Transportation Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Game-Theoretic Framework for Generic Second-Order Traffic Flow Models Using Mean Field Games and Adversarial Inverse Reinforcement Learning

Zhaobin Mo, Xu Chen, Xuan Di, Elisa Iacomini, Chiara Segala, Michael Herty, Mathieu Lauriere

To cite this article:
Zhaobin Mo, Xu Chen, Xuan Di, Elisa Iacomini, Chiara Segala, Michael Herty, Mathieu Lauriere (2024) A Game-Theoretic Framework for Generic Second-Order Traffic Flow Models Using Mean Field Games and Adversarial Inverse Reinforcement Learning. Transportation Science 58(6):1403-1426. https://doi.org/10.1287/trsc.2024.0532

https://pubsonline.informs.org/journal/trsc

TRANSPORTATION SCIENCE
Vol. 58, No. 6, November–December 2024, pp. 1403–1426
ISSN 0041-1655 (print), ISSN 1526-5447 (online)

A Game-Theoretic Framework for Generic Second-Order Traffic Flow Models Using Mean Field Games and Adversarial Inverse Reinforcement Learning

Zhaobin Mo,[a] Xu Chen,[a] Xuan Di,[a,b,*] Elisa Iacomini,[c] Chiara Segala,[d] Michael Herty,[d] Mathieu Lauriere[e]

[a] Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, New York 10027; [b] Data Science Institute, Columbia University, New York, New York 10027; [c] Mathematics and Computer Science Department, University of Ferrara, 44121 Ferrara, Italy; [d] Institut für Geometrie und Praktische Mathematik, RWTH Aachen University, 52062 Aachen, Germany; [e] Institute of Mathematical Sciences, New York University, Shanghai 200122, China

*Corresponding author

Contact: zm2302@columbia.edu, https://orcid.org/0000-0002-0465-8550 (ZM); xc2412@columbia.edu, https://orcid.org/0000-0002-1006-0926 (XC); sharon.di@columbia.edu, https://orcid.org/0000-0003-2925-7697 (XD); elisa.iacomini@unife.it, https://orcid.org/0000-0002-0981-2086 (EI); segala@igpm.rwth-aachen.de, https://orcid.org/0000-0002-6480-3772 (CS); herty@igpm.rwth-aachen.de (MH); mathieu.lauriere@gmail.com (ML)

Received: January 20, 2024
Revised: June 16, 2024
Accepted: June 30, 2024
Published Online in Articles in Advance: August 20, 2024

https://doi.org/10.1287/trsc.2024.0532

Abstract. A traffic system can be interpreted as a multiagent system, wherein vehicles choose the most efficient driving approaches guided by interconnected goals or strategies. This paper aims to develop a family of mean field games (MFG) for generic second-order traffic flow models (GSOM), in which cars control individual velocity to optimize their objective functions. GSOMs do not generally assume that cars optimize self-interested objectives, so such a game-theoretic reinterpretation offers insights into the agents' underlying behaviors. In general, an MFG allows one to model individuals on a microscopic level as rational utility-optimizing agents while translating rich microscopic behaviors to macroscopic models. Building on the MFG framework, we devise a new class of second-order traffic flow MFGs (i.e., GSOM-MFG), which control cars' acceleration to ensure smooth velocity change. A fixed-point algorithm with fictitious play technique is developed to solve GSOM-MFG numerically. In numerical examples, different traffic patterns are presented under different cost functions. For real-world validation, we further use an inverse reinforcement learning approach (IRL) to uncover the underlying cost function on the next-generation simulation (NGSIM) data set. We formulate the problem of inferring cost functions as a min-max game and use an apprenticeship learning algorithm to solve for cost function coefficients. The results show that our proposed GSOM-MFG is a generic framework that can accommodate various cost functions. The Aw Rascle and Zhang (ARZ) and Light-Whitham-Richards (LWR) fundamental diagrams in traffic flow models belong to our GSOM-MFG when costs are specified.

History: This paper has been accepted for the *Transportation Science* Special Issue on ISTTT25 Conference.
Funding: X. Di is supported by the National Science Foundation [CAREER Award CMMI-1943998]. E. Iacomini is partially supported by the Italian Research Center on High-Performance Computing, Big Data and Quantum Computing (ICSC) funded by MUR Missione 4-Next Generation EU (NGEU) [Spoke 1 "FutureHPC & BigData"]. C. Segala and M. Herty thank the Deutsche Forschungsgemeinschaft (DFG) for financial support [Grants 320021702/GRK2326, 333849990/IRTG-2379, B04, B05, and B06 of 442047500/SFB1481, HE5386/18-1,19-2,22-1,23-1,25-1, ERS SFDdM035; Germany's Excellence Strategy EXC-2023 Internet of Production 390621612; and Excellence Strategy of the Federal Government and the Länder]. Support through the EU DATAHYKING is also acknowledged. This work was also funded by the DFG [TRR 154, Mathematical Modelling, Simulation and Optimization Using the Example of Gas Networks, Projects C03 and C05, Project No. 239904186]. Moreover, E. Iacomini and C. Segala are members of the Indam GNCS (Italian National Group of Scientific Calculus).

Keywords: mean field game (MFG) • generic second order traffic flow model • adversarial inverse reinforcement learning (AIRL)

1. Introduction

Traffic flow models are an indispensable tool for urban and suburban traffic management. However, the existing traffic flow models are developed for human drivers. With the advent of autonomous vehicles (AV), the characteristics of traffic flow could be transformed if AVs are designed to drive differently from humans. For example, human driving could be unstable resulting in stop-and-go waves, whereas AVs could be designed ex ante to stabilize traffic. Such a microscopic behavioral

change calls for new macroscopic models to predict traffic flow in the era of autonomy. In this paper, we aim to develop a game-theoretic counterpart of traffic flow models, which allows us to devise the payoff or cost functions for cars on a microscopic scale and transform such a behavior to macroscopic traffic characteristics.

1.1. Motivation

Mean field game (MFG) is a game-theoretic framework that has recently gained growing popularity to design decision making processes in many-agent dynamical systems (Huang, Malhamé, Caines 2006, Lasry and Lions 2007). It has been shown that the first-order traffic flow models can be reinterpreted as an MFG (Huang et al. 2020a). In other words, classical traffic flow models that depict how traffic states (represented by flux, density, and velocity) propagate in time and space are a special class of differential games, in which each car solves an optimal control problem with an objective function, whereas others do so simultaneously. All the cars interact with one another on a microscopic scale through coupled objective functions. When the number of cars becomes large, the macroscopic traffic state evolution is depicted by an equilibrium of the corresponding MFG. However, the existing literature has primarily focused on the first-order traffic flow MFGs, which assume that each car controls its velocity instead of acceleration. Directly controlling velocity could result in sudden changes in speed such as an infinite changing rate of speed, which is unrealistic. Thus, the key research question is, *whether the generic second-order traffic flow models (GSOM)* (Lebacque, Mammar, and Salem 2007) *can be reformulated as MFGs*. It is desirable to establish a second-order mean field game-theoretic framework that enables optimal control of acceleration at a micro level, while scaling up to the tempo-spatial evolution of macroscopic traffic flow quantities. Reformulating GSOM with MFGs would help design new driving acceleration decision making processes for individual cars that could potentially result in smooth traffic flow. More importantly, building on the counterpart of GSOM in MFGs, we could further develop new second-order traffic flow models that encompass desirable mathematical properties.

This paper aims to develop a family of second-order traffic flow MFGs that control vehicle acceleration. First, an equivalence between the classical GSOM, Aw Rascle and Zhang (ARZ) (Aw and Rascle 2000, Zhang 2002) in particular, and MFGs will be established, denoted as GSOM-MFG. Aligned with the conclusion that the GSOM is a family of parametrized first-order models, we discover that the second-order traffic flow MFG is also a family of parametrized first-order traffic flow MFGs. Thus, building on a cost function leading to the ARZ model solution, a family of new second-order traffic flow MFGs, that is, GSOM-MFG, will be developed. We will explore on how the design of various objective functions demonstrate different traffic flow evolution. Numerical examples will be provided to demonstrate the properties of the new games. Moreover, we will establish an adversarial inverse reinforcement learning scheme to uncover the objective function latent in real-world data and validate the value of our proposed GSOM-MFG.

1.2. Literature Review

We will first review literature on GSOMs and MFGs, respectively, and then investigate a series of studies that bridge these two topics. The research gap will be identified on controlling acceleration in MFGs.

In classical traffic flow models, a classification can be made between first-order models and second-order models. The former type of models simplify reality by assuming instantaneous accelerations and describing traffic only in terms of equilibrium conditions. This leads to failing in generating capacity drop, hysteresis, relaxation, platoon diffusion, or spontaneous congestion like stop-and-go waves that are typical features of traffic dynamics (Cristiani and Iacomini 2019, Göttlich, Iacomini, and Jung 2020, Balzotti and Iacomini 2021). To overcome these issues, second-order models have been proposed (Aw and Rascle 2000, Zhang 2002, Lebacque, Mammar, and Salem 2007). They take into account the nonequilibria states, assuming that accelerations are not instantaneous. To do this, the equation that describes the variation of the velocity in time has to be added to the system, replacing the typical given law of the first order models. In this work, we will focus on the generalized second-order models GSOMs that encompass a family of existing models, including LWR and ARZ. The typical GSOM is formulated as

$$[\text{GSOM}] \quad \begin{cases} \rho_t + (\rho u)_x = 0, \\ (\rho\omega)_t + (\rho\omega\, u)_x = \rho\, r(\rho, u, \omega), \\ u = U(\rho, \omega). \end{cases} \tag{1.1}$$

Here, $\rho := \rho(t,x)$ stands for traffic density, $u := u(t,x)$ the speed, and $\omega := \omega(t,x)$ the invariant or Langrangian marker; $r(\cdot)$ is the general relaxation function, and $U(\cdot)$ the velocity function linked to fundamental diagrams.

The GSOM system (1.1) has been studied on variational formulation (Lebacque and Khoshyaran 2013, Li and Zhang 2013, Costeseque and Lebacque 2014), fundamental diagram construction (Seibold et al. 2013), empirical validation (Fan, Herty, and Seibold 2014, Yu, Bayen, and Krstic 2019), vehicle control (Delle Monache, Piccoli, and Rossi 2017, Chiarello, Piccoli, and Tosen 2021, Gong, Piccoli, and Visconti 2021), traffic signal control (Khelifi et al. 2016), and junction modeling (Costeseque, Lebacque, and Khelifi 2015).

MFG is a game-theoretic framework to model complex multiagent dynamics arising from the interactions

of a large population of rational utility-optimizing agents, whose dynamical behaviors are characterized by optimal control problems (Huang, Malhamé, Caines 2006, Lasry and Lions 2007, Cannarsa, Capuani, and Cardaliaguet 2021). By exploiting the "smoothing" effect of a large number of interacting individuals, MFG assumes that each agent only responds to and contributes to the density distribution of the whole population. At mean field equilibria (MFE), an agent's optimal strategy coincides with the population density, characterized by two coupled partial differential equations (PDEs):

1. A backward Hamilton-Jacobi-Bellman (HJB) equation (for representative agent dynamic): Given the density evolution of the population, each agent solves an optimal control problem on a predefined time horizon to reach a minimal cost. For a generic agent, the optimal control problem can be solved by dynamic programming that derives an HJB equation. The equation is solved backward in time.

2. A forward Fokker-Planck equation (for population dynamic): Given individual controls, the population's density evolution resulting from all agents' dynamics is described by a Fokker-Planck equation. The equation is solved forward in time.

MFE is generally challenging to solve due to its forward-backward structure. The numerical methods of solving MFE include fixed-point, Newton's method, and the variational method (Achdou, Camilli, and Capuzzo-Dolcetta 2012, Benamou and Carlier 2015, Chow et al. 2018, Albi et al. 2022, Capuani and Marigonda 2022). The numerical methods require a good initial guess and could fail to converge when the cost structure of the game is complex. Thus, learning based methods, especially neural network based approximation, have emerged in recent years (Guo et al. 2019, Lauriere et al. 2022, Shou et al. 2022, Chen, Liu, and Di 2023a, Fiedler et al. 2023, Zhou et al. 2024). Learning based methods can learn MFEs with complex cost functions and high-dimensional states and policies but could take a long time to train. Thus, in this paper, we will primarily focus on numerical methods, and these methods will be revisited later.

MFG has demonstrated its benefits in modeling dynamic decision making processes of many agents and has become increasingly popular in finance (Lachapelle, Salomon, and Turinici 2010, Guéant, Lasry, and Lions 2011), control (Djehiche, Tcheukam, and Tembine 2016), crowd motion (Lachapelle and Wolfram 2011), autonomous driving (Huang et al. 2020a, 2021; Di and Shi 2021), and mixed traffic stability (Huang et al. 2019, 2020b).

Regarding the linkage between MFGs and classical traffic flow models, the Lighthill-Whitham-Richards (LWR) model (Lighthill and Whitham 1955, Richards 1956) was shown to be a special MFG (Kachroo, Agarwal, and Sastry 2016, Huang et al. 2020a). Chevalier, Le Ny, and Malhamé (2015) generalized the cost function in Kachroo, Agarwal, and Sastry (2016) to one global-in-time and global-in-space to model AVs. MFG was further extended to model lane-change (Festa and Göttlich 2017) and routing on networks (Cristiani and Priuli 2014, Huang et al. 2021, Chen, Liu, and Di 2023b).

To the best of our knowledge, a majority of traffic flow MFG models primarily focus on velocity control. MFGs on acceleration were only proposed recently in Achdou et al. (2020, 2021). The state contains a pair of position and velocity. A separable cost function is used consisting of two terms, a kinetic energy (i.e., the square of velocity) and the square of acceleration. The dynamics of the generic agent is a double integrator. The challenge lies in the state constraint and boundedness to velocity (i.e., part of the state), which leads to neither strictly convex nor coercive Hamilton and potentially unbounded value functions. Achdou et al. (2021) demonstrated that by imposing additional assumptions to the support of the initial state distribution or cost function, the optimal trajectories could form a compact set and guarantee the existence of an equilibrium. In particular, mathematical properties of equilibria can be ensured in a one-dimensional problem with a quadratic running cost function in acceleration. Nevertheless, it still remains unclear as to how these games can be applied to the traffic setting, given that traffic flow has to satisfy certain physics constraints, thus allowing for a narrower set of cost functions. Overall, MFGs on acceleration is still a nascent field that entails a lot of open questions and challenges.

1.3. Contributions of this Paper

This paper aims to develop a game-theoretic framework for GSOM using mean-field approximation. We will start from identifying a counterpart of ARZ-like MFGs, with a physically meaningful cost functions. Building on the MFG formulation of ARZ models, we will modify the cost function and develop new second-order traffic flow MFGs.

In a nutshell, the contributions of this paper include the following:

- Establish a linkage between GSOM and MFG and offer a game-theoretical interpretation of GSOMs.
- Propose a broader class of second-order traffic flow models by manipulating the objective function of GSOM-MFG. Various cost functions are applied on a ring road to demonstrate various traffic pattern evolution.
- Develop a fixed-point algorithm to solve the MFE of GSOM-MFG. Fictitious play technique is applied to stabilize the solution algorithm.

The remainder of the paper is organized as follows. Section 2 states the problem to be solved in this paper. In Section 3, we derive the GSOM-MFG, including continuity equation and HJB equation. In Section 4, we develop a family of second-order traffic flow models of

which objectives are modifications of those for GSOM-MFG to mitigate traffic flow. In Section 5, we develop a numerical solution algorithm to solve the proposed GSOM-MFGs. In Section 6, we aim to validate the proposed framework with inverse reinforcement learning on both synthetic and real-world data. Conclusions and future research directions follow in Section 7.

2. Problem Statement

In this section, we will first revisit the existing first-order traffic flow MFG developed in Huang et al. (2020a) and then establish the problem statement for the second-order traffic flow MFG, the focus of this paper.

We first make the following assumptions for modeling MFGs:

- Each car observes traffic state information on the road from all others.
- Each car plans its velocity control in a time horizon by anticipating others' behaviors.
- Cars act to use their predefined driving costs on the time horizon in a noncooperative way.

The assumptions for taking mean field approximation include the following:

- All cars are indistinguishable.
- All cars have the same form of cost function.

2.1. First-Order Traffic Flow MFG

2.1.1. First-Order *N*-Car Differential Game [DG1]. We consider a time horizon $[0,T]$ (where $0 < T < \infty$ is a finite time horizon), and a platoon of cars indexed by $i \in \{1,2,\dots,N\}$, where N is the total number of cars. The cars are driving in one direction on a closed highway of length L without any entrance nor exit, with initial positions $x_{1,0},\dots,x_{N,0}$, and evolving according to a given velocity function.

Each car aims to select its optimal velocity control by minimizing its driving cost functional defined over $[0,T]$ as

$$J_i^N(x,v_i) = \int_0^T \underbrace{f_i^N(x(t),v_i(t))}_{\text{running cost}}\,dt + \underbrace{V_T(x(T))}_{\text{terminal cost}}, \quad i=1,\dots,N,$$

where for any $i = 1,2,\dots,N$, we suppose car i knows other cars' positions:

$$x(t) = [x_1(t),\dots,x_{i-1}(t),x_i(t),x_{i+1}(t),\dots,x_N(t)]^T.$$

When one car selects its own driving velocity over the planning horizon while everybody else does so simultaneously, we have a so called noncooperative *differential game*. A Nash equilibrium of the N-car differential game is a tuple of controls $v_1^*(t),v_2^*(t),\dots,v_N^*(t)$ satisfying

$$J_i^N(x^*,v_i^*) \le J_i^N(x^*,v_i),\ i=1,\dots,N.$$

At equilibrium, no car can improve its driving cost by unilaterally switching its velocity control.

