_{6,t} = -|v_t - v^{\text{ref}}|, \quad (8)$$

where the reference velocity is typically chosen as the legislated speed limit.

## B. Level- $k$ framework

In a multi-agent traffic scenario, the interactive nature of the decision making process is taken into account by the features,  $\phi_{1,t}$  and  $\phi_{3,t}$ , of the stage reward in (5) that depend on the states of other vehicles. To compute the cumulative reward in (4), for a given sequence of actions of the  $l^{\text{th}}$  autonomous vehicle,  $\gamma_t[l] = \{\gamma_t[l], \gamma_{t+1}[l], \dots, \gamma_{t+N-1}[l]\}$ , it is required to predict the actions of other agents,  $\gamma_t[i] = \{\gamma_t[i], \gamma_{t+1}[i], \dots, \gamma_{t+N-1}[i]\}$ ,  $\forall i \in \mathcal{O}$ , where  $\mathcal{O} = \{i | i \in \mathbb{Z}_{[1,n]}, i \neq l\}$  with  $n$  representing the number of agents, and the corresponding state of the traffic,  $s_{t+j}$ , at prediction steps  $j = 0, 1, \dots, N-1$ , where  $s_t = [x_t[1], y_t[1], v_t[1], \theta_t[1], \dots, x_t[n], y_t[n], v_t[n], \theta_t[n]]^T$ . In this paper, level- $k$  game theory [18], [19] is utilized to model the vehicle-to-vehicle interactions and thus predict the actions of the other agents over the horizon.

In level- $k$  game theory, it is assumed that the decisions taken by the strategic agents are based on the predictions of the actions of the other agents and the agents can have different reasoning levels. The reasoning depth of an agent is indicated by  $k \in \{0, 1, \dots\}$ . The hierarchy begins with agents at level-0, where the agents make instinctive decisions to achieve the objective without accounting for the interactions between other agents. On the other hand, the agents at level- $k \forall k > 0$ , consider the interactions by assuming that the other agents are at level- $(k-1)$  and take decisions accordingly. For instance, a  $l^{\text{th}}$  level-1 agent assumes other agents are at level-0 and predicts their action sequences,  $\gamma_t^{(0)}[i] = \{\gamma_t^{(0)}[i], \gamma_{t+1}^{(0)}[i], \dots, \gamma_{t+N-1}^{(0)}[i]\}$   $\forall i \in \mathcal{O}$ , to compute its own action sequence,  $\gamma_t^{(1)}[l] = \{\gamma_t^{(1)}[l], \gamma_{t+1}^{(1)}[l], \dots, \gamma_{t+N-1}^{(1)}[l]\}$ .

The level- $k$  game theory was adapted to model the vehicle-to-vehicle interactions at an unsignalized four-way intersection in [9]. It is assumed that level-0 vehicles consider the other vehicles in the traffic scenario as stationary obstacles. Therefore, these level-0 drivers implicitly assume the others will yield the right of way, and can be regarded ‘aggressive’.

And level-1 drivers consider other drivers to be aggressive and take ‘cautious’ actions. In [9], the drivers are categorized into level-0, 1 and 2. Since, the behavior of level-2 driver will be similar to that of the level-0 drivers, in this paper, only level-0 and 1 drivers are considered. The stage reward value obtained for  $l^{\text{th}}$  level- $k$  agent at a prediction step  $j$  for an action  $\gamma_{t+j}^{(k)}[l]$ , depend on the current traffic state,  $s_0$ , the ego agent’s actions,  $\{\gamma_t^{(k)}[l], \gamma_{t+1}^{(k)}[l], \dots, \gamma_{t+j-1}^{(k)}[l]\}$ , and the actions of other agents,  $\{\gamma_t^{(k-1)}[i], \gamma_{t+1}^{(k-1)}[i], \dots, \gamma_{t+j-1}^{(k-1)}[i]\} \forall i \in \mathcal{O}$ , is given as

$$R_{t+j}^{(k)}[l] = R_{t+j} \left( \gamma_{t+j}^{(k)}[l] \mid s_0, \gamma_t^{(k)}[l], \gamma_{t+1}^{(k)}[l], \dots, \gamma_{t+j-1}^{(k)}[l], \right. \\ \left. \gamma_t^{(k-1)}[i], \gamma_{t+1}^{(k-1)}[i], \dots, \gamma_{t+j-1}^{(k-1)}[i] \right), \quad (9)$$

and its cumulative reward is

$$\mathcal{R}^{(k)} \left( \gamma_t^{(k)}[l] \right) = \sum_{j=0}^{N-1} \lambda^j R_{t+j}^{(k)}[l]. \quad (10)$$

