

# Dynamic driving and routing games for autonomous vehicles on networks: A mean field game approach

Kuang Huang<sup>a</sup>, Xu Chen<sup>b</sup>, Xuan Di<sup>b,c,\*</sup>, Qiang Du<sup>a,c</sup>

<sup>a</sup> Department of Applied Physics and Applied Mathematics, Columbia University, United States

<sup>b</sup> Department of Civil Engineering and Engineering Mechanics, Columbia University, United States

<sup>c</sup> Data Science Institute, Columbia University, United States

## ARTICLE INFO

### Keywords:

Driving and route choice game  
N-Car differential game  
Mean field game

## ABSTRACT

This paper aims to answer the research question as to optimal design of decision-making processes for autonomous vehicles (AVs), including dynamical selection of driving velocity and route choices on a transportation network. Dynamic traffic assignment (DTA) has been widely used to model travelers' route choice or/and departure-time choice and predict dynamic traffic flow evolution in the short term. However, the existing DTA models do not explicitly describe one's selection of driving velocity on a road link. Driving velocity choice may not be crucial for modeling the movement of human drivers but it is a must-have control to maneuver AVs. In this paper, we aim to develop a game-theoretic model to solve for AVs' optimal driving strategies of velocity control in the interior of a road link and route choice at a junction node. To this end, we will first reinterpret the DTA problem as an  $N$ -car differential game and show that this game can be tackled with a general mean field game-theoretic framework. The developed mean field game is challenging to solve because of the forward and backward structure for velocity control and the complementarity conditions for route choice. An efficient algorithm is developed to address these challenges. The model and the algorithm are illustrated on the Braess network and the OW network with a single destination. On the Braess network, we first compare the LWR based DTA model with the proposed game and find that the driving and routing control navigates AVs with overall lower costs. We then compare the total travel cost without and with the middle link and find that the Braess paradox may still arise under certain conditions. We also test our proposed model and solution algorithm on the OW network.

## 1. Motivation

When all the human-driven vehicles (HVs) on public roads are replaced by autonomous vehicles (AVs), how should we design *decision-making processes for AVs such that they dynamically select driving velocity and route choices on a transportation network to minimize some travel costs over a predefined planning horizon?* The challenge is, however, when a large number of AVs with different origin–destination pairs navigate a transportation network, traffic evolves as time elapses and congestion arises, which could increase AVs' travel costs and in turn influence their decision-making. We are interested in understanding the *equilibrium state* in which each AV cannot minimize its travel cost by unilaterally switching driving velocity and route choices.

\* Corresponding author at: Department of Civil Engineering and Engineering Mechanics, Columbia University, United States.  
E-mail address: [sharon.di@columbia.edu](mailto:sharon.di@columbia.edu) (X. Di).

The traffic assignment problem models one's route choice behavior while interacting with other travelers (which is the travel demand) in order to predict network-wide traffic congestion. While static traffic assignment is proposed for the long-term planning purpose (Di et al., 2013; Di et al., 2014; Di and Liu, 2016; Di and Ban, 2019), dynamic traffic assignment (DTA) is the most popular prescriptive and normative approach to model travelers' route choice or/and departure-time choice and predict dynamic traffic flow evolution in a short term (Friesz et al., 1993; Peeta and Ziliaskopoulos, 2001; Ban et al., 2012). DTA integrates both notions of travel demand (consistent with static traffic assignment) and traffic flow (i.e., dynamic traffic evolution) and is thus appropriate for modeling dynamic movement of human drivers across a network (Friesz et al., 2013). DTA problems bear many variants and interested readers can refer to Peeta and Ziliaskopoulos (2001) and Nie and Zhang (2005b) for a comprehensive overview of DTA models.

The existing DTA models do not explicitly depict one's selection of driving velocity on a road link. Travel speed is either determined by traffic density when macroscopic traffic flow models are used or is simplified when bottleneck or queuing models are used. For example, in the LWR model (Lighthill, 1952; Richards, 1956), one's vehicle velocity on a link is determined by traffic density on that link via fundamental diagrams (Friesz et al., 2013; Han et al., 2016; Han et al., 2019). On a link with bottlenecks, one vehicle moves in the interior of the link without incurring any travel time and only joins a queue at the end of the link (Ban et al., 2012; Osorio et al., 2011). Driving behavioral simplification mentioned above may not be so crucial for the existing models that were primarily developed for human drivers. However, when it comes to design a network-wide control scheme for AVs to navigate a transportation network, optimal selection of driving velocity is a must-have control to maneuver AVs on a link, in addition to departure-time choice at origins and route choice at intermediate junctions.

In this paper, we aim to develop a game-theoretic model to solve for AVs' optimal driving strategies in longitudinal control and route choice. To this end, we will first reinterpret the DTA problem as an  $N$ -car differential game and show that this game can be tackled with a general mean field game-theoretic framework. Then we will generalize the existing DTA problem to a broader context when one can select driving velocity while moving on a network.

In the remainder of the paper, we will first review related literature on DTA and mean field games in Section 2. We will then define one AV's optimal control problem to navigate a transportation network, and then extend it to a differential game when multiple AVs solve their individual optimal control problems while interacting among one another on a congested network in Section 3. Due to complexity of the differential game, we will apply the mean field approximation and derive a mean field game for a large number of AVs in Section 4. We will demonstrate the connection between the mean field game and the classical dynamic user equilibrium concept in Section 5. In Section 6, an efficient algorithm is developed to solve the mean field equilibrium. In Section 7, we will do some numerical experiments to demonstrate the solved mean field equilibrium on Braess networks and the OW network. Conclusions follow in Section 8 with future research directions.

## 2. Literature review

### 2.1. Dynamic user equilibrium

Dynamic user equilibrium (DUE) or dynamic user optimal describes the equilibrium state in which one cannot minimize his or her generalized travel cost by unilaterally switching route choices or departure times simultaneously (Friesz and Han, 2019). Solving DUE usually requires to solve two subproblems: a dynamic network loading (DNL) procedure and a route and/or departure-time choice model.

**The DNL procedure** propagates aggregate traffic flow dynamics and congestion in space and time (depicted by link volume, exit flow, link/path travel time or delay, queue length) given link inflows or path departure rates. It is composed of a link model and a junction model. A *link model* captures three phenomena: flow conservation, flow behavior (to compute travel delay), and flow propagation (how link volume propagates to exit (Bliemer et al., 2017; Yu et al., 2020)). These phenomena are modeled with two types of functions: the delay function model and the exit-flow function model (Nie and Zhang, 2005b). The former computes link delays or queues explicitly using a linear delay function (Friesz et al., 1993; Xu et al., 1999), while the latter implicitly derives link traversal time from link exit flow. All the existing exit-flow functions model link flow on a macroscopic scale, including: M-N model (Merchant and Nemhauser, 1978a; Merchant and Nemhauser, 1978b), whole link model (Xu et al., 1999; Carey and McCartney, 2002; Carey and Ge, 2003; Nie and Zhang, 2005a), traffic flow models without spillback (Friesz et al., 2013) or with spillback (Gentile et al., 2007; Han et al., 2016), discrete-space traffic flow models including cell transmission models (CTM) (Lo, 1999; Lo and Szeto, 2002; Szeto and Lo, 2004; Zhu and Ukkusuri, 2015) and link transmission models (LTM) (Ukkusuri et al., 2012; Gentile, 2015; Batista et al., 2021), point-queue or bottleneck models such as single queue (Ban et al., 2012; Han et al., 2013) or double queues (Osorio et al., 2011), and physical or spatial queue models (Kuwahara and Akamatsu, 2001; Zhang et al., 2013; Doan and Ukkusuri, 2015). Path delay can be computed using an implicit path delay operator (Friesz et al., 2013; Han et al., 2016). A *junction model* determines how flows are distributed at intersections. In traffic flow models on a network, boundary conditions at a junction can become complex and may lead to ill-posed models. Fixed turning ratios or entropy optimization is imposed for a unique flow distribution in LWR models (Han et al., 2016), and demand-supply functions are proposed with sending and receiving flows as functions of traffic states (Lebacque and Khoshyaran, 1999).

**A choice model** stipulates how an individual traveler makes route or departure-time choice, given the prevailing or predictive traffic condition. If the prevailing or present traffic information is used for decision-making, we call it the instantaneous UE (IUE) principle (Wie et al., 1990; Boyce et al., 1995; Lam and Huang, 1995; Ban et al., 2012), otherwise it is ideal DUE (Ran and Boyce, 1996; Ran et al., 1996; Friesz et al., 2013).

The inherent integration of dynamic traffic evolution and route choices enables the DUE problem to be analytically formulated as

differential variational inequalities (DVI) (Friesz and Han, 2019) or differential complementarity systems (DCS) (Ban et al., 2012). Depending on spatial and temporal resolutions, DUE is described by algebraic difference equations or differential–algebraic equations (DAEs), which is a system of (time-delayed) ordinary differential equations (ODEs) or partial differential equations (PDEs) with algebraic constraints. Due to computational intractability of the analytical DUE problem, a large body of literature has resorted to simulations or heuristics (Mahmassani, 2001; Lo and Szeto, 2002; Chiu et al., 2013) to compute the traffic equilibrium state. Levin and Boyles (2016) is one among a few that designs networked traffic controls for AVs using the DUE framework with CTMs. To model the mixed traffic comprised of AVs and human drivers, a multi-class LWR model is developed with a fundamental diagram assuming that AVs and human drivers have different reaction times but the same driving speed. The maximum capacity varies with the penetration rate of AVs. At junctions, intelligent traffic management policy is developed to compute nodal delay for each vehicle class. Simulations are performed to compute optimal controls for AVs and the resultant equilibrium.

## 2.2. Multi-agent differential games and mean field games

The spatio-temporal evolution of aggregate traffic dynamics essentially arises from a large number of individuals' dynamic travel choices and their complex interactions on a network. The fundamental tool to model individuals' decision-making processes in a multi-agent dynamic system is the non-cooperative multi-agent differential game (Dockner et al., 2000; Shou and Di, 2020). In a differential game, each agent solves its own control from an optimal control problem when every other agent in the system does so. Accordingly, one's optimal control problem is coupled with all others through the system state.

Assuming agents are anonymous, instead of solving a long list of highly coupled optimal control problems in a multi-agent differential game, mean field approximation can be applied to exploit the “smoothing” effect of large numbers of interacting individuals. At equilibrium, each agent interacts and reacts only to a “mass” resulting from the aggregate effect of all agents. The mean field game (MFG) is a micro–macro model that allows one to define individual drivers on a microscopic level as rational, utility-optimizing agents while translating their rich microscopic behaviors to a macroscopic scale. It is composed of two coupled components: agent dynamic and mass dynamic. The *agent dynamic* models individuals' dynamics with an optimal control in the form of a backward Hamilton–Jacobi–Bellman (HJB) equation, while the *mass dynamic* models system evolution arising from individuals' choices in the form of a forward Fokker–Planck–Kolmogorov (FPK) equation.

MFGs have become an increasingly popular tool to design new decision-making processes in finance (Guéant et al., 2011; Lachapelle et al., 2010), engineering (Djehiche et al., 2016; Couillet et al., 2012), learning (Gummadi et al., 2012; Iyer et al., 2014), social science (Degond et al., 2014), and pedestrian crowd movement (Lachapelle and Wolfram, 2011; Burger et al., 2013). Djehiche et al. (2016) built a connection between the Wardrop equilibrium and the mean field equilibrium by reinterpreting path flows as the mean field of individuals' route choices. In traffic modeling, there exist only a few studies that employ MFGs on velocity control (Kachroo et al., 2016; Chevalier et al., 2015; Kachroo et al., 2017; Chen et al., 2016) and lane-change (Festa and Göttlich, 2017). The authors' recent studies have employed this tool to solve for optimal velocity control of AVs in pure AV traffic (Huang et al., 2020a) or mixed AV–HV traffic (Huang et al., 2020b; Huang et al., 2019; Di and Shi, 2021). However, these studies are primarily focused on a single road rather than on a network. On a transportation network, Bauso et al. (2016) used the notion of MFG for routing games where static traffic density flow is solved on each link but its temporal evolution is not captured.

This paper aims to fill the above gaps by formulating the AVs' dynamic control problem on a network as a MFG. Due to the integration of driving velocity control and route choices, the formulation and computation of the MFG on a network is expected to be more challenging.

## 3. N-car driving and routing differential game on a network

### 3.1. Problem statement

There are  $N$  cars indexed by  $n \in \{1, 2, \dots, N\}$  who need to move from their initial positions to a destination on a congested transportation network during a predefined time horizon. Their travel mode is autonomous vehicle (AV). We assume these AVs as intelligent agents who select optimal control strategies throughout the network, in order to minimize their pre-programmed travel cost functionals over the predefined planning horizon. In other words, each AV aims to select a minimum-cost driving profile to navigate the transportation network. En-route, one AV needs to select its driving speed continuously; at any junction, the AV needs to choose the next-go-to link. When one AV selects its own driving speed and routing policy while everybody else does so simultaneously, a non-cooperative *differential game* forms. We are interested in an equilibrium control strategy at which no AV can improve its travel cost by unilaterally switching its driving speed or routing policy.

**Assumptions.** To clarify, we first make the following modeling assumptions.

1. Each AV receives the complete information of prevailing traffic, including all other AVs' speeds and route choices.
2. Each AV anticipates future traffic on the predefined time horizon  $[0, T]$  where  $T > 0$ .
3. Each AV controls its driving speed in the interior of a link and next-go-to link choice at a junction to minimize a predefined travel cost functional over a network.
4. There is a vertical queue at each junction. AVs passing through a junction experience a queuing delay when the queue at that junction is non-empty.

5. The ideal DUE principle is employed to solve the traffic equilibrium, which is, the actual traffic information is used for decision-making. In other words, AVs make optimal choices according to what would actually happen in the future.

**Notation.** To present the mathematical formulation of the differential game, we will first introduce the notation that will be used in this paper. A transportation network is represented by a directed graph  $\mathcal{G} = \{\mathcal{N}, \mathcal{L}\}$  that includes a set of nodes  $\mathcal{N}$  and a set of links  $\mathcal{L} \subset \mathcal{N} \times \mathcal{N}$ . For each link  $l = (i, j) \in \mathcal{L}$  where  $i, j \in \mathcal{N}$ , denote its starting point as  $\text{START}(l) = i$  and its end point as  $\text{END}(l) = j$ . Accordingly, one can write  $l = (\text{START}(l), \text{END}(l))$  for each link  $l \in \mathcal{L}$ . We will use symbols  $l$  or  $(i, j)$  interchangeably to represent links in  $\mathcal{L}$ . The length of link  $l \in \mathcal{L}$  is denoted as  $\text{len}(l) \geq 0$ . For each node  $i \in \mathcal{N}$ , denote  $\text{IN}(i) \subset \mathcal{L}$  the set of links whose end point is node  $i$  and  $\text{OUT}(i) \subset \mathcal{L}$  the set of links whose starting point is node  $i$ . Denote  $\mathcal{N}_O \subset \mathcal{N}$  the set of origins and  $s \in \mathcal{N}$  the destination. The intermediate node set is  $\mathcal{N}_I = \mathcal{N} \setminus \mathcal{N}_O \setminus \{s\}$ .

We will first formulate one AV's hybrid optimal control problem on a network in Section 3.2. Then the optimal control problem will be extended to the  $N$ -car differential game in Section 3.3. Throughout this paper, we will use AVs or cars interchangeably to refer to the intelligent agents who navigate a transportation network with optimal driving strategies.

### 3.2. One AV's hybrid optimal control on a network

To formulate one AV's optimal control problem on a network, we first introduce the state and control variables in the interior of a link and at a node.

**State variable.** Consider the  $n^{\text{th}}$  car ( $n = 1, \dots, N$ ) driving on the network, the car's state is described in two scenarios:

- (i) The car is driving in the interior of a link at time  $t \in [0, T]$ . We denote the link the car is on as  $l^{(n)}(t)$  and parameterize the car's position on this link as a real number  $x^{(n)}(t) \in (0, \text{len}(l^{(n)}(t)))$ . In this case, the car's state is specified by:

$$X^{(n)}(t) = (x^{(n)}(t), l^{(n)}(t)) \in \mathbb{R} \times \mathcal{L}, \quad 0 < x^{(n)}(t) < \text{len}(l^{(n)}(t)). \quad (3.1)$$

- (ii) The car is in the queue at a node at time  $t \in [0, T]$ . We denote the node as  $i^{(n)}(t)$ . In this case, the car's state is specified by:

$$X^{(n)}(t) = i^{(n)}(t) \in \mathcal{N}. \quad (3.2)$$

**Control variable.** Correspondingly, the car has two types of decisions to make during its travel on the network: speed control and route choice.

- (i) The  $n^{\text{th}}$  car's driving speed is denoted as  $u^{(n)}(t)$  when the car is in the interior of any link at time  $t \in [0, T]$ . Its feasible region is  $u^{(n)}(t) \in [u_{\min}, u_{\max}]$  where  $u_{\min}$  and  $u_{\max}$  are the minimum and maximum speeds for all cars.
- (ii) The  $n^{\text{th}}$  car's next-go-to link choice is denoted as  $\gamma^{(n)}(i, t)$  when the car exits the queue at node  $i \in \mathcal{N}$  at time  $t \in [0, T]$ . Its feasible region is the outgoing link set  $\text{OUT}(i)$  of node  $i$ .

Summarizing these two types of decisions, we have a driving strategy vector:

$$\alpha^{(n)} = [u^{(n)}(\cdot), \gamma^{(n)}(\cdot, \cdot)], \quad n = 1, \dots, N. \quad (3.3)$$

Because the speed control is implemented as continuous-in-time driving speed profile while the route choice is implemented as discrete-in-time next-go-to link choices, the optimal control problem to solve these two types of decisions is a *hybrid optimal control problem*. With state and control variables defined, we formulate the AV's state dynamics under a control in Section 3.2.1 and define the AV's travel cost functional in Section 3.2.2.

#### 3.2.1. One AV's state dynamics

Knowing the control variables for the  $n^{\text{th}}$  car, we are ready to formulate a dynamical system that stipulates the evolution of the car's state under a hybrid control. There are two types of car state evolution: link dynamic and node transition.

