SCIENCE CHINA
Information Sciences

• REVIEW •

November 2025, Vol. 68, Iss. 11, 210201:1–210201:26
https://doi.org/10.1007/s11432-025-4649-7

Special Topic: Mean-Field Game and Control of Large Population Systems: From Theory to Practice

Mean field games for urban mobility: a review

Xuan DI[1], Zhenhui XU[2*] & Tielong SHEN[3]

[1]Department of Civil Engineering and Engineering Mechanics, Columbia University, New York NY 10027, USA
[2]Department of Systems and Control Engineering, Institute of Science Tokyo, Tokyo 152-8552, Japan
[3]School of Control Science and Engineering, Dalian University of Technology, Dalian 116024, China

Received 2 April 2025/Revised 28 June 2025/Accepted 19 August 2025/Published online 5 November 2025

Abstract This paper aims to review the literature that applies mean field games (MFGs) to various applications in transportation mobility, ranging from traffic flow, large-scale electrical vehicle fleet management, to powertrain speed control of vehicles. MFG models the limiting behavior of N-player who makes sequential optimal decisions. Due to its linkage between the microscopic agent behavior and the macroscopic population dynamics, MFG is a powerful tool to solve an equilibrium outcome of multiagent systems involving a large amount of agents. We strongly believe that MFG has opened up opportunities to capture the complex interactions among various traffic entities, especially those arising from new vehicular technology like connectivity, autonomy, and electrification. Unfortunately, the application of MFG in mobility is relatively understudied, due to the complexity of mobility systems, arising from the participation of a large number of heterogeneous agents with stochasticity in their behaviors and decision making. Through this survey paper, we hope to assemble researchers across disciplines, encompassing transportation engineering, control, mathematics, optimization, and economics, to join the force in these emerging and exciting areas and to solve large-scale problems, for optimal management and policymaking in smart transportation systems.

Keywords mean field game, traffic flow, vehicle routing, speed consensus, EV charging

Citation Di X, Xu Z H, Shen T L. Mean field games for urban mobility: a review. Sci China Inf Sci, 2025, 68(11): 210201, https://doi.org/10.1007/s11432-025-4649-7

1 Introduction

Transportation mobility is essentially a multi-agent system (MAS) that consists of a large number of rational agents, including policymakers, planners, operators, drivers, pedestrians, cyclists, who interact within a traffic environment and make sequential decisions, with the objective to optimize a self-interested or collective goal. Meanwhile, with emerging vehicular technology such as autonomy, connectivity, and electrification, cars are further equipped with the capability of powerful computing and automatic decision making, enabling efficient coordination and optimal control at a large scale. In an intelligent system where cars are connected, automated, and possibly electrified, cars make sequential decisions while interacting strategically with one another, competing for limited resources, such as road space and charging stations. The decisions range from operational decisions like longitudinal and lateral acceleration selection, tactical decisions such as lane change, and power splitting in electrified vehicles for improving efficiency to strategic decisions like destination and routing choices, and when and where to charge batteries. We cannot isolate individual cars' decision making without accounting for others' actions. Accordingly, a game-theoretical framework is a natural tool for modeling such strategic interactions.

Classical game-theoretical methods focus primarily on a small number of interacting agents. As the number of cars grows, it warrants a scalable methodological framework to design (de)centralized control for a network of cars. To model continuous-time multiagent sequential decision making in MAS, we propose an emerging game-theoretical methodology, namely, MFGs [1–4]. MFG is an increasingly important game-theoretic tool to model multi-agent decision making processes, by integrating the state-of-the-art techniques from game theory and dynamic control. It bridges the micro-macro scale between individuals' microscopic control and macroscopic population dynamics. Since its inception, it has been widely applied to finance [5, 6], engineering [7, 8], social science [9], and crowd movement [10, 11]. Its outcomes, i.e., the

* Corresponding author (email: xuzhenhui@eagle.sophia.ac.jp)

game equilibria, hold great potential to inform transportation operation and management, as well as the deployment for practical scalable tools in smart cities.

We strongly believe that the application of MFGs to transportation mobility would offer a rigorous foundation to inform policy and practice as part of the development of the smart transportation ecosystem. However, MFG for mobility is relatively less studied than in other fields like finance. We list some possible conjectures below. First, there are established decades-long developed principled methods, such as car-following models and traffic flow theories, which offer reasonable explanation to existing traffic phenomena. However, they may be inadequate to project future transportation phenomena when new entities like connected and automated vehicles (CAVs) and electric vehicles (EVs) are introduced to public roads. In contrast, MFGs scale naturally to the massive populations of self-interested agents, such as CAVs, and EVs, which interact through shared infrastructure and congestion effects. Second, mobility is a diverse discipline studied by researchers from various areas, including engineering, control, optimization, and economics. These communities focus on different methods and aspects. Nevertheless, cars' powertrain control, velocity control, routing, battery management, and fleet management are correlated and coupled, calling for an integrative paradigm. By linking each agent's optimal control problem to the overall traffic situation, MFGs can overcome fragmented perspectives by offering an integrated framework that connects micro-level decisions to macro-scale outcomes. Third, different from domains like power grid systems or finance, urban mobility is participated by diverse self-interested agents interacting in a complex urban environment. MFGs may enable the design of system-level interventions that remain individually rational, while improving collective efficiency. How to model such interactions in a coherent, data-driven environment could be complicated and requires non-mechanical models.

MFG has opened up research opportunities in transportation, for its potential in offering an analytical foundation for the development of smart transportation systems composed of a large number of heterogeneous agents that interact among one another. It bridges the micro-macro scale by characterizing microscopic behaviors of agents, while inferring macroscopic dynamics of populations. Moreover, MFG provides a tool to (re)invent new theories that would have otherwise been impossible previously. We hope this paper serves as (1) a pointer to lay out a roadmap for transportation researchers who would like to adopt MFG to solve their own transportation problems; (2) a bridge to unify the conversations among researchers in various areas, so as to collaborate collectively towards solving emerging transportation challenges. In a nutshell, the overall contributions of this paper are summarized below.

(1) The first-of-its-kind survey paper summarizing how MFG is applied to various transportation problems.

(2) Potential problems, such as velocity control of CAVs, network-wide vehicle routing, speed consensus with integrated energy management, and EV charging, that also fit the MFG framework.

(3) Open questions to call for interdisciplinary research collaboration to address challenges in modeling, computing, and analysis.

The rest of this paper is organized below. Section 2 introduces the preliminary knowledge of MFG. In Section 3, we review the related work that applies MFG to applications of urban mobility, which are categorized into four types, namely, traffic flow, vehicle routing, electrical fleet management, and powertrain control. Section 4 provides an annotated list of related papers. Section 5 concludes our work and presents potential research directions as well as open questions.

2 Preliminaries on MFG

MFG is a game-theoretic framework to model the interactions of a large population of rational, utility-optimizing agents. Each agent's dynamical behavior is characterized by an optimal decision problem, as introduced by [1,4]. Crucially, MFG leverages the "smoothing" effect of large populations, in which each agent responds to and influences the overall population density, rather than tracking individual opponents directly.

Over the past decade, MFG has proven to be a powerful tool for modeling dynamic decision-making processes in various domains [5,7,9,10]. A fundamental premise of MFG is that agents have the anticipation capability to predict how mass dynamics evolves. At a mean field equilibrium (MFE), the agents' predicted population distribution coincides with the actual mass density distribution.

From a continuous-time perspective, MFE is often characterized by the following pair of coupled partial differential equations (PDEs).

(1) A backward Hamilton-Jacobi-Bellman (HJB) equation (for the representative agent dynamic). Given the density evolution of the population, an agent solves an optimal control problem to reach a minimal cost, by anticipating other agents' choices and future system dynamics. Accordingly, this optimal control problem is represented by an HJB equation, solved backward in time given a terminal state.

(2) A forward Fokker-Planck-Kolmogorov (FPK) equation (for the population dynamic). Given the optimal control of representative agents, the population's density evolution resulting from all agents' dynamics is described by an FPK equation solved forward in time provided the initial state.

2.1 Mathematical formulation

MFGs can be broadly classified into static and dynamic frameworks. In static MFGs, there is no explicit time dependence, whereas in dynamic MFGs, each agent's state evolves over time. Since this review focuses primarily on applications with evolving, time-dependent states, we center our discussion on the dynamic formulation of MFGs. We consider both discrete-time and continuous-time settings and introduce the standard equilibrium concepts employed in MFG theory.

2.1.1 *Discrete-time MFG setting*

Finite population game. Consider an N-player Markov game represented by $\langle \mathcal{S}, \mathcal{A}, m_0, N_T, p, r, \gamma \rangle$, where $\mathcal{S}$ is the state space, $\mathcal{A}$ is the action space, $m_0 \in \Delta_{\mathcal{S}}$ is the initial distribution of states, where $\Delta_{\mathcal{S}}$ denotes the probability simplex over the set $\mathcal{S}$, $N_T \in \mathbb{N} := \{0, 1, 2, \ldots\} \cup \{+\infty\}$ is the time horizon (accommodating both finite and infinite horizons), $p : \mathbb{N} \times \mathcal{S} \times \mathcal{A} \times \Delta_{\mathcal{S}} \to \Delta_{\mathcal{S}}$ is the one-step transition probability, $r : \mathbb{N} \times \mathcal{S} \times \mathcal{A} \times \Delta_{\mathcal{S}} \to \mathbb{R}$ is the one-step reward function, $\gamma \in (0, 1)$ is a discount factor. For each agent $i \in \{1, \ldots, N\}$ at time $n \in \mathbb{N}$, $s_n^i \in \mathcal{S}$ and $a_n^i \in \mathcal{A}$ denote its state and action, respectively. The empirical distribution of states at time n is $\mu_n^N = \frac{1}{N} \sum_{j=1}^{N} \delta_{s_n^j}$, where $\delta_{s_n^j}$ is the Dirac measure at s_n^j. Agent i transitions to the next state according to

$$s_{n+1}^i \sim p_n(\cdot | s_n^i, a_n^i, \mu_n^N),$$

and receives a one-step reward $r_n(s_n^i, a_n^i, \mu_n^N)$.

A policy π^i for agent i specifies a distribution over actions for each time and state, given the relevant states at each time step. Two common information structures for policies are as follows.

• **Centralized.** Each agent observes the entire state vector $(s_n^1, \ldots, s_n^N)$. Then the policy for agent i is $\pi^i : \mathbb{N} \times (\mathcal{S})^N \to \Delta_{\mathcal{A}}$, where $\Delta_{\mathcal{A}}$ is the probability simplex over the set $\mathcal{A}$. So at time n, we have $a_n^i \sim \pi_n^i(\cdot | s_n^1, \ldots, s_n^N)$.

• **Decentralized.** Each agent only observes its own state s_n^i. Then, the policy for agent i is $\pi^i : \mathbb{N} \times \mathcal{S} \to \Delta_{\mathcal{A}}$, and thus $a_n^i \sim \pi_n^i(\cdot | s_n^i)$.

Denote by $\boldsymbol{\pi}^i = \{\pi_n^i\}_{n=1}^{N_T}$ agent i's policy sequence and by $\boldsymbol{\pi}^{-i} := \left\{(\pi_n^1, \ldots, \pi_n^{i-1}, \pi_n^{i+1}, \ldots, \pi_n^N)\right\}_{n=1}^{N_T}$ the collection of the other $N-1$ agents' policies sequence. The total expected discounted reward for agent i is then

$$J^N(\boldsymbol{\pi}^i, \boldsymbol{\pi}^{-i}) = \mathbb{E}\left[\sum_{n=0}^{N_T} \gamma^n r_n(s_n^i, a_n^i, \mu_n^N)\right].$$

Decentralized policies are considered more tractable because the computational burden of solving a centralized control problem in large-scale multi-agent systems, such as transportation networks, grows rapidly with the number of agents. For this reason, our focus is on decentralized approaches. However, they may lead to different Nash equilibria. Below, we introduce the notion of an $\varepsilon(N)$-Nash equilibrium under decentralized policies.

Definition 1 (ε-Nash equilibrium). A set of admissible decentralized policies $\{\pi^{*,i}\}_{i=1}^N$ is an ε-Nash equilibrium for the N players, if for each agent i, we have

$$J^N(\boldsymbol{\pi}^{*,i}, \boldsymbol{\pi}^{*,-i}) \leqslant J^N(\boldsymbol{\pi}^i, \boldsymbol{\pi}^{*,-i}) + \varepsilon, \quad 1 \leqslant i \leqslant N$$

for all admissible alternative policy sequencies $\boldsymbol{\pi}^i$. Here, the parameter $\varepsilon \geqslant 0$.

Mean field game. Taking the limit as $N \to \infty$ leads to the MFG formulation. In this limit, the empirical distribution μ_n^N converges to a deterministic measure μ_n, and the total expected reward becomes

$$J(\boldsymbol{\pi};\boldsymbol{\mu}) = \mathbb{E}\left[\sum_{n=0}^{N_T} \gamma^n r_n(s_n, a_n, \mu_n)\right],$$

subject to

$$s_{n+1} \sim p_n(\cdot|s_n, a_n, \mu_n), \quad a_n \sim \pi_n(\cdot|s_n), \quad s_0 \sim m_0, \quad n \in \mathbb{N}.$$

Let Π^{N_T} denote the N_T-fold product of the one-step policy sets, and let $\Delta_{\mathcal{S}}^{N_T+1}$ denote the (N_T+1)-fold product of $\Delta_{\mathcal{S}}$.

Definition 2 (Mean field equilibrium). A pair $(\boldsymbol{\pi}^*, \boldsymbol{\mu}^*) \in \Pi^{N_T} \times \Delta_{\mathcal{S}}^{N_T+1}$ is a mean field Nash equilibrium if it satisfies the following two conditions.

- **Agent rationality.** $\boldsymbol{\pi}^*$ is a best response to $\boldsymbol{\mu}^*$, where a policy $\boldsymbol{\pi}^*$ is a best response of an agent to the MF population dynamic flow $\boldsymbol{\mu}$ if it maximizes $J(\cdot, \boldsymbol{\mu})$.
- **Population consistency.** for all $t \in \mathbb{N}$, $\boldsymbol{\mu}^*$ is the distribution of s_n^*, starting with the initial distribution m_0 and controlled by policy $\boldsymbol{\pi}^*$.

A fundamental result in the MFG theory states that if an MFE exists in the infinite-population limit, then the corresponding decentralized policies $\boldsymbol{\pi}^*$ constitute an $\varepsilon(N)$-Nash equilibrium for a finite population game, with $\varepsilon(N) \to 0$ as $N \to \infty$.

