Interaction-Aware Merging in Mixed Traffic with Integrated Game-theoretic Predictive Control and Inverse Differential Game

Mohamed-Khalil Bouzidi[1], Ehsan Hashemi[2], *Senior Member, IEEE*

***Abstract*— This paper presents an interaction-aware motion planning and control framework for time-critical traffic scenarios in which interaction with vehicles driven by humans is required. For safe motion planning the proposed method considers interaction between the automated driving system and other vehicles using game theory. The framework includes a novel inverse differential game based on a LSTM to estimate the human driver's objective function online. Then, a game-theoretic predictive controller utilizes these estimates for controlling the automated driving system and predicting the trajectory of the human-driven vehicle. The developed framework is validated in several safety-critical scenarios and testing conditions using *CarSim* high-fidelity simulations including human-in-the-loop case studies with six different test subjects.**

I. Introduction

Automated driving systems (ADS) are being developed for various driving conditions on public roads. One of the main challenges for motion planning of these systems is interaction with other vehicles in mixed traffic and highly dynamic environments (e.g., highway merging) causing the so-called frozen robot problem (FRP) [1]. This FRP, has roots in a decoupled prediction module (which predicts trajectories of dynamic objects in the scene for a horizon) and a planning module (that generates a safe trajectory for the ADS [2]). The surrounding vehicles are therefore considered as dynamic obstacles that do not react to the driven trajectory of the ADS. This could cause the ADS to misjudge the situation and commands a control input that imperils the safety of the passenger or not perform a maneuver at all because possible maneuvers could not be identified (i.e. FRP). To be capable of correctly anticipating the situation, the ADS needs to be aware that the surrounding vehicles interact with the ADS to avoid collisions. This necessitates coupled prediction and planning tasks since drivers adapt their behavior (and generate a control input, in terms of steering and acceleration/brake request) according to the trajectories of the other vehicles and vice versa.

In this regard, Game theory provides a feasible framework to address this simultaneous planning and prediction. However, existing approaches rely on the assumption that the traffic participants are connected and exchange their objectives via V2V (Vehicle-to-Vehicle Communication). Hence, the game is modeled as a cooperative game based on the Pareto-efficient solution (such as in [3], [4]). Nevertheless, in the mixed traffic situation, not all traffic participants are connected via V2V and do not necessarily act cooperatively. Consequently, highway interaction has also been modeled as an uncooperative game, where Nash equilibrium (NE) solutions in differential/dynamic games have been deployed based on iterative linear quadratic regulators (iLQR) [5] or Newton [6]–[8] methods, solving the game in a receding horizon manner. However, these approaches capitalize on the assumption of known driver's objective function to the ADS.

Drivers generally have different goals and driving styles which necessitates the online estimation of the objective function of the human. To estimate objective functions, inverse differential game (IDG) is used in the literature. However, existing methods are either too computationally expensive to run online, such as direct methods [9] which estimate the objective function by forward solving the NE in an inner loop, methods based on Inverse Reinforcement Learning [10], [11], or methods that rely on too strong assumptions derived from Inverse Optimal Control [12].

An interaction-aware motion planning is proposed in [13] which employs a feed-forward network for online IDG and implicit layers that finds a NE in the next step. However, the lack of a vehicle model and limited (less than 50) training data means no guarantee for feasible trajectories. A game-theoretic model predictive control (MPC) is used in [6], [14] to solve the NE; an unscented Kalman filter (UKF) is used in [14] for online estimation of objective function parameters (i.e. a direct method for IDG). A large number of states and objective function parameters (due to more complex dynamic models) causes computational challenges (in real-time) for the proposed UKF-based estimator. More importantly, the underlying physics of the vehicle dynamics has not been included in existing approaches; this may cause the generation of control actions (for mixed traffic scenarios). The existing methods have used the unicycle model as a dynamic model at most which cannot appropriately reflect the vehicle behavior in highway merging maneuvers that are usually high-speed and may subject to uneven torque distribution at each axle. This is especially important in the task of unified Motion planning and Control that we are considering. To address these, this paper proposes an interaction-aware motion planning and control framework based on game theory considering the single track model for generating dynamically feasible trajectories and utilizes a LSTM network for online IDG for merging applications in real-time. Our framework is evaluated in mixed traffic merging scenarios (with a human driver in the loop setup and an

[1] Mohamed-Khalil Bouzidi is with Continental AG and PhD candidate at the Free University Berlin, Germany. Work was performed for the Master's thesis while at the NODE Lab, Department of Mechanical Engineering, University of Alberta, Edmonton, AB, Canada.