**Link dynamic.** When the  $n^{\text{th}}$  car drives in the interior of a link  $l^{(n)}(t) \in \mathcal{L}$  at time  $t$ , its position  $x^{(n)}(t)$  changes continuously according to its controlled speed  $u^{(n)}(t)$ . The car's movement on a link is accordingly modeled by an ODE:  $dx^{(n)}(t)/dt = u^{(n)}(t)$ . Since the car keeps driving on the same link, the link index of the car's state remains the same and the dynamic of the link index becomes:  $dl^{(n)}(t)/dt = 0$ .

**Node transition.** When the car reaches the end point of the link it is driving on, which is  $x^{(n)}(t-) = \text{len}(l^{(n)}(t-))$ , a transition of state happens at time  $t$ . Here  $t-$  represents the left time limit. With the transition, the car moves to the node  $i^{(n)}(t) = \text{END}(l^{(n)}(t-))$  and enters the queue at node  $i^{(n)}(t)$ . We denote the car's exit time from the queue as  $e^{(n)}(t) \in [t, T]$ . During the time interval  $t \leq \tau < e^{(n)}(t)$ , the car's state keeps unchanged, i.e.,  $di^{(n)}(\tau)/d\tau = 0$ . Once the car exits the queue, it leaves the node at time  $e^{(n)}(t)$ . According to the car's next-go-to link choice, it switches to the link  $\gamma^{(n)}(i^{(n)}(t), e^{(n)}(t))$  and its initial position on the new link is reset to zero at time  $e^{(n)}(t)$ . We

denote the set of all time instants of node transition during the  $n^{\text{th}}$  car's trip over the network as  $R^{(n)} \subset [0, T]$ .

**Initial State.** Suppose the  $n^{\text{th}}$  car enters the network through the origin  $o^{(n)} \in \mathcal{N}_O$  at time  $t_0^{(n)} \geq 0$ , the car's initial state is given by  $i^{(n)}(t_0^{(n)}) = o^{(n)}$ . The car's initial state corresponds to its first node transition at time  $t_0^{(n)}$ . That is, the car switches to the link  $\gamma^{(n)}(o^{(n)}, e^{(n)}(t_0^{(n)}))$  when it exits the queue at the origin  $o^{(n)}$  at time  $e^{(n)}(t_0^{(n)})$ .

Summarizing both link dynamic and node transition for the  $n^{\text{th}}$  car ( $n = 1, \dots, N$ ), we now obtain its dynamical system on a network as:

$$(\text{Link dynamic}) \quad \frac{dx^{(n)}(t)}{dt} = u^{(n)}(t), \quad \frac{dl^{(n)}(t)}{dt} = 0, \quad \text{if } 0 < x^{(n)}(t) < \text{len}(l^{(n)}(t)); \quad (3.4a)$$

$$(\text{Node transition}) \quad i^{(n)}(t) = \text{END}(l^{(n)}(t-)), \quad t \in R^{(n)}, \quad \text{if } x^{(n)}(t-) = \text{len}(l^{(n)}(t-)); \quad (3.4b)$$

$$\frac{di^{(n)}(\tau)}{d\tau} = 0 \text{ for } t < \tau < e^{(n)}\left(t\right), \quad (3.4c)$$

$$x^{(n)}(e^{(n)}(t)) = 0, \quad l^{(n)}(e^{(n)}(t)) = \gamma^{(n)}(i^{(n)}(t), e^{(n)}(t)), \quad \text{if } t \in R^{(n)};$$

$$(\text{Initial state}) \quad i^{(n)}(t_0^{(n)}) = o^{(n)}, \quad t_0^{(n)} \in R^{(n)}. \quad (3.4d)$$

Eq. (3.4) is a hybrid dynamical system. The link dynamic is modeled by continuous ODEs, while the node transition happens only at discrete time instants in  $R^{(n)}$ . We denote  $R^{(n)} = \{t_0^{(n)}, t_1^{(n)}, \dots, t_{|R^{(n)}|-1}^{(n)}\}$  where  $|R^{(n)}|$  is the number of times of the  $n^{\text{th}}$  car's node transition and  $t_0^{(n)} < t_1^{(n)} < \dots < t_{|R^{(n)}|-1}^{(n)}$  are the transition time instants. By integrating the  $n^{\text{th}}$  car's dynamical system, we can obtain: (i) the car's trajectory  $x^{(n)}(t)$  in the interior of a link when  $e^{(n)}(t_k^{(n)}) \leq t < t_{k+1}^{(n)}$  for  $k = 0, \dots, |R^{(n)}| - 1$ ; (ii) the car's queuing delay  $e^{(n)}(t_k^{(n)}) - t_k^{(n)}$  at a node for  $k = 0, \dots, |R^{(n)}| - 1$ . If the  $n^{\text{th}}$  car arrives at the destination before time  $T$ , we denote this arrival time as  $t_f^{(n)} < T$ ; otherwise the car is still on the network at the end of the planning horizon  $[0, T]$  and we denote  $t_f^{(n)} = T$ . In both scenarios, the  $n^{\text{th}}$  car's actual travel time horizon on the network is  $[t_0^{(n)}, t_f^{(n)}]$ . Note that in the latter scenario, the car is either on a link or in the queue at a node at time  $t_f^{(n)} = T$ . If the car is in the queue at a node and its exit time  $e^{(n)}(t_{|R^{(n)}|-1}^{(n)}) > T$ , we will make the convention that  $e^{(n)}(t_{|R^{(n)}|-1}^{(n)}) = T$ . We will also make the convention that  $t_{|R^{(n)}|} = t_f^{(n)}$ .

To demonstrate one trajectory realization solved from the above dynamical system, we plot the  $n^{\text{th}}$  car's time-space trajectory (Fig. 1a) on the Braess network (Fig. 1b).

Suppose the car enters the network through the origin node 1 at time  $t_0^{(n)} = 0$  and travels along the path  $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$ . It is assumed that all links (1, 2), (2, 3), (3, 4) have the same length  $l$ . In the interior of a link, the position of the car, denoted as  $x$ , changes continuously with respect to time  $t$  under the influence of its speed control. Suppose the car keeps a constant speed on each link, its time-space trajectory consists of slanted straight lines. The time instants when the car arrives at a non-destination node are  $R^{(n)} = \{t_0^{(n)}, t_1^{(n)}, t_2^{(n)}\}$ . Suppose that the queues at node 1 and 2 are always empty but the queue at node 3 is non-empty at time  $t_2^{(n)}$ , the car exits those queues at time  $e(t_0^{(n)}) = t_0^{(n)} = 0$ ,  $e(t_1^{(n)}) = t_1^{(n)}$  and  $e(t_2^{(n)}) > t_2^{(n)}$ . Exiting a queue, the car switches to the next-go-to link and its position jumps back to zero on the new link. The car reaches the destination  $s = 4$  at time  $t_f^{(n)} < T$  and ends its trip there.

### 3.2.2. One AV's travel cost functional

To select the optimal driving speed and routing policy, the  $n^{\text{th}}$  car solves a hybrid optimal control problem by minimizing a pre-

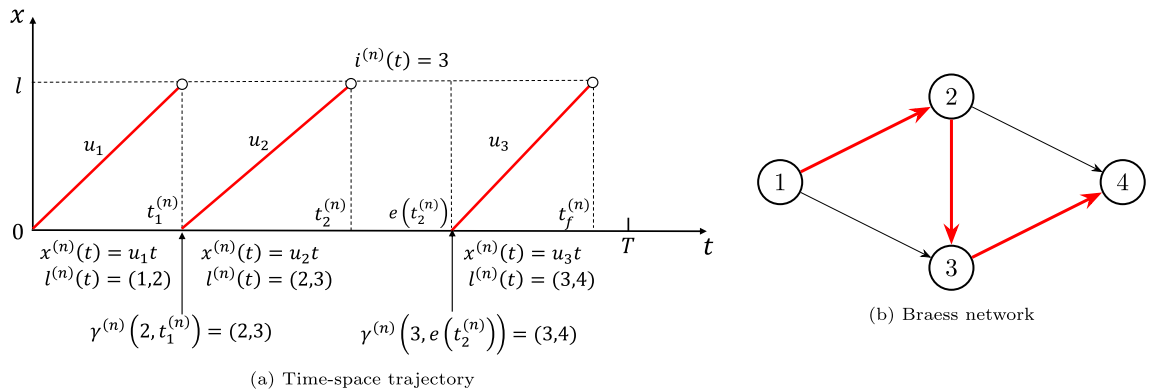


Fig. 1. One AV's state dynamics across the Braess network.

defined travel cost functional. The travel cost functional consists of three parts: link running cost, node queuing cost and terminal cost.

**Link running cost.** Within each time interval  $e^{(n)}(t_k^{(n)}) \leq t < t_{k+1}^{(n)}$  for  $k = 0, \dots, |R^{(n)}| - 1$ , the car drives in the interior of a link. We define the car's running cost on the link as

$$\int_{e^{(n)}(t_k^{(n)})}^{t_{k+1}^{(n)}} f_{\text{run}}^{(n)}(u^{(n)}(t), X^{(n)}(t)) dt,$$

which is an integral of the car's instantaneous running cost  $f_{\text{run}}^{(n)}(u^{(n)}(t), X^{(n)}(t))$  over the time interval. Here  $f_{\text{run}}^{(n)}(\cdot, \cdot)$  is the  $n^{\text{th}}$  car's *running cost function* that quantifies driving objectives such as efficiency and safety.

**Node queuing cost.** Within each time interval  $t_k^{(n)} \leq t < e^{(n)}(t_k^{(n)})$  for  $k = 0, \dots, |R^{(n)}| - 1$ , the car stays in the queue at a node. We define the car's queuing cost at the node as  $f_{\text{que}}^{(n)}(e^{(n)}(t_k^{(n)}) - t_k^{(n)})$ , in which  $e^{(n)}(t_k^{(n)}) - t_k^{(n)}$  is the car's queuing delay at the node. Here  $f_{\text{que}}^{(n)}(\cdot)$  is the  $n^{\text{th}}$  car's *queuing cost function* that quantifies the impact of queuing delay.

**Terminal cost.** The car terminates its trip at the state  $X^{(n)}(t_f^{(n)})$  at time  $t_f^{(n)}$ . We define the car's terminal cost as  $V_{\text{ter}}^{(n)}(t_f^{(n)}, X^{(n)}(t_f^{(n)}))$ . Here  $V_{\text{ter}}^{(n)}(\cdot, \cdot)$  is the  $n^{\text{th}}$  car's *terminal cost function* that represents the car's preference on its terminal state. There are three types of terminal state:

- (i) The car arrives at the destination  $s$  at time  $t_f^{(n)} < T$ . In this case the terminal cost is written as  $V_{\text{ter}}^{(n)}(t_f^{(n)}, s)$  and we will always assume  $V_{\text{ter}}^{(n)}(t_f^{(n)}, s) = 0$ .
- (ii) The car is in the interior of a link at time  $t_f^{(n)} = T$ . In this case the terminal cost is written as  $V_{\text{ter}}^{(n)}(T, x^{(n)}(T), l^{(n)}(T))$ , where  $l^{(n)}(T)$  is the link the car is on and  $x^{(n)}(T)$  is the car's position on the link at time  $T$ .
- (iii) The car is in the queue at a node at time  $t_f^{(n)} = T$ . In this case the terminal cost is written as  $V_{\text{ter}}^{(n)}(T, i^{(n)}(T))$ , where  $i^{(n)}(T)$  is the node at which the car is in the queue at time  $T$ .

Summing up the three types of costs along the  $n^{\text{th}}$  car's trip over the network, the general form of the  $n^{\text{th}}$  car's travel cost functional is defined as:

$$J_n \left( \alpha^{(n)} \right) = \sum_{k=0}^{|R^{(n)}|-1} \left[ \underbrace{\int_{e^{(n)}(t_k^{(n)})}^{t_{k+1}^{(n)}} f_{\text{run}}^{(n)}(u^{(n)}(t), X^{(n)}(t)) dt}_{\text{link running cost}} + \underbrace{f_{\text{que}}^{(n)}(e^{(n)}(t_k^{(n)}) - t_k^{(n)})}_{\text{node queuing cost}} \right] + \underbrace{V_{\text{ter}}^{(n)}(t_f^{(n)}, X^{(n)}(t_f^{(n)}))}_{\text{terminal cost}}. \quad (3.5)$$

Summarizing both state dynamics defined in Eq. (3.4) and the travel cost functional defined in Eq. (3.5), the  $n^{\text{th}}$  car's hybrid optimal control problem on a network is defined as:

$$\alpha^{*(n)} = \operatorname{argmin}_{\alpha^{(n)}} J_n(\alpha^{(n)}), \quad (3.6a)$$

$$\text{s.t.} \begin{cases} \frac{dx^{(n)}(t)}{dt} = u^{(n)}(t), \quad \frac{dl^{(n)}(t)}{dt} = 0, & \text{if } 0 < x^{(n)}(t) < \text{len}(l^{(n)}(t)); \\ i^{(n)}(t) = \text{END}(l^{(n)}(t-)), \quad t \in R^{(n)}, & \text{if } x^{(n)}(t-) = \text{len}(l^{(n)}(t-)); \\ \frac{di^{(n)}(\tau)}{d\tau} = 0 \text{ for } t < \tau < e^{(n)}(t), \quad x^{(n)}(e^{(n)}(t)) = 0, \quad l^{(n)}(e^{(n)}(t)) = \gamma^{(n)}(i^{(n)}(t), e^{(n)}(t)), & \text{if } t \in R^{(n)}; \\ i^{(n)}(t_0^{(n)}) = o^{(n)}, \quad t_0^{(n)} \in R^{(n)}. \end{cases} \quad (3.6b)$$

### 3.3. N-car differential game on a network

The  $n^{\text{th}}$  car ( $n = 1, \dots, N$ ) solves the optimal control problem defined by Eq. (3.6). Remember that there are  $N$  cars moving on the same transportation network. For any  $n = 1, 2, \dots, N$ , suppose the  $n^{\text{th}}$  car knows others' driving strategies:

$$\alpha^{(-n)} = [\alpha^{(1)}, \dots, \alpha^{(n-1)}, \alpha^{(n+1)}, \dots, \alpha^{(N)}]. \quad (3.7)$$

By integrating Eq. (3.6b) for all other cars under their driving strategies, the  $n^{\text{th}}$  car is able to predict the system state:

$$X^{(-n)}(t) = [X^{(1)}(t), \dots, X^{(n-1)}(t), X^{(n+1)}(t), \dots, X^{(N)}(t)], \quad (3.8)$$

on the planning horizon  $t \in [0, T]$ . We make the convention that Eq. (3.8) only includes the cars on the network at time  $t$ .



Due to interactions between cars, the  $n^{\text{th}}$  car's travel cost functional  $J_n(\cdot)$  defined in Eq. (3.5) now depends on others' states and actions. Accordingly, the  $n^{\text{th}}$  car's travel cost functional is reformulated as follows:

$$J_n \left( \alpha^{(n)}, \alpha^{(-n)} \right) = \sum_{k=0}^{|R^{(n)}|-1} \left[ \underbrace{\int_{e^{(n)}(t_k^{(n)})}^{t_{k+1}^{(n)}} f_{\text{run}}^{(n)}(u^{(n)}(t), X^{(n)}(t), X^{(-n)}(t)) dt}_{\text{link running cost}} + \underbrace{f_{\text{que}}^{(n)}(e^{(n)}(t_k^{(n)}), X^{(-n)}(t_k^{(n)})) - t_k^{(n)}}_{\text{node queuing cost}} \right] + \underbrace{V_{\text{ter}}^{(n)}(t_f^{(n)}, X^{(n)}(t_f^{(n)}))}_{\text{terminal cost}}, \quad (3.9)$$

where the  $n^{\text{th}}$  car's running cost function  $f_{\text{run}}^{(n)}(\cdot)$  and exit time function  $e^{(n)}(\cdot)$  also depend on other cars' states.

A Nash equilibrium of the game is a tuple of controls  $\alpha^* = [\alpha^{*(1)}, \alpha^{*(2)}, \dots, \alpha^{*(N)}]$  satisfying:

$$J_n^N(\alpha^{*(n)}, \alpha^{*(-n)}) \leq J_n^N(\alpha^{(n)}, \alpha^{*(-n)}), \quad \forall \alpha^{(n)}, \quad n = 1, \dots, N. \quad (3.10)$$

At equilibrium, no car can improve its travel cost by unilaterally switching its driving speed or routing policy.

When  $N$  cars interact with one another while moving through a transportation network, we need to solve a total of  $N$  optimal control problems that are coupled through states of all cars. This brings challenges in solving the equilibrium of the  $N$ -car differential game as  $N$  becomes large (Cardaliaguet, 2010). To simplify this system, this paper will develop a scalable framework to solve an approximate equilibrium by resorting to the mean field approximation.

#### 4. Mean field driving and routing game on a network

In this section we will derive a mean field driving and routing game from the  $N$ -car differential game on a network. This derivation follows from the approach in Huang et al. (2020a) and the derived mean field game can be viewed as the continuum limit of the  $N$ -car differential game as the number of cars  $N \rightarrow \infty$ .

We will use the *mean field approximation* to bridge between the discrete differential game and continuum mean field game. To apply the mean field approximation, we need the following assumptions:

- (A1) All AVs are homogeneous in the sense that they have the same form of travel cost functional.
- (A2) All AVs are indistinguishable on the road.
- (A3) Each AV's instantaneous link running cost only depends on its driving speed and nearby traffic density; Each AV's instantaneous node queuing cost only depends on its queuing delay.
- (A4) Each AV stays in a queue until all the AVs entering the queue before him leave the queue, i.e., the first-in-first-out (FIFO) principle.
- (A5) The queue at each node has a bottleneck capacity. The outgoing flow rate from the queue reaches this capacity as long as the queue is non-empty.
- (A6) AVs implement their next-go-to link choices in a mixed strategy.

The mean field approximation is applied on both the AVs' speed control and routing policy. The derived mean field game is formulated as the coupled PDE system of a Hamilton–Jacobi–Bellman (HJB) equation and a continuity equation on a network.

To present how a MFG on a network is formulated, we will first apply the mean field approximation to the  $N$ -car differential game and derive a generic car's optimal control problem in Section 4.1. The HJB equation will be derived from the generic car's optimal control problem in Section 4.2 and the continuity equation will be derived from the aggregate traffic evolution of all cars' dynamics in Section 4.3. The derived MFG system is presented in Section 4.4 and reformulated as a mixed complementarity problem (MiCP) in Section 4.5. The problem can be solved numerically on discrete space–time grids, which will be discussed in Section 6. Fig. 2 shows the whole diagram and illustrates connection between different problem formulations.