2.1.2. First-Order Traffic Flow Mean Field Game [LWR-MFG]. If the number of cars goes to infinity $N \to \infty$ in [DG1], we recover the MFG game for the first-order LWR model. In the MFG setting, a representative agent, starting from an initial position x, selects driving speed $v(t), t \in (0,T]$ at time t for the entire time horizon $(0,T]$ to minimize a cumulative cost $f(\cdot)$ with a terminal cost $V_T(\cdot)$. The cost function depends on not only the control of the representative agent $v(t)$, but also traffic density $\rho(t,x)$ at time t and location x. Thus, to solve this optimal control problem, $\rho(t,x)$ is assumed to be known. Accordingly, this problem can be formulated as an optimal control with the constraints that are an ODE for a representative agent and a PDE that is the conservation law or continuity equation (CE) governing the traffic density $\rho(t,x)$:

[Optimal Control of the Representative Agent]

$$u(t,x) = \arg\min_{v(t)} \int_0^T f(v(t);\rho(t,x))\,dt + V_T(x(T)),$$

$$\text{s.t.} \begin{cases} \dot{x}(t) = v(t), & \forall t \in [0,T] \\ x(0) = x, \end{cases} \tag{2.1a}$$

$$\text{[LWR-CE]} \begin{cases} \rho(t,x)_t + (\rho(t,x)u(t,x))_x = 0, \\ \rho(0,x) = \rho_0(x). \end{cases} \tag{2.1b}$$

By referring to Huang et al. (2020a) as a source, we derive the HJB equation to determine the velocity:

$$\text{[LWR-HJB]} \begin{cases} V_t(t,x) + f^*(V_x(t,x),\rho(t,x)) = 0, \\ u(t,x) = f_{\mathcal{V}}^*(V_x(t,x),\rho(t,x)), \end{cases}$$

where the function $f^*(\cdot)$ is the the Legendre transform of $f(\cdot)$ and $f_{\mathcal{V}}^*$ is the derivative with respect to the first argument of f^*. Denote $\mathcal{V} = V_x$. The HJB is meant to be integrated backward in time from the terminal cost $V(x,T) = V_T(x)$. The equilibrium solution of this mean-field game is an optimal velocity field $u^*(x,t)$ and the corresponding optimal density $\rho^*(x,t)$.

We can substitute [LWR-HJB] into [Representative agent's optimal control] and have the LWR-MFG system written in PDEs:

$$\text{[LWR-MFG]} \quad \begin{cases} \text{(CE)} & \rho_t + (\rho u)_x = 0, \\ \text{(HJB)} & V_t + f^*(V_x,\rho) = 0, \\ & u(t,x) = f_{\mathcal{V}}^*(V_x,\rho). \end{cases} \tag{2.2}$$

Here, $\mathcal{V} = V_x$.

2.2. From First- to Second-Order Traffic Flow MFGs

Following a similar procedure as above, we start from a microscopic description to recover later the continuity equations and the HJB equation for the second order case.

2.2.1. Acceleration-Based *N*-Car Differential Game [DG2a]. Also in this case, we consider a total of N cars indexed by $i \in \{1,2,\dots,N\}$ that are running in one direction along a highway of length L. Denote the ith car's position by $x_i(t)$, speed by $v_i(t)$, and acceleration by $a_i(t)$.

Each car aims to control its velocity via selection of optimal acceleration by minimizing its driving cost functional $J(x,v,a_i)$ over a predefined planning horizon $[0,T]$.

A Nash equilibrium of the acceleration-based N-car differential game [DG2a] is a tuple of controls $a_1^*(t)$, $a_2^*(t),\dots,a_N^*(t)$ satisfying

$$J_i^N(x^*,v^*,a_i^*) \le J_i^N(x^*,v^*,a_i), \quad i=1,\dots,N.$$

2.2.2. Parametrized *N*-Car Differential Game [DG2p]. Taking inspiration from the ideas presented in the cited works, namely, Aw and Rascle (2000), Aw et al. (2002), and Lebacque, Mammar, and Salem (2007), we adopt a different approach to handle the acceleration term in the model. Instead of directly dealing with it, we focus on a specific characteristic of drivers, which could be for example their free flow velocity. By considering this attribute, we are able to parameterize both the model and the corresponding mean field game.

This choice of parameterization has several advantages. First, it allows our model to be in line with the existing body of research and knowledge within the traffic community. It ensures that our formulation aligns with established theories and findings in the field. Second, this approach provides a natural extension of the previously mentioned reference (referred to as [DG1]). By incorporating the Lagrangian marker as a parameter, we enhance the flexibility and applicability of the model, enabling it to capture a wider range of traffic scenarios and dynamics.

In the subsequent analysis, we classify vehicles based on their specific attributes, with each category represented by the corresponding class denoted as w. This classification scheme helps us better understand and describe the behavior of vehicles within the context of the model, allowing for more detailed and nuanced analysis of the system.

Assume car i belongs to class $w \in \mathcal{W} \subseteq \mathbb{R}$. The motion of car i in class w over $[0,T]$ is dictated by the following dynamical system for $i=1,2,\dots,N$ (where we assume that there is a functional relation between v_i, w_i and x_i that is specified later):

$$\begin{cases} \dot{x}_i(t) = v_i(t), & x_i(0) = x_{i,0}, \\ \dot{w}_i(t) = r(x,v_i,w_i), & w_i(0) = w_{i,0}. \end{cases}$$

Note here that $\dot{w}_i(t)$ somehow describes the acceleration of car i. Relations between w_i and v_i have been discussed and are known as *generic second-order models* (GSOMs) (Seibold et al. 2013). In the following, we assume the relation

$$v_i = U\left(\frac{1}{x_{i+1}-x_i}, w_i\right). \tag{2.3}$$

We will work with the variables (x_i,v_i,w_i) instead of (x_i,v_i,a_i).

To select a velocity profile, the ith car solves a parametrized optimal control problem over $[0,T]$. Define the new ith car's *driving cost functional* as

$$J_i^N(x,v_i,w_i) = \int_0^T f_i^N(x(t),v_i(t),w_i(t))dt + V_T(x_i(T),w_i(T)).$$

The output of the above optimal control problem is denoted as a best velocity control $v_i^*(t), t \in [0,T]$ for car i of class w. The control of car i depends on cars' velocity of all other classes.

Example 2.1. In the ARZ microscopic model, the acceleration is given by

$$a_i(t) = \dot{v}_i(t) = \frac{v_{i+1}(t)-v_i(t)}{(x_{i+1}(t)-x_i(t))^{\gamma+1}} + r(x,v_i,w_i),$$

for $\gamma \ge 1$ and where

$$r(x,v_i,w_i) = \lambda\left(U\left(\frac{1}{x_{i+1}-x_i}, w_i\right) - v_i\right),$$

with λ a positive constant. We recall that $U(\cdot)$ is the velocity function linked to a fundamental diagram. The relation between v_i and w_i in Equation (2.3) can be expressed, in this particular case, as

$$w_i = v_i + \left(\frac{1}{x_{i+1}-x_i}\right)^{\gamma}.$$

3. GSOM MFGs

In this section, we establish the link from the microscopic second-order traffic model [DG2p] to the macroscopic perspective. This allows us to bridge the gap and transition from the microscopic to the macroscopic level of analysis. By doing so, we can recover the corresponding continuity equation (CE) and HJB.

In preparation for our discussion, it is important to revisit and explore the relationship between the LWR model and the GSOM model. In this way, we can gain a deeper understanding of the connections and similarities between the two traffic flow models.

The LWR model represents traffic flow as a conservation law, where the density and velocity are the primary variables. It assumes a single homogeneous vehicle type and considers the fundamental diagram, which relates traffic density to the velocity and flow rate. To extend the LWR model, we introduced an additional parameter $w \in \mathbb{R}^+$ that accounts for driver behavior and traffic heterogeneity. It allows for variations in parameters such as the free-flow velocity, maximum density, and fundamental diagram shape. These parameters can be adjusted to capture different traffic conditions and driver characteristics. This leads to a continuity equation with parameterized flux that can be seen as the GSOM model where an additional Eulerian variable $\omega = \omega(t,x)$ appears. Indeed, the GSOM model incorporates the concept of traffic heterogeneity and driver behavior by including an extra variable ω, which depends on time and space. By considering this variable, the GSOM model can better capture the evolution of traffic flow. The relationship between the models can be seen as a progression from the basic LWR model to the parametrized LWR model and further to the GSOM model, with increasing levels of complexity and sophistication in modeling driver behavior and traffic heterogeneity.

Remark 3.1. From now on, we will use the following terminologies to refer to w interchangeably, namely, property, class, attribute, or Lagrangian marker.

Building on the LWR-MFG model (2.1), for the second-order case, we will define the optimal control problem for a representative agent and identify the continuity equation.

3.1. Derivation of HJB Equation

As we did for the LWR in Equation (2.1a), we consider the new constrained minimization problem for a representative agent, coupled with the CE of GSOM:

[Optimal Control of the Representative Agent]

$$\begin{aligned}\tilde{u}(t,x,w) = \arg\min_{v(t)} &\int_0^T f(v(t);\rho(t,x),\omega(t,x))\,d\tau \\ &+ V_T(x(T),w(T)),\\ \text{s.t.}&\begin{cases}\dot{x}(t) = v(t),\\ \dot{w}(t) = r(\rho(t,x),v(t),\omega(t,x)),\\ x(0) = x, \quad w(0) = w,\end{cases}\end{aligned} \tag{3.1a}$$

$$\text{[GSOM-CE]}\begin{cases}\rho(t,x)_t + (\rho(t,x)\,u(t,x,w))_x = 0,\\ (\rho(t,x)\omega(t,x))_t\\ \qquad + (\rho(t,x)\omega(t,x)\,u(t,x))_x\\ \qquad = \rho(t,x)\,r(\rho(t,x),u(t,x),\omega(t,x)),\\ \rho(0,x) = \rho_0(x), \quad \omega(0,x) = \omega_0(x).\end{cases} \tag{3.1b}$$

We are now ready to state the theorem that derives the HJB to express the correspondent GSOM model as a MFG.

Theorem 3.1. *Consider the constrained minimization problem (3.1). If we assume the function $f(\rho,v,\omega)$ to be strictly convex with respect to the second argument v, and the relaxation function $r(\rho,u,\omega)$ to be affine linear in u (i.e., $r(\rho,u,\omega) = \alpha u + s(\rho,\omega)$, $\alpha \in \mathbb{R}$), then the corresponding HJB equation can be derived and it reads as*

$$\text{[GSOM-HJB]}\begin{cases}V_t(t,x,w) + f^*(\mathcal{V};\rho(t,x),\omega(t,x))\\ \qquad + s(\rho(t,x),\omega(t,x))\,V_w(t,x) = 0\\ \tilde{u}(t,x,w) = f^*_{\mathcal{V}}(\mathcal{V};\rho(t,x),\omega(t,x)),\\ V(T,x,w) = V_T(x,w),\end{cases} \tag{3.2}$$

where $\mathcal{V} = V_x + \alpha V_w$.

Proof. We define the problem by taking $V(t,x,w)$ to be the optimal cost for a generic car starting from location x and class w at time t:

$$\begin{aligned}V(t,x,w) = \min_v &\left\{\int_t^T f(v(\tau);\rho(\tau,x),\omega(\tau,x))\,d\tau\right\}\\ &+ V_T(x(T),w(T)),\\ \text{s.t.}\quad &\dot{x}(\tau) = v(\tau), \quad x(t) = x,\\ &\dot{w}(\tau) = r(\rho(\tau,x),v(\tau),\omega(\tau,x)), \quad w(t) = w, \quad \tau \in [t,T].\end{aligned} \tag{3.3}$$

Suppose a generic car of class w starts from position x at time t. Consider a small time step Δt, we can divide the driving cost in Equation (3.3) into two parts on $[t,t+\Delta t]$ and $[t+\Delta t,T]$:

$$\begin{aligned}&\int_t^T f(v(\tau);\rho(\tau,x(\tau)),\omega(\tau,x(\tau)))\,d\tau\\ &= \int_t^{t+\Delta t} f(v(\tau);\rho(\tau,x(\tau)),\omega(\tau,x(\tau)))\,d\tau\\ &\quad + \int_{t+\Delta t}^T f(v(\tau);\rho(\tau,x(\tau)),\omega(\tau,x(\tau)))\,d\tau.\end{aligned}$$

Correspondingly, the generic car's decision process is also divided into two stages. First it selects velocity $v(t) = v$ on the horizon $[t,t+\Delta t]$. Then it moves to $x + v\Delta t$ at time $t+\Delta t$ and selects its velocity profile on the rest of the planning horizon $[t+\Delta t,T]$.

The running cost on $[t,t+\Delta t]$ is approximated by

$$\begin{aligned}&\int_t^{t+\Delta t} f(v(\tau);\rho(\tau,x(\tau)),\omega(\tau,x(\tau)))\,d\tau\\ &= f(v;\rho(t,x),\omega(t,x))\Delta t + O(\Delta t^2).\end{aligned}$$

From the new position $x + v\Delta t$, the optimal cost on $[t+\Delta t, T]$ the car can obtain is $V(t+\Delta t, x+v\Delta t, w + r(\rho,v,\omega)\Delta t)$.

By the dynamic programming principle, we have

$$V(t,x,w) = \min_{v}\{f(v;\rho(t,x),\omega(t,x))\Delta t + O(\Delta t^2) + V(t+\Delta t, x+v\Delta t, w+r\Delta t)\}, \quad (3.4)$$

where $r := r(v;\rho,\omega)$.

Take the first-order Taylor's expansion of $V(t+\Delta t, x+v\Delta t, w+r\Delta t)$ near (t,x,w) and denote V_t, V_x, and V_w as the partial derivatives $\partial V/\partial t$, $\partial V/\partial x$, and $\partial V/\partial w$. Equation (3.4) yields

$$V(t,x,w) = \min_{v}\{f(\rho,v,\omega)\Delta t + V + \Delta t V_t + v\Delta t V_x + r\Delta t V_w + O(\Delta t^2)\},$$

where we removed the arguments dependencies to simplify the notation to the reader. Eliminating $V(t,x,w)$ from both sides, dividing both sides by Δt and letting $\Delta t \to 0$, we get

$$V_t + \min_{v}\{f(v;\rho,\omega) + vV_x + rV_w\} = 0.$$

Because r is an affine linear function with respect to v, we can write it as

$$r(\rho,v,\omega) = \alpha v + s(\rho,\omega), \qquad \alpha \in \mathbb{R},$$
$$V_t + \min_{v}\{f(\rho,v,\omega) + v(V_x + \alpha V_w)\} + s(\rho,\omega)V_w = 0. \quad (3.5)$$

Because, by hypothesis, $f(v;\rho,\omega)$ is strictly convex with respect to v, Equation (3.5) can be written as

$$V_t + f^*(\mathcal{V};\rho,\omega) + s(\rho,\omega)V_w = 0, \quad (3.6)$$

where $\mathcal{V} = V_x + \alpha V_w$ and $f^*(\cdot)$ is the Legendre transform of $f(\cdot)$, defined as

$$f^*(\mathcal{V};\rho,\omega) = \min_{v}\{f(v;\rho,\omega) + v\mathcal{V}\}.$$

As a result, the w-specific optimal velocity field $\tilde{u} := \tilde{u}(t,x,w)$ is given by

$$\tilde{u}(t,x,w) = \arg\min_{v}\{f(v;\rho,\omega) + v\mathcal{V}\} = f^*_{\mathcal{V}}(\mathcal{V};\rho,\omega). \quad (3.7)$$

Recall that $\mathcal{V} = V_x + \alpha V_w$, and $f^*_{\mathcal{V}}$ is the derivative with respect to the first argument $\mathcal{V}$ of f^*. We drop the term $s(\rho,\omega)V_w$ because it does not depend on the control v.

This ends the proof. □

We provide Corollary 3.2 to discuss the solution existence of the HJB equation (Equation (3.2)).

Corollary 3.2. *Given $\overline{\rho}(t,x)$ and $\overline{\omega}(t,x)$, $\forall (t,x) \in [0,T]\times\mathcal{X}$ obtained from the continuity equation, the HJB Equation (3.2) admits a solution if the following conditions hold*:

1. *The running cost $f(v(t),\overline{\rho}(t,x),\overline{\omega}(t,x))$ is continuous with regard to v;*
2. *The terminal cost $V_T(x,w)$ is continuous;*
3. *The term $|f(v(t),\overline{\rho}(t,x),\overline{\omega}(t,x))| + |V_T(x,w)|$ is bounded;*
4. *The relaxation function $r(\overline{\rho},v,\overline{\omega})$ to be affine linear in v (i.e., $r(\overline{\rho},v,\overline{\omega}) = \alpha u + s(\overline{\rho},\overline{\omega})$, $\alpha \in \mathbb{R}$).*

Proof. We denote $\tilde{X} = (x,w)$ and $\nabla\tilde{V}(t,\tilde{X}) = (V_x(t,x,w), V_w(t,x,w))$. Given $\overline{\rho}(t,x)$ and $\overline{\omega}(t,x)$, $\forall (t,x) \in [0,T]\times\mathcal{X}$ obtained from the continuity equation, we introduce the following Hamiltonian:

$$H(t,X,\nabla\tilde{V}(t,X)) = f^*(V_x(t,\tilde{X});\overline{\rho}(t,x),\overline{\omega}(t,x)) + s(\overline{\rho}(t,x),\overline{\omega}(t,x))V_w(t,x,w). \quad (3.8)$$

We then reformulate Equation (3.2) as

$$\begin{cases} \partial_t\tilde{V}(t,\tilde{X}) + H(t,X,\nabla\tilde{V}(t,\tilde{X})) = 0 \\ \tilde{V}(T,\tilde{X}) = V_T(x,w). \end{cases} \quad (3.9)$$

The reformulated HJB equation (Equation (3.9)) admits a solution by theorem 3.2 in Nisio (2015). Corollary 3.2 holds. In Section 4, we will demonstrate that conditions (1), (2), and (3) in Corollary 3.2 hold for the running and terminal costs used in this work. □

Before delving into the derivation of continuity equation, we need to define a mean velocity field, which averages over class w, $\forall w \in \mathbb{R}^+$, mathematically,

$$u(t,x) = \int_w \tilde{u}(t,x,w)dw = \int_w f^*_{\mathcal{V}}(\mathcal{V};\rho,\omega)dw. \quad (3.10)$$

The mean velocity field will be fed into CE to propagate the population dynamics.