### C. Multi-model strategy

Human drivers, initially, do not have perfect knowledge about other drivers. However, they gain better understandings of other driver’s characteristics through interactions, and therefore, resolve conflicts effectively. Similarly, the autonomous vehicles in a multi-agent traffic scenario, hold an initial belief about the driver model (level-0 and 1) of the other vehicles as a probability distribution over both models. Subsequently, based on the actual action applied by the other agents, the estimate of the probability distribution is updated at every step.

From the perspective of an  $l^{\text{th}}$  autonomous agent, the probability that the  $i^{\text{th}}$  other agent can be modeled as level- $k$  is represented by  $P_{K_i^l=k}$ . The probability of the model  $k$  is increased when it matches the actual action by

$$k^* = \arg \min_{k \in \{0, 1\}} \|\gamma[i](t) - \gamma_t^{(k)}[i]\| \quad (11a)$$

$$\tilde{P}_{K_i^l=k^*}(t) = P_{K_i^l=k^*}(t-1) + \Delta P \quad (11b)$$

$$P_{K_i^l=k}(t) = \frac{\tilde{P}_{K_i^l=k}(t)}{\sum_{\tilde{k}=0}^1 \tilde{P}_{K_i^l=\tilde{k}}(t)}, \forall k \in \{0, 1\}, \quad (11c)$$

where  $\gamma[i](t)$  and  $\gamma_t^{(k)}[i]$  represent the actual and predicted action taken by  $i^{\text{th}}$  agent assuming level- $k$  model, respectively;  $\Delta P > 0$  is a constant that denotes the rate of increment of the probability;  $\|\gamma[i](t) - \gamma_t^{(k)}[i]\| = |a(t) - a_t^{(k)}| + |\delta_f(t) - \delta_{f,t}^{(k)}|$ . When the input pair of the actual action is equal to that of the predicted action, the probability distribution remains unchanged.

In order to incorporate the multi-model strategy in the decision making process and select the optimal action according to the model of other agents, the expected cumulative reward for the  $l^{\text{th}}$  agent, using (9) and (10), is given by

$$\mathcal{R}_P(\gamma_t[l]) = \sum_{k=0}^1 P_{K_i^l=k}(t) \mathcal{R}^{(k)} \left( \gamma_t^{(k)}[l] \right), \forall i \in \mathcal{O}. \quad (12)$$

### D. Robust decision making

The mismatch between the actual position of the center of mass of the vehicle which is used to determine the rectangular outer approximation of the vehicle (see Section IV-A) and its predictions obtained using the dynamic model in (3) might lead to collision. In the multi-agent traffic scenario under consideration, there are two sources of modeling errors: (i) the uncertainties,  $w^{(m)} = (w_x^{(m)}, w_y^{(m)}) \in \mathcal{W}_m$ , resulting due to the use of a simplified model in (3); and (ii) the uncertainty,  $w^{(d)} = (w_x^{(d)}, w_y^{(d)}) \in \mathcal{W}_d$ , arising due to unknown driver model. Hence, the disturbance set defined in (2) is

$$\mathcal{W} = \mathcal{W}_m \oplus \mathcal{W}_d. \quad (13)$$

Robust approaches can be used to account for these uncertainties while computing the control actions. Since a discrete set of input actions is considered in this work, feedback min-max strategy [12] is utilized in this work for considering the uncertainties originating from the disturbance set  $\mathcal{W}$ . Since the autonomous agents update the driver models of the other agents in the multi-agent traffic scenario at each step according to (11), an adaptive scheme is proposed to incorporate the confidence on the driver models and leverage the fact that level-1 drivers are cautious. The disturbance set considered at each time step is modified of the other agents and leverage .