[2] Ehsan Hashemi is with the Department of Mechanical Engineering, University of Alberta, Edmonton, AB, Canada. ehashemi@ualberta.ca

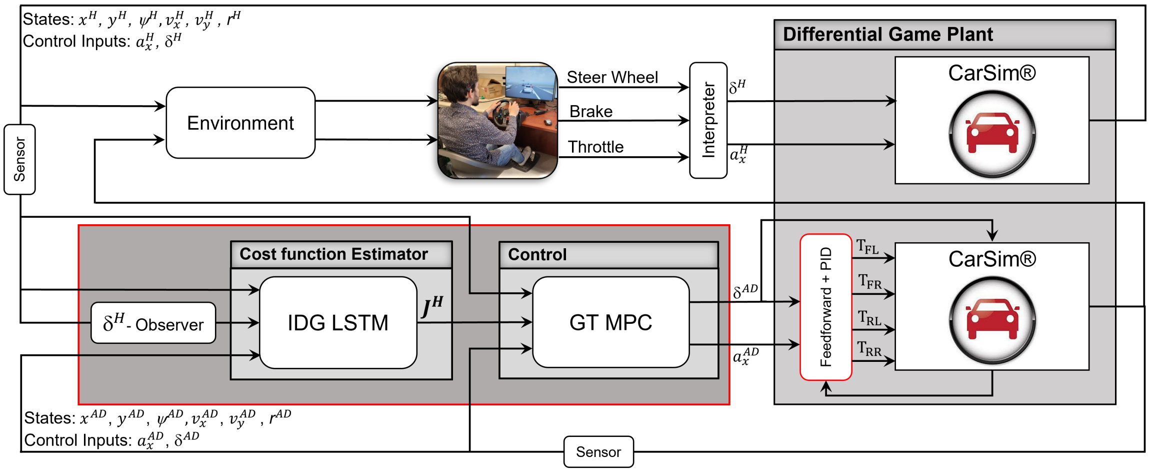

Fig. 1: The interaction-aware game-theoretic motion planning and control framework using unknown input observer

accurate model plant using high-fidelity *CarSim* simulations) which is not feasible for other game-theoretic approaches [7], [13], [14] due to the aforementioned constraints. As a result, the main contributions of our work are summarized as:

- Developing a novel LSTM-based inverse differential game to estimate the human driver's objective function online for a predictive control strategy in real-time applications
- Designing an unknown input observer to estimate the control input of the surrounding agent using onboard sensory measurements
- Devising a game theoretic predictive control that considers the vehicle lateral dynamics to ensure lateral stability and dynamic feasibility for safer Human-ADS interaction.

II. Problem Statement and Model Description

Consider a highway merging scenario in which the ego vehicle (with full ADS feature) is at the highway entrance merging into the traffic. There is a human-driven vehicle (on the left) negotiating/interacting with the ADS accordingly (i.e., Fig. 2). The objective is to plan a feasible collision-free trajectory and to control the vehicle to follow this trajectory in this safety- and time-critical driving scenario.

To grasp the interactive nature of the highway merging, we model this scenario as an *uncooperative game*. We select the NE as the solution concept since in Nash games no player has a structural advantage over other player(s) in contrast to the leader-follower structure of a Stackelberg game.

Definition 1. *The Nash equilibrium is a solution concept that arises if every player/agent i acts simultaneously and optimally with respect to its own objective function J^i and its beliefs of the strategies of the other player j and these beliefs are correct. This corresponds to:*

$$J^i(\boldsymbol{x}^*, \boldsymbol{u}_i^*, \boldsymbol{u}_j^*) < J^i(\boldsymbol{x}^*, \boldsymbol{u}_i, \boldsymbol{u}_j^*) \quad \forall i. \tag{1}$$

Expressed in game-theoretic terms, this highway merging scenario can be described as a scene in which two agents $i \in \{AD, H\}$ (i.e., automated driving system and human-driven vehicle) negotiate in a dynamic game with a finite horizon. To model this scene as a Nash game, we make the following assumption:

Assumption 1. *The Human driver acts rationally to minimize his objective function J^H.*

Rational decision-making and the reaction of the body of the human in driving maneuvers are often modeled in the literature [15] by an optimal controller since humans predict the driving actions of others based on their knowledge and react rationally based on this prediction.

It should be noted that there is no communication between the vehicles such as V2V. Hence, the ADS does not have access to the objective function of the other vehicle (i.e., agent H). However, this objective function is required to calculate the NE. Thus, the objective function will be estimated online using an IDG and the observed trajectories of the vehicles. With the objective function estimation, the solution of the NE is calculated based on the model of the system as shown in Fig. 1). We model the game as a differential game since the vehicle kinematics and dynamics are naturally continuous. The differential games for this problem is described by a dynamical system $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t), \boldsymbol{u}^H(t), \boldsymbol{u}^A(t))$. The unicycle model used in the existing methods does not reflect the actual dynamics of the vehicle (due to not considering tire forces), especially in high-slip cases. This has been addressed in our approach by using a dynamic single-track model to predict trajectories that correctly estimate the maneuvering capabilities (and handling limits) of the vehicles and thus generate dynamically feasible trajectories. Thus, lateral tire forces $F_{y,f}, F_{y,r}$ are taken into account, assuming linear tire forces. The slip angle at the vehicle's front and rear axles are defined by $\alpha_f = \delta - \frac{l_f r + v_y}{v_x}$ and $\alpha_r = \frac{l_r r - v_y}{v_x}$, respectively. δ is the front axle's steering angle and the longitudinal speed, lateral speed, and yaw rate of the vehicle in the body frame (attached to the vehicle center of gravity, CG) is denoted by v_x, v_y, and r, respectively. The vehicle dynamics can then be described as:

$$\dot{v}_x = v_y r + a_x, \quad \dot{v}_y = -v_x r + (\frac{F_{y,f} + F_{y,r}}{m})$$
$$\dot{r} = \ddot{\psi} = (\frac{F_{y,f} l_f - F_{y,r} l_r}{I_z}) \tag{2}$$

where a_x is the longitudinal acceleration (measured by an IMU) in the body frame, m is the vehicle mass, and front/rear axles to CG are denoted by l_f, l_r. The dynamics on the longitudinal/lateral position yields $[\dot{x}, \dot{y}]^\top = R_\psi [v_x, v_y]^\top$ with the rotation matrix R_ψ for both ADS and human-driven vehicle. These models are then stacked, except for the first state (i.e., longitudinal position), to form the dynamical model for the differential game with the control inputs $\boldsymbol{u}^i = [a_x^i, \delta^i]^\top$ for $i \in \{AD, H\}$ and augmented states $\boldsymbol{x} = [e_x, y^{AD}, y^H, \psi^{AD}, \psi^H, v_x^{AD}, v_x^H, v_y^{AD}, v_y^H, r^{AD}, r^H]^\top$, in which the inter-vehicular distance e_x is used as a state variable. The lateral position y of both agents will be used as state variables due to the fact that they could be measured using visual/LiDAR-based navigation methods based on the road features (e.g., lane, curb, light poles). To determine the objective function of the human-driven vehicle H, we use a basis function approach. It is assumed that the objective function of both agents $i \in \{AD, H\}$ can be described by $J^i = \int \boldsymbol{\theta}^i \, \boldsymbol{\phi}^i(t) \, dt$. The parameters $\boldsymbol{\theta}^H$ consisting of the objective weights θ_j^H, the desired velocity $v_{x,d}^H$, and the desired lane position y_d^H have to be identified using IDG. The parameters $\boldsymbol{\theta}^{AD}$ of ADS are tunable. Thus, the objective of the H vehicle is described by (3), in which the term minimizing the lateral velocity thus the side slip angle is crucial to ensure lateral stability.

$$J^i = \int \theta_1^i \left(v_x^i - v_{x,d}^i\right)^2 + \theta_2^i \left(y^i - y_d^i\right)^2 + \theta_3^i (v_y^i)^2$$
$$+ \theta_4^i \tanh \underbrace{\left(\sqrt{\left(e_x^2 + \left(\frac{b}{a}\left(y^H - y^{AD}\right)\right)^2 \right)} - a \right)}_{\Gamma}$$
$$+ \theta_5^i (a_x^i)^2 + \theta_6^i \left(\delta^i\right)^2 \, dt \quad \forall i \tag{3}$$

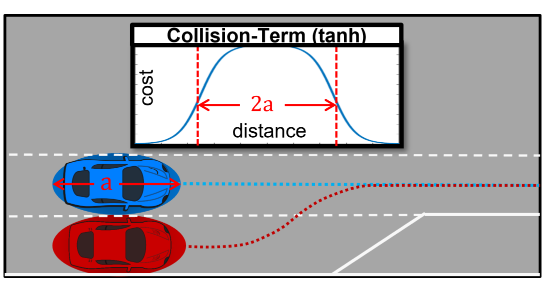

Fig. 2: Illustration of the collision avoidance term in (3) where ellipse semi-axes (a,b) represent the vehicle size.

The objective function (3) can be transformed to the basis function structure with the parameters $\boldsymbol{\theta}^i = [\theta_1^i, v_{x,d}^i, \theta_2^i, y_d^i, \theta_3^i, \theta_4^i, \theta_5^i, \theta_6^i]^\top$ and the basis function

$$\boldsymbol{\phi}^i(t) = [(v_x^i)^2, -2v_x^i, (y^i)^2, -2y^i,$$
$$(v_y^i)^2, \tanh(\Gamma), (a_x^i)^2, (\delta^i)^2]^\top \tag{4}$$

III. Unknown Input Observer Design

The ADS in our control framework is equipped with monocular cameras, LiDARs, and radars for the detection of static/dynamic objects (i.e., surrounding vehicles) and measuring their relative distances/headings. The steering angle of the human-driven vehicle is not measurable; thus, an unknown input observer (UIO) [16] is designed to estimate this input. The steering control input of the human-driven vehicle is estimated in real-time using an unknown input observer and the vehicle linear lateral dynamical model $\dot{\boldsymbol{\zeta}} = A\boldsymbol{\zeta} + B\boldsymbol{u}$ with the states $\boldsymbol{\zeta} = [v_y^H, r^H]^\top$, measurable outputs $\boldsymbol{y} = \boldsymbol{\zeta}$, and the unknown input $\boldsymbol{u} = \delta^H$. The linear state and input matrices A, B are functions of the tire cornering stiffness, longitudinal speed (measured by radars and LiDARs of the ego vehicle, i.e., automated driving system) and nominal values of the vehicle geometry identified by the type of the detected vehicle [17].