3.2. Derivation of the CE

Regarding the GSOM CE, it is well known (Aw et al. 2002) that by passing the macroscopic limit of the system (2.2), we achieve the conservation laws for a macroscopic density quantity $\rho = \rho(t,x)$, a macroscopic velocity quantity $u = u(t,x)$, and $\omega = \omega(t,x)$, which are

$$\text{[GSOM-CE]} \quad \begin{cases} \rho(t,x)_t + (\rho(t,x)\,u(t,x))_x = 0, \\ (\rho(t,x)\omega(t,x))_t + (\rho(t,x)\omega(t,x)\,u(t,x))_x \\ \quad = \rho(t,x)\,r(\rho(t,x),u(t,x),\omega(t,x)). \end{cases} \quad (3.11)$$

The relations between the macroscopic and microscopic quantities are

$$\rho(t,x_i) = \frac{1}{x_{i+1}(t) - x_i(t)}, \quad u(t,x_i) = v_i, \quad \omega(t,x_i) = w_i.$$

Remark 3.2. We would like to pinpoint the difference between the notations $\omega(t,x)$ and $w(t)$; ω is an Eulerian quantity and w the correspondent Lagrangian one. Eulerian coordinates capture the motion of fluid flow through moving points for visualization, whereas Lagrangian coordinates directly track individual fluid particles.

Below we will detail the micro-macro limit. Considering the time-dependent Lagrangian variables $x_i := x_i(t)$, $v_i := v_i(t), w_i := w_i(t)$, the microscopic model related to the generic second order traffic model GSOM is

$$\begin{cases} \dot{x}_i = v_i, & x_i(0) = x_{i,0}, \\ \dot{w}_i = r(x, v_i, w_i), & w_i(0) = w_{i,0}. \end{cases} \tag{3.12}$$

Because in a typical traffic model we have that $r(x, v_i, w_i) = r\big(\Delta X/(x_{i-1} - x_i), v_i, w_i\big)$, we can consider the local density and the specific volume

$$\rho_i(x) = \frac{\Delta X}{x_{i+1} - x_i}, \qquad \varsigma_i(x) = \frac{1}{\rho_i(x)},$$

and rewrite the microscopic system (3.12) as

$$\begin{cases} \dot{\varsigma}_i = \dfrac{1}{\Delta X}(v_{i+1} - v_i), & x_i(0) = x_{i,0}, \\ \dot{w}_i = r(\rho_i, v_i, w_i), & w_i(0) = w_{i,0}. \end{cases}$$

Introducing the Lagrangian "mass" coordinates (X, T), we can then write

$$\begin{cases} \partial_T \varsigma = \partial_X v, \\ \partial_T w = r(\rho, u, \omega), \end{cases} \tag{3.13}$$

because $u(t, x_i) = v_i$, $\omega(t, x_i) = w_i$.

Note that X describes the total space occupied by cars up to point x. Now, we can consider the following relations

$$\partial_X x = \rho^{-1} = \varsigma, \quad \partial_X t = 0, \quad \partial_T x = u, \quad \partial_T t = 0,$$

and we rewrite them as

$$\partial_x X = \rho, \quad \partial_t X = -\rho u, \quad T = t.$$

Zooming in Equation (3.13), that is, multiplying space and time variables by a small parameter ε, we get the limit

$$\begin{cases} \rho_t + (\rho\, u)_x = 0, \\ (\rho\omega)_t + (\rho\omega\, u)_x = \rho\, r(\rho, u, \omega), \end{cases} \tag{3.14}$$

which is the macroscopic system written in the Eulerian coordinate, because the variables like density, velocity, and specific volume do not change in this scaling. Moreover, the microscopic model (3.13) can be seen as a semidiscretization of the macroscopic one (3.14).

3.3. GSOM-MFG Model Summary

Summarizing the HJB equation derived in Equation (3.2) and the CE in Equation (3.11), we present the GSOM-MFG model here. In a nutshell, [GSOM-MFG] is a forward-backward PDE system with initial, boundary, and terminal conditions, summarized below:

$$\text{[GSOM-MFG]} \quad \begin{cases} \text{(CE)} & \rho_t + (\rho\, u)_x = 0, \\ & (\rho\omega)_t + (\rho\omega\, u)_x \\ & \qquad = \rho(t,x)\, r(\rho(t,x), u(t,x), \omega(t,x)), \\ \text{(HJB)} & V_t + f^*(\mathcal{V}; \rho, \omega) + s(\rho, \omega) V_w = 0, \\ & \tilde{u} = f^*_{\mathcal{V}}(\mathcal{V}; \rho, \omega), \\ & u = \displaystyle\int_w f^*_{\mathcal{V}}(\mathcal{V}; \rho, \omega) dw. \end{cases} \tag{3.15}$$

$$\begin{aligned} &\text{(IC)} && \rho(0,x) = \rho_0(x), \ \ \omega(x,0) = \omega_0(x), \\ &\text{(BC)} && \rho(t,0), \rho(t,L), \ \ \omega(t,0), \omega(t,L), \\ &\text{(TC)} && V(T,x,w) = V_T(x,w). \end{aligned}$$

The MFE solution is denoted as

$$\begin{aligned} &\text{SOL}(\text{[GSOM-MFG]}) \\ &= \{\rho(x,t), \omega(x,t), u(t,x), \\ &\qquad \tilde{u}(t,x,w), V(t,x,w)\}_{\forall x \in \mathbb{R}^+,\, t \in [0,T],\, w \in \mathbb{R}^+}. \end{aligned}$$

3.4. First-Order and GSOM Traffic Flow MFGs Comparison

To emphasize the coherence of the computation, we revisit the mean-field game systems of LWR and GSOM. Notably, the models exhibit a comparable structure, with the GSOM model incorporating an extra variable alongside the other ones. To remind, we summarize [LWR-MFG] (2.2) and [GSOM-MFG] (3.15) here:

$$\text{[LWR-MFG]} \quad \begin{cases} \rho_t + (\rho u)_x = 0 \\ V_t + f^*(V_x, \rho) = 0\,, \\ u(t,x) = f^*_{\mathcal{V}}(V_x, \rho) \end{cases}$$

$$\text{where } \mathcal{V} = V_x. \tag{3.16}$$

$$\text{[GSOM-MFG]} \quad \begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho\omega)_t + (\rho\omega\, u)_x = \rho\, r(\rho, u, \omega) \\ V_t + f^*(\mathcal{V}; \rho, \omega) + s(\rho, \omega) V_w = 0 \\ \tilde{u} = f^*_{\mathcal{V}}(\mathcal{V}; \rho, \omega), \\ u = \displaystyle\int_w f^*_{\mathcal{V}}(\mathcal{V}; \rho, \omega) dw. \end{cases},$$

$$\text{where } \mathcal{V} = V_x + \alpha V_w. \tag{3.17}$$

When we set $\alpha = s(\rho, \omega) = 0$ (and as a consequence $r(\rho, u, \omega) = \alpha u + s(\rho, \omega) = 0$) in the GSOM-MFG model, we recover the LWR-MFG model together with the second PDE of the system (3.17). Note that they are decoupled because the LWR-MFG is independent from w.

In the first-order traffic flow MFG, all the equations are defined in terms of the time t and spatial coordinate x variables. However, in the new GSOM system, we encounter a mixed situation. Specifically, two equations are defined in terms of (t, x), whereas the other two equations appear to be defined in (t, x, w), where w represents an additional parameter. However, it is important to note that only the value fit V and the Eulerian velocity u in the GSOM system depend additionally on the parameter w. This setup is reasonable because the GSOM system can be seen as a parameterized model. The parameter w serves as an additional factor influencing the model and allowing for more versatility adaptability. Alternatively, one could consider solving the LWR model for various fluxes, where each flux is determined by a specific value of w. This perspective highlights the possibility of investigating and analyzing the LWR model under different scenarios represented by different flux values.

4. GSOM-MFG Cost Specification

The interpretation of GSOM as an MFG allows us to modify the cost functional and derive a new class of second-order traffic flow models. There are infinite choices of respective cost functions. Below we start from those inspired by traffic flow models.

4.1. ARZ Cost Function

We first specify $U(\rho, \omega)$ as the Greenshields form for the equilibrium velocity that corresponds to the first-order model. Then we get

$$\begin{aligned} U(\rho,\omega) &= U_{LWR}(\rho) + (\omega - U_{LWR}(0)) \\ &= u_m\left(1-\frac{\rho}{\rho_m}\right) + \omega - u_m = u_m\left(\frac{\omega}{u_m} - \frac{\rho}{\rho_m}\right). \end{aligned} \tag{4.1}$$

Note that if $w = u_m$, we recover the LWR Greenshields relation.

Following the similar idea in Huang et al. (2020a), here we choose the following cost function:

$$f_{GSOM} = \frac{1}{2}(U(\rho,\omega) - u)^2,$$

and the ARZ relaxation function

$$r = \lambda(U(\rho,\omega) - u),$$

then [GSOM-MFG] (3.15) becomes

$$\text{[ARZ-MFG]} \quad \begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho\omega)_t + (\rho\omega u)_x = \rho\lambda(U(\rho,\omega) - u) \\ V_t + U(\rho,\omega)V_x - \frac{1}{2}(V_x - \lambda V_w)^2 = 0 \\ \tilde{u} = U(\rho,\omega) - (V_x - \lambda V_w) \\ u = \int_w \tilde{u} dw. \end{cases}$$

Corollary 4.1. *Given cost function*

$$f_{ARZ}(u;\rho,\omega) = \frac{1}{2u_{\mathrm{m}}^2}(U(\rho,\omega) - u)^2, \tag{4.2}$$

where u_m is maximum speed and $U(\rho,\omega)$ is defined in (4.1).

Define

$$r(\rho,u,\omega) = \lambda(U(\rho,\omega) - u). \tag{4.3}$$

In other words, let $\alpha = -\lambda u, s(\rho,\omega) = \lambda U(\rho,\omega)$. [ARZ-MFG] can be reformulated as

[ARZ-MFG]

$$\begin{cases} \text{(CE)} & \rho_t + (\rho u)_x = 0, \\ & (\rho\omega)_t + (\rho\omega u)_x = \rho\lambda(U(\rho,\omega) - u), \\ \text{(HJB)} & V_t + U(\rho,\omega)V_x - \frac{1}{2}(V_x - \lambda V_w)^2 = 0, \\ & \tilde{u} = U(\rho,\omega) - u_{\mathrm{m}}^2(V_x - \lambda V_w). \end{cases} \tag{4.4}$$

This includes the initial condition (IC), the boundary condition (BC), and the terminal condition (TC) specified here:

$$\begin{aligned} &\text{(IC)} && \rho(0,x) = \rho_0(x),\ \omega(0,x) = \omega_0(x), \\ &\text{(BC)} && \rho(t,0), \rho(t,L),\ \omega(t,0), \omega(t,L), \\ &\text{(TC)} && V(T,x,w) = V_T(x,w). \end{aligned}$$

Proof. We first expand the cost function defined in (4.2) using (4.1) as follows:

$$f_{ARZ}(u;\rho,\omega) = \frac{1}{2u_m^2}(U(\rho,\omega) - u)^2 \tag{4.5a}$$

$$= \frac{1}{2u_m^2}\left\{\left[u_m\left(1-\frac{\rho}{\rho_m}\right) + \omega - u_m\right] - u\right\}^2 \tag{4.5b}$$

$$= \frac{1}{2u_m^2}\left[u_m\left(\frac{\omega}{u_m} - \frac{\rho}{\rho_m}\right) - u\right]^2 \tag{4.5c}$$

$$= \frac{1}{2}\left(\frac{u}{u_m}\right)^2 - \left[\left(\frac{\omega}{u_m^2} - \frac{\rho}{u_m\rho_m}\right)\right]u + \frac{1}{2}\left(\frac{\omega}{u_m} - \frac{\rho}{\rho_m}\right)^2. \tag{4.5d}$$

Plugging (4.2) into (3.7), we can compute the optimal velocity field as

$$\tilde{u}(t,x,w) = \arg\min_v\{f_{ARZ}(v;\rho,\omega) + v(V_x + \alpha V_w)\} \tag{4.6a}$$

$$= \arg\min_v \left\{\frac{1}{2}\left(\frac{v}{u_m}\right)^2 - \left[\left(\frac{\omega}{u_m^2} - \frac{\rho}{u_m\rho_m}\right) - (V_x + \alpha V_w)\right]v\right\} \tag{4.6b}$$

$$= u_m\left(\frac{\omega}{u_m} - \frac{\rho}{\rho_m}\right) - u_m^2(V_x + \alpha V_w)$$

$$\triangleq U(\rho,\omega) - u_m^2(V_x + \alpha V_w). \tag{4.6c}$$

In the second line above, we substitute the optimal velocity field into the cost function and drop all the terms that are independent of v without loss of generality.

Then substituting the above optimal velocity field into (3.6), we have

$$V_t + \min_v \{f_{ARZ}(v;\rho,\omega) + v(V_x + \alpha V_w)\} + s(\rho,\omega)V_w = 0 \tag{4.7a}$$

$$\begin{aligned}
\Rightarrow V_t + \Bigg\{ & \frac{1}{2u_m^2}\left[u_m\left(\frac{\omega}{u_m} - \frac{\rho}{\rho_m}\right)(V_x + \alpha V_w)\right]^2 \\
& - \left[u_m\left(\frac{\omega}{u_m} - \frac{\rho}{\rho_m}\right) - (V_x + \alpha V_w)\right]^2 + \frac{1}{2}\left(\frac{\omega}{u_m} - \frac{\rho}{\rho_m}\right)^2 \\
& + v(V_x + \alpha V_w)\Bigg\} + \lambda U(\rho,\omega)V_w = 0 \\
\Rightarrow V_t - & \frac{1}{2u_m^2}\left[u_m\left(\frac{\omega}{u_m} - \frac{\rho}{\rho_m}\right) - (V_x + \alpha V_w)\right]^2 \\
& + \frac{1}{2}\left(\frac{\omega}{u_m} - \frac{\rho}{\rho_m}\right)^2 + \lambda U(\rho,\omega)V_w = 0.
\end{aligned} \tag{4.7b}$$

Below we show the HJB equation in [ARZ-MFG] admits a solution.

Corollary 4.2 *Given $\overline{\rho}(t,x)$ and $\overline{\omega}(t,x)$, $\forall (t,x) \in [0,T]\times\mathcal{X}$ from the continuity equation, the HJB equation with the running cost f_{ARZ} admits a solution if the terminal cost $V_T(x,w)$ is continuous and bounded.*

Proof. Conditions (1) and (2) in Corollary 3.2 hold for the running and terminal costs, respectively. We only need to show the running cost is bounded. We have $f_{ARZ}(u;\rho,\omega) = (U(\rho,\omega)-u)^2/2u_m^2 \leqslant u_m^2/2u_m^2 = 1/2$. Corollary 4.2 holds. □

Summarizing the above game-theoretic interpretation, we can rewrite ARZ models using an MFG representation shown in Equation (3.15), which concludes the proof. □

Theorem 4.3. *The solution of [GSOM] (1.1) is a solution of [ARZ-MFG] (4.4) under the conditions that* (i) *[ARZ-MFG] and [GSOM] have the same initial conditions and boundary conditions, and* (ii) *$V_T(x,w) = C$ where C is an arbitrary constant for [ARZ-MFG].*

Proof. Denote $\rho^*(t,x),\omega^*(t,x),u^*(t,x)$ the solution of [GSOM]. Note that the CE defined in Equation (1.1) is the same as that in Equation (4.4). Now it suffices to show that u^* satisfy the HJB Equations (3.2) for some V^*. Take $V^* \equiv C$, then the terminal condition $V^*(T,x,w) = V_T(x,w) = C$ is satisfied, and Equation (3.2) becomes a single equation $\tilde{u}^*(t,x,w) = U(\rho^*,\omega^*)$, defined in Equation (1.1). Because $\tilde{u}^*(\cdot)$ does not depend on w anymore, we can drop its argument w and define $u^*(t,x) := \tilde{u}^*(t,x,w)$ without loss of generality. In conclusion, ρ^*, ω^*, u^*, and $V^* \equiv C$ is also a solution of [ARZ-MFG]. □

4.2. GSOM Nonseparable Cost Function

Building on the ARZ cost function, we propose a modified version and call it GSOM nonseparable cost function with three terms, representing kinetic energy, driving efficiency with dependence on the Lagrangian marker, and traffic safety; mathematically,

$$\begin{aligned}
& f_{\mathrm{ARZ-NonSep}}(u;\rho,\omega) \\
& = \underbrace{\frac{1}{2u_{\mathrm{m}}^2}(U(\rho,\omega)-u)^2}_{\text{equilibrium speed}} + \underbrace{\frac{1}{2}\left(1-\frac{\omega}{u_{\mathrm{m}}}\right)^2}_{\text{flow heterogeneity}} - \underbrace{\frac{1}{2}\left(1-\frac{\rho}{\rho_{\mathrm{m}}}\right)^2}_{\text{safety}}
\end{aligned} \tag{4.8a}$$

$$\begin{aligned}
& \triangleq \frac{1}{2}\left(\frac{u}{u_{\mathrm{m}}}\right)^2 - U(\rho,\omega)\frac{u}{u_{\mathrm{m}}^2} + \frac{1}{2}\left(\frac{U(\rho,w)}{u_m}\right)^2 \\
& \quad + \frac{1}{2}\left(1-\frac{\omega}{u_{\mathrm{m}}}\right)^2 - \frac{1}{2}\left(1-\frac{\rho}{\rho_{\mathrm{m}}}\right)^2.
\end{aligned} \tag{4.8b}$$

To elaborate, the reason we come up with three terms in line 1 is as follows. First, minimizing $(U(\rho,\omega)-u)^2/2u_{\mathrm{m}}^2$ is equivalent to imposing the velocity field to move toward the equilibrium speed defined by $U(\rho,\omega)$. Second, minimizing $(1-\omega/u_{\mathrm{m}})^2/2$ equals to forcing $\omega(t,x)$ toward the maximum speed u_{m}. Third, minimizing $-(1-\rho/\rho_{\mathrm{m}})^2/2$ equals to pushing $\rho(t,x)$ away from the jam density ρ_{m}. In the second line, we further expand the first term $(U(\rho,\omega)-u)^2$ into three subterms with physical meanings. First, minimizing $u^2/2u_{\mathrm{m}}^2$ equals to minimizing kinetic energy defined by the square of velocity. Second, minimizing $-\omega u/u_{\mathrm{m}}^2$ amounts to increasing ω and u as close to the maximum speed u_{m} as possible. Third, minimizing $\rho u/\rho_{\mathrm{m}}u_{\mathrm{m}}$ equals to when u is high, ρ needs to be low, and vice versa. In other words, speed choice should account for the congestion effect, that is, congestion-aware.

