Introducing Fairness in Lane-Free Traffic

The Application of Karma Games to Enforce Fair Collaboration of CAVs

Author(s):
Chavoshi, Kimia; Ferrara, Antonella; Kouvelas, Anastasios

Publication date:
2024-09-18

Permanent link:
https://doi.org/10.3929/ethz-b-000695771

Originally published in:
TechRxiv, https://doi.org/10.36227/techrxiv.172668503.39089769/v1

Kimia Chavoshi[1], Antonella Ferrara[1], and Anastasios Kouvelas[1]

[1]Affiliation not available

September 18, 2024

Introducing Fairness in Lane-Free Traffic: The Application of Karma Games to Enforce Fair Collaboration of CAVs

Kimia Chavoshi, Antonella Ferrara, and Anastasios Kouvelas

***Abstract*—The present paper proposes a collaborative control strategy for CAV movement in a lane-free environment, where CAVs have different priority values. The priority values are designed to introduce fairness, relying on the interaction history of every CAV. These priority values are the output of the Stationary Nash Equilibrium (SNE) of a modified Karma game. The modification is proposed to tackle a proportional resource allocation problem. The outcome of the modified Karma game is then used to determine the priority of CAVs in their interactions with the other vehicles in their neighborhood, which may constitute potential threats, so-called Threatening Vehicle Cluster (TVC). An MPC is designed at an upper level to enforce an overall fair collaboration among CAVs so as to avoid collisions. Finally, a simulation case study analyzes the proposed control approach regarding performance, efficiency, and fairness.**

***Index Terms*—Connected and Automated Vehicles, Karma games, Lane-free traffic, Interaction-driven traffic.**

I. Introduction

COLLABORATION is an inevitable feature of microscopic traffic control in the era of Connected and Automated Vehicles (CAVs). CAVs are required to collaborate for different purposes, such as reaching a consensus formation (e.g., platoons or swarms), motion planning [1], or avoiding collisions [2]. Although numerous control methods are already presented in the literature for the optimal collaboration of CAVs (such as [3]–[8]), few of them have considered fairness in the collaboration.

The concept of fairness in CAVs collaboration can be illustrated through a simple example. Suppose that two CAVs on a two-lane road are moving at the same speed side by side, each on a different lane. Assume that another CAV with a higher speed approaches them. There are many feasible solutions to avoid collision, such as, for instance, lowering the speed of the third vehicle; manipulating the speed of two front vehicles to create a gap between the two vehicles to allow the third vehicle to overtake; changing the lane of one of the front vehicles to free a lane for the third vehicle to pass. Each of these solutions induces different costs for each CAV. Then, an interesting research question is: Which is a fair distribution of costs?

Kimia Chavoshi and Anastasios Kouvelas are with the Institute for Transport Planning and Systems, Department of Civil, Environmental and Geomatic Engineering, ETH Zurich, CH-8093, Switzerland (e-mail: kimiac@ethz.ch; kouvelas@ethz.ch).
Antonella Ferrara is with the Department of Electrical, Computer and Biomedical Engineering, University of Pavia, 27100 Pavia, Italy (e-mail: antonella.ferrara@unipv.it).

From a broader perspective, each CAV engages in many interactions with different groups of vehicles while moving in traffic. Each interaction induces a cost for the CAV. Thus, it is essential to consider the history of CAVs' interaction to address fairness. In this paper, we use the concept of karma to track the history of interaction between CAVs. In a Karma game, players trade karma to win priority [9] and [10]. Similarly, one can assume that each CAV has a level of karma and urgency. Every time CAVs get involved in an interaction with other vehicles, they play the Karma game to well structure the collaboration. The CAV offers a bid according to its karma and urgency values to win a level of priority (autonomy). The offered bid cannot be higher than the CAV's karma value. According to the outcome of the game, CAV wins a level of priority and earns or pays karma to other players. These priority levels are communicated to an upper-level collaborative control that plans CAV movements.

In the present paper, the illustrated control structure is implemented in a lane-free vehicular traffic framework [11] and [12]. The research aims to design a fair collaborative control algorithm for the safe movement of CAVs in such a framework.

Firstly, we propose a threat detection algorithm to enable CAVs to determine, within their communication area, the vehicles with which they have a risk of collision (namely, the threatening vehicles) and thus require interaction. Many algorithms have been proposed in the literature for obstacle detection (see, for instance, [13] and [14]). Our proposed algorithm has similarities with the algorithms used in reciprocal collision avoidance (such as [14], [15], and [16]). Afterwards, we build the Threatening Vehicle Clusters (TVCs). Implementing the concept of TVCs simplifies the optimization problem solved at the upper level by removing unnecessary relationships among vehicles in the communication range but with no risk of collision. When a TVC is formed, the members play a Karma game at the lower level of control to win priority values. An MPC for controlling CAVs within the TVCs is designed at the upper level. MPC has been widely used for the safe motion planning of autonomous vehicles such as [17] and [18]. The idea of designing a control method for the TVCs is based on the ability and willingness of CAVs to cooperate and take joint actions to avoid collision within each cluster. The priority values of the Karma game are utilized as weights for the MPC objective function. More specifically, a higher priority is associated with a higher weight for the term in the objective function, which accounts for the deviation from the

CAV desired speed. The contributions of this work can be listed as follows:

- Modifying Karma game for proportional resource allocation strategies.
- Designing a threat detection algorithm for CAVs.
- Developing an original control method for fair collaboration of CAVs in the lane-free environment oriented to avoid collisions and promote fairness.

The paper is organized as follows. The next section is devoted to the literature review, briefly reviewing the relevance of lane-free traffic and Karma games. In Section III, the upper-level control strategy is introduced. Therein, the system dynamics are depicted, and the possible threatening relationships between CAVs, with the corresponding concept of TVCs, are introduced. Then, the proposed control approach for fair collaboration of CAVs with the objective of collision avoidance in a lane-free environment is discussed. The modified version of Karma games for a generic proportional resource allocation problem is presented in Section IV, along with a detailed description of the lower-level control strategy. In Section V, the introduced concepts and designed control approach are evaluated by multiple simulations. Finally, the research is concluded, and future developments are discussed in the Conclusions section.

II. LITERATURE REVIEW

Traffic management is an important application field of the control systems (see, for instance, [19], [20], [21], and [22]). The basic traffic principles are designed according to human-driven vehicles. However, the advent of CAVs creates the opportunity to re-think those principles and design an alternative traffic paradigm (see [23]). For instance, conventional lane-based traffic was introduced to ensure safety, considering the lack of communication and cooperation between human-driven vehicles while compromising the higher capacity. Recent research [24] has shown that introducing a lane-free environment for CAV traffic will increase road capacity and smoothen CAV movements. The concept is proposed by [11] and has been quickly extended in various directions. An optimal trajectory planning based on the nudging technique is designed by [25] for CAVs in a lane-free environment. In [26], a cooperative trajectory planning for lane-free autonomous vehicle traffic is proposed and tested by conducting laboratory experiments. To reduce chaos in a lane-free environment, [27] proposed the potential lines, a lateral separation between vehicles with significant differences in speed value. The concept of cruise control for CAVs in lane-free environments is introduced by [28] and [29]. The application of deep reinforcement learning algorithms for autonomous deriving in lane-free traffic focusing on reward function design is studied by [30]. Combining the max-plus algorithm with potential fields, [31] designed a collaborative multiagent decision-making algorithm for lane-free autonomous driving. The internal boundary control problem for lane-free highways is addressed by [32] and [33]. The bidirectional lane-free movement (laneless and directionless) for CAVs is conceptualized in [12] and [34], which additionally removes the conventional direction separation for CAV traffic.

Despite all the research and developments on performance improvement, fairness has not been addressed in lane-free literature. The different concepts of fairness have been used in the context of traffic control. For instance, fair travel time [35] and fair merging time [36] in ramp metering control; fair waiting time [37] or fair travel time [38] in traffic light control; fair service quality [39] for congestion pricing (taxation); and fair travel time [40] for route guidance. Fairness has also been addressed in the CAV collaboration. In [41], a reinforcement learning for CAV movement at the bottleneck has been presented where the reward function has a term related to fair travel time at the bottleneck. In [42], different courtesy strategies are designed to study the gap creation required in lane changing; the fairness in aggregated speed for CAVs is considered as an evaluation measure. A reinforcement learning for cooperative driving of CAVs in lane-based traffic is designed in [43], where a constraint on the fair contribution of CAVs within pairs can foster collaboration among them.

In this paper, we address the fair distribution of priorities in CAV collaboration by implementing a modified version of the Karma game. Game-based approaches have been used extensively for controlling multiagent systems [44]. Different games were developed to address the decision-making process of CAVs [45], [46], [47], and [48]. Karma game belongs to a new category of games, dynamic population games, that is introduced by [9]. Karma is a repeated auction-based game. In each iteration, two players will be randomly selected from the population to compete over a resource. Every player has a level of karma that limits its bid. Moreover, the player has a time-varying urgency level that indicates the importance of acquiring the resource. The player with the higher bid wins the resource and pays karma to the loser or society. The existence of the Stationary Nash equilibrium (SNE) for the Karma game is studied in [10]. In general, dynamic population games have various applications in different disciplines. For instance, [49] employed a dynamic population game to model individual's decision-making behavior under epidemic situations. An early application of the Karma game for intersection management is proposed by [50]. A Karma-based traffic management method for highways is developed by [51], where vehicles play Karma game to access an uncongested priority lane.

III. UPPER-LEVEL CONTROL: MPC FOR CAV COLLABORATION

Suppose a traffic flow of fully connected and automated vehicles (CAVs) in a lane-free environment, where every CAV has a desired speed. The CAVs are connected within a communication radius and communicate information such as location, speed, heading angle, etc. The goal is to design a control method for CAVs in this environment with the two-fold objective of safety and fairness. We design a threat detection algorithm for CAVs. According to the threat relationships, we categorize the CAVs into Threatening Vehicle Clusters (TVCs). The CAVs with no threat will be managed separately. In the next step, a bi-level control strategy is designed. At the upper level, we develop an MPC-based control approach for TVCs with the objective of collision avoidance. A similar

MPC approach is also designed for the CAVs with no threats. Assuming the objective of every CAV is to drive with its desired speed, the priority values are implemented as weights for the deviation of speed from its desired value in the MPC cost function. Thus, CAVs with higher priority values can more strongly impose their private objective on the other members of the TVC. To ensure fairness in the collaborations, we design a game at the lower level of control. For every newly formed TVC, the CAVs play (based on the SNE policy) a modified Karma game to win the priority values. The modified Karma game is designed to fairly assign priority values to the CAVs, considering their collaboration history.

Figure 1 shows the block diagram schematic of the methodology for CAV traffic, where $X^T(k)$ and $U^T(k)$ are the state variables and control signals of all CAVs at time step k. Suppose that M is the total number of possible TVCs in a traffic flow of N CAVs. TVCs will be created at every time step according to the threat detection algorithm. Thus, the MPC modules related to the detected TVCs will be connected (switched on) in the block diagram. The MPC modules for the remaining CAVs (which do not belong to any TVC) will also be switched on. An MPC module for CAV i receives state variables of the CAV, $X_i(k)$, and their desired values, X^d, as inputs and calculates the future control signals $U_i(k{+}1{:}k{+}p)$ ($p{=}m$, the MPC prediction horizon equals control horizon) for the CAV i. Likewise, the MPC module for TVC i receives the related state variables $X_i^C(k)$ and desired values $X^{C,d}$ of the CAVs inside TVC i as inputs; and calculates the future control signals $U_i^C(k{+}1{:}k{+}p)$ for these CAVs. The Karma SNE determines the priority parameters β^{Karma} for the newly formed TVCs. Hence, it is an event-triggered part of the TVC control module. Finally, we update the CAVs state variables by implementing the next step control signals $U^T(k{+}1)$.