2.1.2 *Continuous-time MFG setting*

In the continuous-time scenario, we consider an N-player Nash game where the state process of agent i evolves according to a stochastic differential equation (SDE)

$$\mathrm{d}X_t^i = b(t, X_t^i, a_t^i, \mu_t^N)\mathrm{d}t + \sigma(t, X_t^i)\mathrm{d}W_i(t), \quad X_0^i \sim m_0.$$

The underlying filtered probability space is $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, P)$, where $\{\mathcal{F}_t\}_{t\geqslant 0}$ is a collection of nonnondecreasing σ-algebras. Here, $X_t^i \in \mathbb{R}^n$ and $a_t^i \in \mathbb{R}^m$ are the state and control inputs of agent i, respectively, and $\mu_t^N := \frac{1}{N}\sum_{j=1}^N \delta_{X_t^j}$ is the empirical measure of the agents' states at time t. $\{W^i, 1 \leqslant i \leqslant N\}$ are independent standard Brownian motions, each adapted to $\{\mathcal{F}_t\}_{t\geqslant 0}$ and also independent of initial states.

A joint control strategy is denoted $a = (a^1, \ldots, a^N)$, and $a^{-i} = (a^1, \ldots, a^{i-1}, a^{i+1}, \ldots, a^N)$ refers to the strategy profile excluding agent a^i. The cost functional for agent i is

$$J_i^N(a^i, a^{-i}) = \mathbb{E}\left[\int_0^T f(t, X_t^i, a_t^i, \mu_t^N)\mathrm{d}t + g(x_T^i, \mu_T^N)\right], \quad 1 \leqslant i \leqslant N,$$

where $f(\cdot)$ is a running cost and $g(\cdot)$ is a terminal cost.

Analogous to the discrete-time case, the control policy can be classified into two categories.

- **Centralized control set.** $\mathcal{A}_c^i = \{a^i | a_t^i$ is adapted to $\sigma\{\bigcup_{i=1}^N \mathcal{F}_t^i\}$, where $\mathcal{F}_t^i = \sigma(X_0^i, W_s^i, 0 \leqslant s \leqslant t)$, $i = 1, \ldots, N$.
- **Decentralized control set.** $\mathcal{A}_d^i = \{a^i | a_t^i$ is adapted to $\sigma(X_s^i, 0 \leqslant s \leqslant t)\}$.

Under mild regularity, the sequence $\{\mu_t^N\}_{N\geqslant 1}$ converges as $N \to$ to a deterministic flow μ_t. The limiting control problem for a representative agent is characterized by the coupled forward-backward system

$$-\partial_t V(t,X) = \min_a \left\{ f(t,X,a,\mu_t) + \nabla_X V(t,X) b(t,X,a,\mu_t) + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^{\mathrm{T}}\nabla_{XX}^2] \right\}, V(T,X) = g(X,\mu_T), \quad (1)$$

$$\partial_t \mu_t = -\nabla_X\left(\mu_t b(t,X,a^*,\mu_t)\right)\frac{1}{2}\nabla_X^2\left(\mu_t\sigma\sigma^{\mathrm{T}}\right), \quad \mu_0 = m_0, \quad (2)$$

where $a^*(t,X) = \arg\min_a\{\cdot\}$ is the best-response control obtained from the HJB equation (1). A solution pair (V, μ) to this system is called an MFE. Assume Lipschitz continuity of b, f, σ, g and convexity in a. If (V, μ) solves (1) and (2) and $u^{*,i} = a^*(t, X_t^i)$ for each i, then these control policies form an ε-Nash equilibrium, as defined below.

Definition 3 (**ε-Nash equilibrium**). A set of admissible decentralized control inputs $\{a^{*,i}\}_{i=1}^N$ is an ε-Nash equilibrium for the N players, if for each agent i, we have

$$J_i(a^{*,i}, a^{*,-i}) \leqslant J_i(a^i, a^{*,-i}) + \varepsilon, \quad 1 \leqslant i \leqslant N$$

for all admissible alternative control input a^i, where $\varepsilon \geqslant 0$.

In conclusion, MFE solutions yield computationally tractable and asymptotically optimal control strategies for large-scale multi-agent systems.

2.2 Computational methods

MFG is challenging to solve due to its forward-backward structure. The existing solution methods can be categorized into four classes, namely, fixed-point iteration (FPI), variational method, Newton's method, and learning method. The fixed-point iteration solves the forward and backward equations alternatively till it reaches a fixed point. However, the iterations converge only when T is small, that is, for a short planning horizon. Moreover, there is no theory to estimate how small T should be to guarantee the convergence. The variational method, converting MFG to an optimization problem constrained by FPK equations [12], is efficient for any planning horizon, but relies on the separability of the cost function and potential MFGs. Newton's method [13–15] solves a large single equation system from combined forward and backward equations. It has no requirements on the length of the planning horizon or the cost function. However, Newton's method heavily relies on a good initial guess.

Numerical methods require the spatial-temporal mesh discretization. Instead, neural networks (NN) are a mesh-free scheme [16] and consequently, machine learning (ML) methods have recently gained traction in solving MFGs [17–27]. In particular, Ref. [17] applied actor-critic algorithms to MFGs and Ref. [19] investigated local Nash equilibrium (NE) in stationary MFGs. Ref. [27] developed a unified RL method for mean field game and control problems. Ref. [18] proposed a learning framework for stationary MFGs, while solving MFE still requires a fixed point approach to iteratively update policy learning of the representative agent and the dynamics of population state. Within the linear-quadratic framework, Refs. [28–32] proposed model-free methods to compute MFE and mean field social optima, respectively, without updating the mean field state. Their approaches leverage the integral reinforcement learning techniques to learn the parameters of the control policy with strict convergence guarantees. Readers can refer to [33] for an overview of learning methods for MFG.

Recent years have seen a trend of employing physics-informed machine learning (PIML) to learn PDE solutions [34–36]. Applying it to the PDE formulated MFG becomes natural. The advantage of using PIML is to attain a mess-free continuous-time solution, as vehicles are controlled at high frequency and in a continuous manner. However, since MFG is a coupled forward-backward PDE system, direct application of PIML is non-trivial. Because the PIML procedures applied to both HJB and FPK have to be coupled. In other words, we cannot train two separate PUNNs for HJB and FPK individually, as they are coupled. If we separate the training process for each PDE, the same problem of convergence by training two physics-uninformed NN (PUNN) still persist. Consequently, some studies propose the so-called population-aware MFG [37, 38], which incorporates the population as an input variable when the NN is trained for agent policies in HJB.

Since solving forward-backward dynamic systems could be unstable, the stabilization techniques include fictitious play (FP), which incorporates empirical best responses during the learning process into the decision making [24–26, 39–43], entropy regularization in loss functions [23, 44, 45], policy evaluation [46, 47], (online) mirror descent [25, 48, 49], as well as two-timescale approaches [50–52].

3 Application domains

Transportation mobility encompasses a wide variety of problems, including decision making ranging from microscopic to macroscopic scales, from the vehicle to fleet level, to traffic, and to the network level. Figure 1 summarizes the four application domains to be reviewed in this section in a unified framework.

3.1 Macroscopic traffic flow: velocity control

CAVs are anticipated to improve traffic safety and efficiency. In most literature, CAVs are essentially modeled as human drivers that can "react" faster, "see" farther, and "know" the road environment

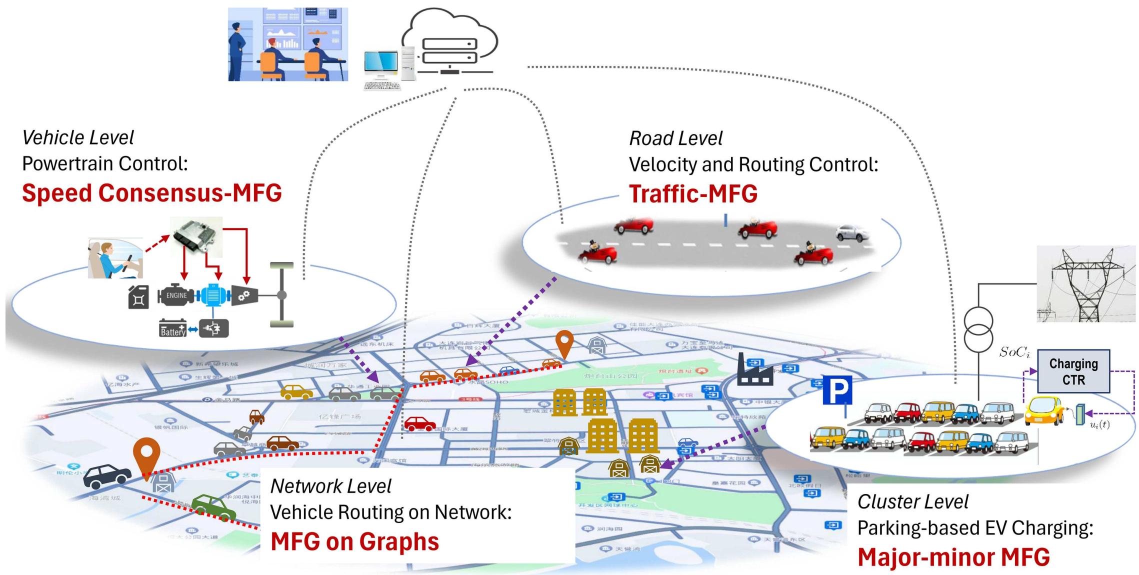


Figure 1 (Color online) Multiscale MFG framework.

better [53]. Mean-field game-theoretic methods provide a different perspective, assuming that CAVs are intelligent agents who aim to optimize a collective or self-interested goal. Under this framework, we have developed a control paradigm, including velocity control on road segments and routing on networks, particularly when traffic is dense and safe operations are vital in urban areas.

Ref. [7] built a connection between the Wardrop equilibrium and the mean field equilibrium by reinterpreting path flows as the mean field of individuals' route choices. In traffic modeling, there exist only a few studies that employ MFGs on velocity control [54–56]. We have employed this tool to solve for optimal velocity control of AVs in pure AV traffic [57] or mixed AV-HV traffic [58, 59].

3.1.1 *Preliminaries on continuum traffic flow models*

In continuum traffic flow models, the traffic state is described by aggregated density and velocity in time and space, and then the traffic system is described by partial differential equations (PDEs). These models implicitly assume that cars move according to hydrodynamics without modeling the underlying rationales of human drivers. The extensively studied continuum traffic flow models include the first-order models like Lighthill-Whitham-Richards (LWR) [60] (see (3a) and (3b)), and the generic second order model (GSOM) [61] like Aw-Rascle-Zhang (ARZ) [62, 63] (see (4b) and (4c)).

Here, the first-order traffic flow model refers to a class of models containing only one equation with one conserved quantity, which is traffic mass. In contrast, the second-order traffic flow model involves two equations with two conserved variables, namely, traffic mass and momentum. On a highway at position x and time t, define traffic density as $\rho(x,t)$, and traffic velocity as $u(x,t)$. The traffic dynamic evolution can be characterized by PDEs, exemplified below.

$$\text{[LWR]}\begin{cases}\rho_t + (\rho u)_x = 0, & \text{(3a)}\\ u = U(\rho), & \text{(3b)}\end{cases}$$

$$\text{[ARZ]}\begin{cases}\rho_t + (\rho u)_x = 0, & \text{(4a)}\\ (\rho w)_t + (\rho u w)_x = 0, & \text{(4b)}\\ u = U(\rho, w), & \text{(4c)}\end{cases}$$

where $U(\rho), U(\rho, w)$ denote traffic velocity functions, mappings from traffic density (and property) to velocity. In particular, these velocity functions show non-increasing relations between traffic density and traffic velocity, namely, the dense traffic is, the slower cars move. Some common functional forms include $U(\rho) := u_m\left(1 - \frac{\rho}{\rho_m}\right), U(\rho, w) := u_m\left(\frac{w}{u_m} - \frac{\rho}{\rho_m}\right)$, where and ρ_m, u_m are maximum traffic density (i.e., maximum number of cars per unit length per lane, at around 150 veh/km/lane) and maximum velocity (i.e., highest speed a car can drive, normally equivalent to posted speed limit of a road segment, for example, 60 miles/h), respectively.

We would like to clarify that $w \in [0, u_m]$, which describes driving behavior, is a macroscopic variable to capture traffic property in terms of behavioral heterogeneity. When $w = u_m$, ARZ is reduced to LWR. In contrast to the assumption that every driver is assumed to follow the same speed limit of u_m in LWR model, in ARZ model, every driver drives at a range of maximum speeds.

In classical continuum traffic flow models, cars are however nonstrategic players, in the sense that their actions are rule-based or determined by some explicit formula. If we assume that cars are strategic players of which the optimal policy is determined by some optimal control problem, a game-theoretical framework is warranted for sequential decision making in driving processes. Moreover, in a traffic environment, one car normally interacts with a large number of other cars present on road simultaneously. Thus, MFG is a natural tool to model the sequential choices of many cars interacting in a traffic environment.

3.1.2 *Problem statement*

Definition 4 (Traffic-MFG)**.** On a highway, there are N cars indexed by $i \in \mathcal{N} := \{1, 2, \dots, N\}$ driving in one direction with initial positions $x_{1,0}, \dots, x_{N,0}$. Their positions at time t are denoted as $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_N(t)]$. The car dynamic is described by

$$\frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = v_i(t), \quad x_i(0) = x_{i0}, \quad i \in \mathcal{N}.$$

Each car aims to select its optimal velocity control by minimizing a driving cost functional defined in

$$J_i^N(v_i, v_{-i}) = \int_0^T \underbrace{f_i^N(v_i(t), x_i(t), x_{-i}(t))}_{\text{cost function}} \,\mathrm{d}t + \underbrace{V_T(x_i(T))}_{\text{terminal cost}}, \ \forall i = 1, \dots, N, \tag{5}$$

where, $\int_0^T f_i^N(v_i(t), x_i(t), x_{-i}(t))\,\mathrm{d}t$ is the running cost over the entire planning horizon, and $f_i^N(\cdot)$ is the cost function that quantifies driving objectives such as efficiency and safety. Here we assume it is a strictly convex function with respect to v_i. $V_T(x_i(T))$ is the terminal cost representing the i-th car's preference on the final position at time T.

A Nash equilibrium of the n-car differential game is a tuple of controls $v_1^*(t), v_2^*(t), \dots, v_N^*(t)$.