Following the same procedure as done for ARZ-MFG derivation, we can plug this cost function into (3.7) and compute the w-specific optimal velocity field as

$$\begin{aligned}
\tilde{u}(t,x,w) &= \arg\min_v \{f_{\mathrm{ARZ-NonSep}}(v;\rho,\omega) + v(V_x + \alpha V_w)\} \\
&= \arg\min_v \left\{\frac{1}{2}\left(\frac{v}{u_{\mathrm{m}}}\right)^2 - U(\rho,\omega)\frac{v}{u_{\mathrm{m}}^2} + v(V_x + \alpha V_w)\right\} \\
&= \arg\min_v \left\{\frac{1}{2u_m^2}v^2 - \left[\frac{\omega}{u_{\mathrm{m}}^2} - \frac{\rho}{u_{\mathrm{m}}\rho_{\mathrm{m}}} - (V_x + \alpha V_w)\right]v\right\} \\
&= u_m\left(\frac{\omega}{u_m} - \frac{\rho}{\rho_m}\right) - u_m^2(V_x + \alpha V_w) \\
&\triangleq U(\rho,\omega) - u_m^2(V_x + \alpha V_w).
\end{aligned}$$

In the second line above, we substitute the optimal velocity field into the cost function and drop all the terms that are independent of v without loss of generality.

Assume $r(\cdot)$ follows (4.3). We substitute the above optimal velocity field into (3.6), the optimal value function is computed as

$$\begin{aligned} &V_t + \min_v\{f_{\text{ARZ-NonSep}}(v;\rho,\omega) + v(V_x + \alpha V_w)\} \\ &\quad + s(\rho,\omega)V_w = 0, \\ &\Rightarrow V_t + \min_v\left\{\frac{1}{2u_m^2}v^2 - \left[\frac{\omega}{u_{\text{m}}^2} - \frac{\rho}{u_{\text{m}}\rho_{\text{m}}} - (V_x + \alpha V_w)\right]v\right\} \\ &\quad + \lambda U(\rho,\omega)V_w + \frac{1}{2}\left(1 - \frac{\omega}{u_{\text{m}}}\right)^2 - \frac{1}{2}\left(1 - \frac{\rho}{\rho_{\text{m}}}\right)^2 = 0. \end{aligned}$$

The second row holds because we substitute the optimal w-specific velocity field into the cost function.

Below we show the HJB equation admits a solution.

Corollary 4.4. *Given $\overline{\rho}(t,x)$ and $\overline{\omega}(t,x)$, $\forall (t,x) \in [0,T]\times\mathcal{X}$ from the continuity equation, the HJB equation with the nonseparable cost admits a solution if the terminal cost $V_T(x,w)$ is continuous and bounded.*

Proof. We only need to show the running cost is bounded. We have $f_{\text{ARZ-NonSep}}(u;\rho,\omega) = (U(\rho,\omega)-u)^2/2u_{\text{m}}^2 + (1-\omega/u_{\text{m}})^2/2 - (1-\rho/\rho_{\text{m}})^2/2 \leqslant 1$. Corollary 4.4 holds. □

In summary, we have

[GSOM-MFG-NonSep]

$$\begin{cases} \text{(CE)} & \rho_t + (\rho u)_x = 0, \\ & (\rho\omega)_t + (\rho\omega u)_x = \lambda(U(\rho,\omega) - u), \\ \text{(HJB)} & V_t - \dfrac{1}{2u_m^2}[U(\rho,\omega) - u_m^2(V_x + \alpha V_w)]^2 \\ & \quad + \lambda U(\rho,\omega)V_w + \dfrac{1}{2}\left(1 - \dfrac{\omega}{u_{\text{m}}}\right)^2 - \dfrac{1}{2}\left(1 - \dfrac{\rho}{\rho_{\text{m}}}\right)^2 = 0, \\ & \tilde{u} = U(\rho,\omega) - u_{\text{m}}^2(V_x - \lambda V_w), \\ & u = \displaystyle\int_w \tilde{u}\,dw. \end{cases} \tag{4.9}$$

4.3. GSOM Separable Cost Function

We propose another modified version and call it GSOM separable cost function with three terms, representing kinetic energy, driving efficiency with dependence on the Lagrangian marker, and traffic safety. Mathematically,

$$f_{\text{ARZ-Sep}}(u;\rho,\omega) \triangleq \frac{1}{2}\left(1 - \frac{u}{u_{\text{m}}}\right)^2 + \frac{1}{2}\left(1 - \frac{\omega}{u_{\text{m}}}\right)^2 - \frac{1}{2}\left(1 - \frac{\rho}{\rho_{\text{m}}}\right)^2. \tag{4.10a}$$

In a comparison with the nonseparable cost, the GSOM separable cost does not have crossing terms among speed choice, Lagrangian marker, and road density. We can plug this cost function into (3.7) and compute the optimal velocity field as

$$\begin{aligned} \tilde{u}(t,x,w) &= \arg\min_v\{f_{\text{ARZ-Sep}}(v;\rho,\omega) + v(V_x + \alpha V_w)\} \\ &= \arg\min_v\left\{\frac{1}{2}\left(1 - \frac{v}{u_{\text{m}}}\right)^2 + v(V_x + \alpha V_w)\right\} \\ &= u_m - u_m^2(V_x + \alpha V_w). \end{aligned}$$

We substitute the above optimal velocity field into (3.6); the optimal value function is computed as

$$\begin{aligned} &V_t + \min_v\{f_{\text{ARZ-Sep}}(v;\rho,\omega) + v(V_x + \alpha V_w)\} \\ &\quad + s(\rho,\omega)V_w = 0, \\ &\Rightarrow V_t + \min_v\left\{\frac{1}{2}\left(1 - \frac{v}{u_m}\right)^2 + (V_x + \alpha V_w)v\right\} \\ &\quad + \lambda U(\rho,\omega)V_w + \frac{1}{2}\left(1 - \frac{\omega}{u_{\text{m}}}\right)^2 \\ &\quad - \frac{1}{2}\left(1 - \frac{\rho}{\rho_{\text{m}}}\right)^2 = 0. \end{aligned}$$

Below we show the HJB equation admits a solution.

Corollary 4.5. *Given $\overline{\rho}(t,x)$ and $\overline{\omega}(t,x)$, $\forall (t,x) \in [0,T]\times\mathcal{X}$ from the continuity equation, the HJB equation with the separable cost admits a solution if the terminal cost $V_T(x,w)$ is continuous and bounded.*

Proof. We only need to show the running cost is bounded. We have $f_{\text{ARZ-Sep}}(u;\rho,\omega) = (1-u/u_{\text{m}})^2/2 + (1-\omega/u_{\text{m}})^2/2 - (1-\rho/\rho_{\text{m}})^2/2 \leqslant 1$. Corollary 4.5 holds. □

In summary, we have

[GSOM-MFG-Sep] (4.11)

$$\begin{cases} \text{(CE)} & \rho_t + (\rho\, u)_x = 0, \\ & (\rho\omega)_t + (\rho\omega\, u)_x = \lambda(U(\rho,\omega) - u), \\ \text{(HJB)} & V_t + \min_v\left\{\dfrac{1}{2}(1 - \dfrac{v}{u_m})^2 + (V_x + \alpha V_w)v\right\} \\ & \quad + \lambda U(\rho,\omega)V_w + \dfrac{1}{2}\left(1 - \dfrac{\omega}{u_{\text{m}}}\right)^2 \\ & \quad - \dfrac{1}{2}\left(1 - \dfrac{\rho}{\rho_{\text{m}}}\right)^2 = 0, \\ & \tilde{u} = u_m - u_m^2(V_x + \alpha V_w), \\ & u = \displaystyle\int_w \tilde{u}\,dw. \end{cases} \tag{4.12}$$

5. MFE Solution Approach and Numerical Results

In this section, we will first develop a fixed-point algorithm to solve [GSOM-MFG] and then demonstrate how different cost functions affect traffic flow patterns on a ring road.

5.1. Related Work

The existing literature primarily uses three types of numerical methods, namely, fixed-point iteration, variational method, and Newton's method. The majority of work on MFGs (Couillet et al. 2012, Chevalier, Le Ny, and Malhamé 2015) are solved using fixed-point method, but it only works well for MFGs with special cost functional structures, for example, when there are no cross terms between representative agent's control (i.e., velocity) and population mass (i.e., traffic density) in cost functions. Traffic flow inspired MFGs, however, exhibit unique characteristics that the cost functional is nonseparable arising from the traffic congestion effect. In other words, vehicles have to slow down when encountering traffic congestion (i.e., lower speed at higher density) and speed up in less congestion areas (i.e., higher speed at low density). Because of the coupling between traffic density and velocity, traffic flow MFGs are generally not potential games and accordingly cannot be solved with variational method either. To solve LWR-MFG, in Huang et al. (2019, 2020a, b) instead, the forward and backward equations are reformulated as a large nonlinear system and finite-difference Newton's method is used. Multigrid preconditioning techniques are further applied to improve and accelerate the convergence. This method, however, involves solving a large system of equations of which the dimensions are determined by the product of discretized time and space dimensions, and are thus not scalable. Recent work on MFG has developed techniques such as fictitious play (FP) (Perrin et al. 2020), which uses historical control information of the representative agent to update the control profile before feeding it into the FPK equation, which is shown to help stabilize policy learning. In this paper, we will use fixed-point method coupled with FP for the solution algorithm.

5.2. Numerical Schemes

Denote $L \in [0,\infty)$ as the length of the road and $T \in [0,\infty)$ as the planning horizon. To numerically solve GSOM-MFG, a space-time meshgrid needs to be first defined. Denote the spatial and temporal step sizes as $\Delta x, \Delta t$, respectively, and the numbers of spatial and temporal points as N_x, N_t, respectively. The relations between the step size and the number of points follow that $\Delta x = L/N_x$ and $\Delta t = T/N_t$, respectively. Accordingly, we can define a sequence of grid points $x_k = k\Delta x, k = 1,\dots,N_x$ and $t_\tau = \tau\Delta t, \tau = 1,\dots,N_t$, respectively. With these grid points, we split the road into a sequence of adjacent cells, denoted as $[x_{k-1}, x_k] = [(k-1)\Delta x,\ k\Delta x],\ k = 1,\dots,N_x$, then, $[0,L] = [(0,\Delta x),(2\Delta x,3\Delta x),\dots,((N_x-1)\Delta x,1)]$. Similarly, we discretize the planning time horizon into a sequence of time intervals $[t_{\tau-1}, t_\tau] = [(\tau-1)\Delta t, \tau\Delta t]$, $\tau = 1,\dots,N_t$, then, $[0,T] = [(0,\Delta t),\dots,((N_t-1)\Delta t,1)]$. To ensure stability of the numerical scheme, N_t is chosen in such a way that Δt respects the Courant–Friedrichs–Lewy (CFL) condition (LeVeque 2002), that is, $v\Delta t \le \Delta x$ should be posed where $v = \max_{k=1,\dots,N_x;\tau=1,\dots,N_t}|u_k^\tau|$. When the MFG has speed constraints $0 \le u \le u_{\max}$, it suffices to ensure $u_{\max}\Delta t \le \Delta x$.

On the two-dimensional meshgrid points, we can now define discretized flow variables for CE. Denote ρ_k^τ, u_k^τ, and ω_k^τ the average density, velocity, and property, respectively, in cell k (i.e., $[(k-1)\Delta x, k\Delta x]$) at time τ, where $\tau = 1,\dots,N_t; k = 1,\dots,N_x; l = 1,\dots,N_w$.

To solve CE, we use a *finite volume Lax-Friedrichs* scheme (LeVeque 2002):

$$\rho_k^{\tau+1} = \frac{1}{2}(\rho_{k-1}^\tau + \rho_{k+1}^\tau) - \frac{\Delta t}{2\Delta x}(\rho_{k+1}^\tau u_{k+1}^\tau - \rho_{k-1}^\tau u_{k-1}^\tau), \tag{5.1a}$$

$$z_k^{\tau+1} = \frac{1}{2}(z_{k-1}^\tau + z_{k+1}^\tau) - \frac{\Delta t}{2\Delta x}(z_{k+1}^\tau u_{k+1}^\tau - z_{k-1}^\tau u_{k-1}^\tau). \tag{5.1b}$$

For the initial condition of CE, we consider $\rho_0(x) = \rho(x,0)$ and $\omega_0(x) = \omega(x,0)$ (z can be recovered as $\rho\omega$) discretized as

$$\rho_k^0 = \frac{1}{\Delta x}\int_{x_{k-1}}^{x_k}\rho_0(x)\,dx, \quad \omega_k^0 = \frac{1}{\Delta x}\int_{x_{k-1}}^{x_k}\omega_0(x)\,dx, \tag{5.2a}$$

$$z_k^0 = \rho_k^0\omega_k^0 = \frac{1}{\Delta x}\int_{x_{k-1}}^{x_k}\rho_0(x)\omega_0(x)\,dx, \qquad k = 1,\dots,N_x. \tag{5.2b}$$

The key difference between GSOM-MFG and LWR-MFG is the additional degree of freedom introduced in the Lagrangian marker w. Accordingly, we need to augment the grid to a three-dimensional meshgrid of (t,x,w). Denote $W \in \mathbb{R}$ as the domain of the Lagrangian marker, the discretized step size as Δw, and the numbers of points as N_w.

On the three-dimensional meshgrid points, we will define discretized value and velocity variables for HJB equations. To solve HJB, we discretize $V_{kl}^\tau = V(\tau, x_k, w_l)$ using the *upwind* scheme in time and the *central difference* scheme in space and property, respectively:

$$V_t(t_\tau, x_k, w_l) = \frac{V_{kl}^{\tau+1} - V_{kl}^\tau}{\Delta t}, \tag{5.3a}$$

$$V_x(t_\tau, x_k, w_l) = \frac{V_{k+1,l}^\tau - V_{k-1,l}^\tau}{2\Delta x}, \tag{5.3b}$$

$$V_w(t_\tau, x_k, w_l) = \frac{V_{k,l+1}^\tau - V_{k,l-1}^\tau}{2\Delta w}. \tag{5.3c}$$

Then HJB Equation (3.6) becomes

$$\frac{V_{kl}^{\tau+1}-V_{kl}^{\tau}}{\Delta t}+f^*\left(\frac{V_{k+1,l}^{\tau}-V_{k-1,l}^{\tau}}{2\Delta x}+\alpha\frac{V_{k,l+1}^{\tau}-V_{k,l-1}^{\tau}}{2\Delta w};\rho_k^{\tau},\omega_k^{\tau}\right)+s(\rho_k^{\tau},\omega_k^{\tau})\frac{V_{k,l+1}^{\tau}-V_{k,l-1}^{\tau}}{2\Delta w}=0.$$

The value function is then computed as

$$\frac{V_{kl}^{\tau+1}-V_{kl}^{\tau}}{\Delta t}=-f^*\left(\frac{V_{k+1,l}^{\tau}-V_{k-1,l}^{\tau}}{2\Delta x}+\alpha\frac{V_{k,l+1}^{\tau}-V_{k,l-1}^{\tau}}{2\Delta w};\rho_k^{\tau},\omega_{kl}^{\tau}\right)-s(\rho_k^{\tau},\omega_k^{\tau})\frac{V_{k,l+1}^{\tau}-V_{k,l-1}^{\tau}}{2\Delta w},\tag{5.4}$$

and the optimal w-specific and mean velocities (3.7)–(3.10) are solved as

$$\tilde{u}_{kl}^{\tau}(t_\tau,x_k,w_l)=f_{\mathcal{V}}^*\left(\frac{V_{k+1,l}^{\tau}-V_{k-1,l}^{\tau}}{2\Delta x}+\alpha\frac{V_{k,l+1}^{\tau}-V_{k,l-1}^{\tau}}{2\Delta w};\rho_{kl}^{\tau},\omega_{kl}^{\tau}\right),\tag{5.5a}$$

$$u_k^{\tau}(t_\tau,x_k)=\frac{\sum_{w_l}\tilde{u}_{kl}^{\tau}(t_\tau,x_k,w_l)}{N_w}.\tag{5.5b}$$

The terminal condition of HJB is discretized as

$$V_{kl}^{N_t}=V_T(x_k,w_l),\quad \forall k=1,\dots,N_x;l=1,\dots,N_w.\tag{5.6}$$

In summary, the total variables include $\{\rho_k^{\tau}\}_{1\le k\le N_x}^{0\le\tau\le N_t}$, $\{\omega_k^{\tau}\}_{1\le k\le N_x}^{0\le\tau\le N_t}$, $\{u_k^{\tau}\}_{1\le k\le N_x}^{1\le\tau\le N_t}$, $\{\tilde{u}_{kl}^{\tau}\}_{1\le k\le N_x,1\le l\le N_w}^{1\le\tau\le N_t}$, and $\{V_{kl}^{\tau}\}_{1\le k\le N_x,1\le l\le N_w}^{1\le\tau\le N_t}$.

5.3. Algorithm

With the discretization scheme defined, we propose a fixed-point method by solving two PDEs iteratively. In other words, CE is solved forward first, and then traffic density and Langrangian marker profiles are passed to HJB, and HJB is solved backward to obtain traffic speed. Traffic speed profile is further fed into CE to propagate traffic density and Langrangian marker. This process repeats until it converges. If convergence is guaranteed, the resultant fixed point would be the MFE of the MFG. As aforementioned, we find that fixed-point method does not normally converge and exhibit instability, due to the nonseparable term between traffic speed and density. Thus, we adopt FP and find that this technique applies here and helps stabilize the equilibrium solution. The algorithm is detailed below.