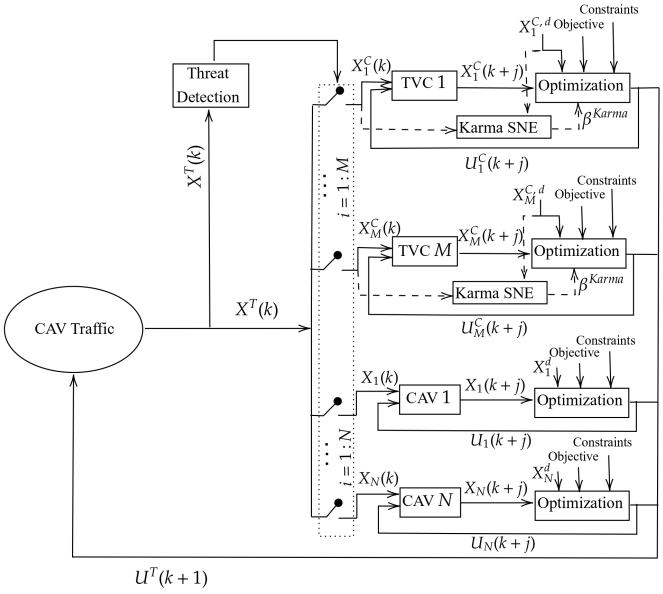


Fig. 1: The block diagram schematic of the methodology.

A. Kinematic Bicycle Model

The CAV movement is modeled with a discrete-time kinematic bicycle model [52].

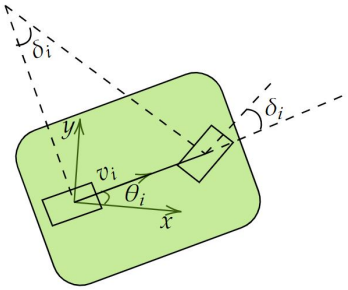


Fig. 2: The diagram of the CAV movement.

$$\begin{cases} x_i(k+1) = x_i(k) + Tv_i(k)\cos(\theta_i(k)) \\ y_i(k+1) = y_i(k) + Tv_i(k)\sin(\theta_i(k)) \\ v_i(k+1) = v_i(k) + Ta_i(k) \\ \theta_i(k+1) = \theta_i(k) + T\frac{v_i(k)\tan(\delta_i(k))}{L} \\ \delta_i(k+1) = \delta_i(k) + Tw_i(k) \end{cases}, \tag{1}$$

where T (s) is the time step for the discretized model. The pair of $(x_i(k), y_i(k))$ represents the location of the CAV centroid in cartesian coordinate for CAV i with length L (m) at time k. Variables $v_i(k)$ (m/s), $\theta_i(k)$ (Rad), and $\delta_i(k)$ (Rad) denote speed, speed angle, and heading angle, respectively. Figure 2 demonstrates these variable states for CAV i. In this model, acceleration, $a_i(k)$ (m/s^2), and steering rate, $w_i(k)$ (Rad/s) are the control signals for CAV i.

B. Threat Detection Algorithm

We assume CAVs within a communication range from each other, so-called Neighbors, communicate their information, such as current variable states and their desired speed and direction. For CAV i, the neighbor set $\mathcal{N}_i$ is defined as:

$$\mathcal{N}_i = \{j | D_{i,j} \leq R\}, \tag{2}$$

$$D_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}\ , \tag{3}$$

where $D_{i,j}$ denotes the distance between CAVs i and j; and R is the communication radius. Some neighbors show potential for conflict and, if not managed correctly, will lead to a collision, a so-called Threat. For CAV i, the threat set $\mathcal{T}_i$ which is a subset of $\mathcal{N}_i$ is defined as follows:

$$\mathcal{T}_i = \left\{ j | j \in \mathcal{N}_i,\ \sin(\eta_{j,i}) \leq \frac{r_{\text{m}}}{D_{j,i}},\ \frac{D_{j,i}}{\text{sgn}(\cos(\eta_{j,i}))S_{j,i}} \leq T_{\text{c}} \right\}, \tag{4}$$

where,

$$\vec{S}_{j,i} = \vec{v}_j - \vec{v}_i. \tag{5}$$

Thus, $S_{j,i}$ is the magnitude of the relative speed between CAVs j and i. The variable $\eta_{j,i}$ denotes the angle between

relative speed vector $\vec{S}_{j,i}$ and distance $D_{j,i}$ as it is shown in Figure 3. The parameter r_{m} represents the safety margin radius. Figure 3 shows an example of two threatening vehicles. Suppose we draw tangent lines from the centroid of vehicle j to the safety margin (the yellow circle around vehicle i). In that case, the area covered between two tangent lines is called the collision spectrum. Vehicle j will enter the safety margin of vehicle i if the relative speed vector is inside the collision spectrum. The parameter T_{c} is look-ahead time and represents the finite time horizon used in the threat detection algorithm. Only the vehicles that may collide within T_{c} time horizon are considered threats. Two conditions are introduced in equation (4) for CAV i to detect threatening vehicles among neighbors. A neighbor j is a threat if the relative speed is inside the collision spectrum and the estimated time to collision is smaller than T_{c}.

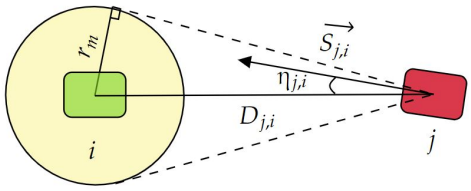


Fig. 3: Threat detection. The yellow circle represents the safety margin for the green CAV. The collision spectrum is determined with dashed lines.

While every CAV can have multiple threats, each of these threats can be involved in the other threat relationship. Considering this chain of threats, the CAVs are grouped as TVCs (Threating Vehicle Clusters). Assume we draw a threat graph with each CAV as a node. The nodes i and j are connected with an edge only if $j \in \mathcal{T}_i$. Two CAVs belong to the same TVC if and only if a path exists in the threat graph connecting the corresponding nodes.

C. Control Approach

We design a receding horizon nonlinear MPC controller for every TVC $\mathcal{G}_l$. At every time step t, a nonlinear optimization problem for the prediction horizon of p is solved and defines the control signals for the control horizon m. The computed control signals for the first time horizon are applied to the system, and variable states are updated for the next step. We assume the same prediction horizon and control horizon for the current problem. The objective function is presented in equation (6), where the decision variables are the acceleration a and steering rate w of CAVs within $\mathcal{G}_l$. The optimization problem is designed as follows:

$$\min_{a_i,w_i} \Bigg(\sum_{k=t+1}^{t+p} \Big(\sum_{i\in\mathcal{G}_l} (\alpha_i(\theta_i(k)-\theta_i^{\mathrm{d}})^2 + \beta_i^{\mathrm{Karma}}(v_i(k)-v_i^{\mathrm{d}})^2 + \gamma_i a_i(k)^2) + \sum_{j=1}^{N_T} \frac{\zeta_j}{\tau_j^c(t)}\epsilon_j(k)\Big)\Bigg) \tag{6}$$

subject to state space model (1) and the following physical constraints:

$$0 \le v_i(k) \le v_{\max} \tag{7}$$

$$y_{\min} + \Delta_{\mathrm{y}} \le y_i(k) \le y_{\max} - \Delta_{\mathrm{y}} \tag{8}$$

$$a_{\min} \le a_i(k) \le a_{\max} \tag{9}$$

$$|a_i(k) - a_i(k-1)| \le \Delta_{\mathrm{a}} \tag{10}$$

$$|\theta_i(k) - \theta_i^{\mathrm{d}}| \le \Delta_\theta \tag{11}$$

$$|\delta_i(k)| \le \Delta_\delta \tag{12}$$

$$|w_i(k)| \le \Delta_{\mathrm{w}} \tag{13}$$

where i denotes the index of the CAVs that belong to $\mathcal{G}_l$. The prediction horizon is denoted by p and is set to 40 time steps. Conducting simulations with lower values for the prediction horizon resulted in poor performance. Thus, the best setup obtained for p equals 40. The road boundaries are denoted by $y_{\min}$ and $y_{\max}$. The objective function presented at (6) consists of four terms, where α_i, β_i^{Karma}, γ_i, and ζ_i are the weights for each term. The first and second terms minimize the deviation of each CAV from its desired speed angle and speed. The third term minimizes acceleration cost.

The last term is designed for collision avoidance. Every threat group $\mathcal{G}_l$ contains N_T number of threat pairs. The following constraints are added to the optimization problem to prevent collision between threat pairs. Assume threat pair number j consists of CAVs e and f. These CAVs collide if their distance gets smaller than the safety margin (see Fig. 3). Thus, the collision avoidance constraint is defined as

$$D_{e,f}(k) \ge r_{\mathrm{m}}, \quad \forall j \in [1,\ N_T] \tag{14}$$

According to equation (4), the relative speed is inside the collision spectrum for the threat CAVs e and f. Considering that, we design an extra soft constraint for collision avoidance.

$$\sin(\eta_{e,f}) - \frac{r_{\mathrm{m}}}{D_{e,f}(k)} \ge -\epsilon_j(k), \tag{15}$$

$$\epsilon_j \ge 0, \quad \forall j \in [1,\ N_T] \tag{16}$$

where ϵ_j is minimized by the cost function. Constraint (15) can be rewritten as

$$D_{e,f}(k) \ge \frac{r_{\mathrm{m}}}{\sin(\eta_{e,f}) + \epsilon_j(k)},$$

where in comparison to (14), put a higher minimum bound for the distance. It forces CAVs to start action earlier, thus smoothening the CAVs' maneuver for collision avoidance. to prioritize the more critical threat pairs, we weigth ϵ_j in the cost function according to predicted time to collision τ_j^c.

$$\tau_j^c(t) = \frac{D_{e,f}(t)}{S_{e,f}(t)\cos(\eta_{e,f}(t))} \tag{17}$$

The weight selection involves normalizing the terms in the objective function. The weights are tuned based on trial and error to prevent unnecessary lateral movements and smoothen the CAV reactions. A higher value is assigned to ζ_i, related to the safety term. The outcome of the Karma game played in the lower level control, β_i^{Karma}, the so-called priority

parameter is implemented as the weight for speed deviation. The higher value of β_i^{Karma} increases the cost of deviation from the desired speed for CAV i. This leads to solutions that put lower costs on CAVs with higher priority parameters. Thus prioritizing the CAVs who achieved higher outcomes in the Karma game.

Finally, to complete the control approach, it is crucial to address the CAVs that do not pose any threat. For every CAV i that does not belong to any TVC at time t, we design the MPC problem formulation as follows:

$$\min_{a_i,w_i} \sum_{k=t+1}^{t+p} \alpha_i(\theta_i(k)-\theta_i^{\text{d}})^2+\beta_i(v_i(k)-v_i^{\text{d}})^2+\gamma_i a_i(k)^2 \quad (18)$$

subject to state space model (1) and physical constraints (7)-(13).