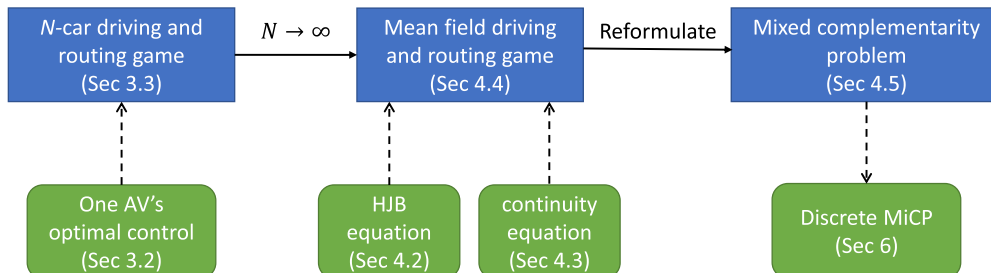


Fig. 2. Connection between different problem formulations.

#### 4.1. Mean field approximation

In this section, we will apply the mean field approximation to the  $N$ -car differential game introduced in Section 3.3 and derive a generic car's optimal control problem. The two steps are:

1. Instead of tracking a total of  $N$  states for every car, we define an aggregate traffic state on the network using the mean field approximation.
2. Instead of solving a total of  $N$  coupled optimal control problems for every car, we derive a generic car's optimal control problem by reformulating the travel cost functional defined in Eq. (3.9) and the equilibrium condition defined in Eq. (3.10) under the aggregate traffic state.

**Mean field approximation.** By the assumption (A2) that all AVs are indistinguishable, each AV only sees the aggregate distribution of all other AVs but cannot distinguish their identity numbers. As a corollary, each AV's travel cost functional only depends on the distribution of all other AVs on the network. When  $N$  is very large, this distribution is well approximated by:

- (i) The link traffic density  $\rho_l(x, t)$ , where  $l \in \mathcal{L}$  is the link index,  $x \in (0, \text{len}(l))$  and  $t \in [0, T]$  are the space and time coordinates;
- (ii) The node queue size  $Q_i(t)$ , where  $i \in \mathcal{N} \setminus \{s\}$  is the node index,  $t \in [0, T]$  is the time coordinate.

We will define the aggregate traffic state on the network as the whole of all link traffic densities and node queue sizes. In other words, the traffic state represents a "mean field" of all AVs on the network. The complex interactions between individual AVs are then decoupled into two components: a generic car's optimal control under the traffic state, and the evolution of the traffic state from all cars' movements.

##### A generic car's optimal control.

In the  $N$ -car differential game, each AV aims to select its optimal control to minimize its travel cost functional defined in Eq. (3.9) in the existence of all other AVs. By the assumption (A1) that all AVs have the same form of travel cost functional, we can consider a generic car's optimal control problem under the traffic state. Let us drop the car index  $n$  for all functions and variables in Eq. (3.9). Then we reformulate the travel cost functional defined in Eq. (3.9) under the traffic state as follows:

1. By the assumption (A3), we can replace the arguments  $X^{(n)}(t), X^{(-n)}(t)$  in the generic car's running cost function  $f_{\text{run}}(\cdot)$  by the local traffic density  $\rho_{l(t)}(x(t), t)$ .
2. By the assumption (A4), the generic car entering the queue at node  $i(t)$  at time  $t$  exits the queue at the time when all cars existed in the queue at time  $t$  exit. By the assumption (A5), the total time for those cars to exit is  $Q_{i(t)}(t)/M_{i(t)}$ , where  $M_{i(t)} > 0$  denotes the bottleneck capacity of the queue at node  $i(t)$ , and the generic car's exit time is  $e(t) = t + Q_{i(t)}(t)/M_{i(t)}$ . When  $t + Q_{i(t)}(t)/M_{i(t)} > T$ , we make the convention that  $e(t) = T$ . Hence the generic car's exit time function is given by:

$$e(t) = \min \left\{ t + \frac{Q_{i(t)}(t)}{M_{i(t)}}, T \right\}. \quad (4.1)$$

Denote  $\alpha = [u(\cdot), \gamma(\cdot, \cdot)]$  as the generic car's hybrid control. It is composed of the speed control  $u(t)$  for  $t \in [t_0, t_f]$  and the routing policy  $\gamma(i, t)$  for  $i \in \mathcal{N}$  and  $t \in [t_0, t_f]$ . The generic car's optimal control problem under the traffic state is formulated as:

$$\begin{aligned} \alpha^* &= \operatorname{argmin}_{\alpha} J(\alpha), \\ &= \operatorname{argmin}_{\alpha} \sum_{k=0}^{|R|-1} \left[ \underbrace{\int_{e(t_k)}^{t_{k+1}} f_{\text{run}}(u(t), \rho_{l(t)}(x(t), t)) dt}_{\text{link running cost}} + \underbrace{f_{\text{que}}(e(t_k) - t_k)}_{\text{node queuing cost}} \right] + \underbrace{V_{\text{ter}}(t_f, X(t_f))}_{\text{terminal cost}}, \end{aligned} \quad (4.2a)$$

$$\text{s.t.} \begin{cases} \frac{dx(t)}{dt} = u(t), \quad \frac{dl(t)}{dt} = 0, & \text{if } 0 < x(t) < \text{len}(l(t)); \\ i(t) = \text{END}(l(t-)), \quad t \in R, & \text{if } x(t-) = \text{len}(l(t-)); \\ \frac{di(\tau)}{d\tau} = 0 \text{ for } t < \tau < e(t), \quad x(e(t)) = 0, \quad l(e(t)) = \gamma(i(t), e(t)), & \text{if } t \in R; \\ i(t_0) = o, \quad t_0 \in R. \end{cases} \quad (4.2b)$$

where:

$[t_0, t_f]$  is the generic car's travel time horizon on the network;

$R = \{t_0, t_1, \dots, t_{|R|-1}\}$  is the set of all time instants of node transition of the generic car, and we make the convention that  $t_{|R|} = t_f$ ;



$e(t)$  is the generic car's exit time function under the traffic state, which is defined by Eq. (4.1);  
 $f_{\text{run}}(\cdot, \cdot)$ ,  $f_{\text{que}}(\cdot)$  and  $V_{\text{ter}}(\cdot, \cdot)$  are the generic car's running, queuing and terminal cost functions, respectively.

Given the traffic state on the network in the planning horizon  $[0, T]$ , we can solve a generic car's optimal control problem defined in Eq. (4.2) with any initial condition. When all cars follow their individual optimal controls, the traffic state evolves as the aggregate behavior of all cars' dynamics. Assuming the traffic state is known, one HJB equation is derived from all cars' optimal controls. Assuming all cars' optimal controls are known, one continuity equation is derived from their dynamic motions. The mean field game (MFG) is then formulated as the coupled system of these two equations.

#### 4.2. Backward Hamilton–Jacobi–Bellman (HJB) equation

With a generic car's optimal control problem introduced, now we move on to derive the HJB equation. Suppose that: (i) the link traffic density  $\rho_l(x, t)$  is known for all  $l \in \mathcal{L}$ ,  $x \in (0, \text{len}(l))$  and  $t \in [0, T]$ ; (ii) the queue size  $Q_i(t)$  is known for all  $i \in \mathcal{N}$  and  $t \in [0, T]$ . The HJB equation determines a set of optimal velocity fields and next-go-to link distributions characterizing a generic car's optimal control.

Our derivation is based on the dynamic programming principle. That is, we consider the following subproblem of the optimal control problem defined in Eq. (4.2): suppose the generic car starts from any state  $X(t^*)$  at any time  $t^* \in [t_0, t_f]$  and aims to minimize its travel cost during  $t \in [t^*, t_f]$ , what is the generic car's optimal control strategy and optimal cost value at time  $t^*$ ?

We will solve the above subproblem in three cases: speed choice on a link, route choice at a node, and the terminal state, which give the HJB equation on a link, its spatial boundary condition at a node, and its terminal condition at the terminal state, respectively.

**Speed choice on a link.** Suppose  $X(t^*) = (x^*, l^*)$  where  $0 < x^* < \text{len}(l^*)$  and  $0 < t^* < T$ . That is, the car is driving in the interior of a link. In this case the car needs to select its optimal speed at time  $t^*$ . We define  $V_{l^*}(x^*, t^*)$  to be the optimal cost value of the car's trip from the state  $(x^*, l^*)$  at time  $t^*$ .

Denote the car's speed at time  $t^*$  as  $u \in [u_{\min}, u_{\max}]$ . We choose a small time step  $\Delta t$  such that the car arrives at position  $x^* + u\Delta t < \text{len}(l^*)$  on link  $l^*$  at time  $t^* + \Delta t < T$ . The car's decision process is then decomposed into two stages: during the time interval  $[t^*, t^* + \Delta t]$  the car's link running cost is  $f_{\text{run}}(u, \rho_{l^*}(x^*, t^*))\Delta t$ ; after time  $t^* + \Delta t$ , the car selects optimal driving speed and routing policy that give the optimal cost value  $V_{l^*}(x^* + u\Delta t, t^* + \Delta t)$ . Summarizing the costs at these two stages, we have:

$$V_{l^*}(x^*, t^*) = \min_{u_{\min} \leq u \leq u_{\max}} \{V_{l^*}(x^* + u\Delta t, t^* + \Delta t) + f_{\text{run}}(u, \rho_{l^*}(x^*, t^*))\Delta t\}. \quad (4.3)$$

Applying the first order Taylor's expansion on  $V_{l^*}(x^* + u\Delta t, t^* + \Delta t)$  and letting  $\Delta t \rightarrow 0$ , we obtain the following HJB equation from Eq. (4.3):

$$\partial_x V_{l^*}(x^*, t^*) + \min_{u_{\min} \leq u \leq u_{\max}} \{u \partial_x V_{l^*}(x^*, t^*) + f_{\text{run}}(u, \rho_{l^*}(x^*, t^*))\} = 0. \quad (4.4)$$

See Huang et al. (2020a) for more details. We denote the car's optimal speed as  $u_{l^*}(x^*, t^*)$ , it is given by the minimizer in Eq. (4.4):

$$u_{l^*}(x^*, t^*) = \text{argmin}_{u_{\min} \leq u \leq u_{\max}} \{u \partial_x V_{l^*}(x^*, t^*) + f_{\text{run}}(u, \rho_{l^*}(x^*, t^*))\}. \quad (4.5)$$

Eqs. (4.4) (4.5) are defined on all links of the network during the planning horizon  $(0, T)$ . We will replace  $l^*$ ,  $x^*$  and  $t^*$  by the general link index  $l \in \mathcal{L}$ , space coordinate  $x \in (0, \text{len}(l))$  and time coordinate  $t \in (0, T)$ .

**Route choice at a node.** Suppose  $X(t^*) = i^*$  and  $t^* \in R$ , where  $i^* \in \mathcal{N} \setminus \{s\}$  and  $0 < t^* < T$ . That is, the car enters the queue at node  $i^*$  at time  $t^*$ . In this case, the car exits the queue at time  $e(t^*) = \min\{t^* + Q_{i^*}(t^*)/M_{i^*}, T\}$  and needs to select its next-go-to link at time  $e(t^*)$ . We define  $\pi_{i^*}(e(t^*))$  to be the optimal cost value of the car's trip after exiting the queue at node  $i^*$  at time  $e(t^*)$ .

By the assumption (A6), the next-go-to link selection is implemented in a mixed strategy. Remember that  $\gamma(i^*, e(t^*)) \in \text{OUT}(i^*)$  refers to the car's next-go-to link choice at node  $i^*$  at time  $e(t^*)$ . But in the mixed strategy, there exists a probability distribution over all feasible choices. We define  $\beta(i^*, l, e(t^*)) \in [0, 1]$  as the probability that the car selects the next-go-to link  $l$  for all  $l \in \text{OUT}(i^*)$ . The probability distribution should satisfy:

$$\sum_{l \in \text{OUT}(i^*)} \beta(i^*, l, e(t^*)) = 1. \quad (4.6)$$

Once the car selects a link  $l \in \text{OUT}(i^*)$ , it moves to the starting point of the link at time  $e(t^*)$  and its optimal cost value is  $V_l(0, e(t^*))$ . Taking the minimum over all  $l \in \text{OUT}(i^*)$ , the car's optimal cost value at the node is given by:

$$\pi_{i^*}(e(t^*)) = \min_{l \in \text{OUT}(i^*)} V_l(0, e(t^*)). \quad (4.7)$$

The equilibrium condition is defined as:

$$\beta(i^*, l, e(t^*)) \geq 0, \quad \pi_{i^*}(e(t^*)) = V_l(0, e(t^*)) \text{ if } \beta(i^*, l, e(t^*)) > 0, \quad l \in \text{OUT}(i^*). \quad (4.8)$$

That is, any used link  $l \in \text{OUT}(i^*)$  is the optimal choice. This is consistent with the classical dynamic user equilibrium definition. Eqs. (4.6)–(4.8) are defined on all non-destination nodes of the network during the planning horizon  $(0, T)$ . We will replace  $i^*$  and  $e(t^*)$  by the general node index  $i \in \mathcal{N} \setminus \{s\}$  and time coordinate  $t \in (0, T)$ .

When  $i^*$  is an intermediate node, the car moves to node  $i^*$  from the end point of some link  $h \in \text{IN}(i^*)$  at time  $t^*$ . With the existence of queuing delay at node  $i^*$ , the car's optimal cost value at the end point of link  $h$  at time  $t^*$ , which is  $V_h(\text{len}(h), t^*)$ , is the nodal optimal cost value  $\pi_{i^*}(e(t^*))$  plus the car's queuing cost  $f_{\text{que}}(e(t^*) - t^*)$ . Note that  $e(t^*) = \min\{t^* + Q_{i^*}(t^*)/M_{i^*}, T\}$ , we obtain:

$$V_h(\text{len}(h), t^*) = \pi_{i^*}\left(\min\left\{t^* + \frac{Q_{i^*}(t^*)}{M_{i^*}}, T\right\}\right) + f_{\text{que}}\left(\min\left\{\frac{Q_{i^*}(t^*)}{M_{i^*}}, T - t^*\right\}\right). \quad (4.9)$$

Eq. (4.9) is defined on all intermediate nodes of the network during the planning horizon  $(0, T)$ . We will replace  $i^*$  and  $t^*$  by the general node index  $i \in \mathcal{N} \setminus \{s\}$  and time coordinate  $t \in (0, T)$ .

**Terminal state.** When  $t^* = t_f$ , the car terminates its trip over the network at time  $t^*$ . There are three cases:

- (i) The car arrives at the destination  $s$  at time  $t^* < T$ . In this case the car's terminal cost gives the optimal cost value at the destination:

$$V_l(\text{len}(l), t^*) = V_{\text{ter}}(t^*, s) = 0, \quad l \in \text{IN}(s). \quad (4.10)$$

- (ii) The car is in the interior of a link at time  $t^* = T$ . In this case the car's terminal cost gives the optimal cost value on a link:

$$V_{l^*}(x^*, T) = V_{\text{ter}}(T, x^*, l^*), \quad (4.11)$$

where  $l^*$  is the link the car is on and  $x^*$  is the car's position on the link at time  $t^* = T$ .

- (iii) The car is in the queue at a non-destination node at time  $t^* = T$ . In this case the car's terminal cost gives the optimal cost value at a node:

$$\pi_{i^*}(T) = V_{\text{ter}}(T, i^*), \quad (4.12)$$

where  $i^*$  is the node at which the car is in the queue at time  $t^* = T$ .

Summarizing the three cases, the terminal conditions of the HJB equation include: (i)  $V_l(\text{len}(l), t) = 0$  for all  $l \in \text{IN}(s)$  and  $t \in (0, T)$ ; (ii)  $V_l(x, T)$  for all  $l \in \mathcal{L}$  and  $x \in (0, \text{len}(l))$ ; (iii)  $\pi_i(T)$  for all  $i \in \mathcal{N} \setminus \{s\}$ . We may choose all  $V_l(x, T)$  and  $\pi_i(T)$  to be the same constant if we do not care about the cars' final positions. Or we can properly define those terminal costs to penalize cars who cannot arrive at the destination in time.

The HJB equation defined by Eqs. (4.4)–(4.9) characterizes a generic car's optimal speed control and routing policy. The optimal velocity fields  $u_l(x, t)$  for  $l \in \mathcal{L}$ ,  $x \in (0, \text{len}(l))$ ,  $t \in (0, T)$  and next-go-to link distributions  $\beta(i, l, t)$  for  $i \in \mathcal{N} \setminus \{s\}$ ,  $l \in \text{OUT}(i)$ ,  $t \in (0, T)$  are solved from the HJB equation, provided the link traffic densities and node queue sizes. The HJB equation is solved backward in time.

### 4.3. Forward continuity equation

When all cars follow the optimal speed control and routing policy, the traffic evolution on the network follows from their dynamic motions. Suppose that: (i) the velocity field  $u_l(x, t)$  is known for all  $l \in \mathcal{L}$ ,  $x \in (0, \text{len}(l))$  and  $t \in (0, T)$ ; (ii) the next-go-to link distribution  $\beta(i, l, t)$  is known for all  $i \in \mathcal{N} \setminus \{s\}$ ,  $l \in \text{OUT}(i)$  and  $t \in (0, T)$ . We derive the continuity equation characterizing the evolution of traffic state on the network.

Our derivation is based on the conservation of cars for both link flow propagation and node flow assignment. We will also introduce the initial condition of the continuity equation.

**Link flow propagation.** In the interior of any link  $l \in \mathcal{L}$ , the conservation of cars is described by the following partial differential equation of traffic density  $\rho_l(x, t)$  and flux  $q_l(x, t)$ :

$$\partial_t \rho_l(x, t) + \partial_x q_l(x, t) = 0, \quad x \in (0, \text{len}(l)), \quad t \in (0, T). \quad (4.13)$$

The flux  $q_l(x, t)$  is computed from traffic density  $\rho_l(x, t)$  and velocity field  $u_l(x, t)$  as follows:

$$q_l(x, t) = \rho_l(x, t) u_l(x, t), \quad x \in (0, \text{len}(l)], \quad t \in (0, T). \quad (4.14)$$

Eqs. (4.13)–(4.14) give the continuity equation on a link.