As $N \to \infty$, standard mean field approximation is applied to approximate the N-car differential game by traffic flow mean field game [Traffic-MFG]. To save space here, we refer the interested readers to [57] for the derivation from differential game to MFG. [Traffic-MFG] is represented as

$$\text{[Traffic-MFG]} \quad \begin{cases} \text{(CE)} \quad & \rho_t + (\rho u)_x = 0, & \text{(6a)} \\ \text{(HJB)} \quad & V_t + f^*(V_x, \rho) = 0, & \text{(6b)} \\ & u = f_p^*(V_x, \rho), & \text{(6c)} \end{cases}$$

where $\rho(x,t), u(x,t), \forall (x,t) \in \mathcal{S} \times \mathcal{T}$ are traffic density and velocity fields, respectively, where $\mathcal{S} := [0, L]$ with $L > 0$ representing the length of the road segment, and $\mathcal{T} := [0, T]$. Denote the MFE solution by $\rho^*(x,t)$ and $u^*(x,t)$, namely,

$$\mathrm{SOL}(\text{[Traffic-MFG]}) = \{\rho^*(x,t), u^*(x,t)\}, x \in \mathcal{S}, t \in \mathcal{T}. \tag{7}$$

The initial and terminal conditions are provided by the initial density $\rho(x,0) = \rho_0(x)$ and the terminal cost $V(x,T) = V_T(x)$, respectively. The choice of boundary conditions depends on traffic scenarios.

(1) When cars drive on a ring road without any entrance or exit, the periodic boundary conditions follow as $\rho(0,t) = \rho(L,t), V(0,t) = V(L,t)$. In other words, on a circular road, traffic density and value function are the same at the beginning and the ending of the road, because these two ends are essentially the same point.

(2) When the road has an entrance at $x = 0$ and an exit at $x = L$, we need to impose the boundary conditions $\rho(0,t) = \rho_{\text{entr}}(t)$ representing traffic inflow at the left boundary, indicating the traffic that originates from the upstream.

Table 1 Summary of Traffic-MFG models. $U(\rho), U(\rho, w)$ denote traffic velocity functions, mappings from traffic density (and property) to velocity. ρ_m, u_m are maximum traffic density (i.e., maximum number of cars per unit length per lane) and maximum velocity (i.e., highest speed a car can drive), respectively. More explanations of the notations can be found in Subsection 3.1.1.

Type	Agent's state $\boldsymbol{s}$	Action $\boldsymbol{a}$	Population's state $\boldsymbol{\mu}$	Model	Cost function $f(\rho, u)$	References
First-order	x	v	ρ	LWR	$\frac{1}{2}\left(U(\rho)-u\right)^2$, where $U(\rho)=u_m\left(1-\frac{\rho}{\rho_m}\right)$	[54, 57–59, 66–69]
				MFG-LWRSep	$\frac{1}{2}(\frac{u}{u_m})^2-\frac{u}{u_m}+\frac{\rho}{\rho_m}$	
				MFG-LWR	$\frac{1}{2u_m^2}\left(U(\rho)-u\right)^2-\frac{1}{2}\left(1-\frac{\rho}{\rho_m}\right)^2$	
Second-order	$\begin{bmatrix} x \\ w \end{bmatrix}$	a	$\begin{bmatrix} \rho \\ \omega \end{bmatrix}$	ARZ	$\frac{1}{2u_m^2}\left(U(\rho,\omega)-u\right)^2$, where $U(\rho,\omega)=u_m\left(\frac{\omega}{u_m}-\frac{\rho}{\rho_m}\right)$	[65]
				MFG-GSOMSep	$\frac{1}{2}\left(1-\frac{u}{u_m}\right)^2+\frac{1}{2}\left(1-\frac{\omega}{u_m}\right)^2-\frac{1}{2}\left(1-\frac{\rho}{\rho_m}\right)^2$	
				MFG-GSOM	$\frac{1}{2u_m^2}\left(U(\rho,w)-u\right)^2+\frac{1}{2}\left(1-\frac{w}{u_m}\right)^2-\frac{1}{2}\left(1-\frac{\rho}{\rho_m}\right)^2$	

3.1.3 *Model elements*

In [57, 64], the non-cooperative control of multiple CAVs is formulated first as an N-car differential game. When N goes to infinity, MFG is shown to be a scalable model for the differential game, as the car population grows. The distributed velocity controller derived from the MFE is shown to be an ϵ-equilibrium of the N-car differential game.

In the longitudinal control of CAVs, each car solves its optimal velocity backward in time, the aggregate effect of which is formulated by an HJB equation; while the mean field approximation derives the evolution of traffic density solved by a transport equation (with many other names like continuity equation, flow conservation equation) forward in time. The distributed velocity controller derived from the MFE is shown to be an ϵ-equilibrium of the N-car differential game.

In the literature, a family of MFGs are proposed equivalent with the LWR (defined in (3a) and (3b)) model in [57] and ARZ (defined in (4a) and (4c)) in [65], respectively. In particular, Ref. [57] has established a connection between an MFG-based macroscopic continuum model and the LWR model. The LWR model has been proven to be a myopic MFG with a specially designed objective function. In conclusion, MFG embodies classical traffic flow models with behavioral interpretation, thereby providing a flexible behavioral foundation and a promising direction to accommodate new traffic entities like CAVs.

How do we define appropriate rewards that could lead to desired outcomes? Ref. [65] has proposed several reward functions, with two listed below and more summarized in Table 1, and explored reward design using a widely used real-world data set, i.e., next generation simulation (NGSIM).

$$f_{\text{MFG-LWR}}(u,\rho)=\underbrace{\frac{1}{2u_m^2}\left(U(\rho)-u\right)^2}_{\text{equilibrium speed}}-\underbrace{\frac{1}{2}\left(1-\frac{\rho}{\rho_m}\right)^2}_{\text{safety}}, \tag{8a}$$

$$f_{\text{MFG-GSOM}}(u,\rho,w)=\underbrace{\frac{1}{2u_m^2}\left(U(\rho,w)-u\right)^2}_{\text{equilibrium speed}}+\underbrace{\frac{1}{2}\left(1-\frac{w}{u_m}\right)^2}_{\text{flow homogeneity}}-\underbrace{\frac{1}{2}\left(1-\frac{\rho}{\rho_m}\right)^2}_{\text{safety}}. \tag{8b}$$

Table 1 summarizes the different types of traffic-MFGs in the literature. For the existing traffic flow models like LWR and ARZ, we can always find a cost function that leads to equivalent formulation of classical PDE models. Inspired by the cost functions identified from these classical traffic flow models, new cost functions have been introduced for traffic flow MFG models. Representative examples are summarized in Table 1. These new cost functions lead to new equilibrium solutions and outcomes.

The comparison of these rewards over the real-world dataset NGSIM is demonstrated in Figure 2, with the heatmaps of the reconstructed velocities along with the real ones. MFG-GSOM outperforms MFG-LWR in terms of estimation accuracy, defined as the mean square error between the observed value and the estimated value of traffic densities using MFG-GSOM.

3.1.4 *Solution methods*

The majority of existing studies are primarily focused on a special class of MFGs or can be reformulated as a potential game in which the reward or cost function satisfies the Lasry-Lions monotonicity conditions [2,

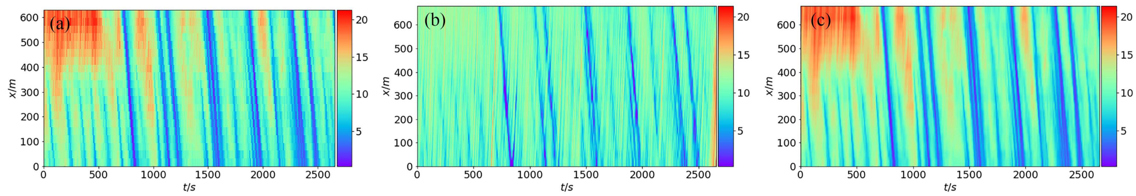


Figure 2 (Color online) (a) Comparison among the real velocity and (b), (c) the reconstructed ones with the proposed cost functions ((b) with the cost function defined in (8a) and (c) with the cost function defined in (8b)).

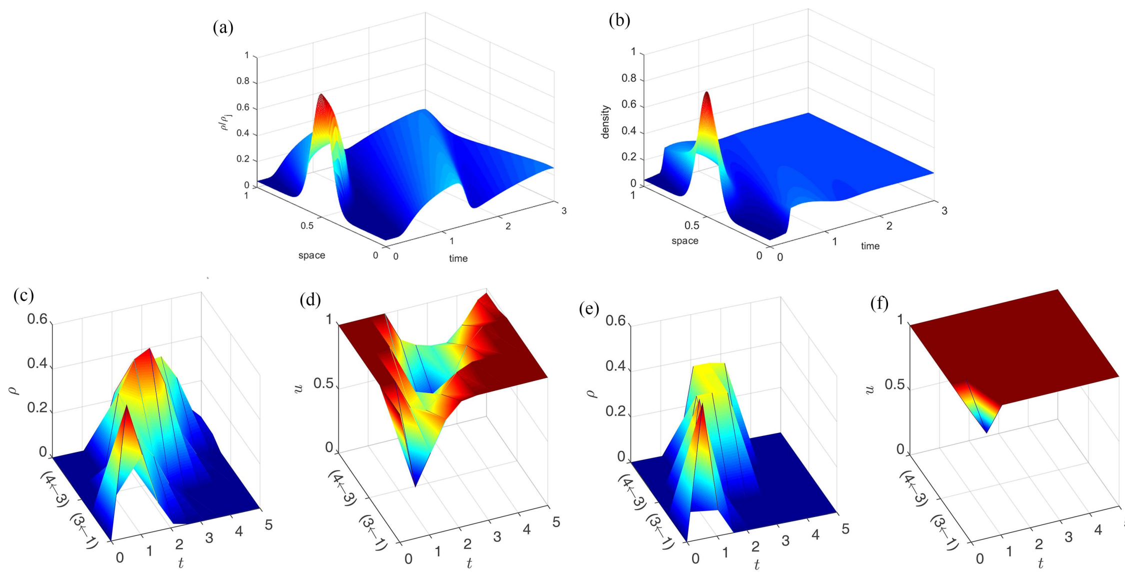


Figure 3 (Color online) Ring road (top) vs. network (bottom). (a) LWR; (b) MFG; (c) LWR density: path $1 \to 3 \to 4$; (d) LWR velocity: path $1 \to 3 \to 4$; (e) MFG density: path $1 \to 3 \to 4$; (f) MFG velocity: path $1 \to 3 \to 4$.

3, 42]. In contrast, Traffic-MFG is non-stationary (i.e., time-dependent control policies), with continuous state and action spaces (i.e., agent state, agent control, and population state are spatiotemporal), as well as non-separable cost functional (i.e., representing the congestion effect that the higher density, the lower velocity, but violating the monotonicity conditions). In other words, Traffic-MFGs is not a potential game and all the above techniques and convergence results cannot be directly applied. To resolve such a challenge, Refs. [70, 71] provided global well-posedness results for second-order master equations in regimes where standard monotonicity fails. Ref. [57] developed a multigrid preconditioned Newton's finite difference algorithm [13–15], coupled with preconditioning techniques to speed up computation. The learning algorithm for this game is explored in [66] to solve HJB. To stabilizie the iteration between HJB and FPK, the fictions play technique is applied.

Since the speed control of CAV is continuous, discrete-time algorithm could lead to less accurate strategies. Leveraging the known dynamics of traffic-MFG, Ref. [66] employed PIML coupled with RL, where PIML offers a grid-free scheme where fine-grained population dynamics can be approximated. In particular, PIML is used to solve FPK, while RL is for HJB. These two processes are performed iteratively till a convergence is reached.

3.1.5 *Results and findings*

Refs. [58, 59] found that the MFG mitigates traffic oscillation faster than LWR, even though the MFG is not intended to stabilize traffic by design. Figure 3 demonstrates the results on a ring road. Around a jam area with symmetric traffic density, vehicles driven by MFG controllers tend to slow down farther upstream before joining the jam and immediately speed up after leaving the jam; in contrast to those driven by LWR controllers whose speed remains symmetric before and after the jam area. This is because LWR's velocity is determined only through traffic density at that location, while that of the MFG depends on the traffic density of the entire horizon.

3.2 Vehicle routing on networks

Modeling MFGs on large graphs or networks is understudied, partly because the traffic dynamics at junction points and edges need to be coupled. Traffic dynamics at junction points could be complex, as there could be spatial queues that accumulate at nodes, which would propagate upstream and affect traffic dynamics of other road edges, and accordingly, the overall dynamics on the entire graph.

3.2.1 *Preliminaries on dynamic traffic assignment models*

Dynamic traffic assignment (DTA) models, consisting of a dynamic network loading (DNL) module and a choice model, aim to describe within-day dynamics of traffic flows on networks [72–74]. Given traffic demand and route choices, a DNL module propagates traffic flow dynamics and traffic congestion forward in time, both on links and at junctions. Route choice and departure-time choice guide travelers to select the next go-to links and departure-time at starting nodes. Travelers could be either cooperative or competitive, leading to dynamic system optimum [75] and dynamic user equilibrium [76, 77], respectively. Normally, these two modules are solved iteratively until they converge to a fixed point (see Figure 4).

DTA models, similar to classical continuum traffic flow models, assume that cars are non-strategic players, they neither select optimal driving velocity while moving on a road segment. Accordingly, extending DTA to accounting for continuous speed selection while assuming that CAVs are strategic players necessitate a game-theoretical framework. On a road network, one car interacts with a large number of other cars present on road simultaneously. Naturally, MFG is considered for such a modeling paradigm.

3.2.2 *Problem statement*

Definition 5 (MFG on graphs)**.** On a directed graph, denoted as $\mathcal{G} = \{\mathcal{N}, \mathcal{L}\}$ where $\mathcal{N}$ is the node set and $\mathcal{L}$ is the edge set, a generic agent moves from its initial position to a destination, aiming to select optimal control to minimize its cost connecting its origin to the destination. For each link $l = (i, j) \in \mathcal{L}$ where $i, j \in \mathcal{N}$, denote its starting point as $\mathrm{START}(l) = i$ and its end point as $\mathrm{END}(l) = j$. The length of link $l \in \mathcal{L}$ is denoted as $\mathrm{len}(l) \geqslant 0$. For each node $i \in \mathcal{N}$, denote $\mathrm{IN}(i) \subset \mathcal{L}$ the set of links whose end point is node i and $\mathrm{OUT}(i) \subset \mathcal{L}$ the set of links whose starting point is node i.

Optimal control of a representative agent. An agent is either in the interior of an edge or at a node.

(1) In the interior of edge $l \in \mathcal{L}$:

- **State** (x, t) is the agent's position on edge l at time t where $x \in [0, \mathrm{len}(l)]$;
- **Action** $u_l(x, t)$ is the velocity of the agent at position x at time t when navigating edge l;
- **Reward** $r_l(u, \rho)$ is the congestion cost arising from the agent population on edge l, which is increasingly monotone in ρ indicating the congestion effect. Define $V_l(x, t)$ as the minimum cost of the representative agent starting from position x at time t, solved by HJB

$$\partial_t V_l(x,t) + \min_u \{r_l(u,\rho) + u\partial_x V_l(x,t)\} = 0, \forall l \in \mathcal{L} \tag{9}$$

with a terminal condition of $V_l(\mathrm{len(l)}, t) = \nu_i(t), l \in \mathrm{IN}(i)$, and ν_i is defined in 10.