The developed optimization problems are solved with a primal-dual interior point algorithm presented in [53].

IV. Lower-level Control: Modified Karma game

To incorporate fairness in the solution, we introduced priority parameters to be utilized in the objective function for TVC at the upper-level control. In this section, we modify the original Karma game (introduced in [10]) to compute the priority parameters for every newly formed TVC. We recommend interested readers study [10] to get familiar with the Karma mechanism and the algorithm for obtaining SNE (Stationary Nash Equilibrium). A summary of variables and functions used in this section is provided in Table I.

We generalize the problem of assigning priority values to CAVs within a TVC. Suppose a population of N agents (CAVs). In every instance, a sub-population (TVC) of $2 \leq N_s \leq N$ agents are matched together in an Auction game for a proportional resource allocation. Suppose the probability of a sub-population with the size N_s appearing, $P_s(N_s)$, is known. Each agent i has a karma of $k_i \in \mathbb{W}$ that limits the bid ($b_i \in \mathbb{W}$) it can make $0 \leq b_i \leq k_i$. Besides, every agent has an urgency of $u_i \in \mathbb{N}$, which shows the internal valuation of the agent for winning a higher share of the resource. According to the offered bids, player i wins a portion c_i (priority value) of a resource C.

$$C = \sum_{i=1}^{N_s} c_i \quad (19)$$

$$c_i = \frac{b_i C}{\sum_{i=1}^{N_s} b_i} \quad (20)$$

Afterward, the karma for agents within the sub-population is updated according to a payment policy [51]. Please note that the karma value belongs to $\mathbb{W}$; thus, the payment policy has two cases as follows.

$$\begin{cases} k_i^+ = k_i - b_i + \lceil \frac{\sum_{i=1}^{N_s} b_i}{N_s} \rceil & w.p \quad f \\ k_i^+ = k_i - b_i + \lfloor \frac{\sum_{i=1}^{N_s} b_i}{N_s} \rfloor & w.p \quad 1-f \end{cases} . \quad (21)$$

where $f = \frac{b_s}{N_s} - \lfloor \frac{b_s}{N_s} \rfloor$ shows the fraction of agents in the sub-population that gets the first case payment.

The objective is to determine the SNE for the game, as mentioned above. Suppose $d[u,k]$ denotes the population probability distribution. The bidding policy for each agent depends on its karma, urgency, and sub-population and is denoted by $\pi(b|u,k,N_s)$. The distribution of player bids in a sub-population N_s is

$$V[b|N_s](d,\pi) = \sum_{u,k} d[u,k]\pi(b|u,k,N_s). \quad (22)$$

sub-population bid is defined as the total bid offered by all the players $b_s = \sum_{i=1}^{N_s} b_i$.

Assume that agents' bids are independent and identically distributed random variables with the PDF of $V[b|N_s]$. Thus, we can estimate the PDF of sub-population bid, $V_s[b_s|N_s](d,\pi)$, by taking N_s fold convolution as follows,

$$V_s[b_s|N_s](d,\pi) = \overbrace{V[b|N_s] * (\cdots * (V[b|N_s] * V[b|N_s]))}^{N_s-1}. \quad (23)$$

In the sub-population N_s, the probability of a player winning c through bidding b, $\gamma[c|b,N_s](d,\pi)$, is equivalent to the probability of $b_s = \frac{Cb}{c}$. Therefore,

$$\gamma[c|b,N_s](d,\pi) = V_s[b_s = \frac{Cb}{c}|N_s](d,\pi). \quad (24)$$

The karma transition function, the conditional probability of updated karma k^+ for a player concerning previous karma k, bid b, and sub-population bid b_s is:

$$K[k^+|k,b,b_s,N_s] = \begin{cases} f & if k^+ = k - b + \lceil \frac{b_s}{N_s} \rceil \\ 1-f & if k^+ = k - b + \lfloor \frac{b_s}{N_s} \rfloor \\ 0 & o.w \end{cases} . \quad (25)$$

where $f = \frac{b_s}{N_s} - \lfloor \frac{b_s}{N_s} \rfloor$. Also, note that we can replace $b_s = \frac{Cb}{c}$, thus derive $K[k^+|k,b,c,N_s]$. The immediate reward function for the agent is

$$\xi[u,b,N_s](d,\pi) = -u \sum_c c\gamma[c|b,N_s](d,\pi). \quad (26)$$

We augment the transition probability at [10] by adding the probability of the future sub-population $P_s(N_s^+)$ as follows

$$\rho[u^+,k^+,N_s^+|u,k,b,N_s](d,\pi) = P_s(N_s^+)\Phi[u^+,u] \sum_c \gamma[c|b,N_s]K[k^+|k,b,c,N_s], \quad (27)$$

where $\Phi[u^+,u]$ is the urgency transition probability derived from the Markov chain process for urgency.

Afterward, we can obtain the same algorithm as [10] to calculate the expected immediate reward $R[u,k|N_s](d,\pi)$, state transition function $P[u^+,k^+,N_s^+|u,k,N_s](d,\pi)$, expected infinite horizon reward $\nu[u,k,N_s](d,\pi)$, and single stage deviation reward $Q[u,k,b,N_s](d,\pi)$ for each sub-population as below.

$$R[u,k|N_s](d,\pi) = \sum_b \pi[b|u,k,N_s]\xi[u,b,N_s]. \quad (28)$$

$$P[u^+,k^+,N_s^+|u,k,N_s](d,\pi) = \sum_b \pi[b|u,k,N_s]\rho[u^+,k^+,N_s^+|u,k,b,N_s]. \quad (29)$$

Terminology	Description
N	Population size.
N_s	Sub-population size.
C	Resource.
k_i	Karma of agent i.
u_i	Urgency of agent i.
b_i	Bid of agent i.
c_i	Outcome (resource portion) for agent i.
$d[u,k]$	Population probability distribution; the fraction of population with urgency u and karma k.
$\pi(b\|u,k,N_s)$	Bidding policy; probability of bidding b for agent with urgency of u and karma of k in sub-population N_s.
$P_s(N_s)$	Probability distribution of sub-population size.
$V[b\|N_s]$	Probability distribution of bids for agent in sub-population N_s.
$V_s[b_s\|N_s]$	Probability distribution of total bid b_s for sub-population N_s.
$\gamma[c\|b,N_s]$	Probability distribution of (resource allocation) outcome c, if agent bids b in subpopulation N_s.
$K[k^+\|k,b,b_s,N_s]$	Karma transition function; probability of updated karma k^+, if agent with karma k bids b in sub-population N_s with total bid b_s.
$\Phi[u^+,u]$	Urgency transition probability; probability of future urgency u^+, if agent has urgency u.
$\rho[u^+,k^+,N_s^+\|u,k,b,N_s]$	Transition function; probability of agent's transition to future urgency u^+, updated karma k^+, and sub-population N_s^+, if agent with urgency of u and karma of k bids b in sub-population N_s.
$P[u^+,k^+,N_s^+\|u,k,N_s]$	State transition function; probability of agent's transition to future urgency u^+, updated karma k^+, and sub-population N_s^+, if agent has urgency of u and karma of k in sub-population N_s.
$\xi[u,b,N_s]$	Immediate reward function.
$R[u,k\|N_s]$	Expected immediate reward function.
$\nu[u,k,N_s]$	Expected infinite horizon reward function.
$Q[u,k,b,N_s]$	Single stage deviation reward.
$\tilde{\pi}[b\|u,k,N_s]$	Perturbed best response policy.

TABLE I: The Summary of variables and functions used in Section IV. Lower-level control.

$$\nu[u,k,N_s](d,\pi) = R[u,k|N_s]+ \\ \alpha\sum_{u^+}\sum_{k^+}\sum_{N_s^+} P[u^+,k^+,N_s^+|u,k,N_s]\nu[u^+,k^+,N_s^+]. \quad (30)$$

$$Q[u,k,b,N_s](d,\pi) = \xi[u,b,N_s]+ \\ \alpha\sum_{u^+}\sum_{k^+}\sum_{N_s^+} \rho[u^+,k^+,N_s^+|u,k,b,N_s]\nu[u^+,k^+,N_s^+]. \quad (31)$$

As explained in [10], the problem of searching SNE is written as finding the rest point for a dynamical system, which represents the evolutionary dynamics for population d and policy π in Karma games. In this system, the policy updates with a rate η based on a mean dynamics function of evolutionary games denoted as H.

$$\dot{\pi}[.|u,k,N_s] = \eta H(Q,\pi) \quad (32)$$

The perturbed best response policy $\tilde{\pi}[b|u,k,N_s]$ is chosen as the mean dynamics to enable the modeling of agents that are not perfectly rational.

$$\tilde{\pi}[b|u,k,N_s](d,\pi) = \frac{\exp(\lambda Q[u,k,b,N_s](d,\pi))}{\sum_b \exp(\lambda Q[u,k,b,N_s](d,\pi))} \quad (33)$$

The parameter λ indicates the agent's rationality.

Finally, the population d and policy π are updated based on the following discretized dynamics, where dt denotes the discretization step size.

$$\pi[b|u,k,N_s] = (1-\eta dt)\pi[b|u,k,N_s]+\eta dt\tilde{\pi}[b|u,k,N_s] \quad (34)$$

$$d[u,k] = (1-dt)d[u,k]+ \\ dt\sum_{N_s^-} P_s(N_s^-)\sum_{N_s}\sum_{u^-}\sum_{k^-} P[u,k,N_s|u^-,k^-,N_s^-]d[u^-,k^-] \quad (35)$$

We repeat these steps (22)-(35) until $d[u,k]$ and $\pi[b|u,k,N_s]$ converge to the SNE.

In every newly formed TVC, CAVs use the bidding policy at SNE to play the Karma game. Accordingly, every CAV wins a c_i value. The priority parameter associated with each CAV, β_i^{Karma}, is calculated as a linear function of the proportional win for CAV i, $\frac{c_i}{C}$.

$$\beta_i^{\text{Karma}} = A\frac{c_i}{C} + B, \quad (36)$$

where A and B are positive and nonzero numbers used for the tuning purpose of the objective function at the upper-level control (6). In this paper, $A = 0.1$ and $B = 5\times10^{-5}$ are used in the simulation case studies.

In the original Karma game, the urgency for agents is described as a private valuation of the agent to win. In this paper, we determine the urgency as $u = g\left((v_i(k) - v_i^{\text{d}})^2\right)$ where $g : \mathbb{R} \to \mathbb{N}$ is an increasing function (see Fig. 1). For every CAV, the higher deviation from the CAV's desired speed increases its urgency level in the Karma game. As a result, the agent is encouraged to offer a higher bid and win more proportion of the resource. Thus, it is expected by increasing the urgency level for agent i, the value of c_i, and correspondingly the priority parameter β_i^{Karma} used in the objective function (6) increase. The higher priority parameter associated with CAV i in upper-level control will decrease $|v_i(k) - v_i^{\text{d}}|$. Therefore, the urgency for CAV i is supposed to decrease for the next Karma game. This procedure is analogous to the negative feedback concept in control theory.

We finalize this section with a discussion about the uniqueness and existence of SNE and the convergence to the SNE. There is no guarantee for the uniqueness of SNE for the original Karma game. However, we can guarantee the existence of SNE for the presented algorithm.