Algorithm 1 (GSOM-MFG Algorithm)

Input
Initial traffic density: $\rho_0(x), x\in[0,L]$; Initial Lagrangian marker: $\omega_0(x), x\in[0,L]$; Terminal value: $V_T(x,w)$, $x\in[0,L], w\in[0,W]$; Boundary condition: $\rho(t,0)$, $\rho(t,L)$, $\omega(t,0), \omega(t,L), t\in[0,T]$; Convergence gap: $Gap=10$; Convergence threshold: $\epsilon=10^{-2}$
Initialization
$\rho_k^{\tau}, k=1,\dots,N_x, \tau=1,\dots,N_t$
$\omega_k^{\tau}, k=1,\dots,N_x, \tau=1,\dots,N_t$
$u_k^{\tau}, k=1,\dots,N_x, \tau=1,\dots,N_t$
$\overline{u}_{kl}^{\tau}=0, k=1,\dots,N_x, l=1,\dots,N_w, \tau=1,\dots,N_t$
$\overline{V}_{kl}^{\tau}=0, k=1,\dots,N_x, l=1,\dots,N_w, \tau=1,\dots,N_t$
while $Gap>\epsilon$ **do** //Check convergence
  iter = 1
  Store the variables from the previous iteration: $\rho_k^{\tau(-)}=\rho_k^{\tau}$, $\omega_k^{\tau(-)}=\omega_k^{\tau}$, $u_k^{\tau(-)}=u_k^{\tau}$.
  Given $\rho_k^0, \omega_k^0, \ \forall k=1,\dots,N_x$ //Forward
  for $\tau\leftarrow 0$ to N_t-1 **do**
    for $k\leftarrow 1$ to N_x **do**
      Given $u_{k-1}^{\tau}, \rho_{k-1}^{\tau}, u_{k+1}^{\tau}, \rho_{k+1}^{\tau}$, propagate $\rho_k^{\tau+1}$ by solving Equation (5.1a)
      Given $u_{k-1}^{\tau}, z_{k-1}^{\tau}, u_{k+1}^{\tau}, \omega_{k+1}^{\tau}$, propagate $z_k^{\tau+1}$ by solving Equation (5.1b)
    end for
  end for
  Given $V_{kl}^{N_t}, \ \forall k=1,\dots,N_x, l=1,\dots,Nw$ //Backward
  for $\tau\leftarrow N_t-1$ to 0 **do**
    for $k\leftarrow 1$ to N_x **do**
      for $l\leftarrow 1$ to N_w **do**
        Given $\rho_k^{\tau}, \omega_k^{\tau}, V_{k+1,l}^{\tau+1}, V_{k-1,l}^{\tau+1}, V_{k,l+1}^{\tau+1}, V_{k,l-1}^{\tau+1}$, obtain $\tilde{u}_{kl}^{\tau}$ by solving Equation (5.5a)
        Given $\rho_k^{\tau}, \omega_k^{\tau}, V_{k+1,l'}^{\tau+1}, V_{k-1,l'}^{\tau+1}, V_{k,l+1}^{\tau+1}, V_{k,l-1}^{\tau+1}, V_{kl}^{\tau+1}$, obtain V_{kl}^{τ} by solving Equation (5.4)
      end for
    end for
  end for
  Fictitious Play (FP): Compute average using all historical values: $\overline{u}_{kl}^{\tau}=(\overline{u}_{kl}^{\tau}+u_{kl}^{\tau})/(iter+1)$, $\overline{V}_{kl}^{\tau}=(\overline{V}_{kl}^{\tau}+V_{kl}^{\tau})/(iter+1)$.
  Update $\tilde{u}_{kl}^{\tau}=\overline{u}_{kl}^{\tau}, V_{kl}^{\tau}=\overline{V}_{kl}^{\tau}, \ k=1,\dots,N_x, l=1,\dots,N_w; \tau=0,\dots,N_t-1$
  Average optimal mean speed from w-specific speed using Equation (3.10): $u_k^{\tau}=\sum_l\tilde{u}_{kl}^{\tau}/N_w$.
  Convergence Gap: $Gap=\|\rho_k^{\tau}-\rho_k^{\tau(-)}\|+\|\omega_k^{\tau}-\omega_k^{\tau(-)}\|+\|u_k^{\tau}-u_k^{\tau(-)}\|$
  iter = iter +1
end while
Output $\rho_k^{\tau}, \omega_k^{\tau}, u_k^{\tau}, \tilde{u}_{kl}^{\tau}, V_{kl}^{\tau}, \ \forall\tau=1,\dots,N_t; k=1,\dots,N_x; l=1,\dots,N_w, l=1,\dots,N_w$

5.4. Numerical Examples

5.4.1. Settings. We investigate traffic flow MFGs with four cost functions summarized below: two first-order traffic flow models and two second-order models.

5.4.1.1. First-Order Traffic Flow MFGs.

1. LWR-MFG:

$$f_{LWR}(u,\rho)=\frac{1}{2}(U(\rho)-u)^2,$$

where $U(\rho)=u_{\mathrm{m}}(1-\rho/\rho_{\mathrm{m}})$.

$$\text{[LWR-MFG]}\quad\begin{cases}\rho_t+(\rho u)_x=0,\\ V_t+U(\rho)V_x-\frac{1}{2}u_{\mathrm{m}}^2V_x{}^2=0,\\ u=U(\rho)-u_{\mathrm{m}}^2V_x.\end{cases}\tag{5.7}$$

with $\rho_0(x)$, $V_T(x)$ and $\rho(t,0),\rho(t,L)$.

2. LWR-MFG-NonSep:

$$f_{\mathrm{NonSep}}(u,\rho)=\frac{1}{2}\left(\frac{u}{u_{\mathrm{m}}}\right)^2-\frac{u}{u_{\mathrm{m}}}+\frac{u\rho}{u_{\mathrm{m}}\rho_{\mathrm{m}}}.$$

$$\text{[LWR-MFG-NonSep]}\quad\begin{cases}\rho_t+(\rho u)_x=0,\\ V_t-\frac{1}{2u_{\mathrm{m}}^2}[U(\rho)-u_{\mathrm{m}}V_x]^2=0,\\ u=U(\rho)-u_{\mathrm{m}}^2V_x.\end{cases}\tag{5.8}$$

with $\rho_0(x)$, $V_T(x)$ and $\rho(t,0),\rho(t,L)$.

Step-by-step derivation for [LWR-MFG] and [LWR-MFG-NonSep] can be found in Appendices A and B, respectively.

5.4.1.2. Second-Order Traffic Flow MFGs.

1. ARZ-MFG: System (4.4) with $\rho_0(x),\omega_0(x)$, $V_T(x,w)$ and $\rho(t,0),\rho(t,L)$, $\omega(t,0),\omega(t,L)$.
2. GSOM-MFG-NonSep: System (4.9) with $\rho_0(x),\omega_0(x)$, $\rho_0(x),\omega_0(x)$, $V_T(x,w)$ and $\rho(t,0),\rho(t,L)$, $\omega(t,0),\omega(t,L)$.

We need to specify the initial condition (IC), the boundary condition (BC), and the terminal condition (TC) as follows.

- IC: We choose the following initial density:

$$\rho_0(x)=\rho_a+(\rho_b-\rho_a)\exp\left[-\frac{(x-L/2)^2}{2\gamma^2}\right],\tag{5.9}$$

where $0\le\rho_a\le\rho_b\le1$ and $\gamma>0$ are constant parameters. Here we choose $\rho_a=0.05,\rho_b=0.95,\gamma=0.35$.

For the second-order traffic flow model, we also need to specify the initial profile for the Lagrangian marker:

$$u_0(x)=\begin{cases}u_a,\ \text{if } x\le\frac{1}{2},\\ u_b,\ \text{if } \frac{1}{2}<x\le1.\end{cases}$$

$$\omega_0(x)=u_0(x)+\rho_0(x)^\gamma,\tag{5.10}$$

where $u_a=0.1,u_b=0.8,\gamma=1$.

- TC: $\begin{cases}V_T(x)=0,\ \text{for the first-order traffic flow model,}\\ V_T(x,w)=w,\text{for the second-order traffic flow model.}\end{cases}$
- BC: $\begin{cases}\rho(t,0)=\rho(t,L),\ \text{for the first-order traffic flow model,}\\ \rho(t,0)=\rho(t,L),\omega(t,0)=\omega(t,L),\text{for the second-order traffic flow model.}\end{cases}$

Set the road length $L=1$ and the planning horizon length $T=1$. Set the maximum value of Lagrangian marker as $W=1$. The meshgrid is discretized with a size of $N_x=N_w=N_t=50$ to fulfill the CFL condition. Set the free flow speed $u_{\mathrm{m}}=1$ and the jam density $\rho_{\mathrm{m}}=1$. Let $\lambda=0$.

Applying Algorithm 1, we compute MFE solutions, including traffic density $\rho^*(t,x)$, Lagrangian marker $\omega^*(t,x)$, mean traffic speed $u^*(t,x)$, w-specific speed $\tilde{u}^*(t,x,w)$, and optimal value $V^*(t,x,w)$.

5.4.2. Results. Below we will present the relevant results for the GSOM-MFG-NonSep model, including algorithmic convergence, MFEs and fundamental diagrams at different values of w, and traffic density and velocity evolution.

Figure 1 illustrates the algorithm's convergence gap against the numerical iterations, indicating that the numerical solver achieves convergence after approximately 10 iterations.

Now we pick two values of Lagrangian marker $w=0$, 1 and demonstrate how velocity u and value functions V, as well as fundamental diagrams, vary at each w in Figures 2 and 3, respectively. We plot the w-specific velocity and optimal value in time and space in Figure 2. The left column displays the values of u and V for $w=0$, whereas the right column displays values for $w=1$. Starting from the same initial velocity profile, we can see how varying w values lead to distinct dynamics in u and V, reflecting diverse microscopic driver behaviors.

To enhance our understanding of how w influences the relationship between traffic density and flux, we

Figure 1. (Color online) GSOM-MFG-NonSep Algorithm Convergence Plot

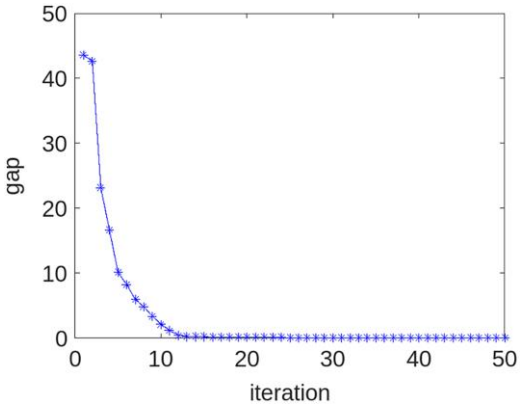
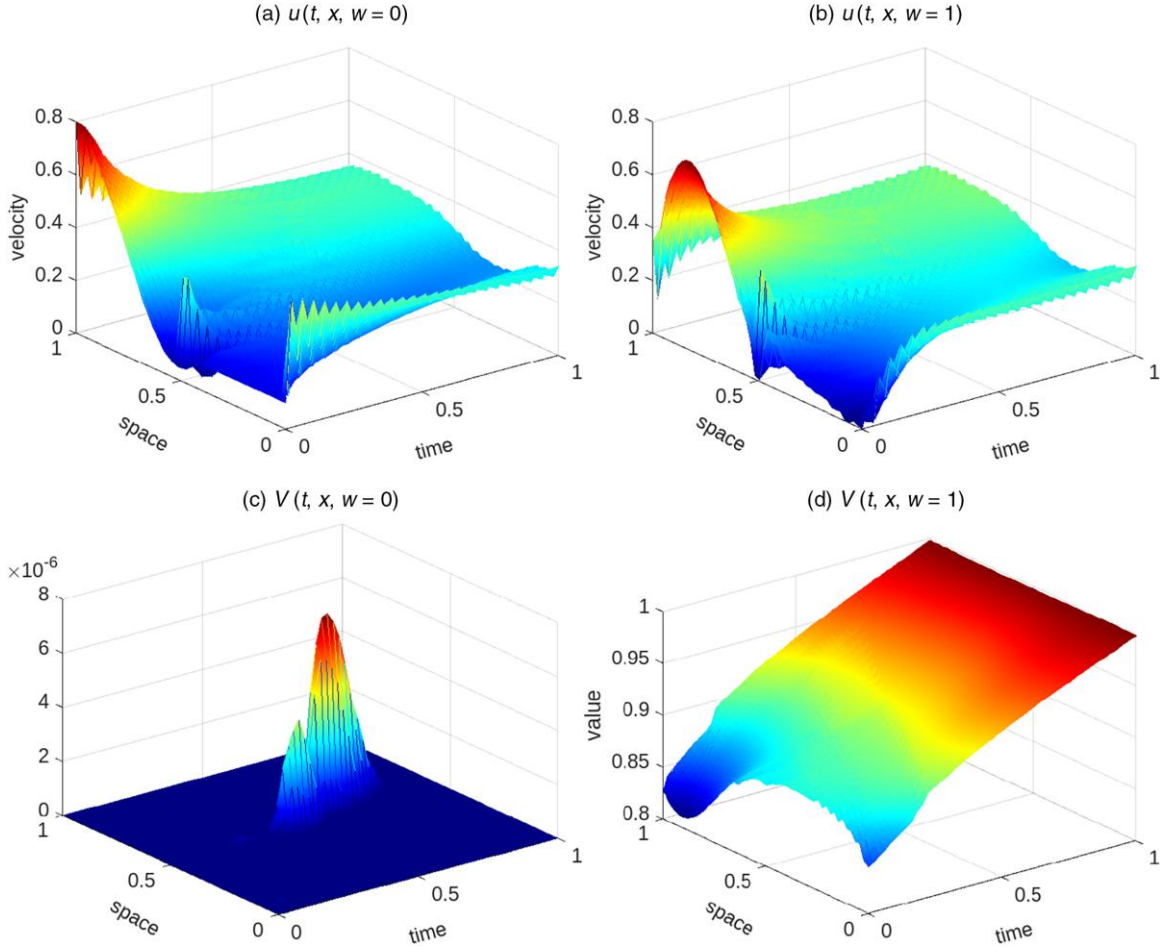


Figure 2. (Color online) GSOM-MFG-NonSep Value and Velocity Profiles at Sample Values of $w = 0, 1$



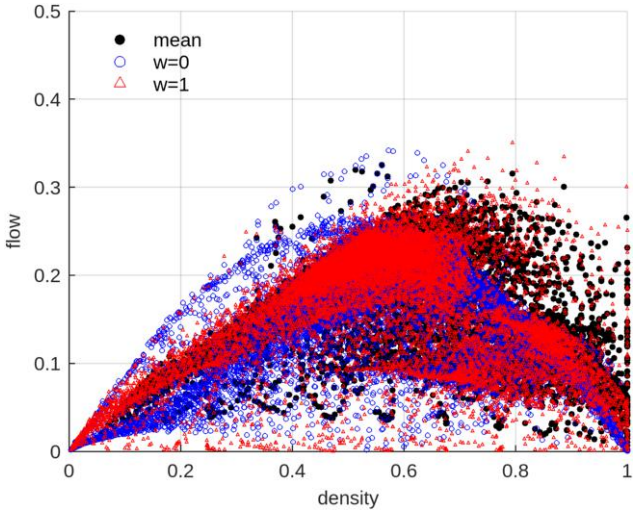
have also depicted the w-specific fundamental diagram in Figure 3. The scatter plots result from multiple numerical experiments with varying initial conditions. The purpose of combining multiple results is to better illustrate the scatter patterns. Black dots represent the traffic density-flux pairs averaged across all w, whereas blue circles and red triangles represent those at $w = 0$ and $w = 1$, respectively. We can observe that dots of different colors display distinct converged fundamental diagrams. In the converged lines, the blue dots ($w = 0$) indicate a higher free-flow speed and maximum flow compared with the red ones ($w = 1$). This is because $w = 0$ represent more aggressive drivers thus leading to higher free-flow speed. This visualization clearly demonstrates that different values of w lead to varying relationships between traffic density and flux.

We further illustrate a whole picture of three variables, namely, traffic density, velocity, and Lagrangian marker (or traffic property), in the entire time and space domain in Figure 4. In particular, Figure 4(c) illustrates the overall dynamics of the Lagrangian marker, identified as a traffic property.

To further illustrate how three key variables, namely, traffic density, velocity, and Lagrangian marker, evolve in space and time, we pick three time snapshots at sampled time instances $t = 0, 0.5, 1$. Figures 5 and 6 illustrate the evolution of these variables for GSOM-MFG-Nonsep and ARZ-MFG, respectively. Starting from exactly the same initials, both games converge to uniform flows where all the variables tend to become constant. GSOM-MFG-Nonsep, however, converges relatively faster than ARZ-MFG because of the cost design in GSOM-MFG-Nonsep.

To further demonstrate the dynamic behavior of four MFG models, we plot their respective fundamental diagrams in Figure 7, where the black dots are generated from data, and the blue and red curves represent the best fitted density-flux relation, following the practice in Fan, Herty, and Seibold (2014) as defined in Equation (6.6). We can see that for both LWR-based MFGs, there exists only one class of cars, so we only need to fit one curve for the fundamental diagram. In contrast, in GSOM-based MFGs, the scatters converge to a set of curves corresponding to different values of w. This

Figure 3. (Color online) w-Specific Fundamental Diagrams



Notes. The black dots represent the mean density and flow. The blue circles represent the density and flow associated with $w = 0$, and the red triangles denote the density and flow for $w = 1$.

shows that GSOM-based MFGs are able to capture different classes of drivers, and thus can capture heterogeneity in the right-hand side of the regime that represents stop-and-go traffic, which is common in real-world observations. Thus, GSOM-based MFGs, like GSOM, are more flexible and robust for real-world validation.