(2) At node $i \in \mathcal{N}$:

- **State** $x(t) = \mathrm{len}(l)$ is agent's position at time t at the ending node of edge l, denoted as node i;
- **Action** $\beta(i, l, t)$ is the probability of choosing the next-go-to edge $l \in \mathrm{OUT}(i)$, and can be interpreted as the proportion of agents selecting edge l (or turning ratio) at node i at time t. We have $\sum_{l \in \mathcal{L}^O(i)} \beta(i, l, t) = 1$;
- **Value** $\nu_i(t)$ is the minimum travel cost starting from node i at time t, satisfying an HJB:

$$\nu_i(t) = \min_{l \in \mathrm{OUT}(i)} V_l(0, t) \tag{10}$$

with a terminal cost at destination s as $\nu_s(t) = 0$, where $V_l(0, t)$ is the minimum cost entering edge l (i.e., $x = 0$) at time t. Eq. (10) can be interpreted as that, the minimum cost from node i to a destination is the minimum among all the value functions over edge l emanating from node i.

Population dynamics.

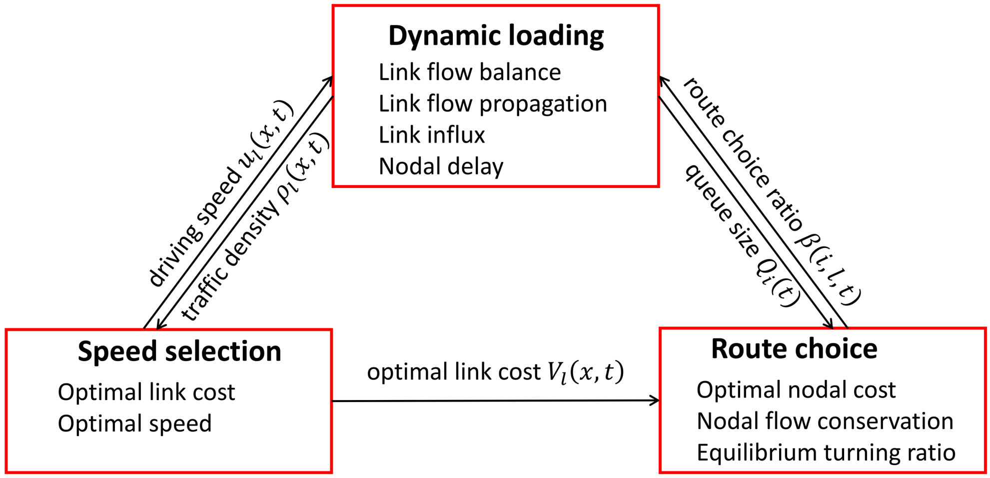


Figure 4 (Color online) Analytical framework for MFG on networks.

(1) **In the interior of edge** $l \in \mathcal{L}$. the population density distribution on edge l, denoted as $\rho_l(x,t), \forall l \in \mathcal{L}$, evolves by an FPK, given the velocity control $u_l(x,t)$ of agents:

$$\partial_t \rho_l(x,t) + \partial_x [\rho_l(x,t) \cdot u_l(x,t)] = 0 \tag{11}$$

with an initial condition $\rho_l(x,0) = \rho_0(x)$.

(2) **At node** $i \in \mathcal{N}$. Let $Q_i(t)$ denote the queue size at node i at time t. Then its dynamic is described as rate of change of queue size = total incoming flow − total outgoing flow.

At the starting node of edge l or the starting position on edge l, agents move to the next-to-go edge based on their route choice. The boundary condition is $\rho_l(0,t)u_l(0,t) = \beta(i,l,t)\{\sum_{h\in \mathrm{IN}(i)} \rho_h(\mathrm{len}(h),t) \cdot u_h(\mathrm{len}(h),t)\}$, where $\rho u =$ flux, in other words, total volume is defined as the product of density and velocity. Its physical meaning is that the traffic flux at the left boundary of an edge equals the sum of inflows into this edge from the right boundary of its adjacent upstream edges.

3.2.3 *Model elements*

Ref. [78] used the notion of MFG for routing games where static traffic density flow is solved on each link but its temporal evolution is not captured. Furthermore, cars need to optimize both longitudinal control on link segments and route choice at junction points to determine their next-go-to links. Accordingly, building on the single road MFG, Ref. [64] developed a networked MFG to describe how CAVs move across a road network. Figure 4 demonstrates the three modules involved in CAVs' decision making components while driving on road networks. The top module depicts the aggregate traffic conditions resulting from agents' choices. The bottom left module represents cars' time-dependent speed choice, and the bottom right module models cars' route choice behavior. These three modules are interacting in a way that cars' driving speed choice on road segments and routing choice at junction points impact how many cars are present on each road segment and at junctions at each time instance.

3.2.4 *Solution methods*

While solving MFG is challenging due to its forward and backward coupling, solving equilibria for MFGs on graph is even more challenging for high-dimensional action space, because each agent has to make decisions when they are at junction nodes or on edges. In [64], MFG is developed to model both velocity and route choices of a large number of intelligent agents on a road network. The numerical methods are challenging given the tight coupling of three modules shown in Figure 4. Thus, we have developed various learning methods to tackle such a computational challenge. We first discretize this game in time and space, resulting in a mixed complementarity problem (MiCP). Fixed-point iteration is employed on the queue size. The algorithm is kicked start with a feasible initialization of the queue size. In each iteration, the current queue size is updated, leading to an updated relaxed MiCP. Accordingly, the solution of the relaxed MiCP gives a set of updated queue sizes. The algorithm iterates with the queue size updated, until the convergence error falls within a predefined threshold.

To resolve the challenge in the convergence of forward and backward iterations, the existing literature employs two techniques, namely, reformulation of MFG as one so-called population MDP, or development of a single-loop algorithm. In particular, Ref. [67] formulated the MFG on graphs as a population MDP, which facilitates the development of MDP-based algorithms. The agent policy and population distribution are updated simultaneously over the entire horizon, without alternating forward and backward processes. In contrast, Refs. [38,79] developed a single-loop scheme that updates the agent policy and the population dynamics simultaneously and found that this scheme is stable by nature, as the computation of the gradient accounts for the descent directions of both the agent and the population. Refs. [66, 68, 69, 80] integrated physics-informed deep learning (PIDL) and reinforcement learning (RL) over graph neural networks to learn the fixed points of the MFG, and the FP technique is used to guarantee convergence.

3.2.5 *Results and findings*

Ref. [64] implemented the mode on Braess network with four nodes and five edges. Comparison results are illustrated in Figures 3(c)–(f). The traffic density for both the LWR and MFG equilibrium along the one path is plotted in a 3D diagram. The x-axis represents the path as a continuous road of length 2, the y-axis is time and the z-axis represents traffic density. The difference demonstrates a similar pattern as observed on a single ring road. At the LWR equilibrium, CAVs drive faster in low density areas and slower in high density areas. This is because the LWR speed is myopic, determined solely by its local traffic density. In contrast, CAVs controlled by MFG drive at relatively high speeds at relatively high density. This is because CAVs deployed with the MFG can look "farther" and adjust their driving speeds in a way to optimize their total travel costs over the network. In other words, even though the cost incurred by high speed and high density is relatively large for the current link where a car is, the car can anticipate the future cost incurred on the subsequent links with the goal of minimizing the cumulative total cost.

3.3 Speed consensus with energy management

The ever-growing concerns about environmental pollution and the impending energy crisis have prompted the automotive industry to explore advanced powertrain technologies that are both efficient and environmentally friendly. As a solution, the electrification of powertrain has shown big potential and so has become a main trend in powertrain technology as predicted by the International Energy Agency (IEA) [81]. Meanwhile, a vital point of focus is cooperative driving at the vehicle level, where improvements in fuel economy derive not only from collaborative traffic flow but also from each HEV's individual energy management strategy (EMS). MFG theory offers a tractable framework for such large-scale coordination problems, wherein the control (e.g., acceleration or speed) of each vehicle is coupled with the collective behavior of the population. By exploiting MFG methods, decentralized control laws can be obtained for real-time powertrain operation, simplifying both the analysis and the computations needed for large fleets of vehicles.

3.3.1 *Preliminaries on HEV powertrain*

A typical architecture of powertrain is shown in Figure 5, which features a planetary gear unit (PGU) connecting three power sources: the internal combustion engine, the generator, and the motor. The PGU consists of three main elements (the carrier, sun gear, and ring gear) and enables the powertrain to function as energy transformation from the fuel and the electricity to the driving power acting to the vehicle wheels. Speed control is usually achieved by decision of the driving torque, and the energy management for improving the energy efficiency is achieved by on-board management of energy flow between the energy source by operating the engine and the electric machines [82, 83].

Longitudinal dynamics is formulated as

$$M\dot{v} = \frac{\tau_{dr}}{R_{tire}} - \left[Mg(\mu_r \cos\theta + \sin\theta) + \frac{1}{2}\rho_a A C_d v^2\right], \tag{12}$$

where v is the speed of vehcile, M, g, R_{tire}, μ_r, θ, ρ_a, A and C_d denote vehicle mass, gravity acceleration, wheel radius, rolling coefficient, slope, air density, frontal area and drag coefficient, respectively. τ_{dr} denotes the driving torque acting on the wheels, which is generated by integration of the power sources

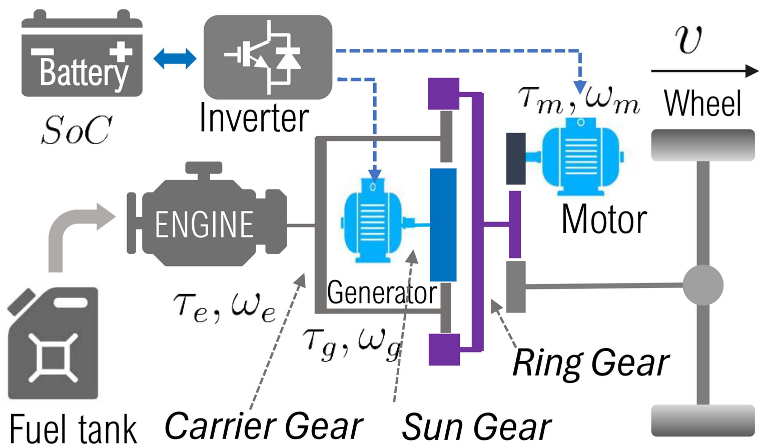


Figure 5 (Color online) Schematic diagram of the powertrain.

connected to the powertrain. Under some reasonable assumptions on the physics of the powertrain, the relationship between the driving torque and the power source is modeled as follows:

$$\frac{\tau_{dr}}{g_f} = \tau_m + \frac{r}{1+r}\tau_e, \tag{13}$$

$$0 = (1+r)\tau_g + \tau_e, \tag{14}$$

where τ_e, τ_m and τ_g are the torques of the engine, motor and generator, respectively. The parameter $r = R_r/R_s$ is determined by R_r and R_s, the radii of the ring gear and sun gear.

This equation is a constraint for deciding the torques split to τ_e, τ_m and τ_g under the driving torque demand, which is usually decided for speed control. Obviously, there still are freedoms in the decision of the three torque, which enable to adjusting the rotational speed of the engine and the generator under another constraint, which is from the mechanical construction of PGU [82, 83]

$$\omega_g = (1+r)\omega_e - r\omega_m, \tag{15}$$

where ω_e, ω_m and ω_g are the rotational speeds of the engine, motor and generator, respectively. The final gear ratio g_f relates the motor speed with the drive shaft speed ω_w by $g_f = \omega_m/\omega_w$, where $\omega_w = v/R_{tire}$. Since the efficiency of the fuel and electricity consumption depends on the operating point coordinated by the generated torque and the speed, see [82–84] for details, the operating points of the engine and the motor could be optimized to achieve high energy efficiency of the powertrain. This issue is referred to as an energy management problem in powertrain control.

Moreover, the state-of-charge (SoC) is also an important factor to be handled when solving the energy management problem, since keeping the SoC in an allowable range is a prerequisite for the function of hybrid electric vehicles (HEVs). In general, an energy management strategy is designed based on the following dynamical model of SoC.

$$\dot{\mathrm{SoC}} = \frac{-U_{oc} + \sqrt{U_{oc}^2 - 4R_bP_b}}{2Q_{b\,\max}R_b}, \tag{16}$$

where U_{oc}, R_b and $Q_{b\,\max}$ denote the battery open-circuit voltage, the battery internal resistance and the maximum charge capacity of the battery, respectively, and $P_b = \eta_m^k\tau_m\omega_m + \eta_g^k\tau_g\omega_g$, where η_m^k and η_g^k are the efficiencies of motor and generator, and $k = 1$ denotes the discharging state while $k = -1$ denotes the charging state.

3.3.2 *Problem statement: speed consensus and energy management*

As illustrated in Figure 6, a two-layer control structure is proposed for fleet-level HEV operation [85].

(1) Upper layer (speed consensus). A population of N vehicles is considered, each having its own speed $v_i(t)$. The longitudinal dynamics are assumed to follow a stochastic differential equation (SDE) of the form

$$\begin{cases} \mathrm{d}v_i(t) = a_i(t)\mathrm{d}t + \sigma\mathrm{d}w_i(t), \\ v_i(0) = v_{i0}, \end{cases} \tag{17}$$

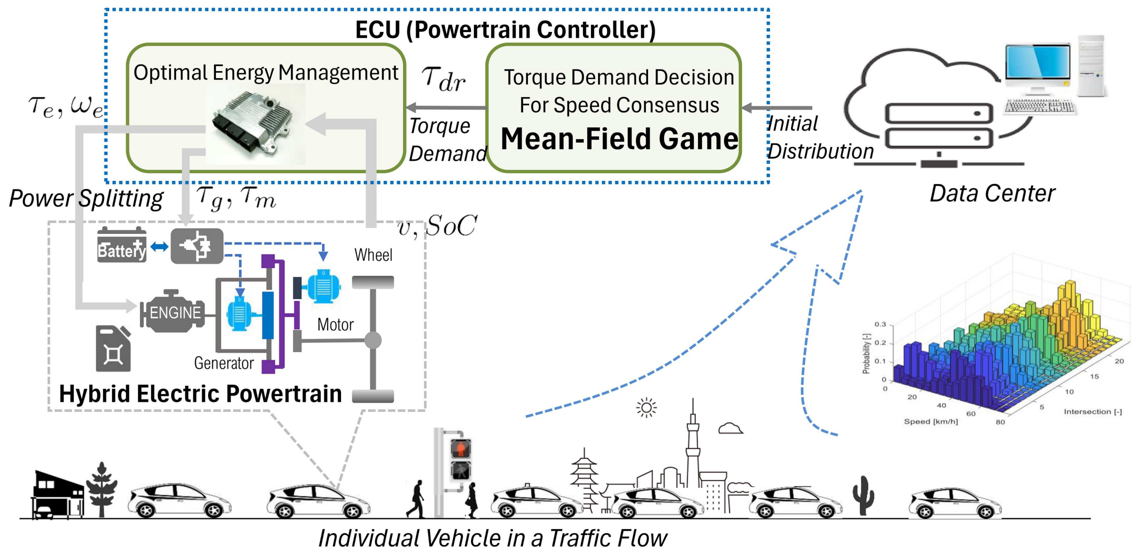


Figure 6 (Color online) Framework structure of two-layer control scheme.

where $a_i(t)$ is the acceleration command (control input) from the i-th vehicle, σ represents modeling uncertainties or road disturbances, and $w_i(t)$ is a standard Brownian motion, aggregating many microscopic uncertainties. Let $\{a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_N\}$ be denoted by a_{-i}, i.e., the controls of all other vehicles except the i-th one.