Proposition 1: An SNE exists for the modified Karma algorithm presented in Section IV.

Proof. Please see the Appendix A.
Due to the complexity of the overall system, we cannot provide a mathematical proof for the convergence of our proposed SNE algorithm.

Algorithm 1 provides a summary of the proposed methodology that was also illustrated in Fig. 1. At every time step, we determine the TVCs by applying the threat detection algorithm. For any newly formed TVC, we utilize the SNE of the designed Karma game in lower-level control to determine the priority parameters. Afterward, we control TVC and remaining CAVs according to the designed MPC methods in upper-level control. Finally, we update the variable states of all CAVs by implementing the control signal for the next time step.

Algorithm 1 Overview to methodology

Input: CAVs variable states, control signals, karma values, urgency level, and model parameters.
for $k = 1 : t_{\mathrm{f}}$ **do**
  # Determine Threat and TVC
  Run threat detection algorithm in III.B
  Output: Clusters $\mathcal{G}_l$, where $l \in [1,\ N_{\mathrm{G}}]$.
  # MPC for TVCs
  for $l = 1 : N_{\mathrm{G}}$ **do**
    # Determine priority parameters
    if $\mathcal{G}_l$ is newly formed utilize the SNE of the modified Karma game in IV.
    Outputs: priority parameters for CAVs in $\mathcal{G}_l$
    Run designed MPC at (6)-(17) and (1)
    Output: control signals at time k for CAVs in $\mathcal{G}_l$
  end for
  # MPC for CAVs with no threat
  for $i = 1 : N_{\mathrm{NT}}$ **do**
    Run designed MPC at (18), (1), and (7)-(13)
    Output: control signals at time k for CAV i
  end for
  # Update variable states
  for $i = 1 : N$ **do**
    insert CAV i control signals to car model (1)
    Output: variable states of CAV i at time $k+1$
  end for
end for=0

V. Simulation results

In this section, we evaluate the proposed methodology's performance by studying the results of the simulation tests. The values of methodology parameters implemented in all tests are summarized in Table II. The simulations are conducted in the Matlab framework. We used the IPOPT solver in the Yalmip toolbox [54] for the MPC problems. The maximum CPU time and maximum number of iterations for the IPOPT solver are set to 1000 s and 1500, respectively.

This section is organized as follows: First, the case study is described, and then the SNE of the modified Karma game for this case study is presented in subsection A. Afterward, in Subsection B, a snapshot of the CAV's traffic simulation presents a detailed picture of CAV's behavior, clusters, and games dynamics. The general overview of results is presented in Subsection C. Subsections D and E focus on safety and computational time. It follows with a detailed analysis of TVC dimension and duration in Subsection F. Finally, the modified Karma game is evaluated based on efficiency and fairness measures in Subsection G.

Parameter	Value	Parameter	Value
T	0.05 s	R	160 m
T_{c}	4 s	L	2 m
r_m	3 m	p	40
α_i	0.25	β_i	0.001
γ_i	1000	ζ_i	0.0001
Δ_{y}	1.5 m	$v_{\max}$	120 km/h
$a_{\min}$	-10.92 m/s^2	$a_{\max}$	5.72 m/s^2
Δ_{a}	0.7 m/s^2	Δ_θ	$\pi/3$ Rad
Δ_δ	$\pi/6$ Rad	Δ_{w}	$2\pi/3$ Rad/s
λ	1000	α	0.9

TABLE II: The value for parameters of the control method implemented in simulation tests.

A. Case Study Description

We simulate group overpass maneuvers as the case study. The case study describes ten vehicles with an initial speed of 80 km/h on a road with a width of 12.5 m. The five vehicles in front have the desired speed of 80 km/h, while the vehicles in the back have a higher desired speed of 100 km/h. We simulate two scenarios, so-called Scenario 1 for uncongested cases and Scenario 2 for congested cases. In the uncongested scenario, the vehicles have an initial time gap of 1 s, while for the congested scenario, the initial time gap is 0.5 s. Each scenario has been tested ten times, and every test has a different initial positioning for vehicles.

The developed methodology, as summarized in Algorithm 1, is implemented to control the CAV movements in these scenarios. Every time a new TVC initiates, we define the priority values with the help of the modified Karma game SNE. The CAVs are ranked according to their deviation from the desired speed. The 50% of vehicles with higher deviation are categorized as urgent, whereas the rest are unurgent. This case study considers two urgency levels of [1 10]. The urgency transition function $\Phi[u^+, u]$ is:

$$\begin{bmatrix} 0.5, 0.5 \\ 0.5, 0.5 \end{bmatrix}$$

where it shows the CAV's uncertainty about its future urgency. For the modified Karma game, the probability of a sub-population N_s appearing, $P_s(N_s)$, or in other words, the probability of a cluster with the dimension of N_s initiates, is shown in Fig. 4. We assume that the probability decreases gradually for bigger sub-populations.

The SNE of the modified Karma game for this case study is presented in Fig. 5 and Fig. 6. As seen in Table II, the future discount factor, α, is set to 0.9 for this case study. Due to this high future awareness, a CAV with low urgency bids 0 unless it has a high karma value. For the larger sub-populations (TVC dimension), the unurgent CAV begins bidding nonzero with a lower karma value. The general trend shows that an urgent

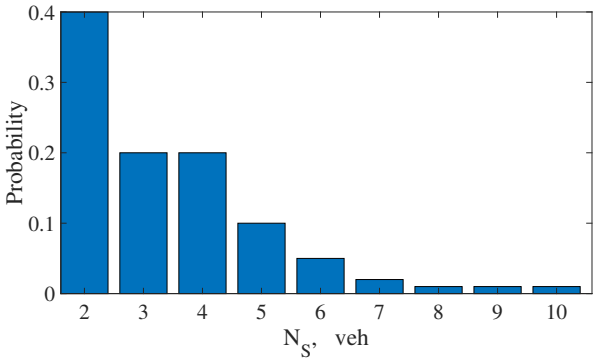


Fig. 4: The probability of a sub-population N_s, $P_s(N_s)$. This probability function is implemented in the modified Karma game for simulation tests.

CAV tends to bid higher values if it belongs to a larger sub-population. As shown in Fig. 6, the probability distribution for both urgent and unurgent parts of the population are the same. This result was expected due to the defined $[u^+, u]$ in this case study. The shares of urgent and unurgent CAVs are always 50%, and for a single CAV, the probability of becoming unurgent or urgent in the future is the same.

The SNE of the modified Karma game for lower future discount factors ($\alpha = 0.5$ and $\alpha = 0.7$) are presented in Appendix B. As one can observe in Fig. 5, Fig. 14a, and Fig. 14b, increasing the future discount factor makes the unurgent CAVs more reluctant to bid nonzero. In particular when the CAV belongs to a smaller sub-population. For $\alpha = 0.5$, this phenomenon is only observed for $N_s = 2$, while for $\alpha = 0.7$, this effect is extended to $N_s \leq 5$ and for $\alpha = 0.9$ it is observed for all the sub-populations. Moreover, CAVs tend to bid smaller values for larger α. Finally, comparing Fig. 6, Fig. 13a, and Fig. 13b shows that increasing α reduces the variance of the population probability distribution at SNE.

B. Snapshot

Figure 7 reports a snapshot of the uncongested scenario, Test 2, at time step 332. The CAVs are numbered according to their initial time setup. CAVs [1-5] belong to the group with the desired speed of 80 km/h, and CAVs [6-10] have the desired 100 km/h speed. The figure indicates three threat clusters with the dimensions of 5,2, and 3, respectively. The time spans for Cluster 1 and Cluster 2 are [331-345] and [287-398], respectively. Cluster 3 appears at time step 332; thus, the CAVs inside Cluster 3 play the Karma game to win the priority values. In this game, CAV 5, 9, and 10 have initial karma of 7, 10, and 10, respectively. CAV 9 is urgent; hence, it is encouraged to bid more than others. The urgent CAV bids eight while others bid 0; Thus, it wins the whole priority value. Afterward, the karma is updated to 10, 5, and 12 for CAVs 5, 9, and 10, respectively. Cluster 3 will last for 33 time steps.

C. Overview of Test Results

The general information of the tests for uncongested and congested scenarios are summarized in Tables III and IV, respectively. The number of (non-repeated) TVCs is denoted by $N_{\rm cl}$. The average cluster duration (the number of simulation times that a cluster lasts) is denoted by $dur_{\rm avg}$. The average cluster dimension (number of CAVs inside a cluster $cl_{\rm dim}$) concerning the cluster duration ($cl_{\rm dur}$) is

$$dim_{\rm avg} = \frac{\sum cl_{\rm dim} cl_{\rm dur}}{\sum cl_{\rm dur}}.$$

Test	$N_{\rm cl}$	$dur_{\rm avg}$ 50 ms	$dim_{\rm avg}$ veh	$v_{\rm rms}$ m/s	$\theta_{\rm rms}$ Rad	$a_{\rm rms}$ m/s^2	t_s s
1	74	16.49	2.38	3.54	0.05	3.49	42.47
2	71	24.28	2.63	4.95	0.14	6.05	48.55
3	92	13.48	2.54	2.98	0.04	3.97	36.44
4	75	16.03	2.66	3.26	0.09	4.47	38.1
5	50	25.46	2.46	2.82	0.07	3.79	33.94
6	61	22.05	2.57	3.25	0.12	4.59	37.23
7	87	16.69	2.60	4.17	0.11	5.26	35.89
8	78	16.72	2.61	2.89	0.05	4.01	36.53
9	68	19.80	2.83	3.15	0.11	4.51	41.66
10	88	18.01	2.64	4.88	0.15	6.36	43.67
Mean	74.4	18.90	2.59	3.59	0.09	4.65	39.45

TABLE III: The simulation results of all the conducted tests for uncongested case, Scenario 1.

Test	$N_{\rm cl}$	$dur_{\rm avg}$ 50 ms	$dim_{\rm avg}$ veh	$v_{\rm rms}$ m/s	$\theta_{\rm rms}$ Rad	$a_{\rm rms}$ m/s^2	t_s s
1	79	7.96	4.03	3.09	0.06	4.56	99.76
2	92	6.13	4.44	2.57	0.04	3.92	94.06
3	53	9.28	4.97	2.73	0.06	4.21	105.89
4	41	10.17	5.61	2.94	0.05	4.03	129.38
5	36	11.63	3.68	4.28	0.09	6.05	81.29
6	57	7.51	5.71	2.90	0.06	4.32	125.43
7	38	11.55	6.20	3.37	0.07	5.10	142.18
8	37	13.03	4.88	3.17	0.07	4.26	98.99
9	53	8.72	5.76	3.25	0.06	4.7	126.67
10	38	15	3.37	3.86	0.05	5.19	63.61
Mean	52.4	10.10	4.87	3.22	0.06	4.63	106.73

TABLE IV: The simulation results of all the conducted tests for congested case, Scenario 2.

In the congested scenario, we observe fewer TVCs with shorter durations while, on average, having higher dimensions than in the uncongested scenario.