**Node flow assignment.** At any node  $i \in \mathcal{N} \setminus \{s\}$ , the conservation of cars is described by the relation:

$$\text{rate of change of queue size} = \text{total incoming flow} - \text{total outgoing flow}.$$

Based on the relation, we are going to compute the total outgoing flow from the rate of change of queue size and the total incoming flow.

Knowing the queue size  $Q_i(t)$  at node  $i$  for  $t \in [0, T]$ , the rate of change of the queue size is given by  $dQ_i(t)/dt$ . The total incoming flow is composed of two parts: the sum of flows from the node's all incoming links and the traffic demand if the node is an origin. The contribution from the former part is given by  $\sum_{h \in \text{IN}(i)} q_h(\text{len}(h), t)$ ; For the latter part, we denote  $d_i(t)$  the traffic demand on node  $i$  at time  $t$  and make the convention that  $d_i(t) = 0$  for all  $t \in [0, T]$  if  $i \notin \mathcal{N}_O$ . Then the total incoming flow at node  $i$  at time  $t$  is given by  $\sum_{h \in \text{IN}(i)} q_h(\text{len}(h), t) + d_i(t)$ . Knowing the rate of change of queue size and the total incoming flow, the total outgoing flow at node  $i$  at time  $t$  is given by  $\sum_{h \in \text{IN}(i)} q_h(\text{len}(h), t) + d_i(t) - \frac{dQ_i(t)}{dt}$ .

The total outgoing flow distributes to the node's all outgoing links. Remember that  $\beta(i, l, t), l \in \text{OUT}(i)$  is the probability distribution over a single car's next-go-to link choices from node  $i$  at time  $t$ . With a large number of cars,  $\beta(i, l, t)$  can be understood as the percentage of cars selecting the outgoing link  $l$  among all cars arriving at node  $i$  at time  $t$ . According to the outgoing flow distribution, we obtain:

$$q_l(0, t) = \beta(i, l, t) \left( \sum_{h \in \text{IN}(i)} q_h(\text{len}(h), t) + d_i(t) - \frac{dQ_i(t)}{dt} \right), \quad l \in \text{OUT}(i), \quad t \in (0, T). \quad (4.15)$$

By the assumption (A5), the total outgoing flow is no greater than  $M_i$  and it equals  $M_i$  as long as  $Q_i(t) > 0$ , which gives the following equations characterizing the evolution of the queue size:

$$\sum_{h \in \text{IN}(i)} q_h(\text{len}(h), t) + d_i(t) - \frac{dQ_i(t)}{dt} \leq M_i, \quad Q_i(t) \geq 0, \quad t \in (0, T); \quad (4.16)$$

$$\sum_{h \in \text{IN}(i)} q_h(\text{len}(h), t) + d_i(t) - \frac{dQ_i(t)}{dt} = M_i, \quad \text{if } Q_i(t) > 0, \quad t \in (0, T). \quad (4.17)$$

Eqs. (4.15)–(4.17) incorporate the Vickrey junction model (Vickrey, 1969) and give the spatial boundary condition of the continuity equation.

**Initial condition.** The network is empty when  $t = 0$ . Hence the initial condition of the continuity equation is given by:

$$\rho_l(x, 0) = 0, \quad l \in \mathcal{L}, \quad x \in (0, \text{len}(l)); \quad (4.18)$$

$$Q_i(0) = 0, \quad i \in \mathcal{N}. \quad (4.19)$$

The continuity equation defined by Eqs. (4.13)–(4.17) characterizes the evolution of traffic state on the network. The link traffic densities  $\rho_l(x, t)$  for  $l \in \mathcal{L}, x \in (0, \text{len}(l)), t \in (0, T)$  and the node queue sizes  $Q_i(t)$  for  $i \in \mathcal{N}, t \in (0, T)$  are solved from the continuity equation, provided the optimal velocity fields and next-go-to link distributions. The continuity equation is solved forward in time.

#### 4.4. Mean field game system

Summarizing all the Eqs. (4.4)–(4.9), (4.13)–(4.17), we obtain the following mean field game system on a network.

[MFGnet]

$$(\text{CE} - \text{Link}) \quad \partial_t \rho_l(x, t) + \partial_x q_l(x, t) = 0, \quad l \in \mathcal{L}; \quad (4.20a)$$

$$q_l(x, t) = \rho_l(x, t) u_l(x, t), \quad l \in \mathcal{L}; \quad (4.20b)$$

$$(\text{CE} - \text{Node}) \quad q_l(0, t) = \beta(i, l, t) \left( \sum_{h \in \text{IN}(i)} q_h(\text{len}(h), t) + d_i(t) - \frac{dQ_i(t)}{dt} \right), \quad i \in \mathcal{N} \setminus \{s\}, \quad l \in \text{OUT}(i); \quad (4.20c)$$

$$\sum_{h \in \text{IN}(i)} q_h(\text{len}(h), t) + d_i(t) - \frac{dQ_i(t)}{dt} \leq M_i, \quad Q_i(t) \geq 0 \quad i \in \mathcal{N} \setminus \{s\}; \quad (4.20d)$$

$$\sum_{h \in \text{IN}(i)} q_h(\text{len}(h), t) + d_i(t) - \frac{dQ_i(t)}{dt} = M_i, \quad \text{if } Q_i(t) > 0, \quad i \in \mathcal{N} \setminus \{s\}; \quad (4.20e)$$

$$(\text{HJB} - \text{Link}) \quad \partial_t V_l(x, t) + \min_{u_{\min} \leq u \leq u_{\max}} \{u \partial_x V_l(x, t) + f_{\text{run}}(u, \rho_l(x, t))\} = 0, \quad l \in \mathcal{L}; \quad (4.20f)$$

$$u_l(x, t) = \arg \min_{u_{\min} \leq u \leq u_{\max}} \{u \partial_x V_l(x, t) + f_{\text{run}}(u, \rho_l(x, t))\}, \quad l \in \mathcal{L}; \quad (4.20g)$$

$$(\text{HJB} - \text{Node}) \quad V_h(\text{len}(h), t) = \pi_i \left( \min \left\{ t + \frac{Q_i(t)}{M_i}, T \right\} \right) + f_{\text{que}} \left( \min \left\{ \frac{Q_i(t)}{M_i}, T - t \right\} \right), \quad i \in \mathcal{N}_I, \quad h \in \text{IN}(i); \quad (4.20h)$$

$$\sum_{l \in \text{OUT}(i)} \beta(i, l, t) = 1, \quad i \in \mathcal{N} \setminus \{s\}; \quad (4.20i)$$

$$\pi_i(t) = \min_{l \in \text{OUT}(i)} V_l(0, t), \quad i \in \mathcal{N} \setminus \{s\}; \quad (4.20j)$$

$$\beta(i, l, t) \geq 0, \quad \pi_i(t) = V_l(0, t) \text{ if } \beta(i, l, t) > 0, \quad i \in \mathcal{N} \setminus \{s\}, \quad l \in \text{OUT}(i). \quad (4.20k)$$

[MFGnet] is a coupled system of the forward continuity equation and the backward HJB equation defined on a network. The following conditions are predetermined for [MFGnet]:

- The initial traffic densities  $\rho_l(x, 0) = 0$  for all  $l \in \mathcal{L}$  and  $x \in (0, \text{len}(l))$ ;
- The initial queue sizes  $Q_i(0) = 0$  for all  $i \in \mathcal{N} \setminus \{s\}$ ;
- The terminal costs at the destination  $V_l(\text{len}(l), t) = 0$  for all  $l \in \text{IN}(s)$  and  $t \in (0, T)$ ;

and the following conditions need to be specified:

- The traffic demands  $d_i(t)$  for all  $i \in \mathcal{N}_O$  and  $t \in [0, T]$ .
- The terminal costs at the final time  $V_l(x, T)$  for all  $l \in \mathcal{L}$  and  $x \in (0, \text{len}(l))$ , and  $\pi_i(T)$  for all  $i \in \mathcal{N} \setminus \{s\}$ .

The mean field equilibrium condition is defined as the following: the evolution of link traffic densities and node queue sizes resulting from the optimal velocity fields and route choices is consistent with the actual traffic evolution. The solution of [MFGnet], which we will refer to as SOL([MFGnet]), gives the mean field equilibrium. The mean field equilibrium contains the following:

- $\rho_l(x, t)$ ,  $u_l(x, t)$  and  $V_l(x, t)$  for all  $l \in \mathcal{L}$ ,  $x \in (0, \text{len}(l))$  and  $t \in [0, T]$ ;
- $\pi_i(t)$  and  $Q_i(t)$  for all  $i \in \mathcal{N} \setminus \{s\}$  and  $t \in [0, T]$ ;
- $\beta(i, l, t)$  for all  $i \in \mathcal{N} \setminus \{s\}$ ,  $l \in \text{OUT}(i)$  and  $t \in [0, T]$ .

The existence of equilibrium solutions depends on both the network topology and cost functions. The following proposition gives a sufficient condition on the solution existence of [MFGnet].

**Proposition 4.1.** *Suppose the network contains only one link, [MFGnet] becomes a mean field game system defined on a single road with only the speed control. In this case, suppose the AVs' running cost function  $f_{\text{run}}(u, \rho) = g(u) + h(\rho)$ , where  $g(u)$  is a strictly convex function of speed  $u$ ,  $h(\rho)$  is a strictly increasing function of density  $\rho$ , and  $g(u)$  and  $h(\rho)$  satisfy certain growth conditions. Then there exists a unique weak solution of [MFGnet] (Cardaliaguet, 2015).*

#### 4.5. Reformulation as a mixed complementarity problem

To facilitate the algorithm development and give meaning to each equation, in this subsection, we will reformulate [MFGnet] as a mixed complementarity problem.

First, the Vickrey Eqs. (4.20d)(4.20e) can be reformulated as the following linear complementarity relation (Ban et al., 2012):

$$0 \leq \frac{dQ_i(t)}{dt} + M_i - \sum_{h \in \text{IN}(i)} q_h \left( \text{len}(h), t \right) - d_i(t) \perp Q_i(t) \geq 0. \quad (4.21)$$

Then let us reformulate the Eq. (4.20i) as a complementarity relation. We will choose all cost functions to be positive, then the cars' travel costs are always positive and we have  $\pi_i(t) > 0$  for all  $i \in \mathcal{N} \setminus \{s\}$  and  $t \in [0, T]$ . In this case, Eq. (4.20i) is equivalent to the following complementarity relation:

$$0 \leq \pi_i(t) \perp \sum_{l \in \text{OUT}(i)} \beta(i, l, t) - 1 \geq 0, \quad (4.22)$$

because the right hand side term  $\sum_{l \in \text{OUT}(i)} \beta(i, l, t) - 1$  must be zero when  $\pi_i(t) > 0$ .

The equilibrium condition defined by Eqs. (4.20j)(4.20k) can be rewritten as the following complementarity relation:

$$0 \leq \beta(i, l, t) \perp V_l(0, t) - \pi_i(t) \geq 0, \quad \forall l \in \text{OUT}(i). \quad (4.23)$$

Using Eqs. (4.21)–(4.23), [MFGnet] can be reformulated as a new system, denoted as [MFG-MiCP], defined in Eq. (4.24) below. It is a mixed complementarity problem (MiCP) composed of partial differential equations, algebraic equations, and complementarity relations (Facchinei and Pang, 2007; Ban et al., 2008; Ban et al., 2012).

**[MFG-MiCP]**

$$[\text{link flow balance}] \quad \partial_t \rho_l(x, t) + \partial_x q_l(x, t) = 0, \quad l \in \mathcal{L}; \quad (4.24a)$$

$$[\text{link flow propagation}] \quad q_l(x, t) = \rho_l(x, t) u_l(x, t), \quad l \in \mathcal{L}; \quad (4.24b)$$

$$[\text{link influx}] \quad q_l(0, t) = \beta(i, l, t) \left( \sum_{h \in \text{IN}(i)} q_h(\text{len}(h), t) + d_i(t) - \frac{dQ_i(t)}{dt} \right), \quad (4.24c)$$

$$i \in \mathcal{N} \setminus \{s\}, \quad l \in \text{OUT}(i);$$

$$[\text{nodal delay}] \quad 0 \leq \frac{dQ_i(t)}{dt} + M_i - \sum_{h \in \text{IN}(i)} q_h(\text{len}(h), t) - d_i(t) \perp Q_i(t) \geq 0, \quad i \in \mathcal{N} \setminus \{s\}; \quad (4.24d)$$

$$[\text{optimal link cost}] \quad \partial_t V_l(x, t) + \min_{u_{\min} \leq u \leq u_{\max}} \{u \partial_x V_l(x, t) + f_{\text{run}}(u, \rho_l(x, t))\} = 0, \quad l \in \mathcal{L}; \quad (4.24e)$$

$$[\text{optimal speed}] \quad u_l(x, t) = \arg \min_{u_{\min} \leq u \leq u_{\max}} \{u \partial_x V_l(x, t) + f_{\text{run}}(u, \rho_l(x, t))\}, \quad l \in \mathcal{L}; \quad (4.24f)$$

$$[\text{optimal nodal cost}] \quad V_h(\text{len}(h), t) = \pi_i \left( \min \left\{ t + \frac{Q_i(t)}{M_i}, T \right\} \right) + f_{\text{que}} \left( \min \left\{ \frac{Q_i(t)}{M_i}, T - t \right\} \right), \quad (4.24g)$$

$$i \in \mathcal{N}, \quad h \in \text{IN}(i);$$

$$[\text{nodal flow conservation}] \quad 0 \leq \pi_i(t) \perp \sum_{l \in \text{OUT}(i)} \beta(i, l, t) - 1 \geq 0, \quad i \in \mathcal{N} \setminus \{s\}; \quad (4.24h)$$

$$[\text{equilibrium turning ratio}] \quad 0 \leq \beta(i, l, t) \perp V_l(0, t) - \pi_i(t) \geq 0, \quad i \in \mathcal{N} \setminus \{s\}, \quad l \in \text{OUT}(i). \quad (4.24i)$$

To solve [MFG-MiCP], we need the same conditions (initial traffic densities and queue sizes, terminal costs, and traffic demands) as those for [MFGnet], which are introduced in Section 4.4.

Now we group the above nine equations and explain their meanings:

1. Flow on links (corresponding to “dynamic loading” in DUE): Eq. (4.24a) describes the flow balance on each link; Eq. (4.24b) prescribes the fundamental relation between traffic density, driving speed and traffic flow; Eq. (4.24c) describes link influx from each node; and Eq. (4.24d) describes the nodal delay at each node.
2. Flow at nodes (corresponding to “route choice” in DUE): Eq. (4.24g) describes how cars’ optimal travel costs are determined at each node; Eq. (4.24h) guarantees flow conservation at each node; and Eq. (4.24i) is the equilibrium condition of cars’ next-go-to link choices at each node.
3. Speed selection: Eq. (4.24e) and Eq. (4.24f) describe how cars’ optimal travel costs and optimal driving speeds are solved from traffic density in the interior of each link.

**Note.** The first two groups of equations are projected to two components of classical DUE, which are dynamic loading and route

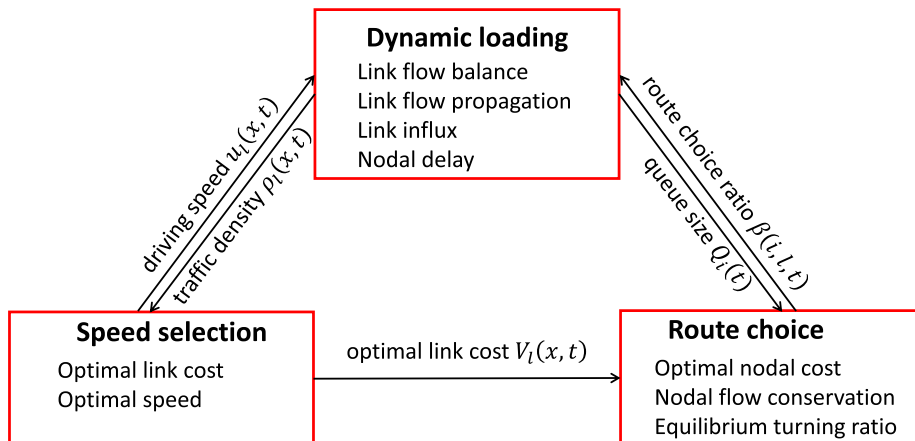


Fig. 3. Components of MFG on networks.

choice. The connection between MFE and DUE will be further established in the next section.

We denote the solution of [MFG-MiCP] as  $\text{SOL}(\text{[MFG-MiCP]})$ , which is the same as  $\text{SOL}(\text{[MFGnet]})$ . In the rest of the paper, we will focus on the system [MFG-MiCP]. Its solution  $\text{SOL}(\text{[MFG-MiCP]})$  describes the mean field equilibrium (MFE), which is the equilibrium dynamic controls of all AVs.

## 5. Connection between MFE and DUE

The system [MFG-MiCP] can be divided into three coupled components: dynamic loading, route choice, and speed selection. The information exchange among these three components is demonstrated in Fig. 3. In the “dynamic loading” procedure, given cars’ driving speeds  $u_l(x, t)$ ,  $\forall l$  and next-go-to link choices  $\beta(i, l, t)$ ,  $\forall i, l$ , the traffic densities  $\rho_l(x, t)$ ,  $\forall l$  and queue sizes  $Q_i(t)$ ,  $\forall i$  are solved from the equations [link flow balance][link flow propagation][link influx][nodal delay], i.e., Eqs. (4.24a)–(4.24d); In the “route choice” procedure, given the queue sizes  $Q_i(t)$ ,  $\forall i$  and cars’ optimal link costs  $V_l(x, t)$ ,  $\forall l$ , the cars’ next-go-to link choices  $\beta(i, l, t)$ ,  $\forall i, l$  are solved from the equations [optimal nodal cost][nodal flow conservation][equilibrium turning ratio], i.e., Eqs. (4.24g)–(4.24i); In the “speed selection” procedure, given the traffic densities  $\rho_l(x, t)$ ,  $\forall l$ , the cars’ driving speeds  $u_l(x, t)$ ,  $\forall l$  and optimal link costs  $V_l(x, t)$ ,  $\forall l$  are solved from the equations [optimal link cost][optimal speed], i.e., Eqs. (4.24e, 4.24f).

The classical DUE concept characterizes a large number of cars’ equilibrium route choices. It does not incorporate cars’ driving speed controls so only the dynamic loading and route choice components are coupled.