To achieve speed consensus (while aiming toward a desired reference v_d), the following cost function is assigned to each vehicle:

$$J_i(a_i, a_{-i}) = \mathbb{E}\left[\int_0^T (\underbrace{K_1(v_i(t) - v_{(N)}(t))^2}_{\text{speed consensus}} + \underbrace{K_2 a_i(t)^2}_{\text{smoother driving}})\mathrm{d}t + \underbrace{K_3\,(v_i(T) - v_d)^2}_{\text{speed regulation}}\right], \tag{18}$$

where $v_{(N)} = \frac{1}{N}\sum_{j=1}^{N} v_j$ is the average speed of the fleet, $K_1, K_2, K_3 > 0$ are weighting coefficients, T is the terminal time, and v_d is a desired reference speed.

The admissible decentralized acceleration strategy set of i-th vehicle is denoted as $\mathcal{U}_i$. In MFG terminology, the aim is to find an ε-Nash equilibrium of the above non-cooperative game (17) and (18), where each vehicle's decentralized control input $a_i(t)$ is chosen so as to minimize J_i given the average effect of all other vehicles.

(2) Lower layer (energy management). Once the upper layer determines the target acceleration a_i (hence the demand torque via (12)), the lower layer allocates power between the engine and electric motor(s) in real-time to maximize overall powertrain efficiency. Formally, one can express this optimization problem as

$$\begin{aligned} \min_{\tau_e, \omega_e, \tau_m} \ & J = \int_0^T \dot{m}_f(t)\mathrm{d}t + \beta|\mathrm{SoC}(t_f) - \mathrm{SoC}(0)| \\ \text{s.t.} \quad & (13)\text{–}(16), \\ & \mathrm{SoC}_{\min} \leqslant \mathrm{SoC}(t) \leqslant \mathrm{SoC}_{\max}, \\ & \tau_{i,\min} \leqslant \tau_i(t) \leqslant \tau_{i,\max}, \\ & \omega_{i,\min} \leqslant \omega_i(t) \leqslant \omega_{i,\max}, \quad i = e, m, g. \end{aligned} \tag{19}$$

Here, $\dot{m}_f(t)$ is the instantaneous fuel consumption usually captured as the brake specific fuel consumption [85], and β is a penalty factor enforcing SoC neutrality (i.e., ensuring the battery ends at or near its initial SoC to avoid solutions that exploit battery depletion without regard for later recharging).

3.3.3 *Solution methods*

In the upper-layer MFG setting, the central objective is to compute an ε-Nash equilibrium for a large population of interacting vehicles. Two mainstream solution methodologies, often referred to as the

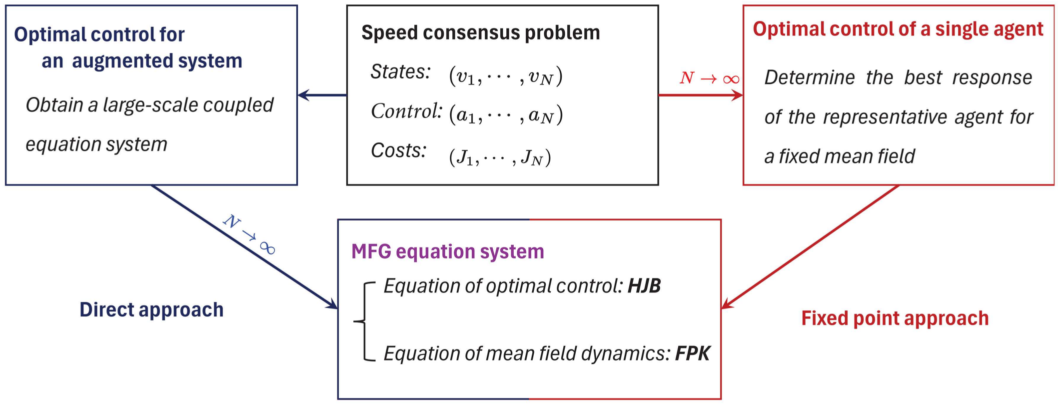


Figure 7 (Color online) Basic idea of MFG methodology.

direct approach and the fixed point approach, can be employed. As shown in Figure 7, the former starts by solving an N-player game directly to obtain a large coupled equation system and then deriving a limit for the solution by taking the number $N \to \infty$. The latter first determines the best response of a representative agent, and then the best responses of all agents regenerate the mean field term. Regardless of which approach is chosen (direct or fixed-point), the resulting equations to be solved often take the form of coupled PDEs (the HJB for optimal control and the KFP for distribution evolution). In linear-quadratic (LQ) settings, such as when cost functionals and dynamics are linear in states/controls and quadratic in costs, these PDEs can be simplified to systems of ordinary differential equations (ODEs). For computing the decentralized control inputs, various advanced numerical methods have been proposed, including: finite difference and finite element schemes, semi-Lagrangian methods, reinforcement learning methods. In the lower-level control problem, focusing on energy management, methods that provide real-time near-optimal solutions include: stochastic dynamic programming [86], equivalent consumption minimization strategy [87] and mixed integer non-linear dynamic programming [88].

3.3.4 *Results and findings*

In [85], it is demonstrated that the application of the mean-field game control law effectively guides large-scale vehicle groups to achieve a target reference speed. The hierarchical optimization strategy coordinates the engine and motor outputs, delivering the required acceleration efficiently. The probability density distributions for crucial vehicle-level and powertrain-level variables, specifically vehicle speed, acceleration, torque distribution rate, engine torque and speed, and motor torque and speed, are illustrated in Figure 8. As shown, the probability density of vehicle speeds significantly concentrates around 120 km/h at 50 s, indicating that vehicles controlled under this scheme reliably reach the maximum allowable speed according to traffic regulations. This results in enhanced road traffic efficiency and improved vehicle throughput capabilities. Additionally, the probability densities for control inputs, namely vehicle acceleration at the higher control layer and torque distribution rate at the lower control layer, become notably narrower over time, reflecting the stable control achieved by the proposed method. These conclusions were verified through numerical simulations with industrial backgrounded physical parameters, confirming the effectiveness and practicality of the proposed two-layer control framework.

3.4 Parking-based EV charging

The integration of EV battery energy into power grids has drawn significant attention for several decades due to the potentially adverse impacts of disordered charging on power systems [89–91]. Large-scale, uncoordinated EV charging can lead to increased peak loads, reduced voltage quality, greater transmission losses, and accelerated transformer aging. On the other hand, the aggregated charging/discharging capability of EVs also offers the opportunity to support grid operations by compensating supply-demand imbalances, thus making EVs a promising distributed energy resource.

However, the collective behavior of thousands or even millions of EVs is inherently uncertain, driven by stochastic mobility demand (e.g., unknown driving times, charging requirements, and battery states). Addressing such large-scale coordination without centralized control is a major challenge. Recent literature focuses on MFG or mean-field control formulations that incorporate aggregated EV behaviors under

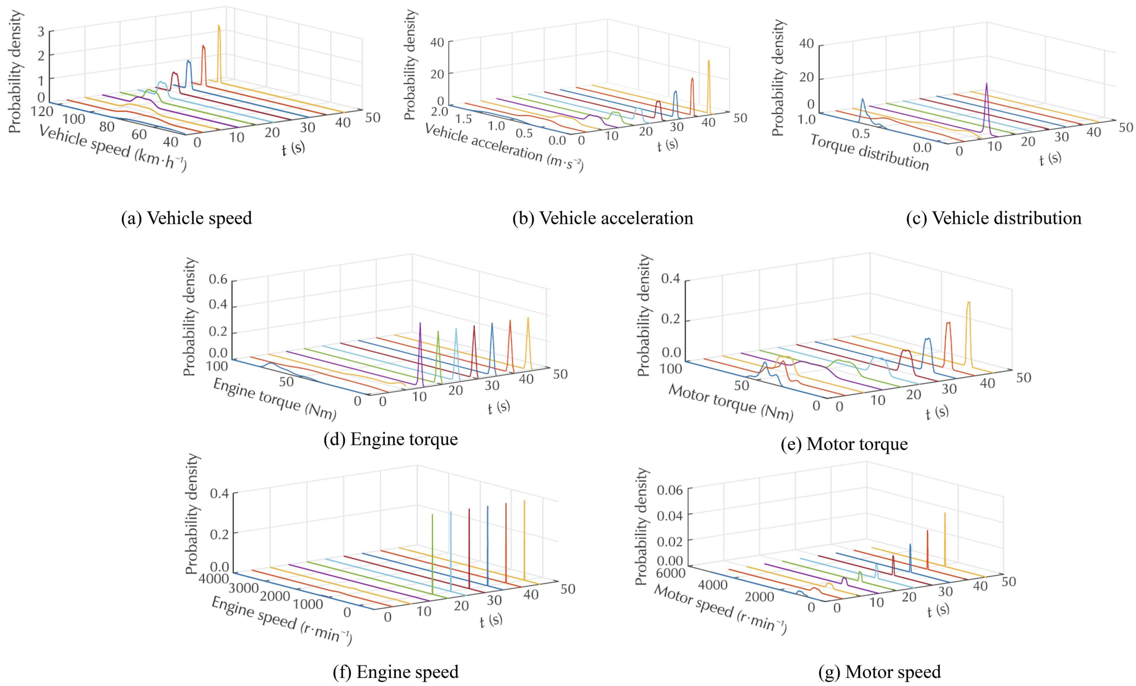


Figure 8 (Color online) Probability density distributions of key vehicle-level and powertrain-level variables under the proposed two-layer framework.

limited or no centralized information [90, 92–95]. In particular, parking-based EV charging schemes have been proposed to leverage the interaction between parking lots, EVs, and the grid to ensure stable and cost-effective energy transactions.

3.4.1 *Preliminaries on parking lot and EV groups integrated into the grid*

A schematic of the proposed parking-based charging architecture is shown in Figure 9. The parking lot is connected bidirectionally to both the power grid and a population of EVs. The parking lot can act as a localized energy hub, purchasing electricity from the grid during off-peak hours at lower prices and potentially selling it back to the grid during peak hours if vehicle-to-grid (V2G) functionality is allowed. In addition, the parking lot may include an energy storage system (ESS) (e.g., stationary battery) that helps smooth out power fluctuations and provides peak-shaving and load-shifting capabilities. On the other hand, EVs arriving at the parking lot can buy electricity (charging) or, when appropriate, sell it back (discharging) to the lot at mutually agreed prices. Energy prices influence both the parking lot's and EVs' decisions, driving the design of optimal policies for charging and discharging.

From a game-theoretic perspective, the parking lot is regarded as a major player (denoted $\mathcal{A}_0$), since its decisions substantially affect the rest of the population. Each EV, denoted $\mathcal{A}_i$, is a minor player whose individual impact is negligible, yet collectively, they contribute significant mean-field coupling terms in the overall system dynamics and costs. Hence, a major-minor stochastic differential game framework can be employed to capture interactions among the parking lot and a large EV population, enabling them to maximize profits or minimize costs in a decentralized setting.

3.4.2 *Problem statement*

Consider a population of $N+1$ players, where the parking lot is the major player ($\mathcal{A}_0$) and N EVs are minor players ($\{\mathcal{A}_i, 1 \leqslant i \leqslant N\}$). The state of charge (SoC) for the parking lot's energy storage device is denoted by $x_0(t)$, and the SoC of the i-th EV is $x_i(t)$. Their dynamics are governed by the following SDEs:

$$\begin{cases} \mathrm{d}x_0 = \left(\dfrac{\alpha_0}{\beta_0}\left(u_0 - N u_{(N)}\right)\right)\mathrm{d}t + \mathrm{d}w_0, \\ x_0(0) = x_{00}, \end{cases} \tag{20}$$

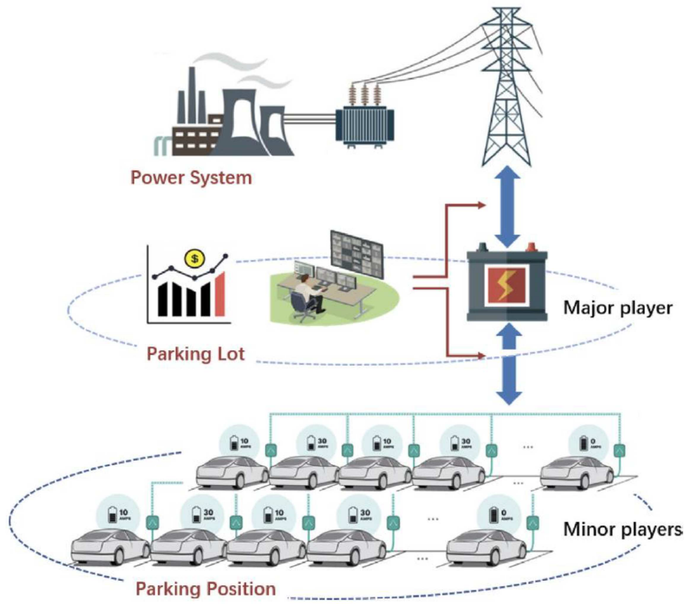


Figure 9 (Color online) Electricity trading model with parking lot and EVs.

$$\begin{cases} \mathrm{d}x_i = \left(\dfrac{\alpha_i}{\beta_i}u_i\right)\mathrm{d}t + \mathrm{d}w_i, \\ x_i(0) = x_{i0}, \quad 1 \leqslant i \leqslant N, \end{cases} \tag{21}$$

where $u_0 \in \mathbb{R}$ represents the power between grid and the parking lot's energy storage (positive u_0 indicates charging from the grid, negative indicates discharging to the grid), and u_i, $1 \leqslant i \leqslant N$, represent power between the parking lot's storage device and the i-th EV. $u_{(N)} = \frac{1}{N}\sum_{j=1}^{N} u_j$ is the average input of minor players. The constant coefficients $\alpha_i \in (0,1]$ and $\beta_i > 0$, $0 \leqslant i \leqslant N$, denote the charging and discharging efficiency and the capacity of battery for the parking lot and EVs. Here, $x_i \in [0,1]$, $0 \leqslant i \leqslant N$, and the noise processes $\{w_i, 0 \leqslant i \leqslant N\}$ are $N+1$ independent standard one-dimensional Brownian motions that reflect the small and unpredictable fluctuations in individual charging rates, as well as the heterogeneous behavior of EVs in the parking lot. The initial states $\{x_{i0}, 0 \leqslant i \leqslant N\}$ are mutually independent and have the same expectation and a finite second moment. These initial states are also independent of $\{w_i, 0 \leqslant i \leqslant N\}$.