As we explained in upper-level control, the objective functions (6 and 18) are designed to minimize the deviation of speed, speed angle, and acceleration from their desired values. Thus, we employ the Root Mean Square Error (RMSE) concept to evaluate the general performance of the method. For variable f with the desired value of f^d, the RMSE error is calculated as follows.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(f(i) - f^d)^2} \tag{37}$$

Accordingly, we define $v_{\rm rms}$ (m/s), $\theta_{\rm rms}$ (rad), and $a_{\rm rms}$ (m/s^2) as the sum of RMSE in terms of speed, speed angle, and acceleration for all CAVs, respectively. Comparing the two scenarios, the difference in terms of $v_{\rm rms}$, $\theta_{\rm rms}$, and $a_{\rm rms}$ are trivial. This indicates the developed method maintains its high performance for the congested scenario. Assume the road width (12.5 m) has four lanes. The observed density for Scenario 1 begins at approximately 13 veh/km/lane and

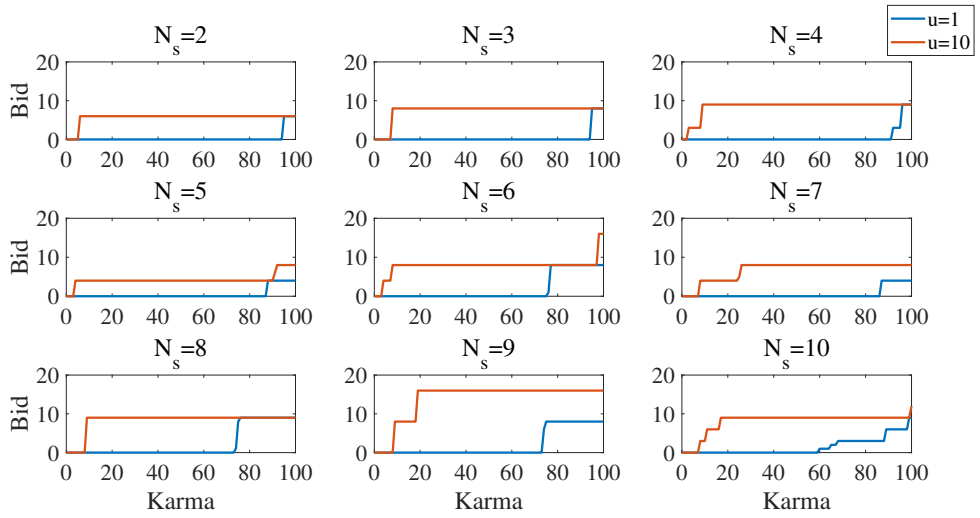


Fig. 5: The policy for different sub-populations at SNE for the modified Karma game with $\alpha = 0.9$.

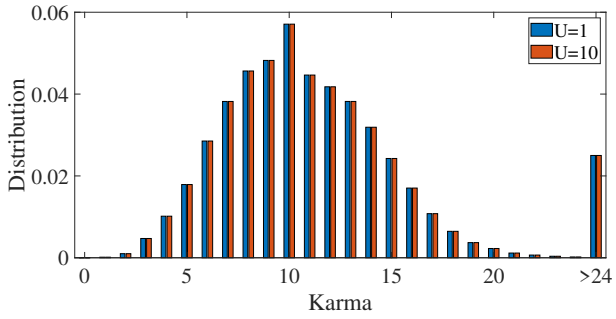


Fig. 6: The population probability distribution at SNE for the modified Karma game with $\alpha = 0.9$.

reaches a maximum of 29 veh/km/lane. For Scenario 2, the initial observed density is approximately 25 veh/km/lane, reaching the maximum of 62 veh/km/lane.

D. Safety

The other crucial objective of the designed method is collision avoidance. To investigate this property, the distance between CAVs is computed for all the tests. The minimum distance of CAVs was never lower than the defined safety margin, $r_{\mathrm{m}} = 3m$ (Table II). Thus, no collision was observed in any of the tests that were conducted. Figure 8 shows a close-up of the CAV's distance for all the tests.

E. Computational Time

The average computational time required for the control of TVCs is denoted by t_s (s) in Tables III and IV. The value of t_s increases significantly for the congested scenario due to the higher value of dim_{avg}. Increasing the number of vehicles in a TVC increases the size of the optimization problem, which requires higher computational effort. Figure 9 confirms this statement. We demonstrate the relation between cluster dimension and the required computational time for MPC by aggregating the data from all the conducted tests for both scenarios in Fig. 9.

To better understand the benefit of clustering, we conduct a simulation with a centralized control method for test 1, Scenario 2. In this centralized method, we implemented the same MPC method developed in this paper, assuming that only one TVC contains all the CAVs. To always obtain the optimal solution by IPOPT solver for the centralized method, the maximum CPU time and maximum number of iterations are increased to 5000 s and 3000, respectively. The average computational time required for the centralized control is 2663.9 s, which is 26.7 times the required time for the developed method with dynamic clustering.

Nevertheless, large-dimension clusters can appear in large-scale scenarios that require higher computational time. Thus, limiting the cluster dimension is crucial for reducing computational costs in large-scale scenarios. It is worth mentioning that the clusters are built based on safety measures and threat relationships. Thus, the threat detection algorithm requires modification to maintain safety while imposing limits on cluster dimensions. For instance, one can introduce Threatening levels for CAV pairs. Therefore, instead of labeling CAV pairs as threatening or non-threatening, each pair is associated with a level according to the risk of collision. It enables us to identify the pairs with lower emergencies that could be excluded from TVC due to the dimension limit.

The real-time implementation of the proposed method is hindered by the required computational time for the upper-level MPC controller. The computational burden for MPC-based strategies was frequently reported in the literature (See [55] and [52]). The recent research [56] proposed a fast implementation of coalitional MPC with the help of machine learning that decreases the computational time by 99%. Inspired by this research, our future focus is investigating the application of machine learning in providing a real-time MPC for upper-level control.

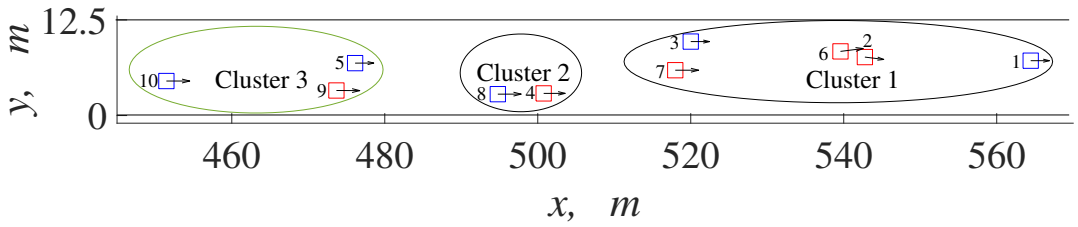


Fig. 7: Snapshot of uncongested scenario, Test 2, time step 332. The urgent CAVs are indicated in red color. The Karma game is played for the green clusters.

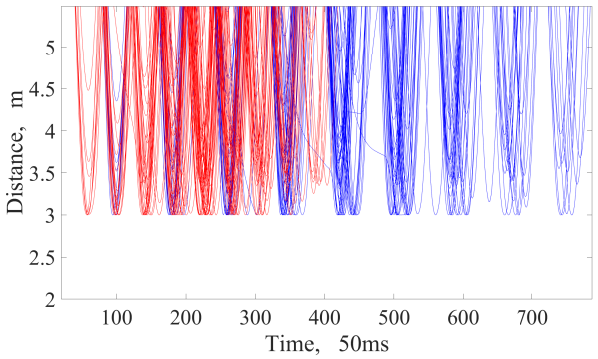


Fig. 8: The distance between CAVs for all the tests. The blue and red patterns belong to the uncongested and congested scenarios, respectively.

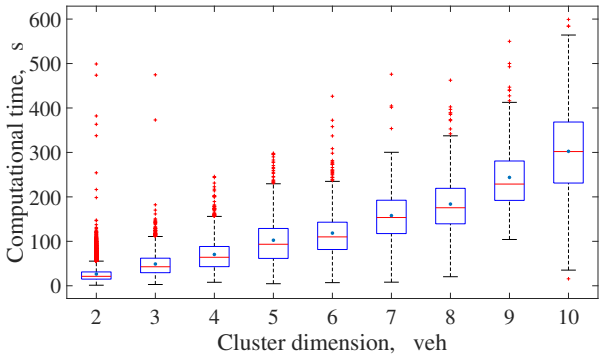


Fig. 9: Boxplot for cluster dimension vs required computational time for MPC. The mean values are shown with the blue stars.

F. TVC Analysis

In this part, TVCs are further analyzed regarding dimension and duration. Figure 10a shows the probability histogram of cluster dimensions (number of CAVs in a cluster) for congested and uncongested scenarios. The figure is created by collecting data from all the tests for every scenario. The cluster dimension lies between 2 and 10. It indicates that the minimum number of CAVs to form a TVC is 2, while the maximum size of a TVC is limited by the number of CAVs in the test. Please note that $P_s(N_s)$ depicted in Fig. 4 is a prediction of the cluster dimension histogram shown in Fig. 10a. The general trend observed in this figure suggests that the TVC with a smaller dimension has a higher probability.

The probability histogram of cluster duration (number of time steps a cluster lasts) is shown in figure 10b. Similarly, the histogram collects data over all tests for every scenario. A bar at the horizontal interval of $[i{-}i{+}10]$ shows the probability of the cluster duration being between $[i - i + 10]$ time steps. According to this figure, the probability decreases for larger TVC durations.

Comparison of the two histograms (Fig. 10a and Fig. 10b) for uncongested (blue color) and congested (red color) scenarios reveals an interesting outcome. For the congested scenario, the probability of the larger TVC dimensions increases. Meanwhile, the TVC duration decreases for the congested scenario. One may question whether increasing the TVC dimension will reduce the TVC duration. To answer this question, we look at the boxplots (Fig. 11a and Fig. 11b) for TVC dimension vs duration for each scenario. Figures show that the higher dimension of TVC is not associated with a lower duration. However, the duration of TVCs with lower dimensions decreases in the more congested scenario.

G. Modified Karma Game Analysis

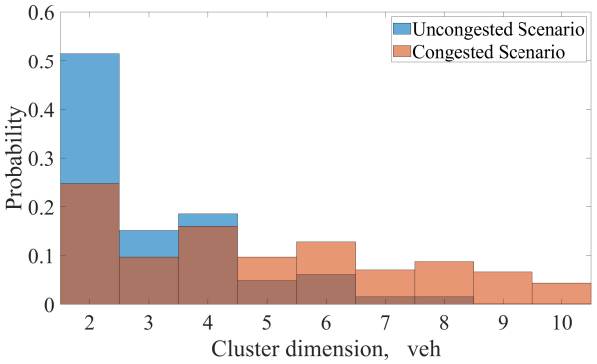
One of the main objectives of this research is to promote fairness in collaboration between CAVs. We evaluate the efficiency and fairness by the criteria introduced in [10]. We compare the SNE policy of the modified Karma game with two other policies, Dictator and Uniform. Dictator policy allocates resources equally only to the urgent agents. Whereas, the Uniform policy distributes the resources equally among all agents.

Suppose u_i denotes the urgency of agent i and β_i^{Poli} is the portion of the resource that player i wins if the policy $Poli$ (modified Karma, Dictator, or Uniform) is applied. The efficiency *eff* is defined as the average of $u_i\beta_i^{Poli}$ over all agents and all games that are played.