The following observer approximates the steering input of the surrounding vehicle, where $\hat{\boldsymbol{\zeta}}$ is the vehicle lateral states (i.e., lateral speed and yaw rate), and $\boldsymbol{z}$ is the state variable for the observer.

$$\dot{\boldsymbol{z}} = F\boldsymbol{z}(t) + K\boldsymbol{y}(t), \quad \hat{\boldsymbol{\zeta}} = \boldsymbol{z}(t) + \bar{H}\boldsymbol{y}(t),$$
$$\hat{\delta}^H = [B^\top B]^{-1} B^\top (\dot{\boldsymbol{z}} - \dot{\hat{\boldsymbol{\zeta}}} - A\hat{\boldsymbol{y}}). \tag{5}$$

Defining the estimation error $\boldsymbol{e} = \boldsymbol{\zeta} - \hat{\boldsymbol{\zeta}}$, we can rewrite the error dynamics with the dynamics as

$$\dot{\boldsymbol{e}} = (A - \bar{H}A - K_1)\boldsymbol{e} + [F - (A - \bar{H}A - K_1)]z \tag{6}$$
$$+ [K_2 - (A - \bar{H}A - K_1)]\boldsymbol{y} + (\bar{H} - I)Bu, \tag{7}$$

which is asymptotically stable if matrices $\bar{H}, F, K = K_1 + K_2$ are designed as $\bar{H} = B[B^\top B]^{-1}B^\top$, $F = (I - \bar{H}C)A - K_1 C$, $K_2 = F\bar{H}$ to decouple the state estimation dynamics from the unknown input and K_1 is selected so that F is Hurwitz [16]. Merging scenarios mostly involve speed variation due to motion planning based on interaction with the surrounding vehicle; this has been considered in the proposed method and state/input matrices are updated using the measured velocity.

IV. Game-theoretic Predictive Control

Using the nonlinear system model $\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u}^H, \boldsymbol{u}^{AD})$ and the objective function of the ADS described in section II, an optimal controller can be designed.

Remark 1. *A feedback NE is computationally intractable for nonlinear problems [18]. On the other hand, an open-loop solution which is only depending on x_0 can be found. In contrast to the feedback solution, the open-loop solution does not result in a NE if there are disturbances.*

Thus, we use a game-theoretic model predictive control (GT MPC) which solves the open-loop problem for a prediction horizon but updates the initial state values x_0 with the actual observed state in each time step. The solution for the open-loop NE (OLNE) can be determined numerically by transforming it into a nonlinear program (NLP). The transformation yields an NLP with the variables: $\boldsymbol{X} = [\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \ldots,\ \boldsymbol{x}_{N_p}\ ]^\top, \boldsymbol{U}^i = [\boldsymbol{u}_1^i,,\ \ldots,\ \boldsymbol{u}_{N_c}^i]^\top$ where N_p corresponds to the prediction horizon and N_c corresponds to the control horizon. We formulate the Lagrangian as in:

$$L^i = J^i + \sum_{k=0}^{N_p} (\boldsymbol{\lambda}_k^i)^\top [\boldsymbol{x}_{k+1} - f_k(\boldsymbol{x}_k, \boldsymbol{u}_k^H,\ \boldsymbol{u}_k^{AD}\ )]\ \forall i \quad (8)$$

Accordingly, for an OLNE the following optimality condition must hold:

$$\begin{aligned} &\boldsymbol{G^i} = \nabla_{\boldsymbol{X},\boldsymbol{U}^i}\ L^i(\boldsymbol{X}, \boldsymbol{U}^{AD}, \boldsymbol{U}^H, \boldsymbol{\lambda}^{AD}, \boldsymbol{\lambda}^H) = \boldsymbol{0} \quad \forall i \\ &\boldsymbol{C} = [\boldsymbol{x}_1 - f_k(\boldsymbol{x}_0, \boldsymbol{u}_0)\ \ldots\ \boldsymbol{x}_{N_p} - f_k(\boldsymbol{x}_{N_p-1}, \boldsymbol{u}_{N_p-1})]^\top = \boldsymbol{0} \end{aligned} \quad (9)$$