Each player's cost function is formulated to reflect the trade-offs between SoC tracking and charging/discharging costs. Let $u_{-i} = \{u_0, \ldots, u_{i-1}, u_{i+1}, \ldots, u_N\}$ be the control inputs of all players except i-th player and $\mathcal{U}_i$ be the admissible decentralized charging/discharging strategy set of i-th player. Then, the parking lot (major player) cost is given by

$$J_0(u_0, \ldots, u_N) = \mathbb{E}\left[\int_0^T \big(\underbrace{q_0(x_0 - \gamma_0 x_{(N)})^2}_{\text{SoC alignment}} + \underbrace{r_0 u_0^2}_{\text{charging/discharging effort}}\big)\mathrm{d}t + \underbrace{q_{0f}(x_0(T) - \eta_{0f})^2}_{\text{SoC tracking}}\right], \tag{22}$$

where $x_{(N)} = \frac{1}{N}\sum_{i=1}^{N} x_i$ is the average SoC of EVs. The coefficients q_0, q_{0f}, r_0 weight the SoC tracking penalty and charging cost of the major player.

The performance of the i-th EV (minor player) is evaluated using the following quadratic cost functional:

$$J_i(u_i, u_{-i}) = \mathbb{E}\left[\int_0^T \big(\underbrace{q_i(x_i + x_{(N)} - \gamma_0 x_0)^2}_{\text{SoC alignment}} + \underbrace{r_i u_i^2}_{\text{charging/discharging effort}}\big)\mathrm{d}t + \underbrace{q_{0f}(x_i(T) - \eta_{if})^2}_{\text{SoC tracking}}\right], \tag{23}$$

where $q_i \geqslant 0, q_{if}, r_i$ are cost coefficients indicating, respectively, how strongly the EV penalizes SoC deviation from desired levels and how aggressively it penalizes charging/discharging efforts.

Definition 6. A set of charging/discharging strategies $u_i^o \in \mathcal{U}_i$, $0 \leqslant i \leqslant N$, for the $N+1$ players is called an ε-Nash equilibrium with respect to the costs J_i, where $\varepsilon \geqslant 0$, if for any i, $0 \leqslant i \leqslant N$, we have

$$J_i(u_i^o, u_{-i}^o) \leqslant J_i(u_i, u_{-i}^o) + \varepsilon, \tag{24}$$

when any alternative $u \in \mathcal{U}_i$ is applied for by player $\mathcal{A}_i$.

The goal, then, is to determine this equilibrium decentralized control policy for all agents in the major-minor LQG game described by (20)–(23).

3.4.3 *Solution methods*

To solve the above parking-based EV charging game, one can use the Nash certainty equivalence (NCE) methodology, a standard approach in LQG mean-field games involving major and minor players. The NCE procedure breaks down the complex, high-dimensional game problem into more tractable sub-problems by imposing consistency conditions among individual control strategies and aggregate effects. The main steps are as follows.

(1) **Limiting two-player model.** Formulate a limit model consisting of just two players: (i) a representative EV (minor) and (ii) the parking lot (major). The representative EV's behavior is assumed to mirror that of the entire EV population in the limit $N \to \infty$.

(2) **Derivation of NCE equation system.** Solve two limit optimal control problems: one for the parking lot, one for the representative EV. From these problems, coupled ordinary differential equations (ODEs) (the NCE system) are obtained, linking each player's optimal strategy to the distribution of states of the other.

(3) **Asymptotic Nash equilibrium.** By solving the NCE ODE system, one obtains feedback control laws $u_0^o(t, x_0)$ for the parking lot and $u_i^o(t, x_i)$ for each EV. These control inputs constitute an asymptotic Nash equilibrium for the finite but large population, where each agent's decision depends only on its own state and a deterministic time-varying parameter (mean field state trajectory). Crucially, this trajectory satisfies the self-consistency condition

In practice, several numerical methods can be employed to solve the resulting ODEs. For LQG settings, closed-form solutions often exist via Riccati equations and linear dynamic systems. When higher nonlinearities are introduced, one may resort to iterative PDE approaches or more advanced solvers (e.g., spectral methods, finite elements, or machine learning-based approximations).

3.4.4 *Results and findings*

In a discrete-time context, Ref. [95] demonstrated that when the number of EVs is sufficiently large and each EV shares similar operational characteristics and objectives, the original game problem can be reformulated as a major-minor MFG. The study derives the fundamental equations describing the mean-field equilibrium and develops a numerical solution for the NCE coupled equation system. This allows for implementing a decentralized charging strategy that achieves global optimality for all participants.

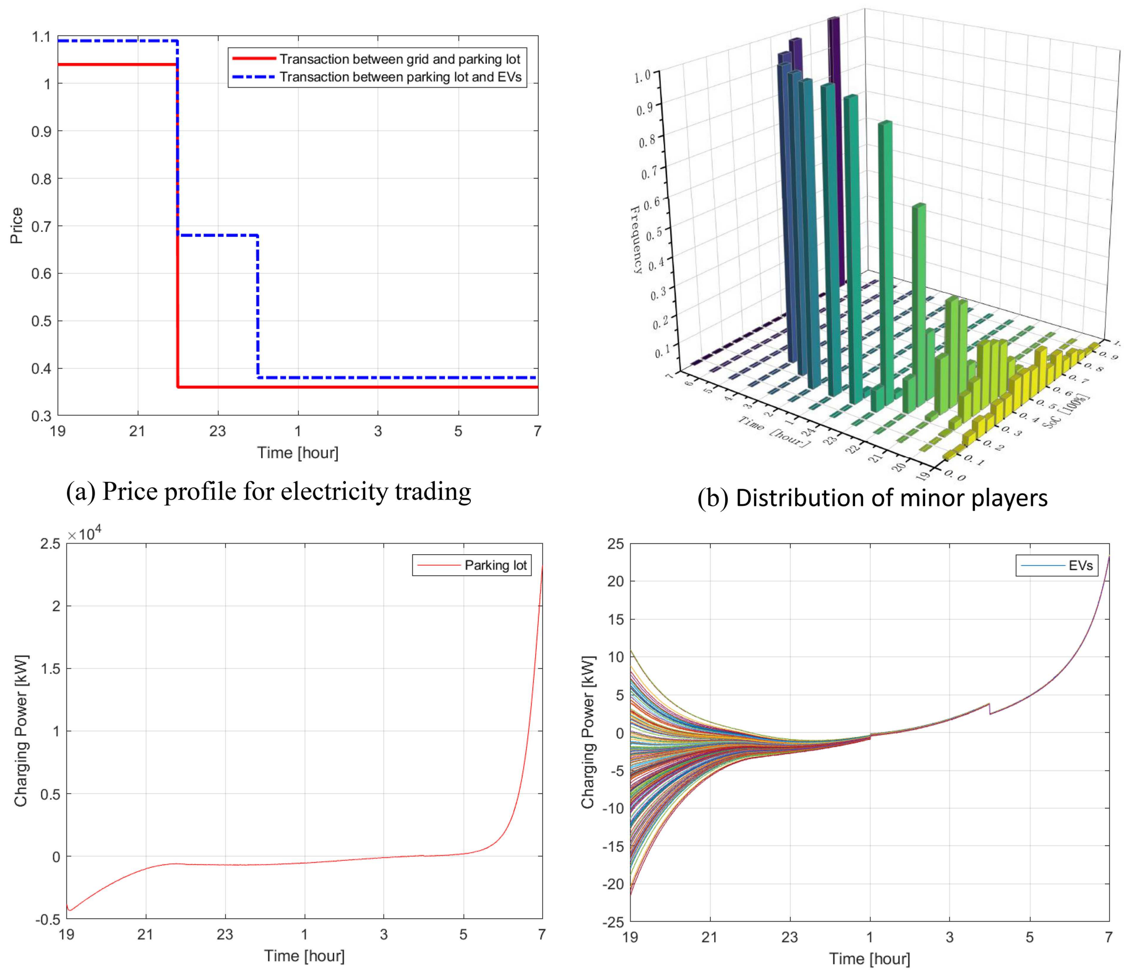
Numerical simulations verify the effectiveness of the proposed strategy by considering a practical scenario involving a parking lot with 200 EVs, operating from 19 : 00 to 7 : 00. This period realistically reflects typical EV owner behavior, where vehicles return in the evening and depart in the morning. Figure 10 specifically presents the electricity price profile, the distribution of EV battery SoC, and the optimized charging dynamics for both the major player (parking lot) and minor players (individual EVs). The results clearly indicate that the proposed strategy effectively maximizes economic benefits for both the parking lot operator and individual EV users, while ensuring all relevant boundary conditions are satisfied.

4 Summary of literature

To iterate how MFG is applied to various urban mobility applications, we summarize the reviewed applications in Table 2.

Table 2 Literature summary on MFG for urban mobility.

Category	References	Agent	State variables	Transition dynamics	Population	Action/ control	Objective function	Algorithm	Findings
Traffic flow	[57, 65]	Cars on ring roads	Location (+ velocity)	Update location based on velocity	Traffic	Velocity control	Kinetic energy + traffic efficiency + safety	Numerical (Newton's method), learning	Classical traffic flow models equivalent to MFG
	[64]	Cars on networks	On edges + at nodes	Move on edge + jump at nodes	Traffic	Velocity control + routing	Kinetic energy + traffic efficiency + safety	Numerical (Newton's method), learning	Augment DTA by speed control
Powertrain control	[84, 85]	Cars on roads	Velocity	Velocity evolves under acceleration	Traffic	Acceleration	Speed consensus + smoother driving + speed regulation	Numerical method	Coordinates engine and motor outputs via an MFG-based fleet-level acceleration strategy
EV management	[90, 95]	Parking lot + EVs	SoC of EVs (+ storage)	SoC evolves under charging/discharging decisions	Vehicle-to-grid network	Charging/ discharging strategies	SoC alignment + charging/discharging effort + SoC tracking	Numerical (backward induction-based, iterative) method	Manages EV charging in an MFG framework for improved efficiency



(a) Price profile for electricity trading

(b) Distribution of minor players

(c) Control law of the major player

(d) Control law of the minor player

Figure 10 (Color online) Charging dynamics of the parking lot (major agent) and EVs (minor players), and the resulting SoC distribution under a major-minor MFG framework with a specified price profile.

5 Conclusion and open questions

We summarize a body of literature using MFG as a new modeling framework for the next-generation transportation control and management systems building from dynamic game theory and machine learning. It will potentially advance our fundamental understanding of the new equilibria of the future transportation ecosystem and the controls needed to guide the ecosystem toward a social optimum. Such a framework would lay a behavioral foundation for the development of a multi-agent simulation platform to inform policy and practice of social good. It will help operators, planners, and regulators better understand the potential consequences of emerging technologies on traffic safety, efficiency, and sustainability, and will, in turn, help prepare them for the transition to smart cities.

5.1 Emerging trends

The game-theoretical framework provides a basis to characterize the dynamic behavior of the mixed traffic system. The outcome of the game examines the optimal policies for the operation, planning, and regulation of infrastructure. Accordingly, MFG provides a rich paradigm for diverse behaviors of mobility entities. Below, we will point out several emerging trends in urban mobility modeling that warrant a paradigm shift in transportation mobility modeling and simulation.

5.1.1 *Lane-free traffic*

MFG has not only been applied to longitudinal control, but lateral control, for example lane-changing maneuver [96], where a voluntary lane-change is initiated according to traffic densities in adjacent lanes. Fundamentally, the control of CAVs is a meta-task maneuver accounting for both longitudinal and lateral movements. The majority of the existing literature assumes that CAVs drive within designated lanes, while some believe that the technology of vehicular connectivity and autonomy could empower lane-free traffic [97, 98], where there does not exist any concept of "lanes", so cars are "not bound to fixed traffic lanes" [97]. MFG holds substantial potential to model lane-free traffic, by not restricting CAVs to move within lanes. Lane-free traffic has been extensively studied using control methods [99–103]. However, these studies primarily assume a limited number of interacting CAVs. Applying MFG to lane-free traffic could open up a new direction that models the collective flocking behavior of cars in the lane-free traffic setting.

5.1.2 *Multiscale coupling: powertrain-vehicle-traffic*

The existing studies have developed MFGs to model how individual vehicles interact on roads, how individual vehicles' powertrain is controlled, and how a network of electrical vehicles is managed across charging stations. These problems are, however, coupled closely. Understanding how powertrain is controlled and when EVs are charged depends on future traffic conditions and routes. In turn, traffic congestion depends on other cars' routing and driving behaviors. Speed and routing control of CAVs would in sequence be affected by traffic conditions ahead in space and advance in time. These decisions are made at different time scales, different frequencies, and on different spatial scales. How to integrate all these decisions in a unified MFG framework would involve multi-scale multifidelity and hierarchical modeling approaches.

5.1.3 *Multimodal transport systems*

In urban mobility systems, normally agents are not homogeneous but interacting on different hierarchical levels. MFG has been applied to various transport modes in an isolated manner, ranging from vehicular traffic [57], pedestrian crowds dynamics [10], to ride-hailing platform [104]. How to leverage MFG for multi-modal transport systems remains a void. MFG could be a valuable tool for such modeling. We conjecture that one option could be to model travelers' mode choice as a game in which each traveler aims to select a travel mode with optimal utility, while the utility of each mode is computed using MFG to account for others' sequential decisions on the operational level. Modeling traveler choice while accounting for microscopic sequential choices has been investigated in the transportation community [105–107]. How to apply MFG and model stage choices, the first stage is mode choice and then the dynamic choice for each travel mode, could be challenging, but a promising direction to explore. For instance, the variants of MFGs, including multipopulation [108, 109], major-minor players [110–114], leader-follower Stackelberge [115–120], and graphon games [121–126] deserve more attention.