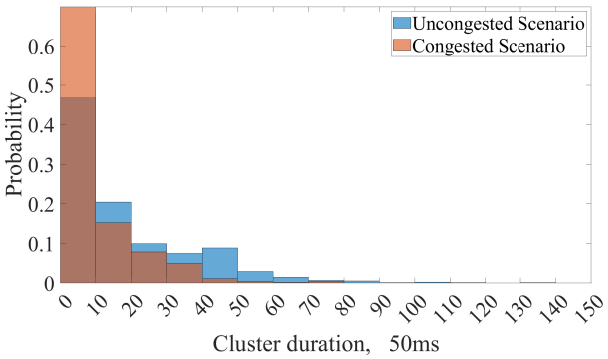
$$eff = \frac{\sum_{i=1}^{N}\sum_{t_i} u_i(t_i)\beta_i^{Poli}(t_i)}{N\sum_{i=1}^{N} tg(i)} \tag{38}$$

where t_i shows the time index where agent i plays game. $tg(i)$ is the total number of games that agent i played. Dictator policy distributes the resource only among agents with high u_i, obtaining the highest efficiency.

Two criteria are proposed by [10] to evaluate fairness in the Karma game: ex-post reward fairness *rf* and ex-post access fairness *af*. Ex-post reward fairness is the standard deviation

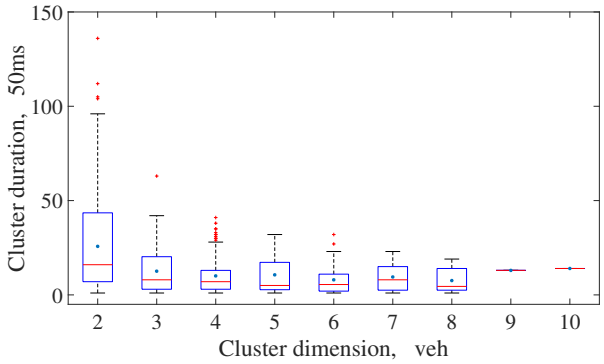


(a) Cluster dimension histogram.

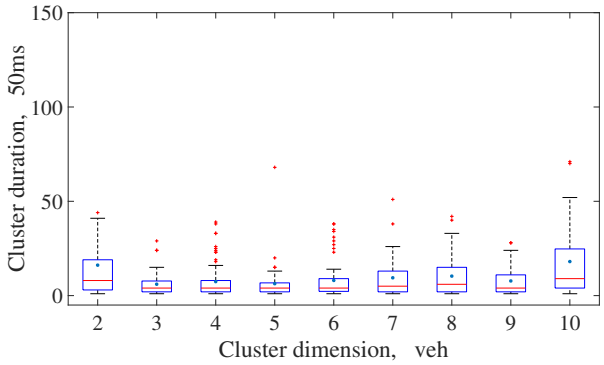


(b) Cluster duration histogram.

Fig. 10: The histograms for cluster dimension and duration. The blue and red colors indicate the uncongested and congested scenarios, respectively.



(a) Boxplot for uncongested case, Scenario 1.



(b) Boxplot for congested case, Scenario 2.

Fig. 11: Boxplot diagrams of cluster dimension vs cluster duration for uncongested and congested scenarios. The mean values are shown with the blue stars.

of the agent's average reward. Ex-post access fairness is the standard deviation of the agent's average access to the resource (the proportion of the resource that is allocated to the agent).

$$rf = -\text{std}\left(\frac{\sum_{t_i} u_i(t_i)\beta_i^{Poli}(t_i)}{tg(i)}\right) \tag{39}$$

$$af = -\text{std}\left(\frac{\sum_{t_i} \beta_i^{Poli}(t_i)}{tg(i)}\right) \tag{40}$$

It is expected that Uniform policy proposes the lowest *af* absolute value. We refer to *eff*, *rf*, and *af* as game criteria.

In Table V, we compare the different policies based on efficiency and fairness criteria for the two scenarios. The numbers represent the average value among all the tests for each scenario. According to the game criteria, the modified Karma game is a balanced approach regarding efficiency and fairness, as shown in Fig. 12). Karma obtains the same efficiency, *eff*, as the Dictator policy (the efficiency benchmark) while gains on fairness values. It promotes a fairness level (*rf* and *af*) close to the Uniform policy (the ex-post access fairness benchmark). For Scenario 1, Karma holds the best *rf* value. These results are summarized in Fig. 12. Figures 12a and 12b show the normalized criteria, hence facilitating the comparison of different policies. To obtain the normalized values, the efficiency criteria are divided by the maximum value obtained among policies (that always belong to the Dictator). The normalized fairness

	Scenario 1			Scenario 2		
Criteria	Karma	Dictator	Uniform	Karma	Dictator	Uniform
eff	287.79	293.28	220.04	195.08	196.66	132.67
rf	-48.12	-59.80	-48.22	-42.11	-55.35	-36.95
af	-4.81	-5.98	-3.79	-5.45	-5.53	-5.31

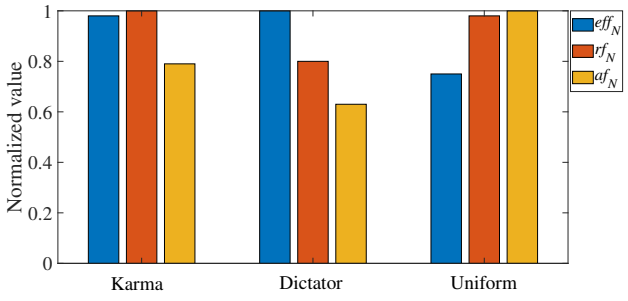
TABLE V: The comparison of different policies regarding efficiency and fairness primary criteria.

criteria equals the maximum value (minimum absolute value) obtained among policies divided by the fairness criteria. Thus, the better results are associated with higher values for all of the normalized criteria, where 1 indicates the best result.

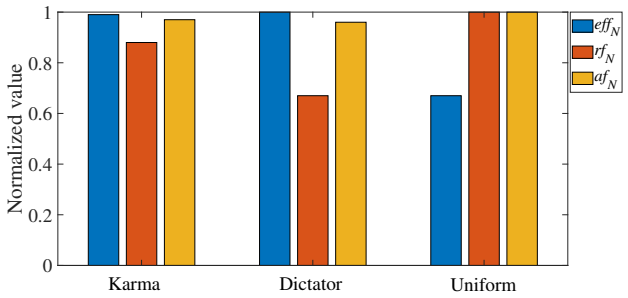
VI. Conclusions

In this paper, we proposed a decentralized bi-level control approach for CAV traffic in a lane-free environment with the three-fold objectives of efficiency, safety, and fairness. We designed a threat detection algorithm to distinguish the pairs of CAVs relying on the collision risk. Afterward, we introduced the notion of TVC by emphasizing crucial CAV connections for safety. This notion enables the division of the CAV traffic system into the subsystems of TVCs. Thus, instead of a centralized control for CAV traffic, we can design decentralized control for TVCs. Hence, it reduces the required computational effort.

An MPC-based approach is designed at the upper-level control for TVCs. To promote fairness in the CAV collabora-



(a) A comparison of policies for uncongested case, Scenario 1.



(b) A comparison of policies for congested case, Scenario 2.

Fig. 12: A comparison of policies in terms of efficiency and fairness normalized criteria for uncongested and congested scenarios. Higher values indicate better performance.

tion, we introduce priority parameters that rely on the CAVs' interaction history. A modified Karma game at the lower level determines priority parameters. The new Karma game is designed for proportional resource allocation within subpopulations (analogous to the TVC). Afterward, the priority parameters are incorporated into the MPC objective function at the upper-level control. The proposed methodology has been evaluated in simulation for a group-overpass case study. The important outcomes of the simulation results are:

- TVCs with higher dimensions require, on average, higher computational time.
- No collision has been observed during the tests.
- The share of the TVC with higher dimensions increases in denser traffic flows.
- Karma game provides a balanced approach regarding efficiency and fairness.

Although the proposed method significantly decreases the required computational time compared to the centralized method, it may not be applicable in real-time large-case scenarios because of the computational requirements of the MPC component. Inspired by [56], we are working to enable real-time implementation. This is indeed one of the objects of our future research. The idea is to rely on the application of machine learning and other strategies to reduce the required computational time for upper-level control.

On the other hand, limiting the possible dimension of clusters can help reduce computational time and required numbers of communication per TVC. This is a crucial measure, especially for the larger number of CAVs. However, it should not compromise safety.

Although the simulation results report no collision incident, providing solid proof of safety for the proposed method requires further investigation. This will be another objective of future development.

Finally, for the real-world application, the robustness of the proposed method toward malicious attacks needs to be investigated. Appropriate tests for evaluating the methodology performance under different sources of malicious attacks must be designed. In case of insufficiency in the method, proper modifications must be included. For instance, the vulnerability of the karma game in the presence of defective agents is an aspect out of the scope of the present paper but worth being addressed in the future. Limiting the obtained karma values for agents by implementing taxation [10] could be a way to prevent the hoarding of karma by defective agents and increase the robustness.

Acknowledgment

The Authors would like to sincerely thank Ezzat Elokda from ETH Zurich for the valuable and fruitful discussion about the Karma game. Antonella Ferrara is partially funded by PRIN 2022 2022BB9JC9 and PRIN PNRR P2022XJ9SX, supported by the European Union – Next Generation EU.

References

[1] W. Wang, G. F. Schuppe, and J. Tumova, "Decentralized multi-agent coordination under mitl specifications and communication constraints," in *2023 31st Mediterranean Conference on Control and Automation (MED)*, 2023, pp. 842–849.

[2] X. Shen and F. Borrelli, "Multi-vehicle conflict resolution in highly constrained spaces by merging optimal control and reinforcement learning," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 3308–3313, 2023, 22nd IFAC World Congress.

[3] M. Zambelli, M. Steinberger, M. Horn, and A. Ferrara, "Two-step asynchronous iterative formation control for heterogeneous vehicles in highway scenarios," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 4, pp. 3119–3128, 2023.

[4] A. Ferrara, G. P. Incremona, E. Birliba, and P. Goatin, "Multi-scale model-based hierarchical control of freeway traffic via platoons of connected and automated vehicles," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 799–812, 2022.

[5] G. Piacentini, P. Goatin, and A. Ferrara, "Traffic control via platoons of intelligent vehicles for saving fuel consumption in freeway systems," *IEEE Control Systems Letters*, vol. 5, no. 2, pp. 593–598, 2020.

[6] M. Zambelli and A. Ferrara, "Robustified distributed model predictive control for coherence and energy efficiency-aware platooning," in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 527–532.

[7] C. Pasquale, S. Sacone, S. Siri, and A. Ferrara, "A new micro-macro metanet model for platoon control in freeway traffic networks," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018, pp. 1481–1486.

[8] C. Daini, P. Goatin, M. L. Delle Monache, and A. Ferrara, "Centralized traffic control via small fleets of connected and automated vehicles," in *2022 European Control Conference (ECC)*. IEEE, 2022, pp. 371–376.

[9] E. Elokda, A. Censi, and S. Bolognani, "Dynamic population games," *arXiv preprint arXiv:2104.14662*, 2021.

[10] E. Elokda, S. Blognani, A. Censi, F. Dörfler, and E. Frazzoli, "A self-contained karma economy for the dynamic allocation of common resources," *Dynamic Games and Applications*, 2023.