$$\bar{\mathcal{W}}_i^l(t) = \mathcal{W}_m \oplus P_{K_i^l=0}(t) \mathcal{W}_d \quad (14)$$

where at time  $t$ ,  $P_{K_i^l=0}(t)$  is the probability that the  $i^{\text{th}}$  agent is a level-0 driver from the perspective of the  $l^{\text{th}}$  autonomous agent, and  $\bar{\mathcal{W}}_i^l(t)$  denotes the disturbance set. It is assumed that, initially, all the agents are level-0 drivers, i.e.,  $P_{K_i^l=0}(0) = 1, \forall i \in \mathcal{O}$ . Essentially, this assumption allows the autonomous agent to be cautious with another interacting agent when there is no/less information about that agent. If an  $i^{\text{th}}$  agent is level-0, as time evolves,  $P_{K_i^l=0}(t)$  will continue to be equal to one, and hence, the autonomous agent take conservative actions (or behave like level-1 driver). On the other hand, when the  $i^{\text{th}}$  agent is a level-1 driver,  $P_{K_i^l=0}(t)$  will decrease, resulting in a reduced disturbance set size, thereby, allowing the autonomous agent to take less conservative actions and adapt to the behavior of the interacting agents while capable of handling the uncertainties arising due to the use of a simple prediction model.

When interacting with an agent  $i \in \mathcal{O}$ , the objective the autonomous agent  $l$  is to maximize the expected cumulative reward (10), while accounting for the effect of the worst-case uncertainty from the possible disturbance realizations. The optimal control sequence,  $\gamma_t^*[l] = \{\gamma_t^*[l], \gamma_{t+1}^*[l], \dots, \gamma_{t+N-1}^*[l]\}$ , is obtained by solving the

following optimization problem

$$\gamma_t^*[l] = \arg \max_{\gamma_t[l] \in \Gamma} \min_{w_{t+j}^p \in \bar{\mathcal{W}}_i^l(t)} \mathcal{R}_P(\gamma_t[l]) \quad (15a)$$

s. t.  $\forall j \in \mathbb{Z}_{0, N-1}$ , (3c), (3d), (3e),  $\forall i \in \mathcal{O}$ ,  $\forall p \in \mathcal{P}$

$$\tilde{x}_{t+j+1} = \tilde{x}_{t+j} + v_{t+j} \cos(\psi_{t+j} + \beta_{t+j}) \Delta t + w_{x,t+j}^p, \quad (15b)$$

$$\tilde{y}_{t+j+1} = \tilde{y}_{t+j} + v_{t+j} \sin(\psi_{t+j} + \beta_{t+j}) \Delta t + w_{y,t+j}^p, \quad (15c)$$

where  $w_{t+j}^p = (w_{x,t+j}^p, w_{y,t+j}^p)$  denote a possible realization of the uncertainty in the global position of the center of mass in  $x$  and  $y$  directions, respectively; and  $\mathcal{P}$  represents the set of indexes of the realizations. The autonomous agent then applies the first element  $\gamma_t^*[l]$  of the optimal control action sequence, i.e.,  $\gamma(t) = \gamma_t^*[l]$ .

## V. SIMULATION RESULTS

The proposed adaptive robust approach is validated for the lane changing maneuver on a numerically simulated three lane highway section<sup>1</sup>. Consider the multi-agent traffic scenario shown in the subfigure [1, 1] (first element represents row and second element represents column) in Fig. 2. The objective of the autonomous agent (blue) is to change from lane II to lane III, while the human agents keep their respective lane. All the human drivers are assumed to be level-1, i.e., they exhibit cautious behavior, however, it is unknown to the AV. The sampling time is set to 0.5 s and two step prediction horizon is considered. The proposed methodology is compared to the nominal, and the robust strategies during the decision making process.

The traffic simulation under the nominal decision making strategy is shown in the subfigures in the first column of Fig. 2. In this case, the disturbance set considered for solving (15) is chosen as an empty set, i.e.,  $\bar{\mathcal{W}}_i^l(t) = \emptyset$ . Since the disturbances are not considered in this case, it can be noted the AV chooses to steer left as soon as the simulation begins which provides the maximum reward. This move is considered to be aggressive. However, the human vehicle 4, being cautious, reacts by steering left at time  $t = 2$  s and returning to the lane center at  $t = 4$  s once the AV has passed. The AV completes the lane change between 60 m and 70 m.