To find a solution that satisfies the optimality conditions, CasADi [19] is used with the following search direction:

$$\begin{aligned} &\boldsymbol{G} = [\boldsymbol{G^{AD}}, \boldsymbol{G^H}, \boldsymbol{C}]^\top \\ &\boldsymbol{H} = \nabla_{(\boldsymbol{X},\boldsymbol{U},\boldsymbol{\lambda})} \boldsymbol{G}(\boldsymbol{X}, \boldsymbol{U^{AD}}, \boldsymbol{U^H}, \boldsymbol{\lambda}^{AD}, \boldsymbol{\lambda}^H) \\ &\boldsymbol{p}_{ND} = -\boldsymbol{H}^{-1}\ \boldsymbol{G} \end{aligned} \quad (10)$$

Proposition 1. *Suppose an OLNE exists and assume that the initial trajectory $\tau_0 = [\boldsymbol{X}_0, \boldsymbol{U}_0^{AD}, \boldsymbol{U}_0^H]^\top$ is sufficiently close to the OLNE $\tau^* = [\boldsymbol{X}^*, \boldsymbol{U}^{AD*}, \boldsymbol{U}^H]^\top$ such that $||\tau_0 - \tau^*|| \leq \epsilon$. Then, the proposed Newton method converges locally to an OLNE at a quadratic rate.*

Proof: We know that the condition for the OLNE is an open-loop strategy $\boldsymbol{u}^i$ (based on $\boldsymbol{x}_0$) for both players i for the minimization program (11) subject to $\boldsymbol{x}_{k+1} = f_k(\boldsymbol{x}_k, \boldsymbol{u}_k^H, \boldsymbol{u}_k^{AD})$ and $\boldsymbol{x}_0 = \boldsymbol{x}(0)$.

$$\min_{\boldsymbol{x},\boldsymbol{u}^i} J^i(\boldsymbol{x}, \boldsymbol{u}_i),\ \forall i, \quad (11)$$

The matrix $\boldsymbol{G}$ is differentiable and $\boldsymbol{H}$ is Lipschitz continuous. Hence, the Newton method will converge locally to the root $\boldsymbol{G} = \boldsymbol{0}$ at a quadratic rate if the previously defined conditions are fulfilled (the root satisfies (11)).

Additionally, we perform a nonlinear transformation for box constraints of the control inputs with a constrained function in the system model $\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u}^i = u_{\text{max}}^i \tanh(c\ \tilde{\boldsymbol{u}}^i))$.

V. Inverse Differential Game

The IDG aims at using previously observed trajectories $\boldsymbol{x}(t)$, $\boldsymbol{u}^H(t)$, and $\boldsymbol{u}^{AD}(t)$ to determine parameters $\boldsymbol{\theta}^H$ of the objective function $J = \int \boldsymbol{\theta}^H \boldsymbol{\phi}^H(t) dt$ with the basis function $\boldsymbol{\phi}^H(t)$ introduced in (4), where eight parameters of the human's objective function need to be identified. However, the problem is ill-posed, i.e., the parameter set $\tilde{\boldsymbol{\theta}} = c\ \boldsymbol{\theta}^*$ with $c \in \mathbb{R}$ also leads to the behavior of $\boldsymbol{\theta}^*$. Due to this non-uniqueness, only seven parameters are sought, since one parameter can be selected arbitrarily. One crucial requirement in our use case is that the estimation has to be done at each sample time. This requirement can't be fulfilled by the Hamilton-based IDG presented in [12] since the accuracy of the estimation depends strongly on the extent to which the observed trajectory represents the complete trajectory [18]. The UKF approach introduced in [14] must solve $2N + 1 = 15$ (N is number of unknown $\boldsymbol{\theta}$) GT MPC problems to estimate the objective function at each time step through the spread of sample points. Considering our dynamical system dimension, real-time capability is critical for the UKF. Thus, a data-driven approach (i.e., LSTM that is a subset of the recurrent neural network for time series problems) for less computational complexity, is used.

The LSTM structure is well suited to estimate states from successively arriving features; the length of the input time series can also vary. Therefore the duration of the maneuver does not have to be known in advance. The objective is to predict the parameter set $\boldsymbol{\theta}_k^H = \boldsymbol{\theta}^H$ (i.e., sequence-to-one-regression problem). The input sequence layer is fed with the current state $\boldsymbol{x}_k$ and the control inputs of the human $\boldsymbol{u}_k^H$ as well as the previous ADS' control input $\boldsymbol{u}_{k-1}^{AD}$. This is due to the fact that the decision of the agent cannot be predicted by observing the human-driven vehicle in isolation since the interaction causes this decision. In Nash games, it is known that each player aims to minimize its basis function in order to minimize the objective function having positive weights of the basis functions (i.e., the solution depends on the basis function values).

Remark 2. *The basis function $\boldsymbol{\phi_k^H}$ at each time step t_k is calculated and used as an input for the neural network. This additional input was chosen to incorporate prior knowledge to improve the generalization performance of the network.*

Hence, the procedure of the network can be described as a continuous map with nonlinear dynamics from input to output $\boldsymbol{\theta}^H = \mathcal{N}(\boldsymbol{x}_k, \boldsymbol{u}_k^H, \boldsymbol{u}_{k-1}^{AD}, \boldsymbol{\phi_k^H})$. First, there are two LSTM layers with 80 and 60 hidden units respectively. The first LSTM outputs a sequence and the second LSTM layer only outputs the last time step. The two LSTM layers are followed by a fully connected layer consisting of 40 neurons and the RELU being the activation function.