6. Inverse Reinforcement Learning for MFG Model Validation

In this section, we validate the game-theoretic framework of GSOMs when observational data become available. This is essentially an inverse problem; namely, given real-world data, the underlying cost function required in the HJB equation can be estimated to help infer the MFE. We will first introduce the methodology of solving such an inverse problem, which is inverse reinforcement learning (IRL). Then we will validate that our IRL approach can correctly identify the underlying cost function that generates the data, when we know the ground-truth MFG. Last but not the least, we will estimate the latent cost function from a real-world data set and compare the estimation errors against various traffic flow models and the proposed GSOM-MFGs. Via the synthetic experiment and the real-world one, we aim to motivate the need of reinterpreting a traffic flow model with game theory.

IRL is proposed to recover the unknown cost functions from the observation of an expert demonstration (Abbeel and Ng 2004). Here the proposed IRL approach executes the following two steps recursively until our prediction error converges: (1) recovering the cost function coefficients and (2) solving the MFE corresponding to these coefficients.

To approximate the cost function f, Abbeel and Ng (2004) assume that the cost function is a linear combination of known cost features. We denote the cost features as $\tilde{\boldsymbol{f}} = [f_1, \ldots, f_M]^T$ and M is the number of cost features. The weight (i.e., cost coefficient) of each cost feature $f_m, m = 1, \ldots, M$ in the cost function is denoted as $\boldsymbol{c} = [c_1, \ldots, c_M]^T, \boldsymbol{c} \in \mathbb{R}^M$. Mathematically, $f = \boldsymbol{c}^T \cdot \tilde{\boldsymbol{f}}$. To estimate the true cost function is to find the cost coefficient $\boldsymbol{c}$ that can recover the expert's policy.

6.1. Game-Theoretic Approach to Adversarial IRL

In this paper, we adopt a game-theoretic approach (Syed and Schapire 2007) to estimate cost functions. The game consists of a *min* player whose goal is to find the strategy $\boldsymbol{c}$ that can minimize the gap between the value of the recovered and expert's policies and a *max* player aiming to find a policy $\hat{\boldsymbol{u}}$ that can minimize the gap between the value of the recovered and expert's policies. The IRL framework using this max-min game structure is further categorized as adversarial IRL (AIRL) (Syed and Schapire 2007, Ruan and Di 2022, Ruan et al. 2023). To simplify the jargon, we will still refer this particular framework as IRL subsequently.

Figure 4. (Color online) GSOM-MFG-NonSep Variable Evolution

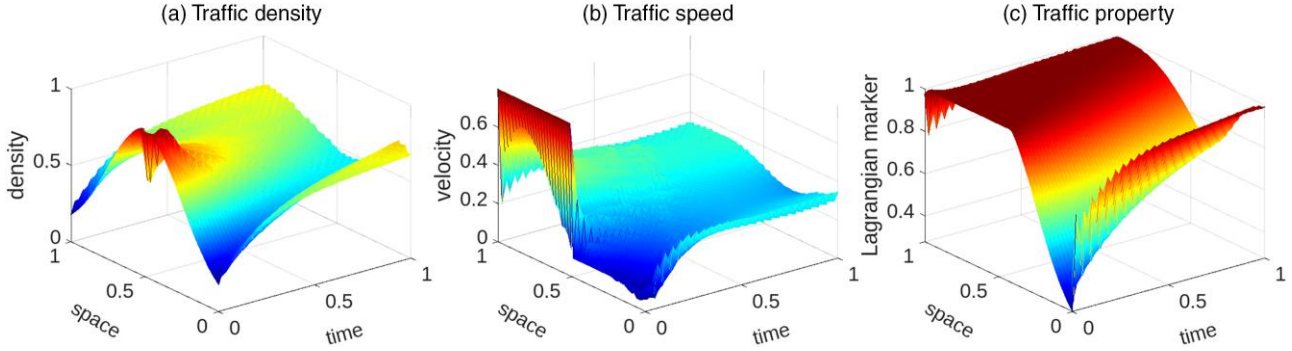
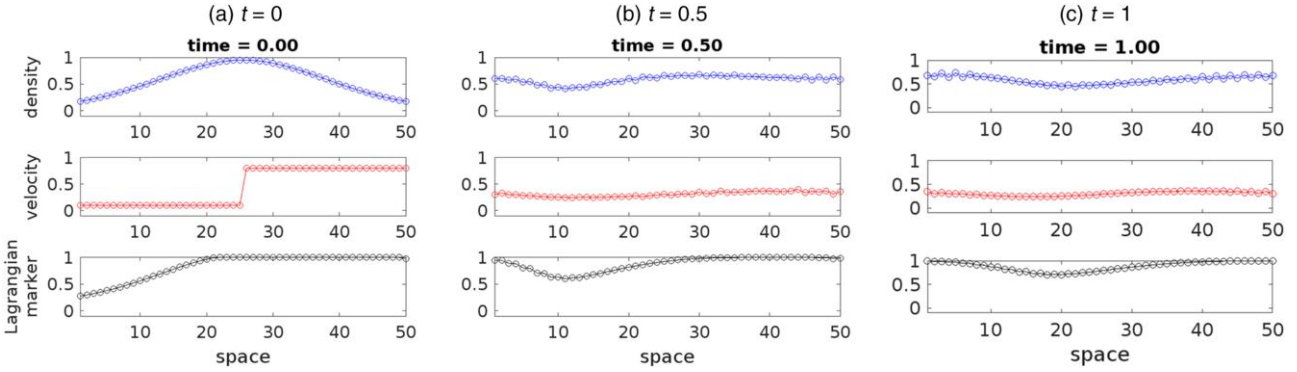


Figure 5. (Color online) GSOM-MFG-NonSep Variable Profiles at Time Instances (0, 0.5, 1.0)



Mathematically, we consider the following optimization

$$\max_{\hat{\boldsymbol{u}}} \min_{\boldsymbol{c}} [\boldsymbol{c} \cdot \boldsymbol{F}_{\hat{\boldsymbol{u}}} - \boldsymbol{c} \cdot \boldsymbol{F}_{\boldsymbol{u}^e}], \tag{6.1}$$

where $\boldsymbol{F}_{\hat{\boldsymbol{u}}} = [F_1^e, \ldots, F_M^e]$ and $\boldsymbol{F}_{\hat{\boldsymbol{u}}} = [\hat{F}_1, \ldots, \hat{F}_M]$ are the value of cost features (i.e., cumulative cost) when executing the recovered $\hat{\boldsymbol{u}}$ and expert's policies $\boldsymbol{u}^e$, respectively. We have

$$F_m^e(x_0, t_0) = \sum_{t=t_0}^{T} f_m(u^e(x,t), \rho^e(x,t), \omega^e(x,t))\Delta t + V_T, \tag{6.2a}$$

$$\hat{F}_m(x_0, t_0) = \sum_{t=t_0}^{T} f_m(\hat{u}(x,t), \rho^e(x,t), \omega^e(x,t))\Delta t + V_T. \tag{6.2b}$$

We use the multiplicative weights for apprenticeship learning (MWAL) algorithm (Syed and Schapire 2007, Ruan et al. 2023) to solve the max-min optimization problem. The algorithm is summarized in Algorithm 2. The input of the algorithm include expert policy $\boldsymbol{u}^e$, population density $\boldsymbol{\rho}^e$, and $\boldsymbol{\omega}^e$. The cost coefficients are first scaled to make sure $\|\boldsymbol{c}\|_1 = 1$ (line 4). The recovered policy $\hat{\boldsymbol{u}}$ is obtained by solving the HJB equation defined in MFG with $\boldsymbol{c}$ (line 5). We calculate the cumulative cost of each feature by sampling trajectories using the recovered and experts' policies. We start from a initial state, denoted by $s_0 = (x_0, t_0)$ from $\mathcal{X} \times \mathcal{T}$. We collect trajectories until the terminal time T as follows:

$$s'_{\hat{\boldsymbol{u}}} = (x'_{\hat{\boldsymbol{u}}}, t'_{\hat{\boldsymbol{u}}}) = (x_{\hat{\boldsymbol{u}}} + \hat{u}(x_{\hat{\boldsymbol{u}}}, t_{\hat{\boldsymbol{u}}}) \cdot \Delta t, t_{\boldsymbol{u}^e} + \Delta t), \tag{6.3a}$$

$$s'_{\boldsymbol{u}^e} = (x'_{\boldsymbol{u}^e}, t'_{\boldsymbol{u}^e}) = (x_{\boldsymbol{u}^e} + u^e(x_{\boldsymbol{u}^e}, t_{\boldsymbol{u}^e}) \cdot \Delta t, t_{\boldsymbol{u}^e} + \Delta t). \tag{6.3b}$$

We then obtain the cumulative cost for each cost features according to sampled trajectories. The cost coefficient of each feature in the cost function is updated according to the error between F_m^e and $\hat{F}_m$ (line 9); η is the step size, which can be regarded as the update rate. We use $\eta = 0.1$ in this work. We use the gap between the recovered and expert's policy to check the algorithm convergence performance.

Algorithm 2 (MWAL-MFG Algorithm)

1: Input: expert policy $\boldsymbol{u}^e$, population density $\boldsymbol{\rho}^e$ and $\boldsymbol{\omega}^e$;
2: Initialize $c_m^{(1)} = 1, \ \forall m = 1, \ldots, M$;
3: **for** $h = 1, \ldots, H$ **do**

Figure 6. (Color online) ARZ-MFG Variable Profiles at Time Instances

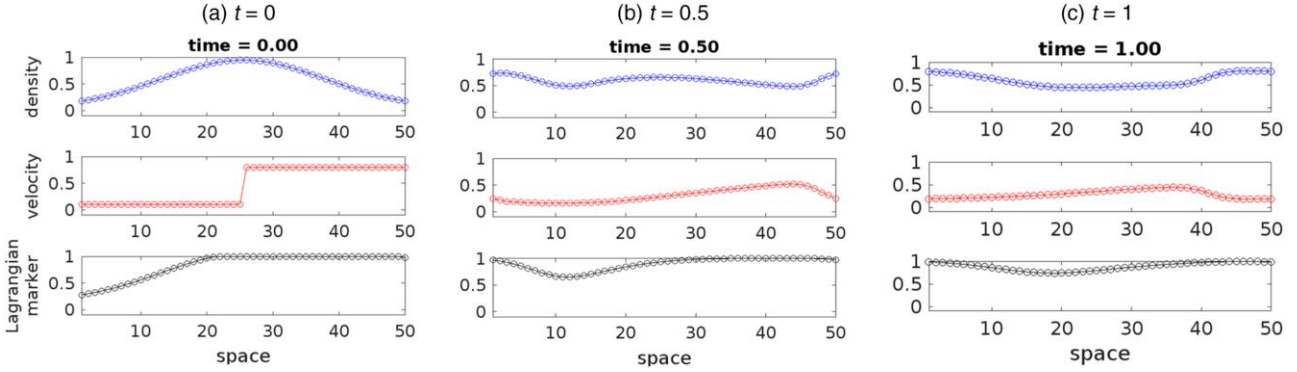
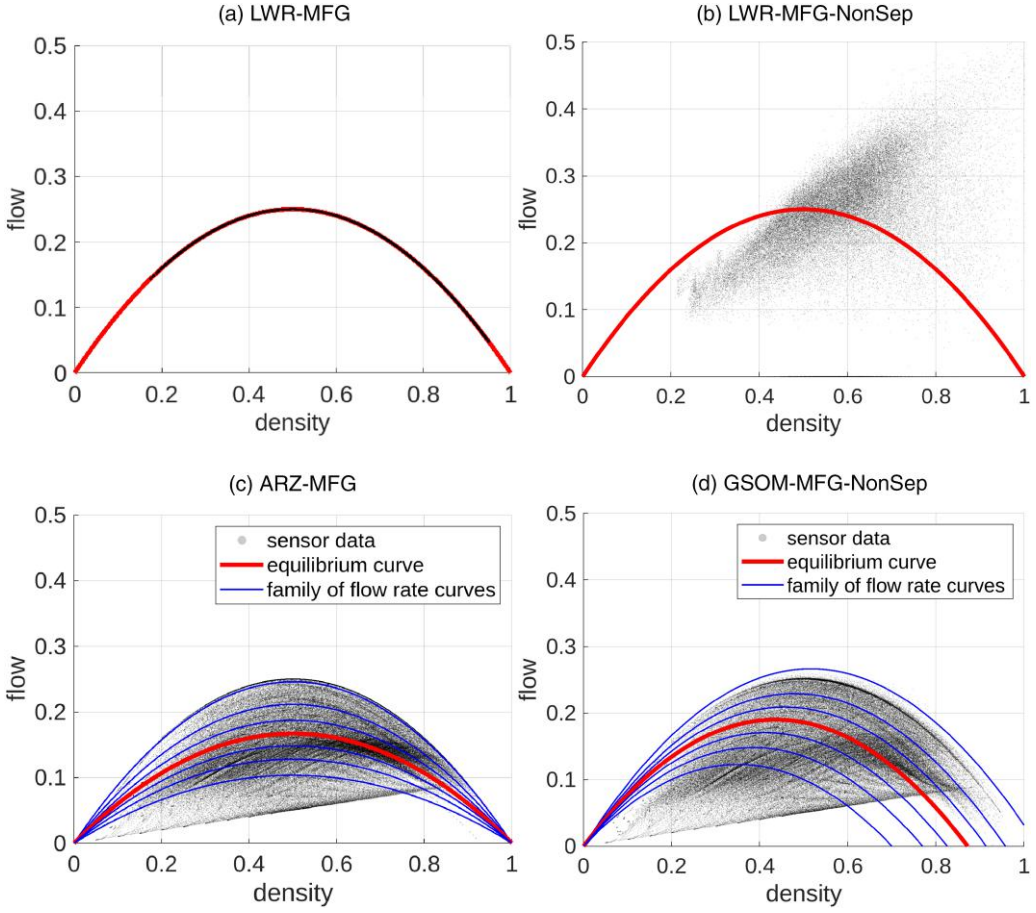


Figure 7. (Color online) Fundamental Diagrams



Notes. The red curve represents the fundamental diagram corresponding to the mean density and flow. The blue curves represent the fundamental diagrams associated with the density and flow at various sample values of w.

4: Set $c_m^{(h)} = c_m^{(h)} / \sum_{m=1}^{M} c_m^{(h)}$, $\forall m = 1, \ldots, M$;
5: Obtain the recovered policy $\hat{\boldsymbol{u}}^{(h)}$ by solving the HJB equation in MFG with cost coefficient $c_m^{(h)} = 1$, $\forall m = 1, \ldots, M$
6: Sample initial state (x_0, t_0) from $\mathcal{X} \times \mathcal{T}$ and collect trajectories according to the recovered policy (Equation (6.3a)) and expert's policy (Equation (6.3b)).
7: Obtain the value of each cost term f_m, $\forall m = 1, \ldots, M$ corresponding to the expert policy (Equation (6.2a));
8: Obtain the value of each cost term f_m, $\forall m = 1, \ldots, M$ corresponding to the policy $u^{(h)}$ (Equation (6.2b));
9: Update cost coefficient: $\forall m = 1, \ldots, M$, $c_m^{(h+1)} = c_m^{(h)} \cdot exp(ln(\beta) \cdot \eta(F_m^e - \hat{F}_m) + 2/4)$, $\beta = (1 + \sqrt{2lnM/H})^{-1}$.
10: **end for**
11: Check convergence
12: Output recovered policy $\hat{\boldsymbol{u}}$

6.2. Model Validation Experiments

6.2.1. Ring Road. In this section, we evaluate the IRL method by showing that it can recover MFE solution of GSOM-MFG in the ring road scenario. We aim to find the cost coefficient in two cost functions: ARZ and ARZ-NonSep by using the MFE results $\boldsymbol{\rho}^*, \boldsymbol{u}^*, \boldsymbol{\omega}^*$ obtained from ARZ-MFG and GSOM-MFG-NonSep. The cost function forms are

$$f_{ARZ}(u; \rho, \omega) = c_1 \left(\frac{u}{u_m}\right)^2 - c_2 \left[\left(\frac{\omega}{u_m^2} - \frac{\rho}{u_m \rho_m}\right)\right] u + c_3 \left(\frac{\omega}{u_m} - \frac{\rho}{\rho_m}\right)^2, \tag{6.4a}$$

$$f_{\mathrm{ARZ-NonSep}}(u; \rho, \omega) = c_1 \frac{1}{u_{\mathrm{m}}^2} (U(\rho, \omega) - u)^2 + c_2 \left(1 - \frac{\omega}{u_{\mathrm{m}}}\right)^2 - c_3 \left(1 - \frac{\rho}{\rho_{\mathrm{m}}}\right)^2. \tag{6.4b}$$

Table 1. Performance Comparison

Metrics	ARZ	ARZ-NonSep
Velocity MSE	0.0017	0.0178
Ground-truth cost coefficients	[0.25, 0.5, 0.25]	[0.33, 0.33, 0.33]
Recovered cost coefficients	[0.2497, 0.5002, 0.2501]	[0.3472, 0.3250, 0.3278]

We measure the reconstruction error by the mean squared error (MSE) between the real and reconstructed velocities:

$$MSE = \frac{1}{N_x N_t N_w} \sum_{k=1}^{N_x} \sum_{\tau=1}^{N_t} \sum_{w=1}^{N_w} (\hat{u}_{k,w}^{\tau} - u_{k,w}^{\tau})^2, \qquad (6.5)$$

where $\hat{u}$ is the recovered velocity and u is the velocity at MFE.

The performance comparison is summarized in Table 1. The ground truth cost coefficients used to obtain MFE corresponding to ARZ and ARZ-NonSep costs are $[0.25, 0.5, 0.25]$ and $[0.33, 0.33, 0.33]$, respectively. It is shown that the IRL method performs well when estimating the cost coefficients. There is barely any difference between the ground truth and recovered cost coefficients in the ARZ cost function.

6.2.2. Real-World Data. We use the next-generation simulation (NGSIM) data set (Coifman and Li 2017), where vehicle trajectories are recorded using bird's-eye view cameras with a sampling frequency of 0.1 seconds. We primarily use data collected from the US-101 highway segment, which provides 45 minutes of data over approximately 630 meters.