5.1.4 *Analytical properties and computational methods*

The mathematical properties of MFE, including existence and uniqueness, for generic cost functions like non-separable cost functional between the action and the population state, remain open questions. These properties are normally discussed for monotonic MFGs, for example, LQ cost functions. Unfortunately, in the majority of the literature on traffic and vehicles, MFGs normally violate monotonicity properties that MFG analysis requires.

MFG systems depend on initial and boundary conditions. Whenever one condition changes, we have to re-compute MFE. Thus, learning methods that could learn the mapping operator instead of a single solution are warranted. Neural operator (NO) is an emerging tool to approximate nonlinear mappings between infinite-dimensional functional spaces using neural networks (NN) [127, 128]. NO could be further encoded with physics knowledge, which is the physics-informed NO (PINO). Learning MFG could potentially be extended to train NOs from the initial and boundary condition to MFE. In this case, we can train an algorithm once, and infer MFEs under different input conditions. A series of studies [68, 69, 129] learn an NO that maps from initial and boundary conditions to an MFE, so that we do not have to retrain a policy network every time a new MFG needs to be solved. To improve the scalability and avoid repeatedly solving MFGs every time their initial state changes, Refs. [69, 129] proposed physics-informed graph neural operators (PIGNO) that utilize a graph neural operator to generate population dynamics, given initial population distributions.

5.2 Open questions

Other than trending areas that call for the game-theoretical framework, there are open questions that are either understudied or have not gained sufficient traction in the MFG community. We would like to bring them to the table and hope that interdisciplinary researchers can join forces to tackle these challenges altogether.

5.2.1 *Stochasticity and uncertainties*

Stochastic effects and uncertainties could potentially arise from various sources, including variability and errors in measurements, dynamic actions of various entities, model biases, and discretization and algorithmic errors. In a transportation system involving many heterogeneous agents, uncertainties are prevalent, both inherently and exogenously. Behavioral stochasticity infiltrates an agent's endogenous decision making process [130], from observations, rewards, transitions, to actions. The presence of multiple agents further renders the traffic environment uncertain and unstable, arising from states and actions of other agents.

Other than behavioral aspects, there are two other types of uncertainty [131], namely, aleatoric uncertainty and epistemic uncertainty. Aleatoric uncertainty (or data uncertainty) is an endogenous property of data and is thus irreducible, coming from measurement noise, incomplete data, mismatch between training and test data. Epistemic uncertainty (or knowledge uncertainty, systematic uncertainty, model discrepancy) is a property of model arising from inadequate knowledge of the modeled system. For example, a single population MFG is applied to model vehicles of various types including passenger cars, motorcycles, and commercial vehicles. They vary in attributes (e.g., mass and length, maximum acceleration and braking rates) and driver behavior (e.g., experience, aggressiveness). The heterogeneity leads to system uncertainty and data noise. Such heterogeneity can lead to insufficiency in establishing a single model that captures diversely manifested behaviors. Other than adopting existing tools like multipopulation MFGs, characterizing uncertainty is critical in MFGs. Uncertainty quantification (UQ) aims to assess robustness of the developed model and bound prediction errors of dynamical systems, by estimating the probability density of quantities with input features and boundary conditions [36, 132]. There is a growing number of literature that applies various UQ techniques, classical or modern, to propagate uncertainties in dynamical systems. In a dynamic game-theoretical setting like MFG, UQ is crucial to understand how perturbations in behavioral parameters, data, and model affect the equilibria outcomes.

5.2.2 *Inverse methods to uncover objective functions*

Other than learning the solutions of PDEs, another direction is to learn a Nash equilibrium based on the observations of agents behaviors that play a Nash equilibrium. If the distribution of individual agents' actions is observed, how do we learn a Nash equilibrium of observed game theory? In other words, how do we validate the game-theoretical framework based on observed data? What is the practicality of applying MFG using observed mobility data? Recent years have seen a growing amount of studies on inverse game theory [133–135], in which the goal is to infer desired behavior from expert demonstrations. However, there still lacks consensus on a formal definition of inverse learning in MFGs, and the connection of inverse MFG to multiagent inverse learning. We hope that researchers in engineering, game theory, robotics, and applied mathematics could collaborate to ground this topic in real-world problems, and leverage cutting-edge tools to solve it. This is not only interesting to theorists, but also important for practitioners, when behaviors of many interacting agents are observable to learn the underlying mechanism of a system, so that we can intervene for socially optimal mechanism design.

5.2.3 *New theory discovery in transportation*

Why is it necessary to apply MFG to model mobility systems? A short answer is, to help invent new theories and to characterize behaviors of new agents. Classical traffic flow theories have been widely applied in the transportation literature, and validated by real-world human driven vehicle datasets. It remains unclear, however, what the macroscopic traffic patterns would be like, when new road users, such as CAVs and EVs, are introduced to public roads. Because the cost function of MFGs can be flexibly manipulated to accommodate a variety of car types and behaviors [65], modeling these new agents using MFGs could be generalizable to these new agents. MFG potentially lays a foundation for systematic assessment of new mobility service. However, traffic and vehicle dynamics have their physical constraints. For example, cars should not move backward. In human-driven traffic, asymmetric hysteresis should still be present [136]. How to incorporate these physical constraints into the game-theoretical setting would be crucial to the adoption of such a methodological framework in the transportation community.

There is a growing body of literature on the interpretation of ML models by MFG, such as generative models [90]. Such a reinterpretation could help open up opportunities to inventing new theories in these areas.

5.2.4 *Convergence of communities*

Researchers of many communities have been actively involved in MFGs, including but not limited to fields including engineering, control, or probability and statistics.

Different communities primarily focus on problems of different forms and goals. For example, researchers in transportation engineering focus on continuous-time non-stationary MFGs, with nonlinear nonseparable cost functions. This is because traffic is highly dynamic and unstable. The goal is to design a time-dependent non-stationary optimal control to adapt to volatile traffic environments. As a sacrifice, since the traffic flow problems are normally not well-regularized, mathematical properties in terms of equilibria existence and uniqueness are challenging to characterize. In contrast, researchers from the control and applied mathematics communities often prioritize analytically tractable scenarios, such as LQG setups. These models ensure desirable mathematical properties, including explicit equilibrium forms, uniqueness, and stability guarantees, however, sometimes at the expense of real-world relevance. Despite these differences, interdisciplinary collaboration is gaining momentum. Transportation researchers are increasingly adopting advanced analytical tools from mathematical and control communities, while scholars in control theory are progressively incorporating more realistic, data-driven scenarios traditionally emphasized more in transportation engineering. To sustain and amplify this convergence, initiatives such as establishing shared datasets and benchmarks, and unified modeling frameworks are essential. Through such coordinated efforts, the research community can develop MFG methodologies that are both mathematically rigorous and practically effective in addressing urban mobility challenges.

Acknowledgements This work of Xuan DI was supported by National Science Foundation (Grant No. CMMI-1943998).

References

1 Caines P E, Huang M, Malhamé R P. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. Commun Inf Syst, 2006, 6: 221–252

2 Lasry J M, Lions P L. Jeux á champ moyen. I—Le cas stationnaire. Comptes Rendus Mathématique, 2006, 343: 619–625

3 Lasry J M, Lions P L. Jeux á champ moyen. II—Horizon fini et contrôle optimal. Comptes Rendus Mathématique, 2006, 343: 679–684

4 Lasry J M, Lions P L. Mean field games. Jpn J Math, 2007, 2: 229–260

5 Guéant O, Lasry J M, Lions P L. Mean field games and applications. In: Proceedings of Paris-Princeton Lectures on Mathematical Finance, 2011. 205–266

6 Lachapelle A, Salomon J, Turinici G. Computation of mean field equilibria in economics. Math Model Methods Appl Sci, 2010, 20: 567–588

7 Djehiche B, Tcheukam A, Tembine H. Mean-field-type games in engineering. ArXiv:1605.03281

8 Couillet R, Perlaza S M, Tembine H, et al. Electrical vehicles in the smart grid: a mean field game analysis. IEEE J Sel Areas Commun, 2012, 30: 1086–1096

9 Degond P, Liu J G, Ringhofer C. Large-scale dynamics of mean-field games driven by local Nash equilibria. J Nonlinear Sci, 2014, 24: 93–115

10 Lachapelle A, Wolfram M T. On a mean field game approach modeling congestion and aversion in pedestrian crowds. Transp Res Part B-Methodological, 2011, 45: 1572–1589

11 Burger M, Di Francesco M, Markowich P, et al. Mean field games with nonlinear mobilities in pedestrian dynamics. ArXiv:1304.5201

12 Benamou J D, Carlier G, Santambrogio F. Variational mean field games. In: Proceedings of Active Particles, 2017. 141–171

13 Achdou Y, Capuzzo-Dolcetta I. Mean field games: numerical methods. SIAM J Numer Anal, 2010, 48: 1136–1162

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

15 Achdou Y, Perez V. Iterative strategies for solving linearized discrete mean field games systems. NHM, 2012, 7: 197–217

16 Ruthotto L, Osher S J, Li W, et al. A machine learning framework for solving high-dimensional mean field game and mean field control problems. Proc Natl Acad Sci USA, 2020, 117: 9183–9193

17 Mguni D, Jennings J, Munoz de Cote E. Decentralised learning in systems with many, many strategic agents. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2018

18 Guo X, Hu A, Xu R, et al. Learning mean-field games. In: Proceedings of Advances in Neural Information Processing Systems, 2019

19 Subramanian J, Mahajan A. Reinforcement learning in stationary mean-field games. In: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems, 2019. 251–259

20 Fouque J P, Zhang Z. Deep learning methods for mean field control problems with delay. Front Appl Math Stat, 2020, 6: 11

21 Carmona R, Lauriére M. Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games I: the ergodic case. SIAM J Numer Anal, 2021, 59: 1455–1485
22 Germain M, Mikael J, Warin X. Numerical resolution of McKean-Vlasov FBSDEs using neural networks. Methodol Comput Appl Probab, 2022, 24: 2557–2586
23 Cui K, Koeppl H. Approximately solving mean field games via entropy-regularized deep reinforcement learning. In: Proceedings of International Conference on Artificial Intelligence and Statistics, 2021. 1909–1917
24 Perrin S, Lauriére M, Pérolat J, et al. Generalization in mean field games by learning master policies. AAAI, 2022, 36: 9413–9421
25 Xie Q, Yang Z, Wang Z, et al. Learning while playing in mean-field games: convergence and optimality. In: Proceedings of International Conference on Machine Learning, 2021. 11436–11447
26 Lauriere M, Perrin S, Girgin S, et al. Scalable deep reinforcement learning algorithms for mean field games. In: Proceedings of the 39th International Conference on Machine Learning, 2022. 12078–12095
27 Angiuli A, Fouque J P, Lauriére M. Unified reinforcement Q-learning for mean field game and control problems. Math Control Signals Syst, 2022, 34: 217–271
28 Xu Z, Shen T, Huang M. Model-free policy iteration approach to NCE-based strategy design for linear quadratic Gaussian games. Automatica, 2023, 155: 111162
29 Xu Z, Shen T. Decentralized ϵ-Nash strategy for linear quadratic mean field games using a successive approximation approach. Asian J Control, 2024, 26: 565–574
30 Xu Z, Wang B C, Shen T. Mean field LQG social optimization: a reinforcement learning approach. Automatica, 2025, 172: 111924
31 Xu Z, Chen J, Wang B C, et al. Robust mean field social control: a unified reinforcement learning framework. ArXiv:2502.20029
32 Xu Z, Chen J, Wang B C, et al. Data-driven mean field equilibrium computation in large-population LQG games. ArXiv:2502.19993
33 Laurière M, Perrin S, Geist M, et al. Learning mean field games: a survey. ArXiv:2205.12944
34 Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys, 2019, 378: 686–707
35 Karniadakis G E, Kevrekidis I G, Lu L, et al. Physics-informed machine learning. Nat Rev Phys, 2021, 3: 422–440
36 Di X, Shi R, Mo Z, et al. Physics-informed deep learning for traffic state estimation: a survey and the outlook. Algorithms, 2023, 16: 305
37 Wu Z, Laurière M, Chua S J C, et al. Population-aware online mirror descent for mean-field games by deep reinforcement learning. ArXiv:2403.03552
38 Zhang C, Chen X, Di X. Stochastic semi-gradient descent for learning mean field games with population-aware function approximation. ArXiv:2408.08192
39 Cardaliaguet P, Hadikhanloo S. Learning in mean field games: the fictitious play. ESAIM-COCV, 2017, 23: 569–591
40 Perrin S, Perolat J, Laurière M, et al. Fictitious play for mean field games: continuous time analysis and applications. In: Proceedings of the 34th International Conference on Neural Information Processing Systems, 2020
41 Perrin S, Laurière M, Pérolat J, et al. Mean field games flock! The reinforcement learning way. ArXiv:2105.07933
42 Elie R, Perolat J, Laurière M, et al. On the convergence of model free learning in mean field games. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2020. 7143–7150
43 Perrin S, Pérolat J, Laurière M, et al. Fictitious play for mean field games: continuous time analysis and applications. In: Proceedings of Advances in Neural Information Processing Systems, 2020. 13199–13213
44 Anahtarci B, Kariksiz C D, Saldi N. Q-learning in regularized mean-field games. Dyn Games Appl, 2023, 13: 89–117
45 Guo X, Xu R Y, Zariphopoulou T. Entropy regularization for mean field games with learning. In: Proceedings of Mathematics of Operations Research, 2022
46 Hadikhanloo S. Learning in anonymous nonatomic games with applications to first-order mean field games. ArXiv:1704.00378
47 Pérolat J, Perrin S, Elie R, et al. Scaling mean field games by online mirror descent. In: Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems, 2022. 1028–1037
48 Perolat J, Perrin S, Elie R, et al. Scaling up mean field games with online mirror descent. ArXiv:2103.00623
49 Yardim B, Cayci S, Geist M, et al. Policy mirror ascent for efficient and independent learning in mean field games. In: Proceedings of International Conference on Machine Learning, 2023. 39722–39754
50 Angiuli A, Fouque J P, Laurière M, et al. Convergence of multi-scale reinforcement q-learning algorithms for mean field game and control problems. ArXiv:2312.06659
51 Zaman M A U, Koppel A, Bhatt S, et al. Oracle-free reinforcement learning in mean-field games along a single sample path. In: Proceedings of International Conference on Artificial Intelligence and Statistics, 2023. 10178–10206
52 Mao W, Qiu H, Wang C, et al. A mean-field game approach to cloud resource management with function approximation. In: Proceedings of Advances in Neural Information Processing Systems, 2022. 36243–36258
53 Di X, Shi R. A survey on autonomous vehicle control in the era of mixed-autonomy: from physics-based to AI-guided driving policy learning. Transp Res Part C-Emerg Technol, 2021, 125: 103008
54 Kachroo P, Agarwal S, Sastry S. Inverse problem for non-viscous mean field control: example from traffic. IEEE Trans Automat Contr, 2016, 61: 3412–3421
55 Chevalier G, Le Ny J, Malhamé R. A micro-macro traffic model based on mean-field games. In: Proceedings of American Control Conference (ACC), 2015. 1983–1988
56 Kachroo P, Agarwal S, Piccoli B, et al. Multiscale modeling and control architecture for V2X enabled traffic streams. IEEE Trans Veh Technol, 2017, 66: 4616–4626
57 Huang K, Di X, Du Q, et al. A game-theoretic framework for autonomous vehicles velocity control: bridging microscopic differential games and macroscopic mean field games. Discrete Cont Dyn Syst-B, 2020, 25: 4869–4903
58 Huang K, Di X, Du Q, et al. Stabilizing traffic via autonomous vehicles: a continuum mean field game approach. In: Proceedings of IEEE Intelligent Transportation Systems Conference (ITSC), 2019. 3269–3274
59 Huang K, Di X, Du Q, et al. Scalable traffic stability analysis in mixed-autonomy using continuum models. Transp Res Part C-Emerging Technol, 2020, 111: 616–630
60 Lighthill M J, Whitham G B. On kinematic waves II. A theory of traffic flow on long crowded roads. Proc Royal Soc London Ser A Math Phys Sci, 1955, 229: 317–345
61 Lebacque J P, Mammar S, Salem H H. Generic second order traffic flow modelling. In: Proceedings of Transportation and Traffic Theory, 2007
62 Aw A, Klar A, Rascle M, et al. Derivation of continuum traffic flow models from microscopic follow-the-leader models. SIAM J Appl Math, 2002, 63: 259–278
63 Zhang H M. A non-equilibrium traffic model devoid of gas-like behavior. Transp Res Part B-Methodol, 2002, 36: 275–290
64 Huang K, Chen X, Di X, et al. Dynamic driving and routing games for autonomous vehicles on networks: a mean field game approach. Transp Res Part C-Emerg Technol, 2021, 128: 103189
65 Mo Z, Chen X, Di X, et al. A game-theoretic framework for generic second-order traffic flow models using mean field games