The results obtained by using the robust strategy that considers  $\bar{\mathcal{W}}_i^l(t) = \mathcal{W}$ , for the decision making process, is shown in the subfigures in the third column of Fig. 2. The disturbance set  $\mathcal{W}$  is defined as in (13). As mentioned in Section IV-D, initially, all the human drivers are considered to be level-0 from the perspective of the AV, while they are assigned to be level-1 drivers. As seen in Fig. 3, due to interactions with vehicle 4, as time evolves, the AV is able to reduce the probability of vehicle 4 being level-0. It can also be noted from subfigure [3, 3] in Fig. 2, robust control actions are taken to account for the set,  $s \oplus \bar{\mathcal{W}}_i^l(t)$ , dash-dotted box surrounding the human drivers shown in the Fig. 2. As a

<sup>1</sup>The code is made publicly available at <https://github.com/gokulsivasankar/RobustDecisionMaking>

result, the AV is conservative in performing the lane change maneuver by completing it between 90 m and 100 m.

Finally, the simulation results for the adaptive control strategy is shown in the subfigures in the second column of Fig. 2. The initial disturbance set of all the human agents are similar to the previous case. However, by using the adaptive disturbance set in (14) for computing the control actions, the AV is able to complete the lane change around 70 m. Also, it can be observed that the size of the set,  $s \oplus \bar{\mathcal{W}}_i^l(t)$ ,  $\forall i \in \{1, 3\}$ , remains constant because, the level-0 and 1 actions are the same for these two vehicles and therefore, according to (11), the driver model is not updated. The proposed strategy allows the AV to behave cautiously when there is uncertainty in the driver model and adapt its behavior according to the estimate of the model of the interacting driver.

## VI. CONCLUSIONS

An adaptive robust decision making strategy has been proposed for the autonomous vehicles sharing the road with human drivers in a multi-agent traffic scenario. The interactions between vehicles are modeled using a level-k game-theoretic framework. The proposed robust control strategy accounts for the uncertainties in the vehicle dynamic model and the driver model estimation. The autonomous vehicle estimates the driver model of the other agents at each time step and is shown to use it to adapt its behavior through numerical simulations of a lane changing maneuver.

## ACKNOWLEDGEMENT

We would like to express our appreciation to Mr. Nan Li and Dr. Anouck Girard who provided expertise that greatly assisted the research.

## REFERENCES

- [1] R. Okuda, Y. Kajiwara, and K. Terashima, “A survey of technical trend ofadas and autonomous driving,” in *Technical Papers of 2014 International Symposium on VLSI Design, Automation and Test*, pp. 1–4, IEEE, 2014.
- [2] D. A. Lazar, R. Pedarsani, K. Chandrasekher, and D. Sadigh, “Maximizing road capacity using cars that influence people,” in *2018 IEEE Conference on Decision and Control (CDC)*, pp. 1801–1808, IEEE, 2018.
- [3] T. L. Griffiths, F. Lieder, and N. D. Goodman, “Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic,” *Topics in cognitive science*, vol. 7, no. 2, pp. 217–229, 2015.
- [4] L. Claussmann, A. Carvalho, and G. Schildbach, “A path planner for autonomous driving on highways using a human mimicry approach with binary decision diagrams,” in *2015 European Control Conference (ECC)*, pp. 2976–2982, IEEE, 2015.
- [5] S. Brechtel, T. Gindel, and R. Dillmann, “Probabilistic decision-making under uncertainty for autonomous driving using continuous pomdps,” in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 392–399, IEEE, 2014.
- [6] D. Sadigh, S. Sastry, S. A. Seshia, and A. D. Dragan, “Planning for autonomous cars that leverage effects on human actions.,” in *Robotics: Science and Systems*, vol. 2, Ann Arbor, MI, USA, 2016.
- [7] D. O. Stahl, “Evolution of smartn players,” *Games and Economic Behavior*, vol. 5, no. 4, pp. 604–617, 1993.
- [8] 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 Transactions on control systems technology*, vol. 26, no. 5, pp. 1782–1797, 2017.