If we assume that cars do not optimize their driving speeds but choose the equilibrium speeds characterized by the LWR model (Lighthill and Whitham, 1955; Richards, 1956):

$$u_l(x, t) = U(\rho_l(x, t)), \quad l \in \mathcal{L}, \quad x \in (0, \text{len}(l)), \quad t \in (0, T). \quad (5.1)$$

That is, the driving speed of every car on the network is totally determined by its nearby local density. Accordingly, we present a special case of MFG using the LWR speed, denoted as [MFG-LWR-MiCP], which is formulated below.

### [MFG-LWR-MiCP]

$$[\text{link flow balance}] \quad \partial_t \rho_l(x, t) + \partial_x q_l(x, t) = 0, \quad l \in \mathcal{L}; \quad (5.2a)$$

$$[\text{link flow propagation}] \quad q_l(x, t) = \rho_l(x, t) u_l(x, t), \quad l \in \mathcal{L}; \quad (5.2b)$$

$$[\text{link influx}] \quad q_l(0, t) = \beta(i, l, t) \left( \sum_{h \in \text{IN}(i)} q_h(\text{len}(h), t) + d_i(t) \right) - \frac{dQ_i(t)}{dt}, \quad (5.2c)$$

$$i \in \mathcal{N} \setminus \{s\}, \quad l \in \text{OUT}(i);$$

$$[\text{nodal delay}] \quad 0 \leq \frac{dQ_i(t)}{dt} + M_i - \sum_{h \in \text{IN}(i)} q_h(\text{len}(h), t) - d_i(t) \perp Q_i(t) \geq 0, \quad i \in \mathcal{N} \setminus \{s\}; \quad (5.2d)$$

$$[\text{optimal link cost}] \quad \partial_t V_l(x, t) + \min_{u_{\min} \leq u \leq u_{\max}} \{u \partial_x V_l(x, t) + f_{\text{run}}(u, \rho_l(x, t))\} = 0, \quad l \in \mathcal{L}; \quad (5.2e)$$

$$[\text{LWR speed}] \quad u_l(x, t) = U(\rho_l(x, t)), \quad l \in \mathcal{L}; \quad (5.2f)$$

$$[\text{optimal nodal cost}] \quad V_h(\text{len}(h), t) = \pi_i \left( \min \left\{ t + \frac{Q_i(t)}{M_i}, T \right\} \right) + f_{\text{que}} \left( \min \left\{ \frac{Q_i(t)}{M_i}, T - t \right\} \right), \quad (5.2g)$$

$$i \in \mathcal{N}_I, \quad h \in \text{IN}(i);$$

$$[\text{nodal flow conservation}] \quad 0 \leq \pi_i(t) \perp \sum_{l \in \text{OUT}(i)} \beta(i, l, t) - 1 \geq 0, \quad i \in \mathcal{N} \setminus \{s\}; \quad (5.2h)$$

$$[\text{equilibrium turning ratio}] \quad 0 \leq \beta(i, l, t) \perp V_l(0, t) - \pi_i(t) \geq 0, \quad i \in \mathcal{N} \setminus \{s\}, \quad l \in \text{OUT}(i). \quad (5.2i)$$

In [MFG-LWR-MiCP], one can choose any running cost function  $f_{\text{run}}(\cdot)$ , not restricted to the total travel time as in classical DUE. Even under the same cost function, the optimal speed solved from the HJB equation may have a large deviation from the LWR speed defined in Eq. (5.1). We may solve very different equilibria from [MFG-MiCP] and [MFG-LWR-MiCP]. The two equilibria will be compared on Braess networks in Section 7.

## 6. Solution algorithm

In this section, we will introduce the solution algorithm for solving the mean field equilibrium. We will first discretize the [MFG-MiCP] system on a mesh grid. The discretized system is a finite dimensional MiCP. Then we can solve this discretized MiCP using the



PATH solver built in GAMS (Rosenthal, 2007).

Denote  $\Delta x$  and  $\Delta t$  the spatial and temporal mesh sizes. For the time discretization, we divide the whole time horizon  $[0, T]$  into  $N_t + 1$  time steps:

$$0 = t_0 < t_1 < \dots < t_{N_t} = T, \quad (6.1)$$

where  $t_k = k\Delta t$  ( $k = 0, 1, \dots, N_t$ ), and we assume that  $N_t = T/\Delta t$  is a positive integer.

For the spatial discretization, we divide each link  $l = (i, j) \in \mathcal{L}$  into  $N_x^l = \text{len}(l)/\Delta x$  sublinks (we will always assume  $N_x^l$  is a positive integer):

$$(i, i_1), (i_1, i_2), \dots, (i_{N_x^l-2}, i_{N_x^l-1}), (i_{N_x^l-1}, j). \quad (6.2)$$

As an example, Fig. 4 shows the discretization of the Braess networks with four and five links. It is assumed that all links in the networks have length 1 and  $\Delta x = 0.5$ .

We observe that the spatial discretization of a network  $\mathcal{G} = (\mathcal{N}, \mathcal{L})$  is equivalent to adding auxiliary nodes into  $\mathcal{N}$  and replacing each link  $l \in \mathcal{L}$  by a group of sublinks. Denote the set of auxiliary nodes by  $\mathcal{N}'$  and the set of all nodes  $\mathcal{N}^D = \mathcal{N} \cup \mathcal{N}'$ . Denote the set of all sublinks by  $\mathcal{L}^D$ . The discretized network is then represented by  $\mathcal{G}^D = (\mathcal{N}^D, \mathcal{L}^D)$ . The discretized network  $\mathcal{G}^D$  keeps the same topological structure as the original network  $\mathcal{G}$  but has more nodes and links.

Now we are ready to discretize all variables and equations in [MFG-MiCP] on the discretized network  $\mathcal{G}^D$  and at the time steps defined in Eq. (6.1). We define:

- $\rho_{ij}^k$ : the average traffic density on link  $(i, j) \in \mathcal{L}^D$  at time  $t_k, k = 0, 1, \dots, N_t$ .
- $p_{ij}^k$ : the entry flow rate entering link  $(i, j) \in \mathcal{L}^D$  from node  $i \in \mathcal{N}^D \setminus \{s\}$  at time  $t_k, k = 0, 1, \dots, N_t - 1$ .
- $q_{ij}^k$ : the exit flow rate leaving link  $(i, j) \in \mathcal{L}^D$  from node  $j \in \mathcal{N}^D$  at time  $t_k, k = 0, 1, \dots, N_t - 1$ .
- $u_{ij}^k$ : the average driving speed of cars on link  $(i, j) \in \mathcal{L}^D$  at time  $t^k, k = 0, 1, \dots, N_t - 1$ .
- $\beta_{ij}^k$ : the percentage of cars selecting link  $(i, j) \in \mathcal{L}^D$  among all cars arriving at node  $i \in \mathcal{N}^D \setminus \{s\}$  at time  $t^k, k = 0, 1, \dots, N_t - 1$ .
- $V_{ij}^k$ : the optimal travel cost if a car exits the queue at node  $i \in \mathcal{N}^D \setminus \{s\}$  and selects the link  $(i, j) \in \mathcal{L}^D$  at time  $t^k, k = 0, 1, \dots, N_t$ .
- $\pi_i^k$ : the optimal travel cost if a car exits the queue at node  $i \in \mathcal{N}^D \setminus \{s\}$  and selects the minimum-cost link from node  $i$  at time  $t^k, k = 0, 1, \dots, N_t$ .
- $\lambda_i^k$ : the optimal travel cost if a car enters the queue at node  $i \in \mathcal{N}^D$  at time  $t^k, k = 0, 1, \dots, N_t$ .
- $Q_i^k$ : the queue size at node  $i \in \mathcal{N}^D \setminus \{s\}$  at time  $t^k, k = 0, 1, \dots, N_t$ .

Provided the listed discrete variables, we then discretize each equation in [MFG-MiCP]:

- The equations [link flow balance][optimal link cost][optimal speed] are discretized by the upwinding scheme (Huang et al., 2020a). Note that the driving speed is in the range  $[u_{\min}, u_{\max}]$ , the choice of the spatial and temporal mesh sizes should satisfy the CFL condition  $u_{\max} \Delta t \leq \Delta x$  to guarantee numerical stability (LeVeque, 2002).
- In the equation [link influx], the term  $\frac{dQ_i(t)}{dt}$  is discretized as  $\frac{Q_i^{k+1} - Q_i^k}{\Delta t}$  at time  $t^k$ .
- For the equation [nodal delay], the forward Euler discretization is not well-defined (Han et al., 2013). Here we use the following backward Euler scheme to discretize the equation [nodal delay], which is suggested by Ban et al. (2012):

$$0 \leq \frac{Q_i^{k+1} - Q_i^k}{\Delta t} + M_i - \sum_{(m,i) \in \mathcal{L}^D} q_{mi}^k - d_i^k \perp Q_i^{k+1} \geq 0. \quad (6.3)$$

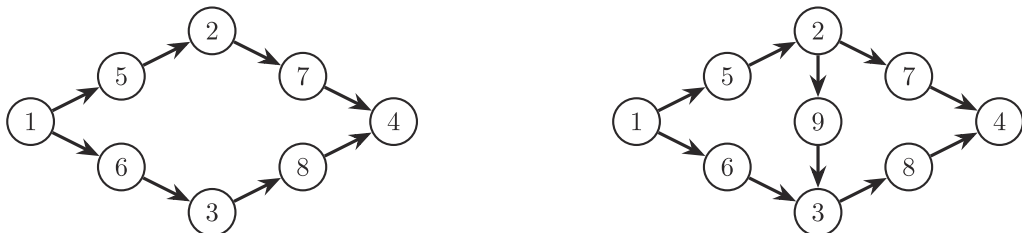


Fig. 4. Discretized Braess networks, all links have length 1 and  $\Delta x = 0.5$ .

- In [optimal nodal cost], the term  $\pi_i \left( \min \left\{ t + \frac{Q_i(t)}{M_i}, T \right\} \right) + f_{\text{que}} \left( \min \left\{ \frac{Q_i(t)}{M_i}, T - t \right\} \right)$  at time  $t^k$  is discretized using the following interpolation scheme (Ban et al., 2008):

$$\sum_{k'=k}^{N_t-1} \mathbf{1}_{k' \leq k + Q_i^k / (M_i \Delta t) < k'+1} \cdot \left[ \left( k' + 1 - \frac{Q_i^k}{M_i \Delta t} - k \right) \pi_i^{k'} + \left( \frac{Q_i^k}{M_i \Delta t} + k - k' \right) \pi_i^{k'+1} \right] + \mathbf{1}_{k + Q_i^k / (M_i \Delta t) \geq N_t} \cdot \pi_i^{N_t} + f_{\text{que}} \left( \min \left\{ \frac{Q_i^k}{M_i}, T - k \Delta t \right\} \right). \quad (6.4)$$

- The equations [link flow propagation][nodal flow conservation][equilibrium turning ratio] do not involve space or time derivatives, therefore they can be directly discretized.

Summarizing all above, the discretized system of [MFG-MiCP] is:

$$\left\{ \begin{array}{l} \rho_{ij}^{k+1} = \rho_{ij}^k + \frac{\Delta t}{\Delta x} (p_{ij}^k - q_{ij}^k), \quad (i, j) \in \mathcal{L}^D, 0 \leq k \leq N_t - 1; \quad (6.5a) \\ q_{ij}^k = \rho_{ij}^k u_{ij}^k, \quad (i, j) \in \mathcal{L}^D, 0 \leq k \leq N_t - 1; \quad (6.5b) \\ p_{ij}^k = \beta_{ij}^k \left( \sum_{(m,i) \in \mathcal{L}^D} q_{mi}^k + d_i^k - \frac{Q_i^{k+1} - Q_i^k}{\Delta t} \right), \quad (i, j) \in \mathcal{L}^D, 0 \leq k \leq N_t - 1; \quad (6.5c) \\ 0 \leq \frac{Q_i^{k+1} - Q_i^k}{\Delta t} + M_i - \sum_{(m,i) \in \mathcal{L}^D} q_{mi}^k - d_i^k \perp Q_i^{k+1} \geq 0, \quad i \in \mathcal{N}^D \setminus \{s\}, 0 \leq k \leq N_t - 1; \quad (6.5d) \\ V_{ij}^{k+1} - V_{ij}^k + \Delta t \left( u_{ij}^k \frac{\lambda_j^{k+1} - V_{ij}^{k+1}}{\Delta x} + f \left( u_{ij}^k, \rho_{ij}^k \right) \right) = 0, \quad (i, j) \in \mathcal{L}^D, 0 \leq k \leq N_t - 1; \quad (6.5e) \\ u_{ij}^k = \operatorname{argmin}_{u_{\min} \leq u \leq u_{\max}} \left\{ u \frac{\lambda_j^{k+1} - V_{ij}^{k+1}}{\Delta x} + f \left( u, \rho_{ij}^k \right) \right\} = 0, \quad (i, j) \in \mathcal{L}^D, 0 \leq k \leq N_t - 1; \quad (6.5f) \\ \lambda_i^k = \sum_{k'=k}^{N_t-1} \mathbf{1}_{k' \leq k + Q_i^k / (M_i \Delta t) < k'+1} \cdot \left[ \left( k' + 1 - \frac{Q_i^k}{M_i \Delta t} - k \right) \pi_i^{k'} + \left( \frac{Q_i^k}{M_i \Delta t} + k - k' \right) \pi_i^{k'+1} \right] \\ + \mathbf{1}_{k + Q_i^k / (M_i \Delta t) \geq N_t} \cdot \pi_i^{N_t} + f_{\text{que}} \left( \min \left\{ \frac{Q_i^k}{M_i}, T - k \Delta t \right\} \right), \quad i \in \mathcal{N}^D \setminus \{s\}, 0 \leq k \leq N_t; \quad (6.5g) \\ 0 \leq \pi_i^k \perp \sum_{(ij) \in \mathcal{L}^D} \beta_{ij}^k - 1 \geq 0, \quad i \in \mathcal{N}^D \setminus \{s\}, 0 \leq k \leq N_t - 1; \quad (6.5h) \\ 0 \leq \beta_{ij}^k \perp V_{ij}^k - \pi_i^k \geq 0, \quad (i, j) \in \mathcal{L}^D, 0 \leq k \leq N_t - 1. \quad (6.5i) \end{array} \right.$$

The given conditions are:  $\rho_{ij}^0 = 0$  for all  $(i, j) \in \mathcal{L}^D$ ,  $Q_i^0 = 0$  for all  $i \in \mathcal{N}^D \setminus \{s\}$ , and  $\lambda_s^k = 0$  for all  $0 \leq k \leq N_t$ . To solve the discretized system we also need to specify the following: (i) the traffic demand  $d_i^k = d_i(t_k)$  for all  $i \in \mathcal{N}^D$  and all time steps  $t_k$  ( $k = 0, \dots, N_t$ ); (ii) the terminal cost  $V_{ij}^{N_t}$  for all  $(i, j) \in \mathcal{L}^D$  at the final time; (iii) the terminal cost  $\pi_i^{N_t}$  for all  $i \in \mathcal{N}^D$  at the final time.

Now we are going to present a solution algorithm to solve the discretized [MFG-MiCP] system. The algorithm employs fixed-point iterations on the queue sizes  $\mathbf{Q} = (Q_i)_{i \in \mathcal{N}^D}$ , following the approach in Ban et al. (2008). We will start from any feasible initialization of queue sizes  $\mathbf{Q}^{(0)}$ . In each iteration, we substitute the current queue sizes  $\mathbf{Q}^{(n)}$  into the summation term in Eq. (6.5g), which gives a relaxed MiCP. The solution of the relaxed MiCP gives a set of updated queue sizes  $\mathbf{Q}^{(n+1)}$ . We keep doing the iteration and updating the queue sizes until the iteration error:

$$E_{\mathbf{Q}}^{(n)} = \|\mathbf{Q}^{(n+1)} - \mathbf{Q}^{(n)}\|_2 \leq \varepsilon, \quad (6.6)$$

where  $\|\cdot\|_2$  is the 2-norm and  $\varepsilon$  is a predefined threshold.

In each iteration of the solution algorithm, the relaxed MiCP is solved by the PATH solver built in GAMS. The version of GAMS used in this paper is GAMS 24.9. The [MFG-LWR-MiCP] system can be discretized and solved in the same way.

## 7. Numerical experiments

In this section we will demonstrate our mean field game model and solution algorithm on Braess networks and the OW network. The basic experimental settings are illustrated in Section 7.1. Then four experiments are designed for different purposes.

1. In the first experiment, we will validate numerical convergence of our solution algorithm (Section 7.2).
2. In the second experiment, we will compare the equilibria solved from [MFG-MiCP] and [MFG-LWR-MiCP] with the same running and queuing cost functions on a three-path Braess network to demonstrate the role of driving speed control (Section 7.3).
3. In the third experiment, we will demonstrate the occurrence of Braess paradox in the context of AVs by comparing their travel costs solved from [MFG-MiCP] on Braess networks with two and three paths (Section 7.4).
4. In the fourth experiment, we will test the performance of the presented solution algorithm on the OW network, and compare the equilibria solved from [MFG-MiCP] and [MFG-LWR-MiCP] with the same running and cost functions (Section 7.5).

### 7.1. Experimental settings

We will work on three networks: the two-path network with four links, as shown in Fig. 5a, the three-path network with five links, as shown in Fig. 5b, and the OW network that will be shown in Section 7.5. The length of every link is assumed to be one if not particularly indicated, which is  $\text{len}(l) = 1, \forall l \in \mathcal{L}$ .

We define the AV's link running cost function as:

$$f_{\text{run}} \left( u, \rho \right) = \underbrace{\frac{c_1}{2} \left( \frac{u}{u_{\max}} \right)^2}_{\text{kinetic energy}} + \underbrace{c_2 \frac{\rho}{\rho_{\text{jam}}}}_{\text{safety}} + \underbrace{c_3}_{\text{efficiency}}, \quad (7.1)$$

where  $c_1, c_2$  and  $c_3$  are positive coefficients. The first term represents the kinetic energy; the second term quantifies driving safety by taking traffic densities as a penalty term, meaning that AVs tend to avoid congested areas; the third term quantifies driving efficiency, whose integration over one AV's travel time horizon  $[t_0, t_f]$  on a network gives:

$$\int_{t_0}^{t_f} c_3 dt = c_3 (t_f - t_0), \quad (7.2)$$

which is proportional to the AV's total travel time on a network.