A. Synthetic Training Data Generation

The acquisition of training data remains a challenge in scenarios involving complex vehicle interactions. Furthermore, there is no dataset available which is labeled with the ground truth objective function parameters J^H that we need for our LSTM-based IDG. To overcome these we leverage our Differential Game model to simulate a diverse range of merging scenarios. This offers the advantage of generating an arbitrary number of training data with different complexity/diversity. This process starts with the *i)* Parameter Sampling: Initial states of the differential game and the objective function parameter of both vehicles are randomly sampled to define the conditions of the merging scenario, governed by a uniform distribution with defined resolutions and parameter ranges. *ii)* Using the sampled parameters, the GT MPC generates trajectories for the interacting vehicles, assuming that the trajectories generated by GT MPC are sufficiently consistent with real driving scenarios involving humans. This has been confirmed by our comprehensive testing in the last section. *iii)* The generated trajectories undergo validation to ensure scenarios resulting in collisions, unsuccessful merging, or other undesirable outcomes are filtered out.

Consequently, different lengths of trajectories for specific periods (e.g., $t_n \leq t \leq t_m$) are used for training such that LSTM estimates the parameter vector after observing a fraction of the trajectory (checked with RMS error of the vehicle trajectory for several human-in-the-loop experiments).

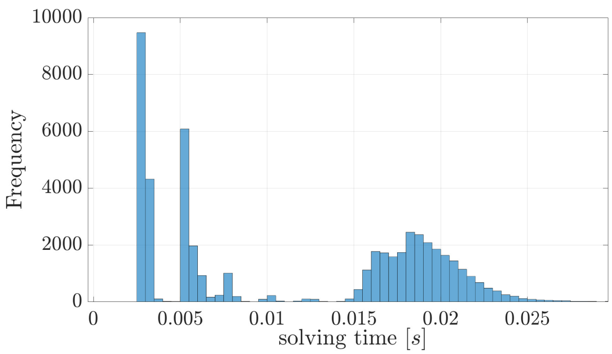

Fig. 3: Solving time histogram of the GT MPC in every time step in the MCA (average = 11.7 ms)

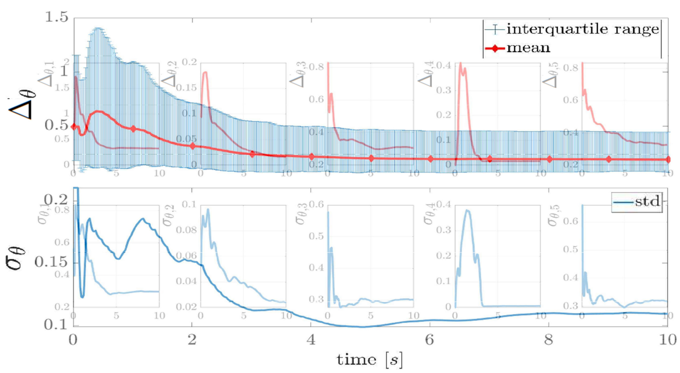

Fig. 4: Objective function parameter estimation error of all samples in the MCA.

VI. Performance Evaluation

First, it is shown that the framework exhibit reliable performance when the traffic participant acts rationally (i.e., the human is replaced by a GT predictive control). Real-time performance of the proposed controller is illustrated in Fig. 3 using a histogram of the solving time. A Monte-Carlo-Analysis (MCA) is conducted using 100 random merging scenarios using high-fidelity simulations in *CarSim* (in a 3.80 GHz AMD Ryzen 9 3900X 12-Core processor with 32 GB memory). The normalized estimation error of the objective function's parameters is defined by $\Delta_{\theta,i}(t) = \frac{||\hat{\theta}_i(t)-\theta_i^{tr}||}{\theta_i^n}$ is calculated from the ground truth value θ_i^{tr}, the estimated parameter $\hat{\theta}_i$, and the normalization factor θ_i^n which is the average value of θ_i^t in all sampled test runs. As a result, the estimation error mean and standard deviation are compared for various parameters in Fig. 4.