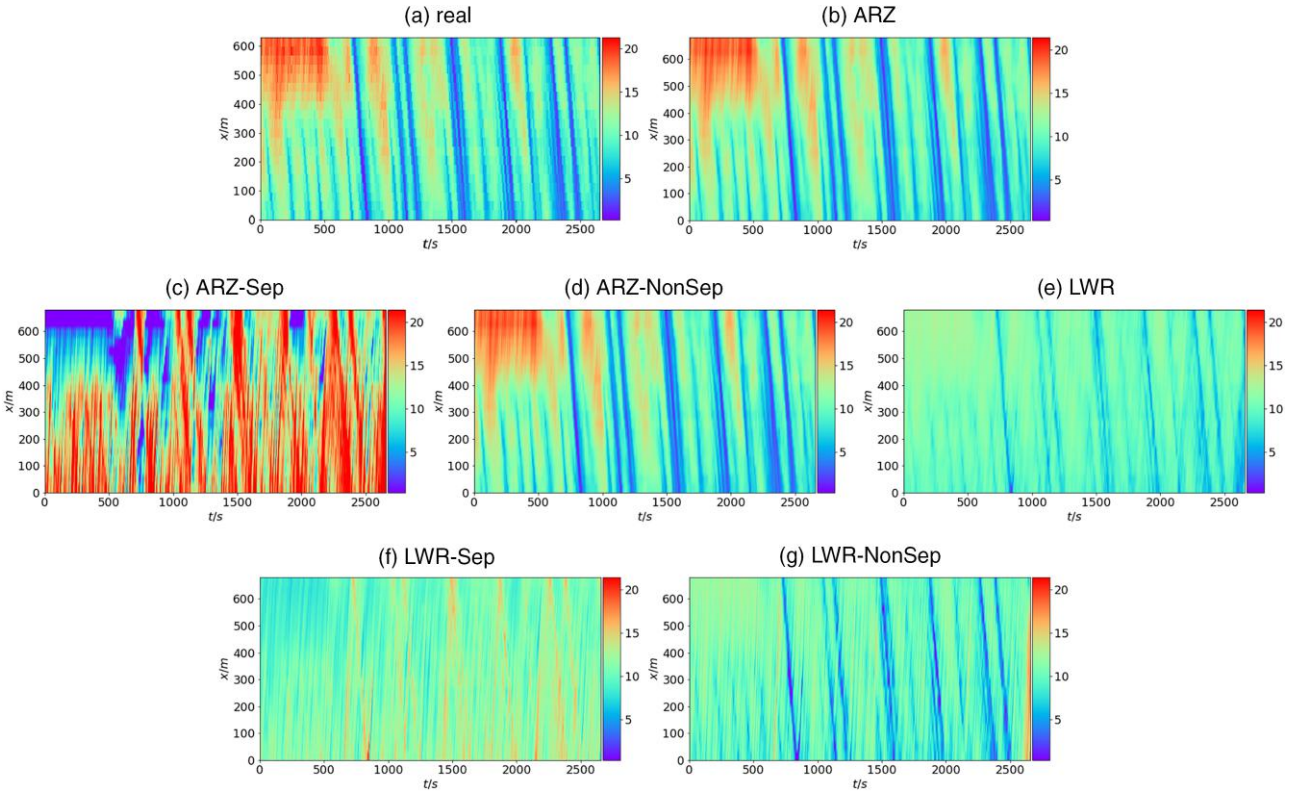
On this data set, we evaluate our method by showing that it can reproduce data by assembling real-world observations. In particular, we will show that ARZ-MFG reproduce densities and velocities with a better agreement with real traffic data, as opposed to traffic flow model like LWR or MFGs using ARZ-Nonsep as the cost function.

6.2.2.1. Data Processing. The spatial and temporal domains are partitioned into intervals of 30 meters and 1.5 seconds, respectively. These intervals were chosen to ensure a sufficient number of samples while adhering to the CFL condition and to minimize the effects of data noise by avoiding excessively small grid sizes. Consequently, the processed data includes space-mean density $\{\rho_k^{\tau}\}_{1\le k\le N_x}^{0\le \tau\le N_t}$ and space-mean velocity $\{u_k^{\tau}\}_{1\le k\le N_x}^{1\le \tau\le N_t}$,

Table 2. Model Validation with NGSIM Data

Scenario	Approach	Case	Cost function	Error (MSE)	Message
Game-theoretic setup	IRL for GSOMMFGs	ARZ	ARZ	3.34×10^{-23}	1. The ARZ-NonSep cost structure can better represent NGSIM data compared with other approaches. 2. Compared with PIDL and Kalman filter approaches, the IRL approach does not require knowledge of the physics regarding fundamental diagrams and can easily accommodate various cost structures. 3. PIDL training is limited to the time frames when data can be observed, whereas the IRL approach is not. Once the cost coefficient is estimated from observations, it can be applied to solve MFGs over any finite time horizon. In summary, the IRL approach has stronger generalizability than other methods. Our proposed GSOM-MFG is a generic framework that can accommodate various cost functions. The ARZ and LWR fundamental diagrams belong to the GSOM-MFG when the cost structures are specified.
			ARZ-NonSep	9.47×10^{-27}	
			ARZ-Sep	9.90×10^{1}	
		LWR	LWR	6.49×10^{0}	
			LWR-NonSep	8.49×10^{0}	
			LWR-Sep	2.84×10^{1}	
Deep learning setup	PIDL	ARZ	—	2.25×10^{0}	
		LWR	—	2.80×10^{0}	
Classic setup	Kalman filter	ARZ	—	1.65×10^{1}	
		LWR	—	2.11×10^{1}	

Figure 8. (Color online) Comparison Among the Real Velocity (a) and Reconstructed Ones with Different Cost Functions Using IRL (b–g)



where $N_x = 630 \text{ meters}/30 \text{ meters} = 21$ and $N_t = (45 \times 60) \text{ seconds}/1.5 \text{ seconds} - 10 = 1770$. We exclude the initial and final five time steps from our analysis. This is due to incomplete data collection at the beginning and end of the recording period. Given the processed densities and velocities, we then calculate Lagrangian markers following the same practice as in Fan, Herty, and Seibold (2014):

$$\hat{\omega} = \hat{u} + h(\hat{\rho}) = \hat{u} + \frac{\hat{\rho} u_m}{\rho_m}. \tag{6.6}$$

6.2.2.2. Velocity Error. We measure the reconstruction error between the real and reconstructed velocities by MSE, mathematically,

$$MSE = \frac{1}{N_x N_t} \sum_{k=1}^{N_x} \sum_{\tau=1}^{N_t} (\hat{u}_k^\tau - u_k^\tau)^2, \tag{6.7}$$

where $\hat{u}$ is the reconstructed velocity.

6.2.2.3. Performance Comparison. We compare the reconstruction performance of four different loss functions, namely, ARZ, ARZ-NonSep, LWR, LWR-NonSep, and LWR-Sep. We also include a deep learning–based method, that is, physics-informed deep learning (PIDL) (Shi, Mo, and Di 2021a, Shi et al. 2021b, Mo, Fu, and Di 2022a, Mo et al. 2022b, Di et al. 2023), and a classic method, that is, Kalman filter (Wang and Papageorgiou 2005, Di, Liu, and Davis 2010), for comparison. Table 2 presents the MSE of using each cost function of the MFG when reconstructing the velocity. Figure 8 illustrates the heatmaps of the reconstructed velocities along with the real ones. Figure 9 presents the reconstructed velocity of PIDL using ARZ and LWR as the physics, respectively. We can see that using ARZ and ARZ-NonSep lead to significantly better reconstruction accuracy compared with the use of LWR cost functions.

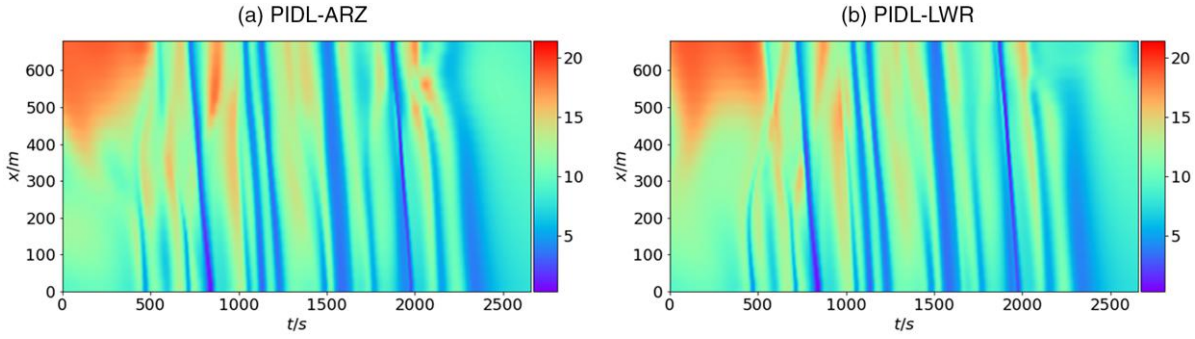
6.2.2.4. IRL Policy Evaluation. Because the ARZ-MFG achieves the minimum estimation error among all the comparative models, we will further demonstrate its performance from the perspective of IRL. To illustrate the policy learned by ARZ-MFG, we calculate the transition probability of moving to the next state s' after time step Δt given the current state $s = (x_k, t_\tau)$; namely, $Pr(s' | s = (x_k, t_\tau))$, using the following equations:

$$Pr(s' = (x_{k+1}, t_{\tau+1}) | s = (x_k, t_\tau)) = \frac{\hat{u}_k^\tau \Delta t}{\Delta x}, \tag{6.8a}$$

$$Pr(s' = (x_k, t_{\tau+1}) | s = (x_k, t_\tau)) = 1 - Pr(s' = (x_{k+1}, t_{\tau+1}) | s = (x_k, t_\tau)), \tag{6.8b}$$

where $\Delta x = L/N_x$ and $\Delta t = T/N_t$; $x_k = k\Delta x, k = 1, \ldots, N_x$; and $t_\tau = \tau \Delta t, \tau = 1, \ldots, N_t$, respectively. Because of the

Figure 9. (Color online) Reconstructed Velocity Using PIDL



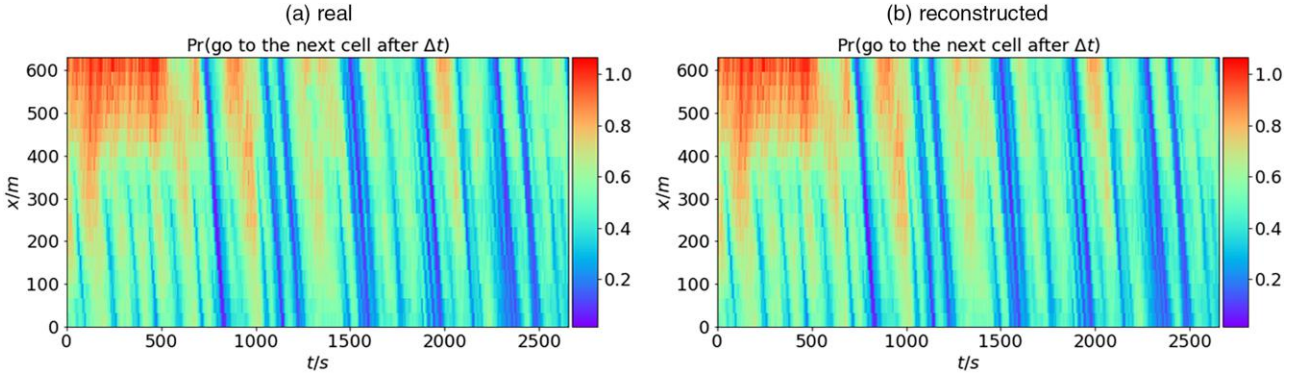
CFL condition, no other s' is allowed. The real-world transition probability can also be calculated by substituting $\hat{u}$ with the actual value u. Figure 10 presents a comparison between the transition probability matrix derived from the learned policy and the actual data.

Figure 11 depicts the frequency of visits to each state for both the real and constructed data, calculated through Monte Carlo simulations starting from a uniformly distributed initial state. We can see both the transition probability matrix and state visitation frequency of the IRL policy closely resemble the real ones.

7. Conclusions and Future Research

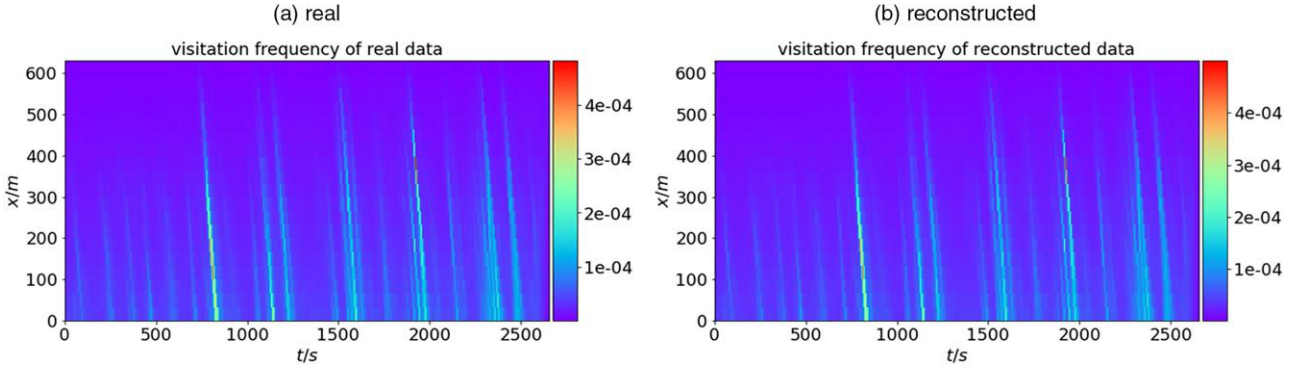
This paper establishes a game-theoretic interpretation of generalized second-order traffic flow models. Such an interpretation not only manifests an equivalence between GSOM and MFG but also allows us to modify the cost functional of the optimal control problem for the representative agent and devise other forms of MFGs and traffic flow models. We start from the first-order LWR-MFGs consisting of a representative agent's optimal control problem (depicting the generic agent's driving velocity choice, aka HJB equation), and a continuity equation (propagating density distribution of the population, aka FPK equation). We then generalize LWR-MFGs to generalized second-order traffic flow MFGs, by introducing an additional variable, denoted as w, that describes how a car's class evolves. This variable is equivalent to acceleration if it is a function of speed v, like in the ARZ traffic flow model. In our proposed generalized second-order traffic flow MFGs, GSOM-MFG, w could be a free variable that augments the dimension of the HJB equation, including the value function and the optimal velocity profile. Essentially, the HJB equation for GSOM-MFG is a parametrized version of that for LWR-MFG. The proposed GSOM-MFG encompasses first- and second-order traffic flow models, when cost function of the representative agent's optimal control problem is carefully chosen. Moreover, we propose new cost functions that lead to new parametrized velocity profiles. We develop a forward-backward fixed-point numerical method along with fictitious play to stabilize the solution algorithm. The results are demonstrated with four types of cost functions on a ring road. Furthermore, to demonstrate the validity of the proposed GSOM-MFGs, we propose to solve an inverse problem

Figure 10. (Color online) Comparison of the Transition Probability Matrices for Real (a) and Reconstructed (b) Data



Note. The color in each cell indicates the probability of transitioning to the subsequent cell in space x and time t.

Figure 11. (Color online) Comparison of State Visitation Frequencies for Real (a) and Reconstructed (b) Data



where observed individual driving trajectories are collected but the underlying cost function remains unknown. Adversarial inverse reinforcement learning is employed to estimate the coefficients involved in the cost function for all four types of games. Using the NGSIM data, we have found that ARZ-MFG generates more accurate traffic density and velocity profiles, compared with other MFG and traffic flow models.

This work can be extended in several directions: (1) investigate mathematical properties of GSOM-MFG, including existence and uniqueness of equilibria and (2) extend GSOM-MFG to networks where cars need to choose routing decisions (Huang et al. 2021). To the best of our knowledge, there is a void in the state-of-the-art that could prove such properties for the MFGs proposed in this paper, because the nonseparable cost functions do not demonstrate neither monotonic nor contraction properties. Thus, exploring the mathematical properties for traffic flow MFGs is important and urgent. Second, dynamic traffic assignment and traffic flow models on networks are primarily focused on routing choice of cars, whereas MFGs on networks allow cars to select both driving control on edges and route choice at junction points. How to formulate this problem on large-scale networks and solve it efficiently would be a challenge and left for future work. Nevertheless, in this paper, the agent's impact is assumed to be localized in space. In other words, one agent's driving control at location x can only affect traffic density at that location. Thus, we call it "local MFG traffic flow models." There is a branch of literature on nonlocal traffic flow models that model the impact of a single agent on traffic density further down- or upstream, which is beyond the scope of this paper. Nonlocal MFGs and nonlocal traffic flow models is another direction worth of exploration.

Acknowledgments

Moreover, E. Iacomini and C. Segala are members of the Indam GNCS (Italian National Group of Scientific Calculus). Z. Mo and X. Chen contributed equally to this work.