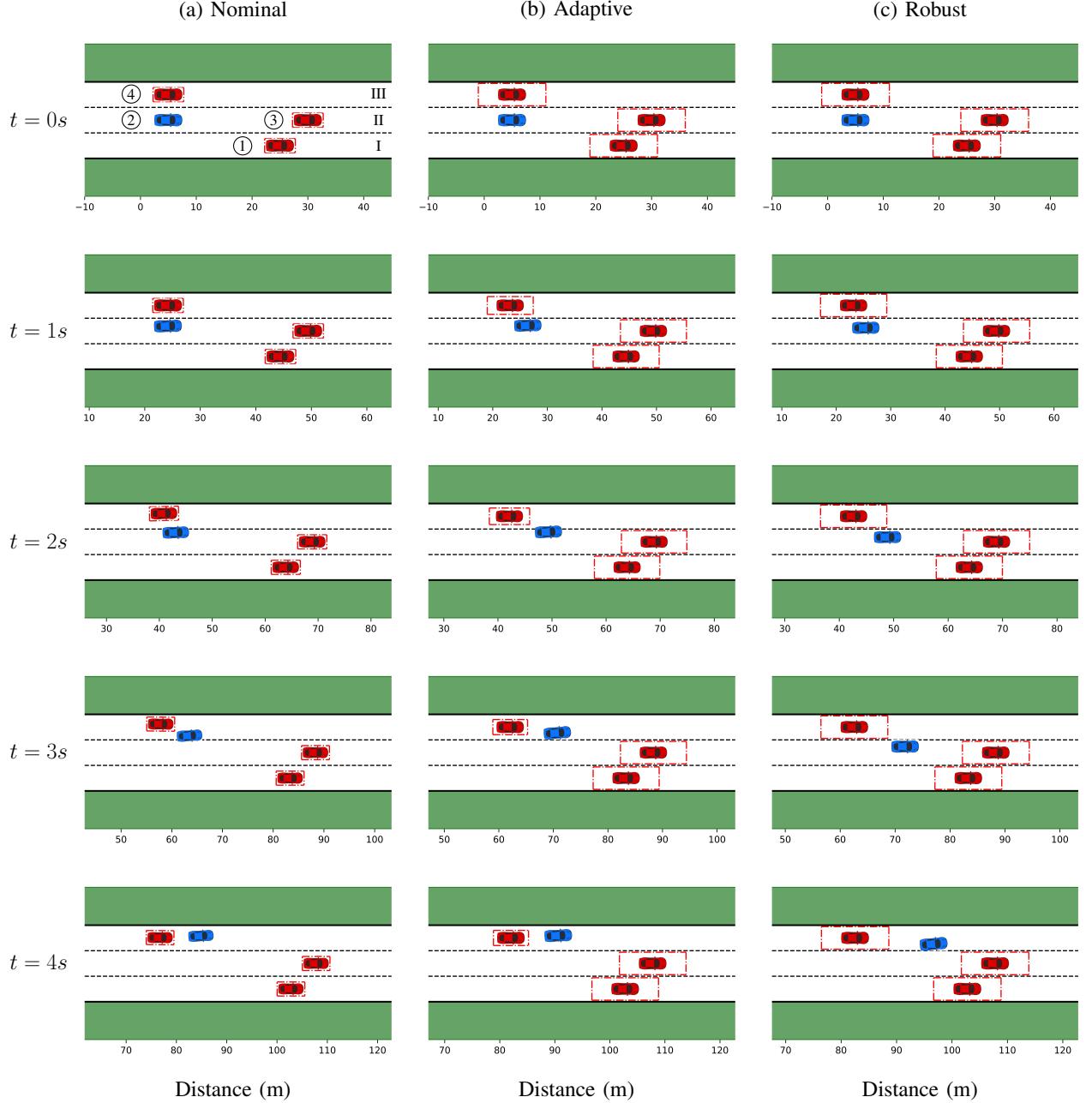


Fig. 2. A four second simulation sequence with a one second time interval (see along the column) showing the lane changing maneuver performed by the autonomous vehicle under (a) nominal; (b) adaptive; and (c) robust decision making strategies in a multi-traffic scenario. The circled numbers indicate the vehicle ids and the Roman numerals denote the lane number in the subfigure [1, 1]. The dash-dotted lines indicate the set  $s \oplus \bar{\mathcal{W}}_i^l(t)$ ,  $\forall i \in \mathcal{O}$  from the perspective of the autonomous agent.