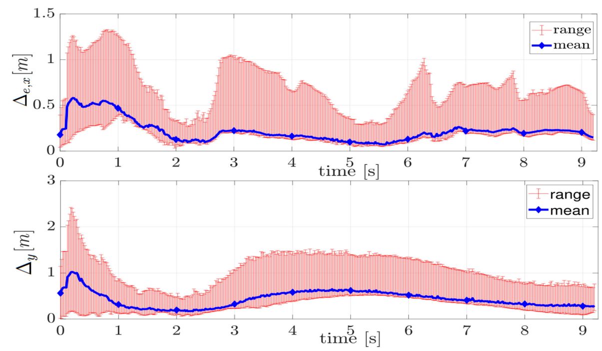

Fig. 5: Evaluation of the longitudinal distance and the lateral position prediction error of all subjects in the case study

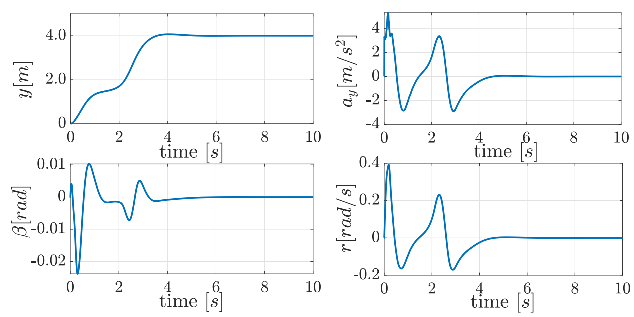

Fig. 6: Lateral response of the ADS (including sideslip angle β and yaw rate r) for a merging

Assumption 1 is investigated in detail through several experiments in the following. We conduct a case study with 6 different subjects/drivers who are not familiar with the algorithm. Subjects drive the human-in-the-loop control system on the high-fidelity driving simulator twice. The first half of the participants start with a conventional planner similar to [20] and the second half with the new planner. We selected a challenging scenario for the ADS where both cars start with the same initial speed of $v_{x,0} = 20$m/s and the longitudinal distance of $e_{x,0} = 1$m. In all cases, the novel approach was able to perform a safe merging. The front merging ratio was 4/6. The average merging time was 4.26s. The baseline planner was able to complete the merging maneuver in $5/6$ of the cases and the ratio merging in front was $3/6$. The average merging time was 7.18s.

Longitudinal distance and lateral postion prediction errors are defined by $\Delta_{e_x,k} \triangleq \frac{1}{N_p}\sum_{i=1}^{N_p}[\sqrt{(\hat{e}_{xk,i})^2 - (e^{tr}_{x,k+i})^2}]$ and $\Delta_{y,k} \triangleq \frac{1}{N_p}\sum_{i=1}^{N_p}[\sqrt{(\hat{y}_{k,i})^2 - (y^{tr}_{k+i})^2}]$, where $\hat{e}_{xk}, \hat{y}_k$ are the prediction at time step k of the states e_x, y for the future timestep $k+i$. Then, they are compared in Fig. 5 using human-in-the-loop control tests with six subjects; the proposed framework demonstrates reliable performance for all scenarios. In Fig. 6 one example is depicted where ADS controls the vehicle safely, maintains lateral stability despite the deviation between assumed single track model and *CarSim*. Fig. 8 shows the effectiveness of the method in a merging scenario in which the human driver slows down slightly for ADS to merge. The developed interaction-aware framework anticipates the other driver's reaction for safe merging (e.g. 7). The baseline planner, however, waits

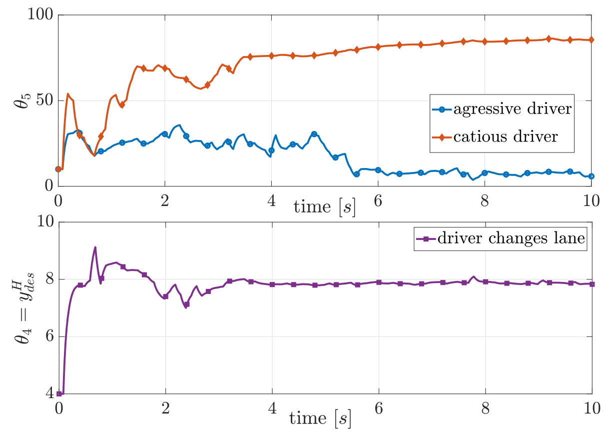

Fig. 7: Comparison of the weight of collision avoidance term estimation in J^H between an aggressive and a cautious driver; and detection of the desired lane change of the human

for the human to brake stronger because it is not updated on the reaction of the human to ADS. As a result, as the baseline controller does not react, the human-driven vehicle accelerates. This is a significant challenge (i.e., frozen robot) in dense traffic in which the succeeding vehicles act similarly.