We define the AV's node queuing cost function as:

$$f_{\text{que}}(t_{\text{delay}}) = c_4 t_{\text{delay}}, \quad (7.3)$$

which is proportional to the AV's queuing delay  $t_{\text{delay}}$  at a node.

The speed-density function for [MFG-LWR-MiCP] is set to be:

$$U(\rho) = u_{\max} \left( 1 - \frac{\rho}{\rho_{\text{jam}}} \right). \quad (7.4)$$

We will assume  $u_{\min} = 0, u_{\max} = 1$  and  $\rho_{\text{jam}} = 1$  if not particularly indicated. To simulate a numerical experiment, we still need to specify the following:

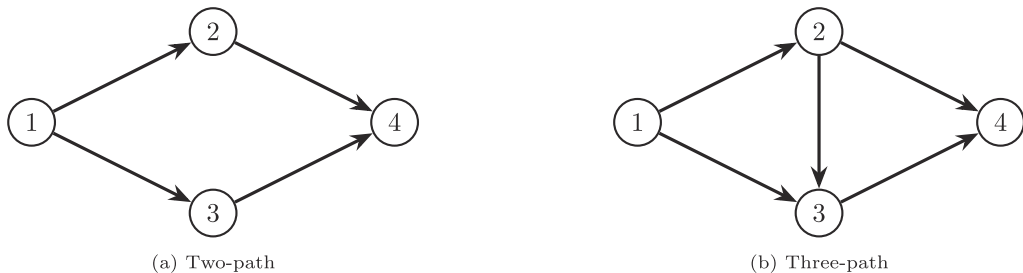


Fig. 5. Braess networks.

- The simulation time  $T$ ;
- The coefficients  $c_1, c_2, c_3$  and  $c_4$  in the cost functions defined in Eq. (7.1) and Eq. (7.3);
- The bottleneck capacities  $M_i$  for all  $i \in \mathcal{N} \setminus \{s\}$ ;
- the traffic demand  $d_i(t)$  for all  $i \in \mathcal{N}_O$  and  $t \in [0, T]$ ;
- The terminal cost  $V_l(x, T)$  for all  $l \in \mathcal{L}$  and  $\pi_i(T)$  for all  $i \in \mathcal{N} \setminus \{s\}$  at the final time.

## 7.2. Numerical convergence

To ensure the credibility of numerical solutions solved from the algorithm presented in Section 6, it is necessary to validate numerical convergence. That is, when the discretization mesh sizes  $\Delta x$  and  $\Delta t$  go to zero, the numerical solution converges to the true solution of the continuum system.

In this subsection, we will validate numerical convergence with the [MFG-MiCP] system on the two-path network. The simulation time  $T = 3$  and the cost function coefficients  $c_1 = c_2 = 1, c_3 = 0.5, c_4 = 1$ . The bottleneck capacity  $M_i = 1$  for every  $i = 1, 2, 3$ . Cars enter the network through the single origin  $o = 1$  and the traffic demand is given by:

$$d_1(t) = \begin{cases} 0.5, & t \in [0, 0.5]; \\ 0, & t \in (0.5, 3]. \end{cases} \quad (7.5)$$

At the final time, the terminal costs  $V_l(x, T)$  and  $\pi_i(T)$  are set to be zero for all  $l \in \mathcal{L}$  and  $i \in \mathcal{N} \setminus \{s\}$ .

Since the explicit solution of [MFG-MiCP] is not available, we will validate numerical convergence by comparing numerical solutions on multi-grid. We solve [MFG-MiCP] with different mesh sizes  $\Delta x$  and  $\Delta t$ , but keep the ratio  $\Delta x/\Delta t = 1$ . For each  $\Delta x$ , we compare the coarse solution with mesh size  $\Delta x$  and the refined solution with mesh size  $\Delta x/2$ , which are denoted as  $\text{SOL}([\text{MFG-MiCP}], \Delta x)$  and  $\text{SOL}([\text{MFG-MiCP}], \Delta x/2)$ , respectively. Note that these two numerical solutions are defined on different mesh grids, we will first project the coarse solution  $\text{SOL}([\text{MFG-MiCP}], \Delta x)$  onto fine grids with mesh size  $\Delta x/2$  using the piecewise constant interpolation.

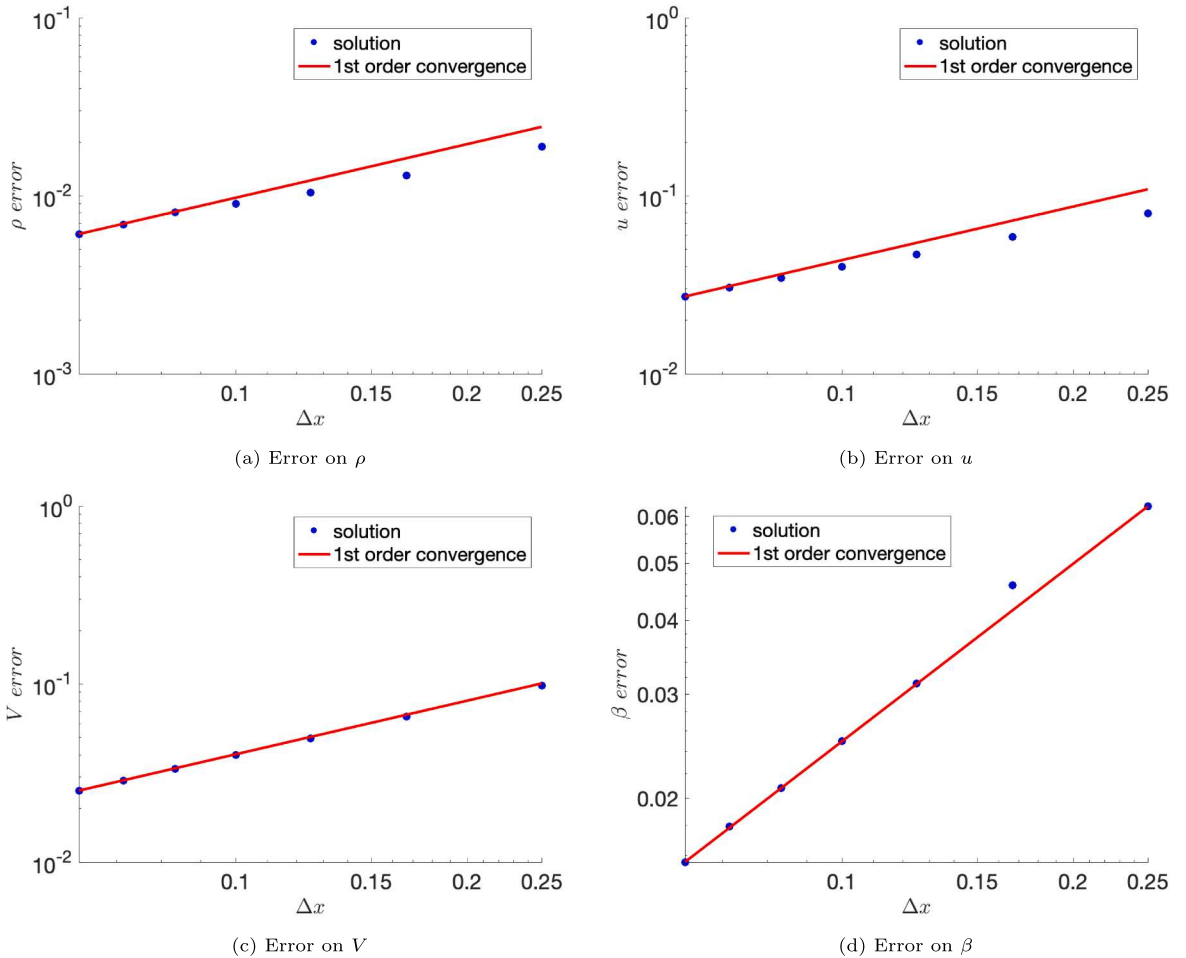


Fig. 6. Numerical convergence.

Denote the projected solution as  $\text{PROJ}(\text{SOL}([\text{MFG-MiCP}], \Delta x), \Delta x/2)$ . Then the solutions  $\text{SOL}([\text{MFG-MiCP}], \Delta x/2)$  and  $\text{PROJ}(\text{SOL}([\text{MFG-MiCP}], \Delta x), \Delta x/2)$  are on the same mesh grids. We compute the mean absolute error between the two solutions on variables  $\rho, u, V$  and  $\beta$ . The errors are plotted with respect to  $\Delta x$  in Fig. 6.

In Fig. 6, the x-axis is the mesh size  $\Delta x$  and the y-axis is the numerical error. The blue dots represent errors computed from different values of  $\Delta x$ . The red line represents the perfect first order convergence, i.e., the error is proportional to  $\Delta x$ . We observe that the errors on all variables go to zero as  $\Delta x \rightarrow 0$ . This means the change on the numerical solution will go to zero if we keep refining the mesh grids, which validates the numerical convergence (LeVeque, 2007). In particular, it is shown in the figure that the numerical solution has first order convergence.

### 7.3. Comparison between MFG and LWR equilibria on the three-path network

In this subsection, we would like to compare the MFG and LWR-based optimal controls of AVs across the three-path network, and demonstrate the role of driving speed control. We solve both [MFG-MiCP] and [MFG-LWR-MiCP] with the same running and queuing cost functions on the three-path network. The solved equilibria are compared in terms of density evolution, velocity profiles and queuing delays.

The settings of these experiments are as follows:

- The simulation time  $T = 12$ .
- The links (1, 2) and (1, 3) have the same running cost function coefficients  $c_1 = c_2 = c_3 = 1$ ; the link (2, 4) has the running cost function coefficients  $c_1 = 5, c_2 = 5$  and  $c_3 = 5$ ; the link (3, 4) has the running cost function coefficients  $c_1 = 0.5, c_2 = 0$  and  $c_3 = 0.25$ ; the link (2, 3) has the running cost function coefficients  $c_1 = 0.5, c_2 = 0$  and  $c_3 = 0.1$ . With these coefficients, the links (1, 2) and (1, 3) are symmetric.

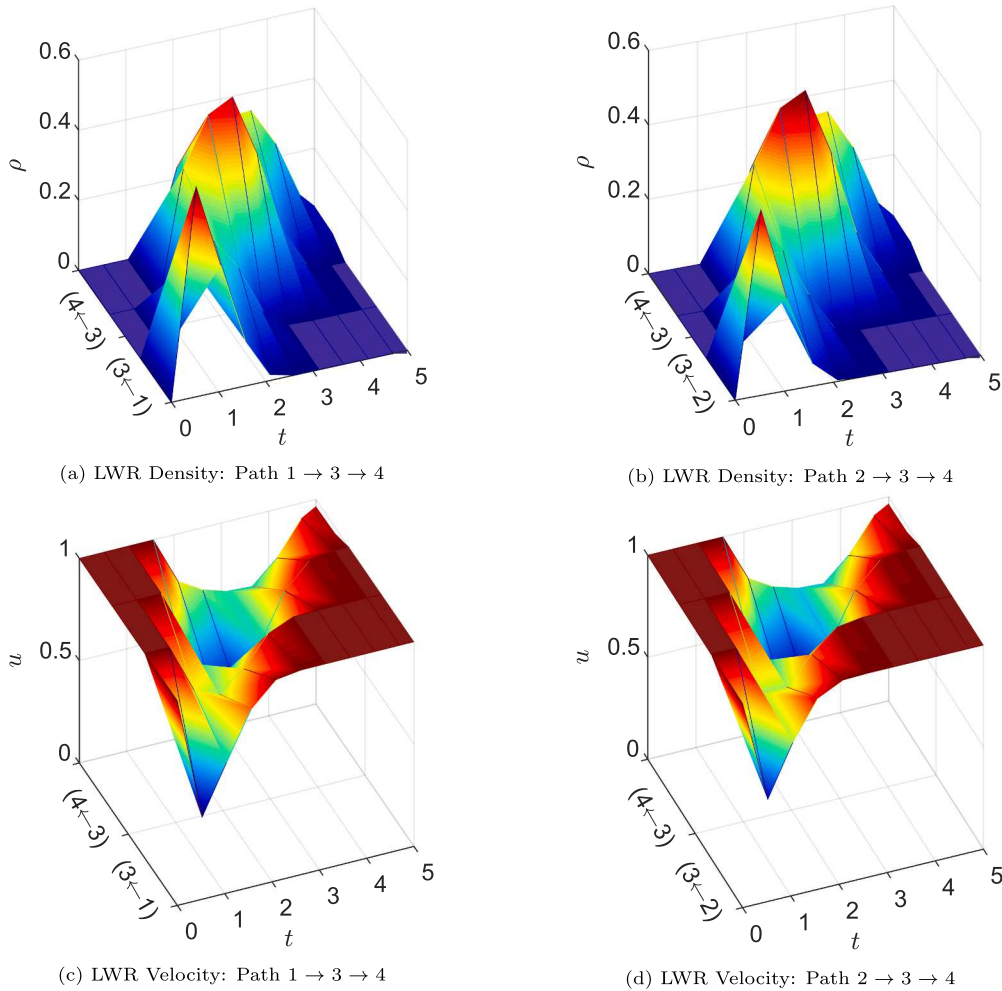


Fig. 7. LWR equilibrium traffic on the three-path network.

- The bottleneck capacity  $M_1 = M_2 = 1$  and  $M_3 = 0.35$ . The queuing cost function coefficient  $c_4 = 1$ .
- AVs enter the network through origin nodes 1 and 2. The traffic demands are given by:

$$d_1(t) = \begin{cases} 0.6, & t \leq 0.5; \\ 0, & t > 0.5; \end{cases}, \quad d_2(t) = \begin{cases} 0.5, & t \leq 0.5; \\ 0, & t > 0.5. \end{cases} \quad (7.6)$$

- The AVs' terminal costs at time  $T$  are given in the following way: we first specify the nodal terminal costs  $\pi_1(T) = 2, \pi_2(T) = 1.33, \pi_3(T) = 0.67$ ; the link terminal cost  $V_l(x, T)$  is given by linear interpolation between the nodal terminal costs.

For both the solved MFG and LWR equilibria, the link  $(2, 4)$  turns out not to be selected, likely because its running cost coefficients are large. Hence cars entering the network through node 1 drive along the path  $1 \rightarrow 3 \rightarrow 4$  while cars entering the network through node 2 drive along the path  $2 \rightarrow 3 \rightarrow 4$ . The flows from the two paths merge at node 3 and form a non-empty queue.

In Fig. 7 the traffic density and velocity for the LWR equilibrium along the two paths is plotted in a 3D diagram. The x-axis represents the path as a continuous road of length 2, the y-axis is time and the z-axis represents traffic density or velocity. In Fig. 8 the traffic density and velocity for the MFG equilibrium is plotted in the same way.

By comparing Fig. 8 to Fig. 7, we observe the difference between MFG and LWR equilibria on driving speeds: In the LWR equilibrium, AVs drive faster in low density areas and slower in high density areas. In other words, the LWR speed is totally determined by its local traffic density and thus this strategy is “myopic”. By contrast, in the MFG equilibrium, AVs may drive at relatively high speeds even if the local density is relatively high. That is, AVs deployed with the MFG equilibrium controls can look “farther” and adjust their driving speeds in a way to optimize their total travel costs over the network. In other words, even the cost incurred by high speed and high density is relatively large for the current link where an AV is, the AV can anticipate the future cost incurred on the subsequent links with the goal of minimizing the cumulative total cost.

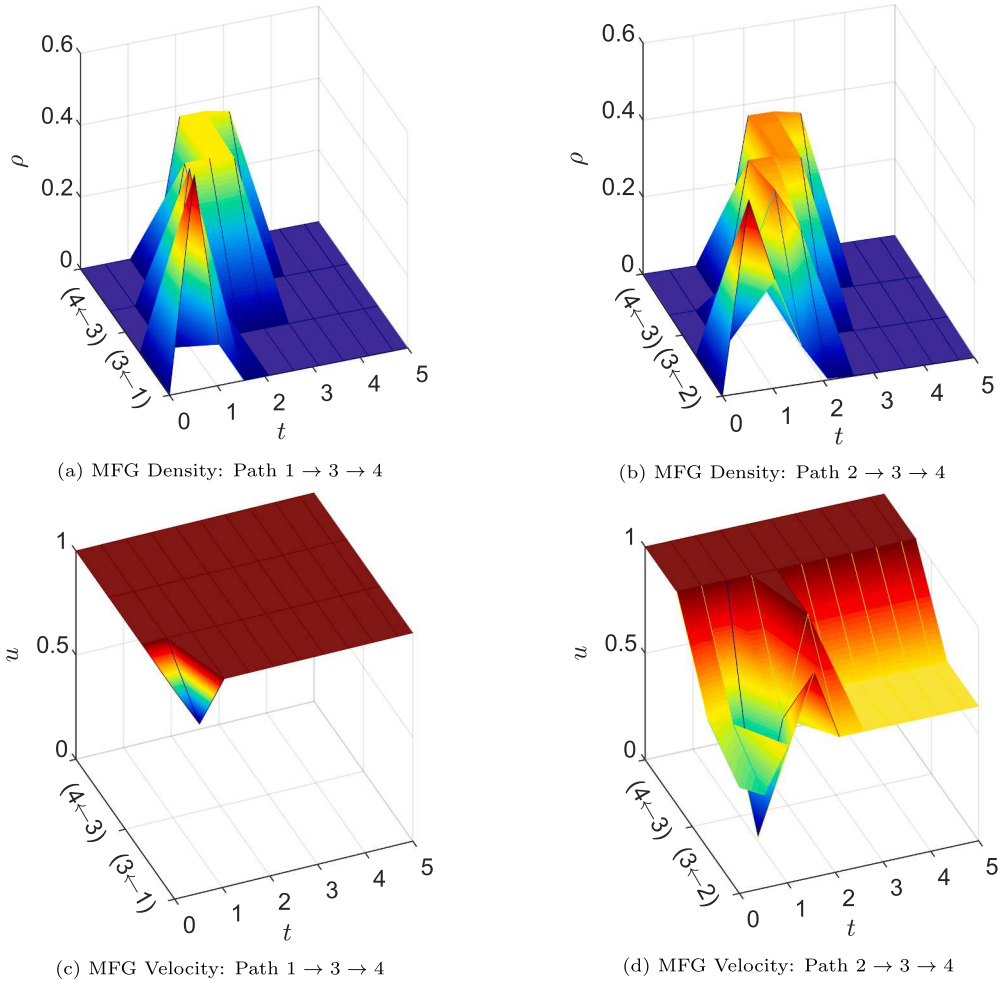


Fig. 8. MFG equilibrium traffic on the three-path network.