- [9] N. Li, I. Kolmanovsky, A. Girard, and Y. Yildiz, "Game theoretic modeling of vehicle interactions at unsignalized intersections and application to autonomous vehicle control," in *2018 Annual American Control Conference (ACC)*, pp. 3215–3220, IEEE, 2018.
- [10] R. Tian, S. Li, N. Li, I. Kolmanovsky, A. Girard, and Y. Yildiz, "Adaptive game-theoretic decision making for autonomous vehicle control at roundabouts," in *2018 IEEE Conference on Decision and Control (CDC)*, pp. 321–326, IEEE, 2018.
- [11] G. I. Jin, S. Bastian, M. M. Richard, and A. Matthias, "Risk-aware motion planning for automated vehicle among human-driven cars," in *2019 American Control Conference (ACC)*, pp. –, IEEE, 2019.
- [12] P. O. Scokaert and D. Mayne, "Min-max feedback model predictive control for constrained linear systems," *IEEE Transactions on Automatic control*, vol. 43, no. 8, pp. 1136–1142, 1998.
- [13] D. Q. Mayne, M. M. Seron, and S. Raković, "Robust model predic-
- itive control of constrained linear systems with bounded disturbances," *Automatica*, vol. 41, no. 2, pp. 219–224, 2005.
- [14] A. Richards and J. P. How, "Model predictive control of vehicle maneuvers with guaranteed completion time and robust feasibility," in *American Control Conference, 2003. Proceedings of the 2003*, vol. 5, pp. 4034–4040 vol.5, June 2003.
- [15] G. S. Sankar, R. C. Shekhar, C. Manzie, T. Sano, and H. Nakada, "Fast calibration of a robust model predictive controller for diesel engine airpath," *IEEE Transactions on Control Systems Technology*, pp. 1–15, 2019.
- [16] G. S. Sankar, R. C. Shekhar, C. Manzie, T. Sano, and H. Nakada, "Model predictive controller with average emissions constraints for diesel airpath," *Control Engineering Practice*, vol. 90, pp. 182 – 189, 2019.
- [17] J. Kong, M. Pfeiffer, G. Schildbach, and F. Borrelli, "Kinematic and

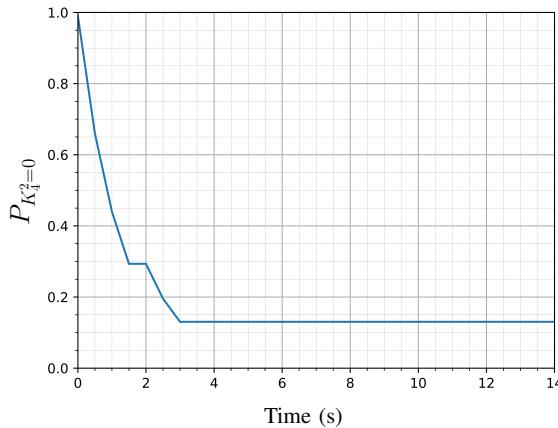


Fig. 3. Probability that the vehicle to the left of the autonomous vehicle can be modeled as level-0 from the perspective of the autonomous vehicle.

- dynamic vehicle models for autonomous driving control design,” in *2015 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1094–1099, IEEE, 2015.
- [18] M. A. Costa-Gomes and V. P. Crawford, “Cognition and behavior in two-person guessing games: An experimental study,” *American Economic Review*, vol. 96, no. 5, pp. 1737–1768, 2006.
  - [19] M. A. Costa-Gomes, N. Iriberri, and V. P. Crawford, “Comparing models of strategic thinking in van huyck, battalio, and beil’s coordination games,” *Journal of the European Economic Association*, vol. 7, no. 2/3, pp. 365–376, 2009.