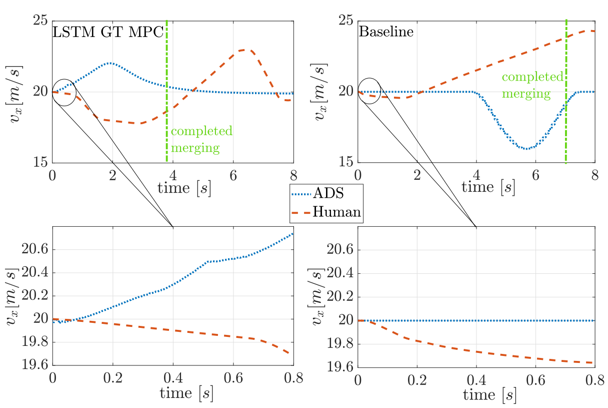

Fig. 8: Effectiveness of the interaction-aware framework compared to a baseline planner during an arduous merging

VII. CONCLUSIONS

An interaction-aware motion planning and control framework, which is able to interact with a human-driven vehicle, was designed for automated driving systems. The performance of the method was evaluated in safety-critical highway merging scenarios under different driving conditions. The merging scenario was first framed as Nash differential games using the vehicle single-track model. Then, a game-theoretic predictive controller was devised and tested. The predictive controller uses human objectives estimated in real-time using a novel inverse differential game. The experimental evaluation in *CarSim* with a human-in-the-loop setup demonstrated that the game-theoretic modeling of such scenarios and the interaction-aware control framework is able to handle different driving styles of the human-driven (i.e., surrounding) vehicles with good computational efficiency.

REFERENCES

[1] P. Trautman and A. Krause, "Unfreezing the robot: Navigation in dense, interacting crowds," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2010, pp. 797–803.

[2] S. Dixit, S. Fallah, U. Montanaro, M. Dianati, A. Stevens, F. Mccullough, and A. Mouzakitis, "Trajectory planning and tracking for autonomous overtaking: State-of-the-art and future prospects," *Annual Reviews in Control*, vol. 45, pp. 76–86, 2018.

[3] Viana, H. Kanchwala, K. Ahiska, and N. Aouf, "A comparison of trajectory planning and control frameworks for cooperative autonomous driving," *Journal of Dynamic Systems Measurement and Control*, vol. 143, 01 2021.

[4] M. Hruszczak, B. Lowe, F. Schrodel, M. Freese, and N. Bajcinca, "Game theoretical decision making approach for a cooperative lane change," *IFAC-PapersOnLine*, vol. 53, pp. 15 247–15 252, 01 2020.

[5] D. Fridovich-Keil, E. Ratner, L. Peters, A. Dragan, and C. Tomlin, "Efficient iterative linear-quadratic approximations for nonlinear multiplayer general-sum differential games," 05 2020, pp. 1475–1481.

[6] S. L. Cleac'h, M. Schwager, and Z. Manchester, "ALGAMES: A fast solver for constrained dynamic games," in *Robotics: Science and Systems XVI*. Robotics: Science and Systems Foundation, 2020.

[7] W. Schwarting, A. Pierson, J. Alonso-Mora, S. Karaman, and D. Rus, "Social behavior for autonomous vehicles," *Proceedings of the National Academy of Sciences*, vol. 116, no. 50, pp. 24 972–24 978, 2019.

[8] 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, pp. 1–9.

[9] T. Molloy, "Inverse noncooperative dynamic games," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 11 788–11 793, 2017.

[10] N. Mehr, M. Wang, and M. Schwager, "Maximum-entropy multi-agent dynamic games: Forward and inverse solutions," 2021.

[11] J. Inga Charaja, E. Bischoff, F. Köpf, M. Flad, and S. Hohmann, "Inverse cooperative and non-cooperative dynamic games based on maximum entropy inverse reinforcement learning," *ArXiv*, vol. abs/1911.07503, 2019.

[12] T. L. Molloy, J. Inga, M. Flad, J. J. Ford, T. Perez, and S. Hohmann, "Inverse open-loop noncooperative differential games and inverse optimal control," *IEEE Transactions on Automatic Control*, vol. 65, no. 2, pp. 897–904, 2020.

[13] P. Geiger and C.-N. Straehle, "Learning game-theoretic models of multiagent trajectories using implicit layers," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35.

[14] S. Le Cleac'h, M. Schwager, and Z. Manchester, "Lucidgames: Online unscented inverse dynamic games for adaptive trajectory prediction and planning," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5485–5492, 2021.

[15] C. C. MacAdam, "Understanding and modeling the human driver," *Vehicle System Dynamics*, vol. 40, pp. 101 – 134, 2003.

[16] M. Darouach, M. Zasadzinski, and S. J. Xu, "Full-order observers for linear systems with unknown inputs," *IEEE Transactions on Automatic Control*, vol. 39, no. 3, pp. 606–609, 1994.

[17] E. Hashemi, X. He, and K. H. Johansson, "A dynamical game approach for integrated stabilization and path tracking for autonomous vehicles," in *2020 American Control Conference*. IEEE, 2020, pp. 4108–4113.

[18] J. Inga, *Inverse Dynamic Game Methods for Identification of Cooperative System Behavior*. Karlsruher Institut für Technologie, 2020.

[19] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "CasADi – A software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1–36, 2019.

[20] Y. Ding, W. Zhuang, L. Wang, J. Liu, L. Guvenc, and Z. Li, "Safe and optimal lane-change path planning for automated driving," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 235, no. 4, pp. 1070–1083, 2021.