and adversarial inverse reinforcement learning. Transp Sci, 2024, 58: 1403–1426
66 Chen X, Liu S, Di X. A hybrid framework of reinforcement learning and physics-informed deep learning for spatiotemporal mean field games. In: Proceedings of the International Conference on Autonomous Agents and Multiagent Systems, 2023. 1079–1087
67 Chen X, Liu S, Di X. Learning dual mean field games on graphs. In: Proceedings of ECAI, 2023. 421–428
68 Chen X, Yongjie F, Liu S, et al. Physics-informed neural operator for coupled forward-backward partial differential equations. In: Proceedings of the 1st Workshop on the Synergy of Scientific and Machine Learning Modeling, 2023
69 Liu S, Chen X, Di X. Scalable learning for spatiotemporal mean field games using physics-informed neural operator. Mathematics, 2024, 12: 803
70 Gangbo W, Mészáros A R, Mou C, et al. Mean field games master equations with nonseparable Hamiltonians and displacement monotonicity. Ann Probab, 2022, 50: 2178–2217
71 Mou C, Zhang J. Mean field game master equations with anti-monotonicity conditions. J Eur Math Soc, 2025, 27: 4469–4499
72 Friesz T L, Bernstein D, Smith T E, et al. A variational inequality formulation of the dynamic network user equilibrium problem. Oper Res, 1993, 41: 179–191
73 Peeta S, Ziliaskopoulos A K. Foundations of dynamic traffic assignment: the past, the present and the future. Netws Spatial Economics, 2001, 1: 233–265
74 Ban X J, Pang J S, Liu H X, et al. Continuous-time point-queue models in dynamic network loading. Transp Res Part B-Methodol, 2012, 46: 360–380
75 Ziliaskopoulos A K. A linear programming model for the single destination system optimum dynamic traffic assignment problem. Transp Sci, 2000, 34: 37–49
76 Friesz T L, Han K. The mathematical foundations of dynamic user equilibrium. Transp Res Part B-Methodol, 2019, 126: 309–328
77 Friesz T L, Han K, Neto P A, et al. Dynamic user equilibrium based on a hydrodynamic model. Transp Res Part B-Methodol, 2013, 47: 102–126
78 Bauso D, Zhang X, Papachristodoulou A. Density flow in dynamical networks via mean-field games. IEEE Trans Automat Contr, 2016, 62: 1342–1355
79 Zhang C, Chen X, Di X. A single online agent can efficiently learn mean field games. ArXiv:2405.03718
80 Chen X, Liu S, Di X. Physics-informed graph neural operator for mean field games on graph: a scalable learning approach. Games, 2024, 15: 12
81 Abergel T, Brown A, Cazzola P, et al. Energy Technology Perspectives 2017: Catalysing Energy Technology Transformations. Paris: OECD, 2017
82 Zhang J, Shen T. Real-time fuel economy optimization with nonlinear MPC for PHEVs. IEEE Trans Contr Syst Technol, 2016, 24: 2167–2175
83 Chen J, Kuboyama T, Shen T. Collective behavior information-based design approach to energy management strategy for large-scale population of HEVs. Appl Energy, 2025, 377: 124530
84 Chen J, Xu Z, Shen T. Traffic prediction-based long-term energy management approach incorporating engine transient control for HEVs. Energy, 2025, 320: 135249
85 Fu Q, Xu F, Shen T, et al. Distributed optimal energy consumption control of HEVs under MFG-based speed consensus. Control Theor Technol, 2020, 18: 193–203
86 Opila D F, Wang X, McGee R, et al. Real-time implementation and hardware testing of a hybrid vehicle energy management controller based on stochastic dynamic programming. J Dynamic Syst Measurement Control, 2013, 135: 021002
87 Paganelli G, Delprat S, Guerra T M, et al. Equivalent consumption minimization strategy for parallel hybrid powertrains. In: Proceedings of the 55th Vehicular Technology Conference, 2002. 2076–2081
88 East S, Cannon M. Fast optimal energy management with engine on/off decisions for plug-in hybrid electric vehicles. IEEE Control Syst Lett, 2019, 3: 1074–1079
89 Zhang M, Tian L, Yang S, et al. Influence of electric vehicle charging load distribution on power grid. Power Syst Protect Control, 2014, 42: 86–92
90 Zhang B J, Katsoulakis M A. A mean-field games laboratory for generative modeling. ArXiv:2304.13534
91 Fu Q, Xu Z, Takai K, et al. MFG-based decentralized charging control design of large-scale PEVs with consideration of collective consensus. IEICE Trans Fundament, 2022, E105.A: 1038–1048
92 Tchuendom R F, Malhamé R, Caines P. A quantilized mean field game approach to energy pricing with application to fleets of plug-in electric vehicles. In: Proceedings of the 58th Conference on Decision and Control, 2019. 299–304
93 Tajeddini M A, Kebriaei H. A mean-field game method for decentralized charging coordination of a large population of plug-in electric vehicles. IEEE Syst J, 2018, 13: 854–863
94 Fu Q, Xu Z, Takai K, et al. Parking-based EV collective charging control strategy design: a mean-field game approach. In: Proceedings of Applied Energy Symposium 2021: Low Carbon Cities and Urban Energy Systems, 2021
95 Lin R, Xu Z, Huang X, et al. Optimal scheduling management of the parking lot and decentralized charging of electric vehicles based on Mean Field Game. Appl Energy, 2022, 328: 120198
96 Festa A, Göttlich S. A mean field games approach for multi-lane traffic management. ArXiv:1711.04116
97 Papageorgiou M, Mountakis K S, Karafyllis I, et al. Lane-free artificial-fluid concept for vehicular traffic. Proc IEEE, 2021, 109: 114–121
98 Sekeran M, Rostami-Shahrbabaki M, Syed A A, et al. Lane-free traffic: history and state of the art. In: Proceedings of the 25th International Conference on Intelligent Transportation Systems (ITSC), 2022. 1037–1042
99 Malekzadeh M, Papamichail I, Papageorgiou M, et al. Optimal internal boundary control of lane-free automated vehicle traffic. Transp Res Part C-Emerg Technol, 2021, 126: 103060
100 Troullinos D, Chalkiadakis G, Papamichail I, et al. Collaborative multiagent decision making for lane-free autonomous driving. In: Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems, 2021. 1335–1343
101 Levy R, Haddad J. Path and trajectory planning for autonomous vehicles on roads without lanes. In: Proceedings of IEEE International Intelligent Transportation Systems Conference (ITSC), 2021. 3871–3876
102 Rostami-Shahrbabaki M, Weikl S, Niels T, et al. Modeling vehicle flocking in lane-free automated traffic. Transp Res Record-J Transp Res Board, 2023, 2677: 499–512
103 Malekzadeh M, Troullinos D, Papamichail I, et al. Internal boundary control in lane-free automated vehicle traffic: comparison of approaches via microscopic simulation. Transp Res Part C-Emerg Technol, 2024, 158: 104456
104 Li Y, Dimakis A, Courcoubetis C A. Repositioning, ride-matching, and abandonment in on-demand ride-hailing platforms: a mean field game approach. ArXiv:2504.02346
105 (Jeff) Ban X, Dessouky M, Pang J S, et al. A general equilibrium model for transportation systems with e-hailing services and flow congestion. Transp Res Part B-Methodol, 2019, 129: 273–304
106 Chen X, Di X. A unified network equilibrium for e-hailing platform operation and customer mode choice. ArXiv:2203.04865
107 Chen X, Di X. A network equilibrium model for integrated shared mobility services with ride-pooling. Transp Res Part C-Emerg Technol, 2024, 167: 104837

108 Cirant M. Multi-population mean field games systems with Neumann boundary conditions. J de Mathématiques Pures Appliquées, 2015, 103: 1294–1315
109 Ren L, Jin Y, Niu Z, et al. Hierarchical cooperation in LQ multi-population mean field game with its application to opinion evolution. IEEE Trans Netw Sci Eng, 2024, 11: 5008–5022
110 Ma Y, Huang M. Linear quadratic mean field games with a major player: the multi-scale approach. Automatica, 2020, 113: 108774
111 Firoozi D, Caines P E. ϵ-Nash equilibria for major-minor LQG mean field games with partial observations of all agents. IEEE Trans Automat Contr, 2020, 66: 2778–2786
112 Firoozi D, Jaimungal S, Caines P E. Convex analysis for LQG systems with applications to major-minor LQG mean-field game systems. Syst Control Lett, 2020, 142: 104734
113 Şen N, Caines P E. Mean field game theory with a partially observed major agent. SIAM J Control Optim, 2016, 54: 3174–3224
114 Fu G, Graewe P, Horst U, et al. A mean field game of optimal portfolio liquidation. Math OR, 2021, 46: 1250–1281
115 Nourian M, Caines P E, Malhame R P, et al. Mean field LQG control in leader-follower stochastic multi-agent systems: likelihood ratio based adaptation. IEEE Trans Automat Contr, 2012, 57: 2801–2816
116 Bensoussan A, Chau M H M, Lai Y, et al. Linear-quadratic mean field Stackelberg games with state and control delays. SIAM J Control Optim, 2017, 55: 2748–2781
117 Wang G, Wang Y, Zhang S. An asymmetric information mean-field type linear-quadratic stochastic Stackelberg differential game with one leader and two followers. Optim Control Appl Methods, 2020, 41: 1034–1051
118 Shokri M, Kebriaei H. Leader-follower network aggregative game with stochastic agents' communication and activeness. IEEE Trans Automat Contr, 2020, 65: 5496–5502
119 Wang G, Zhang S. A mean-field linear-quadratic stochastic Stackelberg differential game with one leader and two followers. J Syst Sci Complex, 2020, 33: 1383–1401
120 Lin Y, Jiang X, Zhang W. An open-loop Stackelberg strategy for the linear quadratic mean-field stochastic differential game. IEEE Trans Automat Contr, 2018, 64: 97–110
121 Caines P E, Huang M. Graphon mean field games and their equations. SIAM J Control Optim, 2021, 59: 4373–4399
122 Gao S, Tchuendom R F, Caines P E. Linear quadratic graphon field games. Commun Inf Syst, 2021, 21: 341–369
123 Aurell A, Carmona R, Lauriére M. Stochastic graphon games: II. The linear-quadratic case. Appl Math Optim, 2022, 85: 39
124 Cui K, Koeppl H. Learning graphon mean field games and approximate Nash equilibria. In: Proceedings of International Conference on Learning Representations, 2022
125 Tangpi L, Zhou X. Optimal investment in a large population of competitive and heterogeneous agents. Finance Stoch, 2024, 28: 497–551
126 Zhou F, Zhang C, Chen X, et al. Graphon mean field games with a representative player: analysis and learning algorithm. In: Proceedings of the 41st International Conference on Machine Learning, 2024
127 Lu L, Jin P, Pang G, et al. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. Nat Mach Intell, 2021, 3: 218–229
128 Li Z, Kovachki N, Azizzadenesheli K, et al. Fourier neural operator for parametric partial differential equations. In: Proceedings of the International Conference on Learning Representations, 2021
129 Chen X, Liu S, Di X. Bridging agent dynamics and population behaviors: scalable learning for mean field games on graph via neural operators. In: Proceedings of AAAI Conference on Artificial Intelligence, 2024
130 Sastry P S, Phansalkar V V, Thathachar M A L. Decentralized learning of Nash equilibria in multi-person stochastic games with incomplete information. IEEE Trans Syst Man Cybern, 1994, 24: 769–777
131 Abdar M, Pourpanah F, Hussain S, et al. A review of uncertainty quantification in deep learning: Techniques, applications and challenges. Inf Fusion, 2021, 76: 243–297
132 Smith R C. Uncertainty Quantification: Theory, Implementation, and Applications. Philadelphia: SIAM, 2024
133 Fu J, Tacchetti A, Perolat J, et al. Evaluating strategic structures in multi-agent inverse reinforcement learning. J Artif Intell Res, 2021, 71: 925–951
134 Ramponi G, Kolev P, Pietquin O, et al. On imitation in mean-field games. In: Proceedings of Advances in Neural Information Processing Systems, 2023. 40426–40437
135 Tang J, Swamy G, Fang F, et al. Multi-agent imitation learning: value is easy, regret is hard. In: Proceedings of Advances in Neural Information Processing Systems, 2024. 27790–27816
136 Zhang H M. A mathematical theory of traffic hysteresis. Transp Res Part B-Methodol, 1999, 33: 1–23