Appendix A. LWR-MFG Derivation

Recall that the LWR-MFG cost function is defined as

$$f_{\mathrm{LWR}}(u,\rho)=\frac{1}{2}(U(\rho)-u)^2=\frac{1}{2}\left[u_{\mathrm{m}}\left(1-\frac{\rho}{\rho_{\mathrm{m}}}\right)-u\right]^2, \quad \text{(A.1a)}$$

where the speed relation is characterized by the Greenshields fundamental diagram:

$$U(\rho)=u_{\mathrm{m}}\left(1-\frac{\rho}{\rho_{\mathrm{m}}}\right). \quad \text{(A.2)}$$

Using Equation (2.2), we can compute the optimal speed as

$$\begin{aligned}u&=\arg\min_{v}\{f(v;\rho)+vV_x\}\\&=\arg\min_{v}\left\{\frac{1}{2}v^2-\left[u_{\mathrm{m}}\left(1-\frac{\rho}{\rho_{\mathrm{m}}}\right)-V_x\right]v+\frac{u_{\mathrm{m}}^2}{2}\left(1-\frac{\rho}{\rho_{\mathrm{m}}}\right)^2\right\}\\&=u_{\mathrm{m}}\left(1-\frac{\rho}{\rho_{\mathrm{m}}}\right)-u_{\mathrm{m}}^2V_x=U(\rho)-u_m^2V_x.\end{aligned}$$

We then compute the optimal value as

$$\begin{aligned}&V_t+\min_{v}\{f(v;\rho)+vV_x\}=0\\&\Rightarrow V_t+\left\{\frac{1}{2}\left[u_m\left(1-\frac{\rho}{\rho_m}\right)-V_x\right]^2-\left[u_m\left(1-\frac{\rho}{\rho_m}\right)-V_x\right]\right.\\&\qquad\left.\times\left[u_m\left(1-\frac{\rho}{\rho_m}\right)-V_x\right]+\frac{u_m^2}{2}\left(1-\frac{\rho}{\rho_m}\right)^2\right\}=0\\&\Rightarrow V_t-\frac{1}{2}\left[u_m\left(1-\frac{\rho}{\rho_m}\right)-V_x\right]^2+\frac{u_m^2}{2}\left(1-\frac{\rho}{\rho_m}\right)^2=0\\&\Rightarrow V_t=\frac{1}{2}\left[u_m\left(1-\frac{\rho}{\rho_m}\right)-V_x\right]^2-\frac{u_m^2}{2}\left(1-\frac{\rho}{\rho_m}\right)^2=0\\&\Rightarrow V_t=-u_m\left(1-\frac{\rho}{\rho_m}\right)V_x+\frac{1}{2}u_m^2V_x^2\\&\Rightarrow V_t=-\left(U(\rho)-\frac{1}{2}V_x\right)u_m^2V_x.\end{aligned}$$

Substituting the above optimal velocity and value functions into Equation (2.2) leads to the LWR-MFG system (5.7).

Appendix B. LWR-MFG-NonSep Derivation

The nonsepable cost function defined for [LWR-MFG-NonSep] can be interpreted as below:

$$\begin{aligned} f_{\text{NonSep}}(u,\rho) &= \underbrace{\frac{1}{2}\left(\frac{u}{u_{\text{m}}}\right)^2}_{\text{kinetic energy}} - \underbrace{\frac{u}{u_{\text{m}}}}_{\text{efficiency}} + \underbrace{\frac{u\rho}{u_{\text{m}}\rho_{\text{m}}}}_{\text{safety}} \\ &= \frac{1}{2u_m^2}u^2 + \left(\frac{\rho}{u_{\text{m}}\rho_{\text{m}}} - \frac{1}{u_{\text{m}}}\right)u \\ &= \frac{1}{2u_{\text{m}}^2}(U(\rho)-u)^2 - \frac{1}{2}\left(1-\frac{\rho}{\rho_{\text{m}}}\right)^2, \end{aligned}$$

where $U(\rho)$ is defined in Equation (A.2).

Using Equation (2.2), we can compute the optimal speed as

$$\begin{aligned} u &= \arg\min_v\{f(v;\rho) + v\,V_x\} \\ &= \arg\min_v\left\{\frac{1}{2u_{\text{m}}^2}v^2 + \left(\frac{\rho}{u_{\text{m}}\rho_{\text{m}}} - \frac{1}{u_{\text{m}}} + V_x\right)v\right\} \\ &= u_{\text{m}}\left(1-\frac{\rho}{\rho_{\text{m}}} - u_{\text{m}}V_x\right) \\ &= u_{\text{m}}\left(1-\frac{\rho}{\rho_{\text{m}}}\right) - u_{\text{m}}^2V_x = U(\rho) - u_{\text{m}}^2V_x. \end{aligned}$$

We then compute the optimal value as

$$\begin{aligned} &V_t + \min_v\{f(v;\rho) + vV_x\} = 0 \\ &\Rightarrow V_t + \min_v\left\{\frac{1}{2u_{\text{m}}^2}v^2 + \left(\frac{\rho}{u_{\text{m}}\rho_{\text{m}}} - \frac{1}{u_{\text{m}}} + V_x\right)v\right\} = 0 \\ &\Rightarrow V_t + \left\{\frac{1}{2u_m^2}\left[u_{\text{m}}\left(1-\frac{\rho}{\rho_m}\right) - u_{\text{m}}^2V_x\right]^2\right. \\ &\qquad \left. + \left(\frac{\rho}{u_{\text{m}}\rho_{\text{m}}} - \frac{1}{u_{\text{m}}} + V_x\right)\left[u_{\text{m}}\left(1-\frac{\rho}{\rho_{\text{m}}}\right) - u_{\text{m}}^2V_x\right]\right\} = 0 \\ &\Rightarrow V_t + \left\{\frac{1}{2u_{\text{m}}^2}\left[u_{\text{m}}\left(1-\frac{\rho}{\rho_{\text{m}}}\right) - u_{\text{m}}V_x\right]^2\right. \\ &\quad \left. - \frac{1}{u_{\text{m}}^2}\left[u_{\text{m}}\left(1-\frac{\rho}{\rho_{\text{m}}}\right) - u_{\text{m}}V_x\right]\left[u_{\text{m}}\left(1-\frac{\rho}{\rho_{\text{m}}}\right) - u_{\text{m}}V_x\right]\right\} = 0 \\ &\Rightarrow V_t - \frac{1}{2u_{\text{m}}^2}\left[u_{\text{m}}\left(1-\frac{\rho}{\rho_{\text{m}}}\right) - u_{\text{m}}V_x\right]^2 = 0 \\ &\Rightarrow V_t = \frac{1}{2u_{\text{m}}^2}\left[u_{\text{m}}\left(1-\frac{\rho}{\rho_{\text{m}}}\right) - u_{\text{m}}V_x\right]^2 \triangleq \frac{1}{2u_{\text{m}}^2}[U(\rho) - u_{\text{m}}V_x]^2. \end{aligned}$$

Substituting the above optimal velocity and value functions into Equation (2.2) leads to the LWR-MFG-NonSep system (5.8).

References

Abbeel P, Ng AY (2004) Apprenticeship learning via inverse reinforcement learning. *Proc. 21st Internat. Conf. Machine Learn.* (Association for Computing Machinery, New York), 1–8.

Achdou Y, Camilli F, Capuzzo-Dolcetta I (2012) Mean field games: Numerical methods for the planning problem. *SIAM J. Control Optim.* 50(1):77–109.

Achdou Y, Mannucci P, Marchi C, Tchou N (2020) Deterministic mean field games with control on the acceleration. *Nonlinear Differential Equations Appl. NoDEA* 27(3):1–32.

Achdou Y, Mannucci P, Marchi C, Tchou N (2021) Deterministic mean field games with control on the acceleration and state constraints. Preprint, submitted April 15, https://arxiv.org/abs/2104.07292.

Albi G, Herty M, Kalise D, Segala C (2022) Moment-driven predictive control of mean-field collective dynamics. *SIAM J. Control Optim.* 60(2):814–841.

Aw A, Rascle M (2000) Resurrection of "second order" models of traffic flow. *SIAM J. Appl. Math.* 60(3):916–938.

Aw A, Klar A, Rascle M, Materne T (2002) Derivation of continuum traffic flow models from microscopic follow-the-leader models. *SIAM J. Appl. Math.* 63(1):259–278.

Balzotti C, Iacomini E (2021) Stop-and-go waves: A microscopic and a macroscopic description. *Mathematical Descriptions of Traffic Flow: Micro, Macro and Kinetic Models* (Springer, Berlin), 63–78.

Benamou JD, Carlier G (2015) Augmented Lagrangian methods for transport optimization, mean field games and degenerate elliptic equations. *J. Optim. Theory Appl.* 167(1):1–26.

Cannarsa P, Capuani R, Cardaliaguet P (2021) Mean field games with state constraints: From mild to pointwise solutions of the PDE system. *Calculation Variable Partial Differential Equations* 60(3): 108.

Capuani R, Marigonda A (2022) Constrained mean field games equilibria as fixed point of random lifting of set-valued maps. *IFAC PapersOnLine* 55(30):180–185.

Chen X, Liu S, Di X (2023a) A hybrid framework of reinforcement learning and physics-informed deep learning for spatiotemporal mean field games. *Proc. Internat. Conf. Autonomous Agents Multiagent Systems* (International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC), 1079–1087.

Chen X, Liu S, Di X (2023b) Learning dual mean field games on graphs. *Proc. 26th Eur. Conf. Artificial Intelligence (Kraków, Poland).*

Chevalier G, Le Ny J, Malhamé R (2015) A micro-macro traffic model based on mean-field games. *Proc. Amer. Control Conf.* (IEEE, Piscataway, NJ), 1983–1988.

Chiarello FA, Piccoli B, Tosin A (2021) Multiscale control of generic second order traffic models by driver-assist vehicles. *Multiscale Modeling Simulations* 19(2):589–611.

Chow YT, Li W, Osher S, Yin W (2018) Algorithm for Hamilton-Jacobi equations in density space via a generalized Hopf formula. Preprint, submitted May 4, https://arxiv.org/abs/1805.01636.

Coifman B, Li L (2017) A critical evaluation of the next generation simulation (NGSIM) vehicle trajectory data set. *Transportation Res. Part B Methodological* 105:362–377.

Costeseque G, Lebacque JP (2014) A variational formulation for higher order macroscopic traffic flow models: Numerical investigation. *Transportation Res. Part B Methodological* 70:112–133.

Costeseque G, Lebacque JP, Khelifi A (2015) Lagrangian GSOM traffic flow models on junctions. *IFAC PapersOnLine* 48(1):147–152.

Couillet R, Perlaza SM, Tembine H, Debbah M (2012) Electrical vehicles in the smart grid: A mean field game analysis. *IEEE J. Selected Areas Comm.* 30(6):1086–1096.

Cristiani E, Iacomini E (2019) An interface-free multi-scale multi-order model for traffic flow. *Discrete Continuous Dynamic Systems Ser. B* 25(11):6189–6207.

Cristiani E, Priuli FS (2014) A destination-preserving model for simulating wardrop equilibria in traffic flow on networks. Preprint, submitted September 1, https://arxiv.org/abs/1409.0350.

Delle Monache ML, Piccoli B, Rossi F (2017) Traffic regulation via controlled speed limit. *SIAM J. Control Optim.* 55(5):2936–2958.

Di X, Shi R (2021) A survey on autonomous vehicle control in the era of mixed-autonomy: From physics-based to ai-guided driving

policy learning. *Transportation Res. Part C Emerging Tech.* 125: 103008.

Di X, Liu HX, Davis GA (2010) Hybrid extended Kalman filtering approach for traffic density estimation along signalized arterials: Use of global positioning system data. *Transportation Res. Rec.* 2188(1):165–173.

Di X, Shi R, Mo Z, Fu Y (2023) Physics-informed deep learning for traffic state estimation: A survey and the outlook. *Algorithms (Basel)* 16(6):305.

Djehiche B, Tcheukam A, Tembine H (2016) Mean-field-type games in engineering. Preprint, submitted May 11, https://arxiv.org/abs/1605.03281.

Fan S, Herty M, Seibold B (2014) Comparative model accuracy of a data-fitted generalized Aw-Rascle-Zhang model. *Networks Heterog. Media* 9(2):239–268.

Festa A, Göttlich S (2017) A mean field games approach for multilane traffic management. Preprint, submitted November 11, https://arxiv.org/abs/1711.04116.

Fiedler C, Herty M, Rom M, Segala C, Trimpe S (2023) Reproducing kernel Hilbert spaces in the mean field limit. *Kinetic Related Models* 16(6):850–870.

Gong X, Piccoli B, Visconti G (2021) Mean-field of optimal control problems for hybrid model of multilane traffic. *IEEE Control Systems Lett.* 5(6):1964–1969.

Göttlich S, Iacomini E, Jung T (2020) Properties of the LWR model with time delay. *Networks Heterog. Media* 16(1):31–47.

Guéant O, Lasry JM, Lions PL (2011) Mean field games and applications. *Paris-Princeton Lectures on Mathematical Finance 2010* (Springer, Berlin), 205–266.

Guo X, Hu A, Xu R, Zhang J (2019) Learning mean-field games. *Adv. Neural Inform. Processing Systems (NeurIPS 2019)*, vol. 32 (Curran Associates, Inc., Red Hook, NY), 4963–4974.

Huang M, Malhamé RP, Caines PE (2006) Large population stochastic dynamic games: Closed-loop Mckean-Vlasov systems and the Nash certainty equivalence principle. *Comm. Inform. Systems* 6(3): 221–252.

Huang K, Chen X, Di X, Du Q (2021) Dynamic driving and routing games for autonomous vehicles on networks: A mean field game approach. *Transportation Res. Part C Emerging Tech.* 128:103189.

Huang K, Di X, Du Q, Chen X (2019) Stabilizing traffic via autonomous vehicles: A continuum mean field game approach. *Proc. IEEE Intelligent Transportation Systems Conf.* (IEEE, Piscataway, NJ), 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 Dynamic Systems Ser. B* 25(12):4869–4903.

Huang K, Di X, Du Q, Chen X (2020b) Scalable traffic stability analysis in mixed-autonomy using continuum models. *Transportation Res. Part C Emerging Tech.* 111:616–630.

Kachroo P, Agarwal S, Sastry S (2016) Inverse problem for nonviscous mean field control: Example from traffic. *IEEE Trans. Automated Control* 61(11):3412–3421.

Khelifi A, Haj-Salem H, Lebacque JP, Nabli L (2016) Lagrangian discretization of generic second order models: Application to traffic control. *Appl. Math. Inform. Sci. Internat. J.* 10(4):1243–1254.

Lachapelle A, Wolfram MT (2011) On a mean field game approach modeling congestion and aversion in pedestrian crowds. *Transportation Res. Part B Methodological* 45(10):1572–1589.

Lachapelle A, Salomon J, Turinici G (2010) Computation of mean field equilibria in economics. *Math. Models Methods Appl. Sci.* 20(04): 567–588.

Lasry JM, Lions PL (2007) Mean field games. *Japanese J. Math.* 2(1): 229–260.

Lauriere M, Perrin S, Girgin S, Muller P, Jain A, Cabannes T, Piliouras G, et al. (2022) Scalable deep reinforcement learning algorithms for mean field games. *Proc. 39th Internat. Conf. Machine Learn.*, Proceedings of Machine Learning Research, vol. 162 (PMLR, New York), 12078–12095.

Lebacque JP, Khoshyaran MM (2013) A variational formulation for higher order macroscopic traffic flow models of the GSOM family. *Proc. Soc. Behav. Sci.* 80:370–394.

Lebacque JP, Mammar S, Salem HH (2007) Generic second order traffic flow modelling. *Transportation and Traffic Theory 2007*, 755–776.

LeVeque RJ (2002) *Finite Volume Methods for Hyperbolic Problems*, vol. 31 (Cambridge University Press, Cambridge, UK).

Li J, Zhang H (2013) The variational formulation of a non-equilibrium traffic flow model: Theory and implications. *Proc. Soc. Behav. Sci.* 80:327–340.

Lighthill MJ, Whitham GB (1955) On kinematic waves II. A theory of traffic flow on long crowded roads. *Proc. Roy. Soc. London A Math. Phys. Sci.* 229(1178):317–345.

Mo Z, Fu Y, Di X (2022a) Quantifying uncertainty in traffic state estimation using generative adversarial networks. *Proc. IEEE 25th Internat. Conf. Intelligent Transportation Systems* (IEEE, Piscataway, NJ), 2769–2774.

Mo Z, Fu Y, Xu D, Di X (2022b) Trafficflowgan: Physics-informed flow based generative adversarial network for uncertainty quantification. *Proc. Joint Eur. Conf. Machine Learn. Knowledge Discovery Databases* (Springer, Berlin), 323–339.

Nisio M (2015) *Viscosity Solutions for HJB Equations* (Springer Japan, Tokyo).

Perrin S, Perolat J, Laurière M, Geist M, Elie R, Pietquin O (2020) Fictitious play for mean field games: Continuous time analysis and applications. *Proc. 34th Internat. Conf. Neural Inform. Processing Systems* (ACM, New York).

Richards PI (1956) Shock waves on the highway. *Oper. Res.* 4(1):42–51.

Ruan K, Di X (2022) Learning human driving behaviors with sequential causal imitation learning. *Proc. Conf. AAAI Artificial Intelligence* 36:4583–4592.

Ruan K, Zhang J, Di X, Bareinboim E (2023) Causal imitation learning via inverse reinforcement learning. *Proc. 11th Internat. Conf. Learn. Representations* (ICLR, Appleton, WI).

Seibold B, Flynn MR, Kasimov AR, Rosales RR (2013) Constructing set-valued fundamental diagrams from Jamiton solutions in second order traffic models. *Networks Heterog. Media* 8:745–772.

Shi R, Mo Z, Di X (2021a) Physics-informed deep learning for traffic state estimation: A hybrid paradigm informed by second-order traffic models. *Proc. Conf. AAAI Artificial Intelligence* 35:540–547.

Shi R, Mo Z, Huang K, Di X, Du Q (2021b) A physics-informed deep learning paradigm for traffic state and fundamental diagram estimation. *IEEE Trans. Intelligent Transportation Systems* 23(8):11688–11698.

Shou Z, Chen X, Fu Y, Di X (2022) Multi-agent reinforcement learning for Markov routing games: A new modeling paradigm for dynamic traffic assignment. *Transportation Res. Part C Emerging Tech.* 137:103560.

Syed U, Schapire RE (2007) A game-theoretic approach to apprenticeship learning. *Proc. 20th Internat. Conf. Neural Inform. Processing Systems* (Curran Associates, Inc., Red Hook, NY), 1449–1456.

Wang Y, Papageorgiou M (2005) Real-time freeway traffic state estimation based on extended Kalman filter: A general approach. *Transportation Res. Part B Methodological* 39(2):141–167.

Yu H, Bayen AM, Krstic M (2019) Boundary observer for congested freeway traffic state estimation via Aw-Rascle-Zhang model. *IFAC PapersOnLine* 52(2):183–188.

Zhang HM (2002) A non-equilibrium traffic model devoid of gas-like behavior. *Transportation Res. Part B Methodological* 36(3): 275–290.

Zhou F, Zhang C, Chen X, Di X (2024) Graphon mean field games with a representative player: Analysis and learning algorithm. Preprint, submitted May 8, https://arxiv.org/abs/2405.08005.