It is also interesting to compare the queue size at node 3 of MFG and LWR equilibria during the time horizon  $t \in [0, 5]$ , which is shown in Fig. 9a. We do not plot the queue size for later times because the network is already empty at  $t = 5$  for both equilibria. We observe from Fig. 9a that the queue size of the MFG equilibrium is larger than that of the LWR equilibrium. This is because AVs drive at higher speeds on both links (1, 3) and (2, 3) when  $t \leq 2$  in the MFG equilibrium than in the LWR equilibrium. Hence the node 3 is more congested and more cars stuck in the queue at the node in the MFG equilibrium. However, although AVs have longer queuing delays at node 3, we observe from Figs. 8a and b that all AVs arrive at the destination at approximately  $t = 4$  in the MFG equilibrium, while that arrival time for AVs in the LWR equilibrium is approximately  $t = 5$  from Figs. 7a and b.

In Fig. 9b, the solution algorithm presented in Section 6 is validated by showing convergence of the fixed-point iteration on the queue size. The iteration error, which is defined in Eq. (6.6), converges to zero in two iterations for the LWR equilibrium and in six iterations for the MFG equilibrium. It verifies the efficiency of the solution algorithm.

#### 7.4. Braess paradox

This subsection aims to demonstrate the occurrence of the Braess paradox. We will first solve a MFE on the two-path network. Then the middle link is added and a new MFE is solved on the three-path network. We artificially create the occurrence of the paradox, which is when travel cost with the middle link is higher than that without.

The settings of this experiment are as follows:

- The simulation time  $T = 6$ .
- The links (1, 2) and (3, 4) have the same running cost function coefficients  $c_1 = 1$ ,  $c_2 = 5$  and  $c_3 = 0$ ; The links (1, 3) and (2, 4) have the same running cost function coefficients  $c_1 = 1$ ,  $c_2 = 0$  and  $c_3 = 3$ . It means that (1, 2) and (3, 4) are bottleneck links on which AVs' travel costs highly depend on traffic density. In contrast, AVs can move freely on links (1, 3) and (2, 4) to minimize their travel times no matter how congested the links are.
- The three-path network is constructed by adding the middle link (2, 3) into the two-path network. The middle link (2, 3) has a length of 0.25 and its running cost function coefficients are  $c_1 = 1$ ,  $c_2 = 5$  and  $c_3 = 0$ . Since this link is very short, AVs can traverse it in a short time and incur a smaller travel cost than other links.
- The bottleneck capacity  $M_i = 0.8$  for every  $i = 1, 2, 3$  on both two-path and three-path networks. The queuing cost function coefficient  $c_4 = 1$ .
- For both cases, AVs enter the network through the single origin  $o = 1$  and the traffic demand is given by:

$$d_1(t) = \begin{cases} 0.75, & t \in [0, 1]; \\ 0, & t \in (1, 6]. \end{cases} \quad (7.7)$$

- The AVs' terminal costs at time  $T$  are given by:

$$V_l(x, T) = \begin{cases} 2 - x, & l = (1, 2), (1, 3), (2, 3); \\ 1 - x, & l = (2, 4), (3, 4), \end{cases} \quad (7.8)$$

and:

$$\pi_i(T) = \begin{cases} 2, & i = 1; \\ 1, & i = 2, 3. \end{cases} \quad (7.9)$$

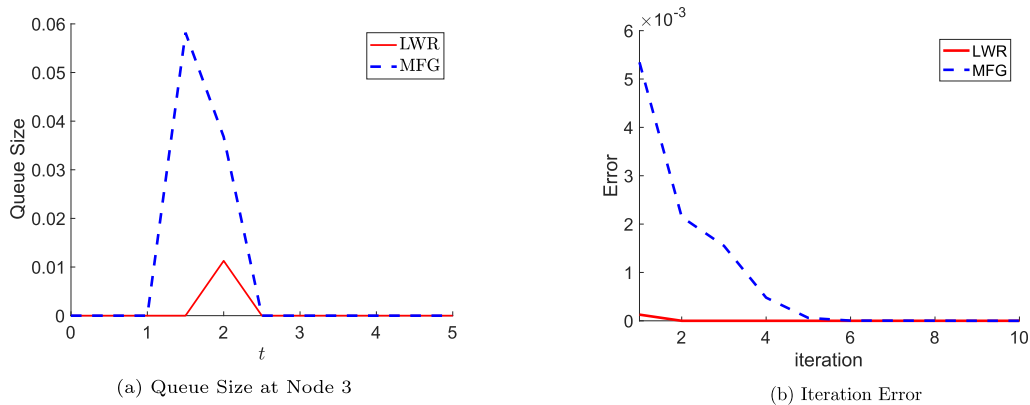


Fig. 9. Queuing information at node 3 on the three-path network.

It means that there will be a penalty if an AV does not reach the destination before time  $T$ , and the penalty cost depends on the distance between the AV's position on the network and the destination.

In Fig. 10, the traffic density is plotted on all links of the two networks at the time snapshots  $t = 1.75$  and  $t = 2.5$ , respectively, to demonstrate the AVs' dynamic route choices on the two networks. We observe from Figs. 10a and b that both paths are used in the two-path network, but more AVs choose the top path  $1 \rightarrow 2 \rightarrow 4$ . This is different from the classical static Braess paradox in which cars are split equally on the two paths. In the latter situation, cars have the same travel time on links  $(1, 2)$  and  $(3, 4)$  because the two links have the same link travel time function and the total flow passing the two links is the same. Same for links  $(1, 3)$  and  $(2, 4)$ . So the total travel time is the same on the two paths. However, equally split flow is no longer an equilibrium in the dynamic case. Even if the links  $(1, 2)$  and  $(3, 4)$  have the same cost function coefficients and the total flow passing the two links is the same, the travel costs on the two links are in general different because the costs also depend on the distribution of the total flow with respect to time. The flow distribution entering the link  $(3, 4)$  at different times depends on the AVs' speed controls on the link  $(1, 3)$ , which is in general different from the flow distribution entering the link  $(1, 2)$ . For the same reason, AVs have different travel costs on links  $(1, 3)$  and  $(2, 4)$ , and the two paths give different travel costs. To summarize, the existence of AVs' speed controls break the symmetry between the two paths.

We observe from Figs. 10c and d that a majority of AVs choose to travel along the path  $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$  on the three-path network. At  $t = 1.75$ , only 22 percent of AVs are on link  $(1, 3)$  and  $(2, 4)$ ; at  $t = 2.5$ , all AVs are on the path  $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$ . This is because AVs experience a small travel cost on the middle link  $(2, 3)$ , so most of them choose to go along the path  $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$  and benefit from the middle link.

In Fig. 11, we plot the optimal nodal cost  $\pi_1(t)$ , which is the optimal travel cost of AVs entering the network through node 1 at time  $t$ , for  $t \in [0, 1]$ . We observe that when  $t < 0.5$ , AVs will have a lower travel cost on the three-path network than on the two-path network; but when  $t > 0.5$ , AVs will have a higher cost on the three-path network. In other words, AVs can reduce their travel costs by utilizing the middle link if they enter the network early; but AVs entering the network at later times fall victim to the seemingly low-cost middle link. It is because at early times, there are a few cars on the network and the traffic density  $\rho$  on each link is very small, so the links  $(1, 2)$  and  $(3, 4)$  are superior to links  $(1, 3)$  and  $(2, 4)$ ; at later times, there are more cars on the network and the links  $(1, 2)$  and  $(3, 4)$  become bottlenecks since AVs' travel costs on these two links are highly sensitive to their respective traffic densities.

### 7.5. OW Network

In this subsection, we would like to test the solution algorithm presented in Section 6 on the OW network, which has 13 nodes and 24 links. Both [MFG-MiCP] and [MFG-LWR-MiCP] are solved on the OW network with the same running and queuing cost functions, and we compare the solved equilibria in terms of average velocity, occupied links and nodal optimal costs.

The settings of this experiment are as follows:

- The simulation time  $T = 12$ .

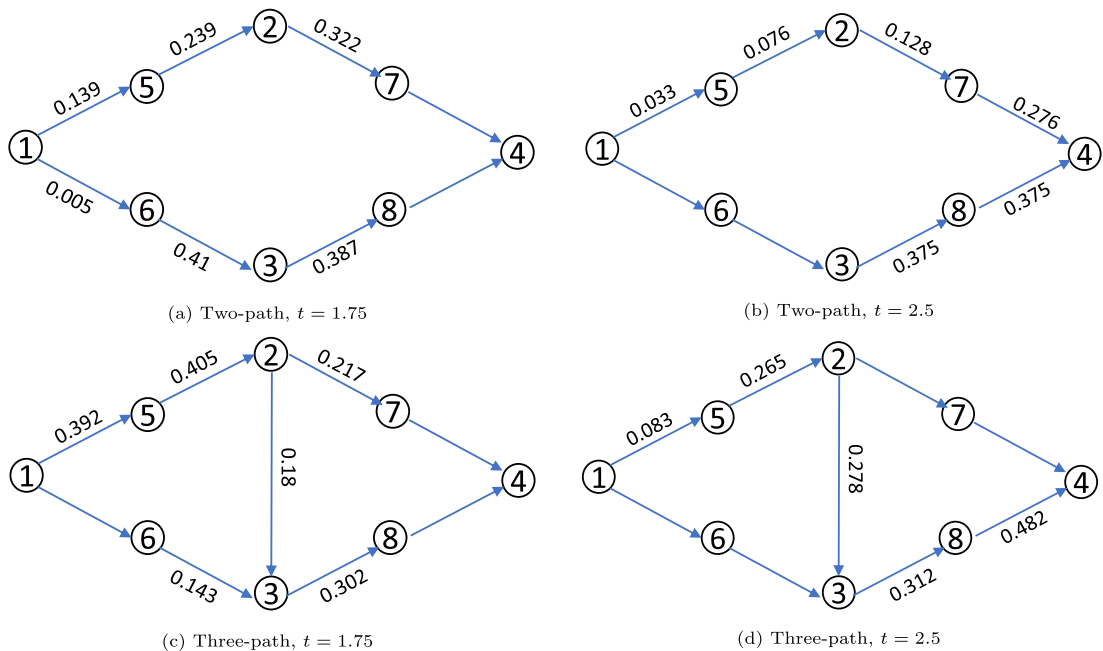


Fig. 10. Density evolution on the two-path and three-path networks.

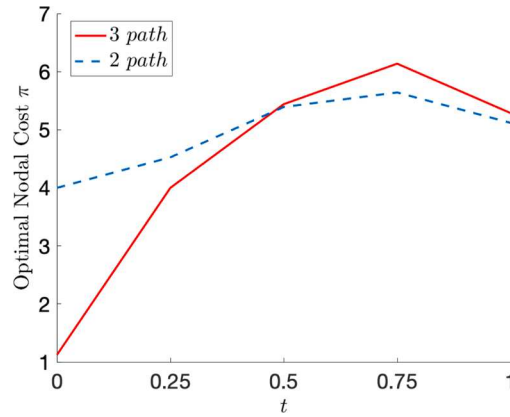


Fig. 11. Optimal travel cost at the origin node on the two-path and three-path networks.

- All links other than (4, 5) and (5, 9) have the same running cost function coefficients  $c_1 = c_2 = 1$  and  $c_3 = 0.5$ ; The links (4, 5) and (5, 9) have the same running cost function coefficients  $c_1 = c_2 = 2$  and  $c_3 = 5$ .
- The bottleneck capacity  $M_i = 1$  for all nodes  $i$ . The queuing cost function coefficient  $c_4 = 1$ .
- AVs enter the network through the single origin  $o = 1$ . The traffic demand is given by:

$$d_1(t) = \begin{cases} 0.75, & t \in [0, 2]; \\ 0, & t \in (2, 12]. \end{cases} \quad (7.10)$$

- The AVs' terminal costs at time  $T$  for all nodes are given according to their distance from destination  $s = 13$  with  $\pi_{13}(T) = 0$ .

For both [MFG-MiCP] and [MFG-LWR-MiCP], we solve the equilibria by the developed solution algorithm. The initial values of queue sizes are set to zero. The computation times for the MFG and LWR equilibria are 62.24s and 28.06s, respectively, on a laptop. That is, the presented solution algorithm is still efficient on the OW network.

Figs. 12a and b show the density evolution of the solved MFG equilibrium at time  $T = 2$  and  $T = 5$ , respectively. The black nodes are those in the OW network and the blue ones are auxiliary nodes introduced for link discretization.

Figs. 13a shows the average velocity of all AVs on the network and Fig. 13b shows the number of AVs' occupied links as a function of time  $t$  for both MFG and LWR equilibria. We observe from the figures that AVs tend to select fewer links and drive at higher speeds in the MFG equilibrium than in the LWR equilibrium. Such a phenomenon is similar to what we observe on the three-path network: AVs deployed with MFG equilibrium controls can freely select their driving speeds to optimize their cumulative travel costs over a network.

In Fig. 13c, the optimal nodal cost  $\pi_1(t)$  at the origin 1 of the two equilibria is compared during the time horizon  $t \in [0, 1.5]$ . We observe that before time  $t = 0.5$ , the MFG and LWR equilibria have the same optimal nodal cost, possibly because AVs entering the network at early times experience little congestion and AVs in both MFG and LWR equilibria can have high speeds. While when  $t > 0.5$ , cars entering the network at time  $t$  may go into high density areas, and AVs in the LWR equilibrium are restricted to drive at very low speeds. Because AVs in the MFG equilibrium have more control power on the driving speed control, they can adjust their speeds and achieve lower travel costs.

## 8. Conclusions and future directions

This paper illustrates the application of mean field games to the modeling of AVs' driving and routing game on a network. Mean field games (MFGs) have been widely used to design decision-making processes in finances and economics, with rather limited research in traffic modeling. Its application to AVs' decision choices on a network is a natural generalization of the classical DTA models. DTA models are popular tools to prescribe human drivers' en-route path choices and traffic flow evolution in the short term. However, they do not stipulate cars' velocity control while moving inside a road segment. Velocity control of AVs becomes increasingly important to navigate AVs across a congested traffic network. Therefore a new modeling paradigm is warranted. To model the movement of AVs in a transportation network, two choices are involved: driving speed control in the interior of a link, and next-go-to link choice at a junction node. To implement MFG, we decompose the problem into a backward HJB equation and a forward continuity equation. The derived MFG system is then reformulated as a MiCP. Based on the MiCP formulation, we illustrate the connection between MFG and classical DUE concept. An efficient numerical algorithm is developed to solve the MiCP, and the solved AVs' optimal controls are illustrated on the Braess network and the OW network. On the Braess network, by comparing the equilibria under optimal speed selection and the LWR speed, the advantage of AVs' optimal velocity controls is demonstrated. Braess paradox is discovered in the AV traffic, indicating that we need to carefully design AV control algorithms to prevent the paradox from happening. On the OW network, by showing that our solution algorithm can successfully solve the equilibrium solution in a short computation time, we verify

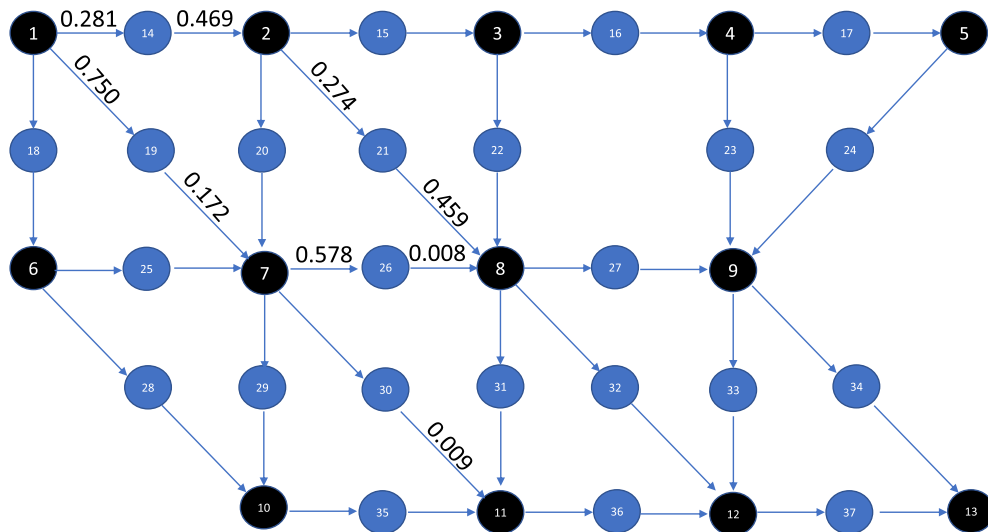
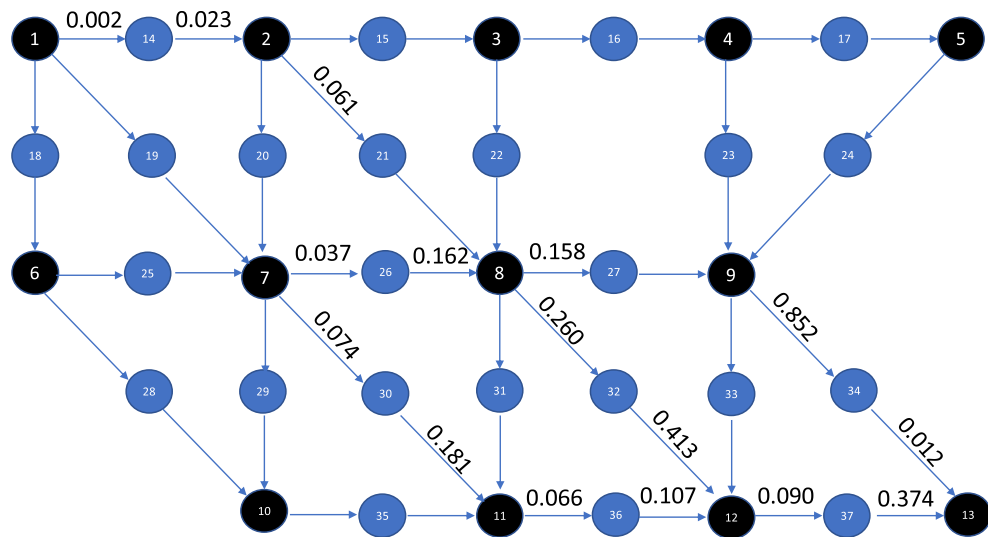
(a) Traffic Density at  $T = 2$ (b) Traffic Density at  $T = 5$ 

Fig. 12. MFG equilibrium density evolution on the OW network.