[11] M. Papageorgiou, K.-S. Mountakis, I. Karafyllis, I. Papamichail, and Y. Wang, "Lane-free artificial-fluid concept for vehicular traffic," *Proceedings of the IEEE*, vol. 109, no. 2, pp. 114–121, 2021.

[12] K. Chavoshi and A. Kouvelas, "Cooperative distributed control for laneless and direction-less movement of autonomous vehicles on highway networks," in *9th Symposium of the European Association for Research in Transportation (hEART 2020)*, 2020.

[13] S. Topan, K. Leung, Y. Chen, P. Tupekar, E. Schmerling, J. Nilsson, M. Cox, and M. Pavone, "Interaction-dynamics-aware perception zones for obstacle detection safety evaluation," in *2022 IEEE Intelligent Vehicles Symposium (IV)*, 2022, pp. 1201–1210.

[14] J. Alonso-Mora, A. Breitenmoser, M. Rufli, P. Beardsley, and R. Siegwart, "Optimal reciprocal collision avoidance for multiple non-holonomic robots," in *Distributed autonomous robotic systems: The 10th international symposium*. Springer, 2013, pp. 203–216.

[15] D. Bareiss and J. Van den Berg, "Generalized reciprocal collision avoidance," *The International Journal of Robotics Research*, vol. 34, no. 12, pp. 1501–1514, 2015.

[16] S. Wang, X. Hu, J. Xiao, and T. Chen, "Repulsion-oriented reciprocal collision avoidance for multiple mobile robots," *Journal of Intelligent & Robotic Systems*, vol. 104, pp. 1–21, 2022.

[17] I. Batkovic, M. Zanon, and P. Falcone, *Model Predictive Control for Safe Autonomous Driving Applications*. Cham: Springer International Publishing, 2023, pp. 255–282.

[18] M. Geurts, A. Katriniok, E. Silvas, and W. Heemels, "Model predictive control for lane merging automation with recursive feasibility guarantees," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 4858–4864, 2023, 22nd IFAC World Congress.

[19] A. Ferrara, S. Sacone, and S. Siri, *Freeway traffic modelling and control*. Springer, 2018, p. 90.

[20] S. Siri, C. Pasquale, S. Sacone, and A. Ferrara, "Freeway traffic control: A survey," *Automatica*, vol. 130, p. 109655, 2021.

[21] V. P. Chellapandi, L. Yuan, C. G. Brinton, S. H. Żak, and Z. Wang, "Federated learning for connected and automated vehicles: A survey of existing approaches and challenges," *IEEE Transactions on Intelligent Vehicles*, 2023.

[22] K. Chavoshi, A. Ferrara, and A. Kouvelas, "A feedback linearization approach for coordinated traffic flow management in highway systems," *Control Engineering Practice*, vol. 139, p. 105615, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0967066123001843

[23] K. Chavoshi, A. Genser, and A. Kouvelas, "A pairing algorithm for conflict-free crossings of automated vehicles at lightless intersections," *Electronics*, vol. 10, no. 14, 2021. [Online]. Available: https://www.mdpi.com/2079-9292/10/14/1702

[24] M. Sekeran, M. Rostami-Shahrbabaki, A. A. Syed, M. Margreiter, and K. Bogenberger, "Lane-free traffic: History and state of the art," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, 2022, pp. 1037–1042.

[25] V. K. Yanumula, P. Typaldos, D. Troullinos, M. Malekzadeh, I. Papamichail, and M. Papageorgiou, "Optimal trajectory planning for connected and automated vehicles in lane-free traffic with vehicle nudging," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 3, pp. 2385–2399, 2023.

[26] R. Levy and J. Haddad, "Cooperative path and trajectory planning for autonomous vehicles on roads without lanes: A laboratory experimental demonstration," *Transportation Research Part C: Emerging Technologies*, vol. 144, no. 103813, 2022.

[27] M. Rostami-Shahrbabaki, H. Zhang, M. Sekeran, and K. Bogenberger, "Increasing the capacity of a lane-free beltway for connected and automated vehicles using potential lines," in *102nd Annual Meeting Transportation Research Board*, 2023.

[28] D. Theodosis, I. Karafyllis, and M. Papageorgiou, "Cruise controllers for lane-free ring-roads based on control lyapunov functions," *Journal of the Franklin Institute*, vol. 360, no. 9, pp. 6131–6161, 2023.

[29] I. Karafyllis, D. Theodosis, and M. Papageorgiou, "Lyapunov-based two-dimensional cruise control of autonomous vehicles on lane-free roads," *Automatica*, vol. 145, no. 110517, 2022.

[30] A. Karalakou, D. Troullinos, G. Chalkiadakis, and M. Papageorgiou, "Deep reinforcement learning reward function design for autonomous driving in lane-free traffic," *Systems*, vol. 134, no. 3, 2023.

[31] D. Troullinos, G. Chalkiadakis, I. Papamichail, and M. Papageorgiou, "Collaborative multiagent decision making for lane-free autonomous driving," in *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, ser. AAMAS '21. International Foundation for Autonomous Agents and Multiagent Systems, 2021, p. 1335–1343.

[32] M. Malekzadeh, I. Papamichail, and M. Papageorgiou, "Linear–quadratic regulators for internal boundary control of lane-free automated vehicle traffic," *Control Engineering Practice*, vol. 115, no. 104912, 2021.

[33] M. Malekzadeh, V. K. Yanumula, I. Papamichail, and M. Papageorgiou, "Overlapping internal boundary control of lane-free automated vehicle traffic," *Control Engineering Practice*, vol. 133, no. 105435, 2023.

[34] K. Chavoshi and A. Kouvelas, "Distributed control for laneless and directionless movement of connected and automated vehicles," in *21st Swiss Transport Research Conference (STRC 2021)*, 2021.

[35] Z. He, H. Pei, Y. Guo, D. Yao, and L. Li, "Theoretical trade-off between fairness and efficiency in the cooperative driving problem for cavs at on-ramps," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 5, pp. 41–54, 2024.

[36] M. Kunibe, H. Asahina, H. Shigeno, and I. Sasase, "A scheduling scheme for autonomous vehicle highway merging with an outflow traffic and fairness analysis," *IEEE Access*, vol. 9, pp. 49 219–49 232, 2021.

[37] C. Li, X. Ma, L. Xia, Q. Zhao, and J. Yang, "Fairness control of traffic light via deep reinforcement learning," in *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE)*, 2020, pp. 652–658.

[38] Y. Ye, J. Ding, T. Wang, J. Zhou, X. Wei, and M. Chen, "Fairlight: Fairness-aware autonomous traffic signal control with hierarchical action space," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2022.

[39] B. L. Ferguson and J. R. Marden, "The impact of fairness on performance in congestion networks," in *2021 American control conference (ACC)*. IEEE, 2021, pp. 4521–4526.

[40] L. Zhang, M. Khalgui, Z. Li, and Y. Zhang, "Fairness concern-based coordinated vehicle route guidance using an asymmetrical congestion game," *IET Intelligent Transport Systems*, vol. 16, no. 9, pp. 1236–1248, 2022.

[41] A. Athalye and S. Nayak, "Fairness and robustness of mixed autonomous traffic control with reinforcement learning," 2021.

[42] L. Jiang, Y. Xie, and N. G. Evans, "A simulation study of cooperative and autonomous vehicles (cav) considering courtesy, ethics, and fairness," *Plos one*, vol. 18, no. 5, p. e0283649, 2023.

[43] S. Chen, M. Wang, W. Song, Y. Yang, and M. Fu, "Multi-agent reinforcement learning-based twin-vehicle fair cooperative driving in dynamic highway scenarios," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2022, pp. 730–736.

[44] D. Paccagnan, R. Chandan, and J. R. Marden, "Utility and mechanism design in multi-agent systems: An overview," *Annual Reviews in Control*, vol. 53, pp. 315–328, 2022.

[45] G. Arslan, J. R. Marden, and J. S. Shamma, "Autonomous vehicle-target assignment: A game-theoretical formulation," 2007.

[46] M. Liu, I. Kolmanovsky, H. E. Tseng, S. Huang, D. Filev, and A. Girard, "Potential game-based decision-making for autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 8, pp. 8014–8027, 2023.

[47] N. Li, Y. Yao, I. Kolmanovsky, E. Atkins, and A. R. Girard, "Game-theoretic modeling of multi-vehicle interactions at uncontrolled intersections," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 2, pp. 1428–1442, 2022.

[48] Q. Zhang, R. Langari, H. E. Tseng, D. Filev, S. Szwabowski, and S. Coskun, "A game theoretic model predictive controller with aggressiveness estimation for mandatory lane change," *IEEE Transactions on Intelligent Vehicles*, vol. 5, no. 1, pp. 75–89, 2020.

[49] A. R. Hota, U. Maitra, E. Elokda, and S. Bolognani, "Learning to mitigate epidemic risks: A dynamic population game approach," *Dynamic Games and Applications*, pp. 1–24, 2023.

[50] A. Censi, S. Bolognani, J. G. Zilly, S. Sadat Mousavi, and E. Frazzoli, "Today me, tomorrow thee: Efficient resource allocation in competitive settings using karma games," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019, pp. 686–693.

[51] E. Elokda, C. Cendese, K. Zhang, A. Censi, J. Lygeros, and E. Frazzoli, "Karma priority lanes for fair and efficient bottleneck congestion management," in *2023 31st Mediterranean Conference on Control and Automation (MED)*, 2023, pp. 458–463.

[52] R. Levy and J. Haddad, "Path and trajectory planning for autonomous vehicles on roads without lanes," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 2021, pp. 3871–3876.

[53] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 106, no. 1, pp. 25–57, 2006.

[54] J. Löfberg, "Yalmip : A toolbox for modeling and optimization in matlab," in *In Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[55] M. Schuurmans, A. Katriniok, C. Meissen, H. E. Tseng, and P. Patrinos, "Safe, learning-based mpc for highway driving under lane-change uncertainty: A distributionally robust approach," *Artificial Intelligence*, vol. 320, p. 103920, 2023.

[56] E. Masero, S. Ruiz-Moreno, J. R. D. Frejo, J. M. Maestre, and E. F. Camacho, "A fast implementation of coalitional model

predictive controllers based on machine learning: Application to solar power plants," *Engineering Applications of Artificial Intelligence*, vol. 118, p. 105666, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095219762200656X

Kimia Chavoshi (S'20) received the B.Sc. and M.Sc. degrees in Control Systems from the Electrical Engineering Department of Sharif University of Technology, Tehran, Iran, in 2015 and 2018, respectively. She is currently working towards a Ph.D. degree in Traffic Engineering and Control at the Institute for Transport Planning and Systems (IVT), Department of Civil, Environmental and Geomatic Engineering, ETH Zurich (ETHZ), Switzerland. Her research lies at the intersection of control theory and traffic engineering and is funded by the Swiss National Science Foundation (SNSF). Her research interests include developing control strategies for CAVs' collaboration as well as macroscopic traffic flow control and optimization.

Antonella Ferrara (Fellow, IEEE) received the M.Sc. degree in electronic engineering and the Ph.D. degree in computer science and electronics from the University of Genoa, Italy, in 1987 and 1992, respectively. Since 2005, she has been a Full Professor of automatic control with the University of Pavia, Italy. She is the author and coauthor of more than 450 publications, including more than 160 journal papers, 2 monographs and 1 edited book. Her main research interests include sliding mode control and other nonlinear control methodologies applied to traffic systems, intelligent vehicles, power networks, and robotics. She is currently a Senior Editor of the IEEE Open Journal of Intelligent Transportation Systems and an Associate Editor of Automatica. She was a Senior Editor of the IEEE Transactions on Intelligent Vehicles and an Associate Editor of the IEEE Transactions on Control Systems Technology, IEEE Transactions on Automatic Control, IEEE Control Systems Magazine, and International Journal of Robust and Nonlinear Control. She has been the EUCA Conference Editorial Board Chair since 2018. She is a member of the IEEE Control Systems Society and IEEE Intelligent Transportation Systems Society, as well as of several Technical Committees among which the IEEE TC on Automotive Control, IEEE TC on Smart Cities, IEEE TC on Variable Structure Systems, IFAC Technical Committee on Nonlinear Control Systems, IFAC TC on Transportation Systems, and IFAC Technical Committee on Intelligent Autonomous Vehicles. She is an IEEE Fellow and an IFAC Fellow.

Anastasios Kouvelas (M'12–SM'21) is the Director of the research group Traffic Engineering and Control at the Institute for Transport Planning and Systems (IVT), Dept. of Civil, Environmental and Geomatic Engineering, ETH Zurich (since August 2018). He has received the Diploma, M.Sc. and Ph.D. degrees from the Department of Production & Management Engineering (Operations Research), Technical University of Crete, Greece, in 2004, 2006, and 2011, respectively, specializing in modeling, control, and optimization of large-scale transport systems. Prior to joining IVT, he was a Research Scientist at Urban Transport Systems Laboratory (LUTS), EPFL (2014-2018), and a Postdoctoral Fellow with Partners for Advanced Transportation Technology (PATH) at University of California, Berkeley (2012-2014). Before this, he was appointed Adjunct Professor at Technical University of Crete (2011) and Research Associate at the Information Technologies Institute, Centre for Research & Technology Hellas (CERTH), Thessaloniki, Greece (2011-2012). In 2009, he was a Doctoral Visiting Scholar at the Center for Advanced Transportation Technologies (CATT) of the Viterbi School of Engineering, Dept. of Electrical Engineering, University of Southern California, Los Angeles. Dr. Kouvelas has been awarded with the 2012 Best IEEE ITS PhD Dissertation Award from IEEE Intelligent Transportation Systems Society. He is a member of the Transportation Research Board (TRB) Standing Committee (AHB15) on Intelligent Transportation Systems, since 2019. Dr. Kouvelas serves as Associate Editor in IET Intelligent Transport Systems journal since 2017 and has guest-edited several special issues in transportation journals.

APPENDIX A
PROOF OF THE PROPOSITION 1

Proof. According to Theorem 1 at [10], if the state transition function ρ is continuous in (d,π) and the karma is preserved inside the population, then an SNE exists for the Karma algorithm.

Based on equation (27), the continuity of $\rho[u^+,k^+,N_s^+|u,k,b,N_s]$ in (d,π), is equivalent to the continuity of $\gamma[c|b,N_s]$ in (d,π) as $\gamma[c|b,N_s](d,\pi)$ is the only term which depends on (d,π). According to equations (24):

$$\gamma[c|b,N_s](d,\pi)=V_s[b_s=\frac{Cb}{c}|N_s](d,\pi).$$

Therefore, $\gamma[c|b,N_s]$ is continuous in (d,π) if $V_s[b_s|N_s](d,\pi)$ is continuous.

Looking at equation (23), $V_s[b_s|N_s](d,\pi)$ is an N_s fold convolution of $V[b|N_s](d,\pi)$. Thus, $V_s[b_s|N_s](d,\pi)$ is continuous provided that $V[b|N_s](d,\pi)$ is continuous. Based on equation (22):

$$V[b|N_s](d,\pi)=\sum_{u,k}d[u,k]\pi(b|u,k,N_s),$$

which is continuous in (d,π). Thus we can conclude the continuity of $\rho[u^+,k^+,N_s^+|u,k,b,N_s]$ in (d,π).

The karma is preserved inside the population if the total value of karma remains the same. In our modified Karma game, it is equivalent to $\sum k^+=\sum k$. In order to investigate this condition, we can use the karma transition function in equation (25) to calculate the sum of updated karma ($\hat{}k^+$) as follows:

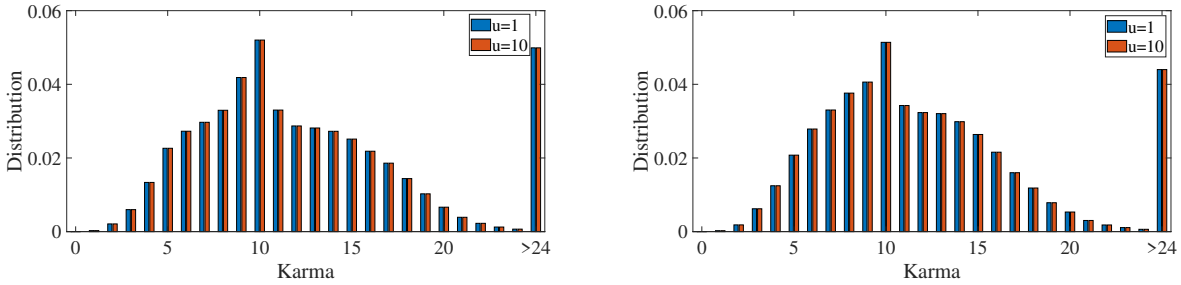
$$\begin{aligned}\sum_{i=1}^{N_s}k_i^+ &= \sum_{i=1}^{N_s}\Bigg(k_i-b_i+\left(\frac{b_s}{N_s}-\lfloor\frac{b_s}{N_s}\rfloor\right)\left(\lceil\frac{b_s}{N_s}\rceil\right)\\ &\quad+\left(1-\frac{b_s}{N_s}+\lfloor\frac{b_s}{N_s}\rfloor\right)\left(\lfloor\frac{b_s}{N_s}\rfloor\right)\Bigg)\\ &=\sum_{i=1}^{N_s}k_i-\sum_{i=1}^{N_s}b_i+\sum_{i=1}^{N_s}\Bigg(\left(\frac{b_s}{N_s}-\lfloor\frac{b_s}{N_s}\rfloor\right)\left(1+\lfloor\frac{b_s}{N_s}\rfloor\right)\\ &\quad+\left(1-\frac{b_s}{N_s}+\lfloor\frac{b_s}{N_s}\rfloor\right)\left(\lfloor\frac{b_s}{N_s}\rfloor\right)\Bigg)\\ &=\sum_{i=1}^{N_s}k_i-b_s+\sum_{i=1}^{N_s}\frac{b_s}{N_s}=\sum_{i=1}^{N_s}k_i.\end{aligned}$$

Therefore, karma is preserved inside the population for the presented algorithm.

As both conditions in Theorem 1 at [10] are fulfilled, we can conclude the existence of SNE for the modified karma game presented in Section IV. □

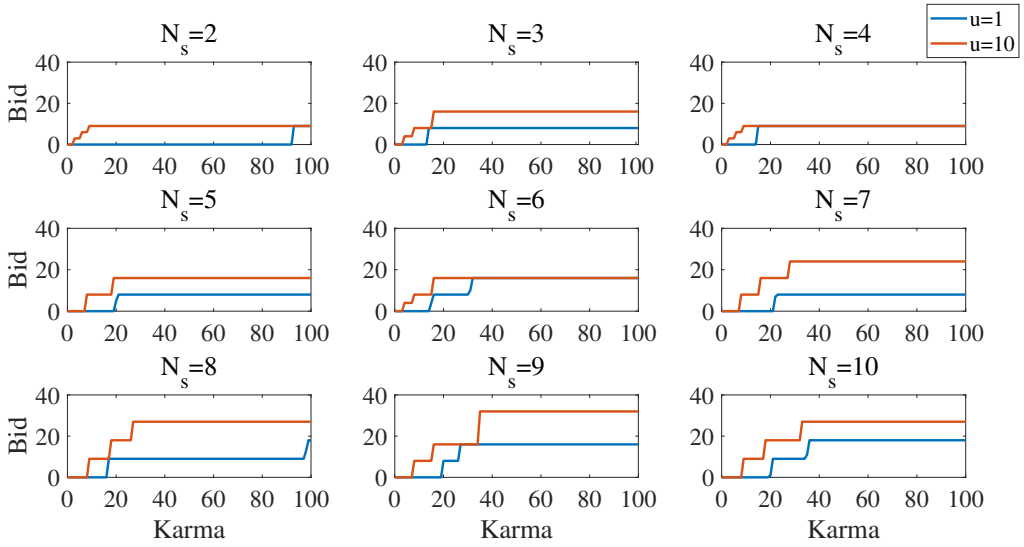
APPENDIX B
IMPACT OF FUTURE DISCOUNT FACTOR ON SNE

The SNE of the modified Karma game with the future discount factor $\alpha=0.5$ and $\alpha=0.7$ are shown in Fig. 13 and Fig. 14.

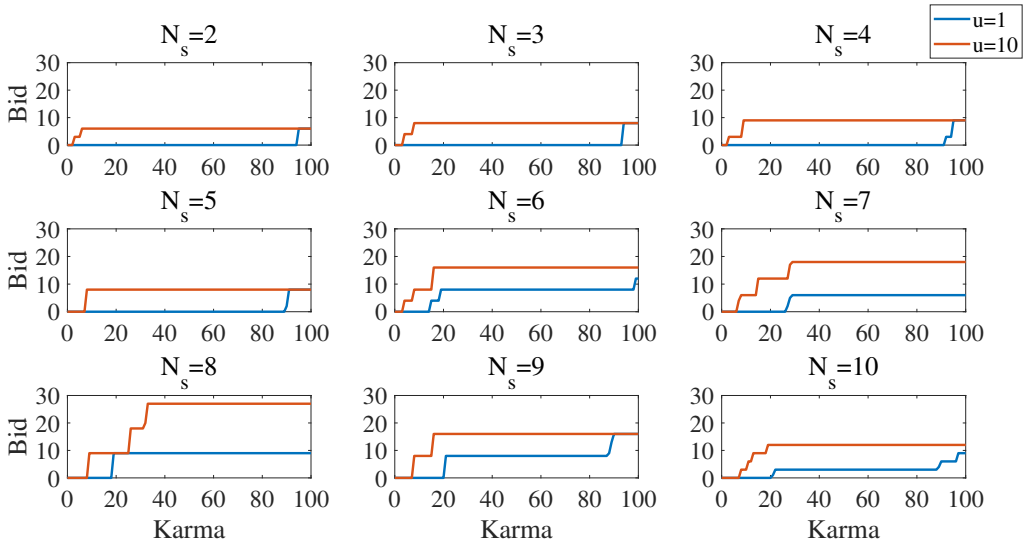


(a) The population probability distribution at SNE, $\alpha = 0.5$. (b) The population probability distribution at SNE, $\alpha = 0.7$.

Fig. 13: The population probability distribution at SNE for the modified Karma game with $\alpha = 0.5$ and $\alpha = 0.7$.



(a) The policy for different sub-populations at SNE, $\alpha = 0.5$.



(b) The policy for different sub-populations at SNE, $\alpha = 0.7$.

Fig. 14: The bidding policy for the modified Karma game with $\alpha = 0.5$ and $\alpha = 0.7$.