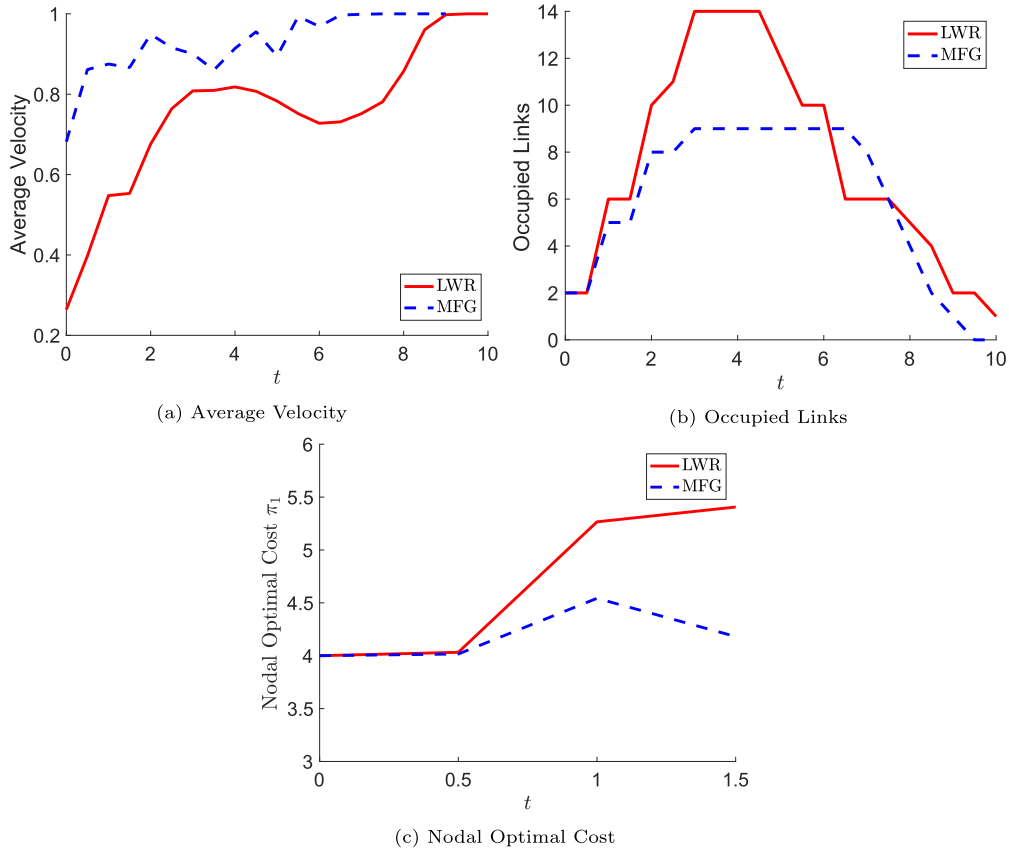


Fig. 13. Comparison between MFG and LWR equilibria on the OW network.

the efficiency of the solution algorithm.

This work can be extended as follows. (1) The proposed MFG system is defined for the many-to-one scenario. A multi-class MFG can be developed to accommodate multiple destinations. (2) The equilibrium existence is given for a particular network and a family of cost functions. We would like to extend the theoretical analysis to more general networks and cost functions. (3) We assume all cars to be AVs. When both AVs and human-driven vehicles co-exist on a transportation network, a multi-class mean field game can be developed to study the mixed traffic equilibrium. Given the goal of this paper is the first-of-its-kind to lay the theoretical foundation for a scalable control framework of AVs' driving and routing game on a network, all the aforementioned analytical and computational issues will be left for future research.

#### CRedit authorship contribution statement

**Kuang Huang:** Conceptualization, Formal analysis, Methodology, Visualization, Writing - original draft. **Xu Chen:** Validation, Visualization. **Xuan Di:** Conceptualization, Methodology, Supervision, Writing - review & editing. **Qiang Du:** Supervision, Writing - review & editing.

#### Acknowledgments

The authors would like to thank Data Science Institute from Columbia University for providing a seed grant for this research. The third author acknowledges support from the National Science Foundation under the award number CMMI-1943998. The fourth author acknowledges support from the National Science Foundation under the award numbers CCF-1704833 and DMS-2012562.

#### References

- Ban, X.-J., Liu, H.-X., Ferris, M.C., Ran, B., 2008. A link-node complementarity model and solution algorithm for dynamic user equilibria with exact flow propagations. *Transport. Res. Part B: Methodol.* 42 (9), 823–842.
- Ban, X.-J., Pang, J.-S., Liu, H.-X., Ma, R., 2012. Continuous-time point-queue models in dynamic network loading. *Transport. Res. Part B: Methodol.* 46 (3), 360–380.
- Batista, S., Leclercq, L., Menéndez, M., 2021. Dynamic traffic assignment for regional networks with traffic-dependent trip lengths and regional paths. *Transport. Res. Part C: Emerg. Technol.* 127, 103076.
- Bauso, D., Zhang, X., Papachristodoulou, A., 2016. Density flow in dynamical networks via mean-field games. *IEEE Trans. Autom. Control* 62 (3), 1342–1355.

- Bliemer, M.C.J., Raadsen, M.P.H., Brederode, L.J.N., Bell, M.G.H., Wismans, L.J.J., Smith, M.J., 2017. Genetics of traffic assignment models for strategic transport planning. *Transport. Res. Part B: Methodol.* 37 (1).
- Boyce, D.E., Ran, B., Leblanc, L.J., 1995. Solving an instantaneous dynamic user-optimal route choice model. *Transport. Sci.* 29 (2), 128–142.
- Burger, M., Di Francesco, M., Markowich, P., Wolfram, M.-T., 2013. Mean field games with nonlinear mobilities in pedestrian dynamics. *arXiv preprint arXiv:1304.5201*.
- Cardaliaguet, P., 2010. Notes on mean field games. *Tech. Rep.*
- Cardaliaguet, P., 2015. Weak solutions for first order mean field games with local coupling. In: *Analysis and Geometry in Control theory and its Applications*. Springer, pp. 111–158.
- Carey, M., Ge, Y., 2003. Comparing whole-link travel time models. *Transport. Res. Part B: Methodol.* 37 (10), 905–926.
- Carey, M., McCartney, M., 2002. Behaviour of a whole-link travel time model used in dynamic traffic assignment. *Transport. Res. Part B: Methodol.* 36 (1), 83–95.
- Chen, Z., He, F., Yin, Y., 2016. Optimal deployment of charging lanes for electric vehicles in transportation networks. *Transport. Res. Part B: Methodol.* 91, 344–365.
- Chevalier, G., Le Ny, J., Malhamé, R., 2015. A micro-macro traffic model based on mean-field games. In: *2015 American Control Conference (ACC)*. IEEE, pp. 1983–1988.
- Chiu, Y.-C., Khani, A., Noh, H., Bustillos, B., Hickman, M., 2013. Technical report on shrp 2 c10b version of dynust and fast-trips. *Tech. Rep.*
- Couillet, R., Perlaza, S.M., Tembine, H., Debbah, M., 2012. Electrical vehicles in the smart grid: A mean field game analysis. *IEEE J. Sel. Areas Commun.* 30 (6), 1086–1096.
- Degond, P., Liu, J.-G., Ringhofer, C., 2014. Large-scale dynamics of mean-field games driven by local nash equilibria. *J. Nonlinear Sci.* 24 (1), 93–115.
- Di, Xuan, Ban, Xuegang (Jeff), 2019. A unified equilibrium framework of new shared mobility systems. *Transp. Res. Part B* 129, 50–78.
- Di, X., He, X., Guo, X., Liu, H.X., 2014. Braess paradox under the boundedly rational user equilibria. *Transp. Res. Part B* 67, 86–108.
- Di, X., Liu, H.X., 2016. Boundedly rational route choice behavior: A review of models and methodologies. *Transp. Res. Part B* 85, 142–179.
- Di, X., Liu, H.X., Pang, J.-S., Ban, X.J., 2013. Boundedly rational user equilibria (BRUE): Mathematical formulation and solution sets. *Transp. Res. Part B* 57, 300–313.
- Djehiche, B., Techeukam, A., Tembine, H., 2016. Mean-field-type games in engineering. *arXiv preprint arXiv:1605.03281*.
- Di, Xuan, Shi, Rongye, 2021. A survey on autonomous vehicle control in the era of mixed-autonomy: from physics-based to AI-guided driving policy learning. *Transp. Res. Part C* 125. <https://doi.org/10.1016/j.trc.2021.103008>.
- Doan, K., Ukkusuri, S.V., 2015. Dynamic system optimal model for multi-od traffic networks with an advanced spatial queuing model. *Transport. Res. Part C: Emerg. Technol.* 51, 41–65.
- Dockner, E.J., Jorgensen, S., Van Long, N., Sorger, G., 2000. *Differential Games in Economics and Management Science*. Cambridge University Press.
- Facchinei, F., Pang, J.-S., 2007. *Finite-dimensional variational inequalities and complementarity problems*. Springer Science & Business Media.
- Festa, A., Göttlich, S., 2017. A mean field games approach for multi-lane traffic management. *arXiv preprint arXiv:1711.04116*.
- Friesz, T.L., Bernstein, D., Smith, T.E., Tobin, R.L., Wie, B.-W., 1993. A variational inequality formulation of the dynamic network user equilibrium problem. *Oper. Res.* 41 (1), 179–191.
- Friesz, T.L., Han, K., 2019. The mathematical foundations of dynamic user equilibrium. *Transport. Res. Part B: Methodol.* 126, 309–328.
- Friesz, T.L., Han, K., Neto, P.A., Meimand, A., Yao, T., 2013. Dynamic user equilibrium based on a hydrodynamic model. *Transport. Res. Part B: Methodol.* 47, 102–126.
- Gentile, G., 2015. Using the general link transmission model in a dynamic traffic assignment to simulate congestion on urban networks. *Transport. Res. Procedia* 5, 66–81.
- Gentile, G., Meschini, L., Papola, N., 2007. Spillback congestion in dynamic traffic assignment: a macroscopic flow model with time-varying bottlenecks. *Transport. Res. Part B: Methodol.* 41 (10), 1114–1138.
- Guéant, O., Lasry, J.-M., Lions, P.-L., 2011. Mean field games and applications. In: *Paris-Princeton Lectures on Mathematical Finance 2010*. Springer, pp. 205–266.
- Gummadi, R., Johari, R., Yu, J.Y., 2012. Mean field equilibria of multi armed bandit games. In: *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, pp. 1110–1110.
- Han, K., Eve, G., Friesz, T.L., 2019. Computing dynamic user equilibria on large-scale networks with software implementation. *Networks Spatial Econ.* 1–34.
- Han, K., Friesz, T.L., Yao, T., 2013. A partial differential equation formulation of vickrey's bottleneck model, part i: Methodology and theoretical analysis. *Transport. Res. Part B: Methodol.* 49, 55–74.
- Han, K., Piccoli, B., Friesz, T.L., 2016. Continuity of the path delay operator for dynamic network loading with spillback. *Transport. Res. Part B: Methodol.* 92, 211–233.
- Huang, K., Di, X., Du, Q., Chen, X., 2019. Stabilizing traffic via autonomous vehicles: A continuum mean field game approach. In: *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, pp. 3269–3274.
- Huang, K., Di, X., Du, Q., Chen, X., 2020a. A game-theoretic framework for autonomous vehicles velocity control: Bridging microscopic differential games and macroscopic mean field games. *Discrete Continuous Dyn. Sys.- Series B* 25 (12), 4869–4903.
- Huang, K., Di, X., Du, Q., Chen, X., 2020b. Scalable traffic stability analysis in mixed-autonomy using continuum models. *Transport. Res. Part C: Emerg. Technol.* 111, 616–630.
- Iyer, K., Johari, R., Sundararajan, M., 2014. Mean field equilibria of dynamic auctions with learning. *Manage. Sci.* 60 (12), 2949–2970.
- Kachroo, P., Agarwal, S., Piccoli, B., Özbay, K., 2017. Multiscale modeling and control architecture for v2x enabled traffic streams. *IEEE Trans. Veh. Technol.* 66 (6), 4616–4626.
- Kachroo, P., Agarwal, S., Sastry, S., 2016. Inverse problem for non-viscous mean field control: Example from traffic. *IEEE Trans. Autom. Control* 61 (11), 3412–3421.
- Kuwahara, M., Akamatsu, T., 2001. Dynamic user optimal assignment with physical queues for a many-to-many od pattern. *Transport. Res. Part B: Methodol.* 35 (5), 461–479.
- Lachapelle, A., Salomon, J., Turinici, G., 2010. Computation of mean field equilibria in economics. *Mathe. Models Methods Appl. Sci.* 20 (04), 567–588.
- Lachapelle, A., Wolfram, M.-T., 2011. On a mean field game approach modeling congestion and aversion in pedestrian crowds. *Transport. Res. Part B: Methodol.* 45 (10), 1572–1589.
- Lam, W.H., Huang, H.-J., 1995. Dynamic user optimal traffic assignment model for many to one travel demand. *Transport. Res. Part B: Methodol.* 29 (4), 243–259.
- Lebacque, J.-P., Khoshyaran, M., 1999. Modelling vehicular traffic flow on networks using macroscopic models. *Finite Vol. Complex Appl. II* 551–558.
- LeVeque, R.J., 2002. *Finite volume methods for hyperbolic problems*. vol. 31. Cambridge University Press.
- LeVeque, R.J., 2007. *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems*. SIAM.
- Levin, M.W., Boyles, S.D., 2016. A multiclass cell transmission model for shared human and autonomous vehicle roads. *Transport. Res. Part C: Emerg. Technol.* 62, 103–116.
- Lighthill, M.J., 1952. On sound generated aerodynamically I. general theory. *Proc. R. Soc. Lond. A* 211 (1107), 564–587.
- Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves II. A theory of traffic flow on long crowded roads. *Proc. R. Soc. Lond. A* 229 (1178), 317–345.
- Lo, H., 1999. A dynamic traffic assignment formulation that encapsulates the cell-transmission model. In: *14th International Symposium on Transportation and Traffic Theory/Transportation Research Institute*.
- Lo, H.K., Szeto, W.Y., 2002. A cell-based variational inequality formulation of the dynamic user optimal assignment problem. *Transport. Res. Part B: Methodol.* 36 (5), 421–443.
- Mahmassani, H.S., 2001. Dynamic network traffic assignment and simulation methodology for advanced system management applications. *Networks Spatial Econ.* 1 (3–4), 267–292.
- Merchant, D.K., Nemhauser, G.L., 1978a. A model and an algorithm for the dynamic traffic assignment problems. *Transport. Sci.* 12 (3), 183–199.
- Merchant, D.K., Nemhauser, G.L., 1978b. Optimality conditions for a dynamic traffic assignment model. *Transport. Sci.* 12 (3), 200–207.
- Nie, X., Zhang, H., 2005a. Delay-function-based link models: their properties and computational issues. *Transport. Res. Part B: Methodol.* 39 (8), 729–751.
- Nie, X., Zhang, H.M., 2005b. A comparative study of some macroscopic link models used in dynamic traffic assignment. *Networks Spatial Econ.* 5 (1), 89–115.



- Osorio, C., Flötteröd, G., Bierlaire, M., 2011. Dynamic network loading: a stochastic differentiable model that derives link state distributions. *Transport. Res. Part B: Methodol.* 45 (9), 1410–1423.
- Peeta, S., Ziliaskopoulos, A.K., 2001. Foundations of dynamic traffic assignment: The past, the present and the future. *Networks Spatial Econ.* 1 (3–4), 233–265.
- Ran, B., Boyce, D.E., 1996. A link-based variational inequality formulation of ideal dynamic user-optimal route choice problem. *Transport. Res. Part C: Emerg. Technol.* 4 (1), 1–12.
- Ran, B., Hall, R.W., Boyce, D.E., 1996. A link-based variational inequality model for dynamic departure time/route choice. *Transport. Res. Part B: Methodol.* 30 (1), 31–46.
- Richards, P.I., 1956. Shock waves on the highway. *Oper. Res.* 4 (1), 42–51.
- Rosenthal, R.E., 2007. Gams: A user's guide. GAMS Development Corporation.
- Shou, Zhenyu, Di, Xuan, 2020. Multi-agent reinforcement learning for dynamic routing games: a unified paradigm. arXiv doi:2011.10915. Submitted for publication.
- Szeto, W., Lo, H.K., 2004. A cell-based simultaneous route and departure time choice model with elastic demand. *Transport. Res. Part B: Methodol.* 38 (7), 593–612.
- Ukkusuri, S.V., Han, L., Doan, K., 2012. Dynamic user equilibrium with a path based cell transmission model for general traffic networks. *Transport. Res. Part B: Methodol.* 46 (10), 1657–1684.
- Vickrey, W.S., 1969. Congestion theory and transport investment. *Am. Econ. Rev.* 59 (2), 251–260.
- Wie, B.-W., Friesz, T.L., Tobin, R.L., 1990. Dynamic user optimal traffic assignment on congested multidestination networks. *Transport. Res. Part B: Methodol.* 24 (6), 431–442.
- Xu, Y., Wu, J.H., Florian, M., Marcotte, P., Zhu, D., 1999. Advances in the continuous dynamic network loading problem. *Transport. Sci.* 33 (4), 341–353.
- Yu, Y., Han, K., Ochieng, W., 2020. Day-to-day dynamic traffic assignment with imperfect information, bounded rationality and information sharing. *Transport. Res. Part C: Emerg. Technol.* 114.
- Zhang, H., Nie, Y., Qian, Z., 2013. Modelling network flow with and without link interactions: the cases of point queue, spatial queue and cell transmission model. *Transportmetrica B: Transport Dyn.* 1 (1), 33–51.
- Zhu, F., Ukkusuri, S.V., 2015. A linear programming formulation for autonomous intersection control within a dynamic traffic assignment and connected vehicle environment. *Transport. Res. Part C: Emerg. Technol.* 55, 363–378.