

# Stabilizing Traffic via Autonomous Vehicles: A Continuum Mean Field Game Approach

Kuang Huang<sup>1</sup>, Xuan Di<sup>2,3</sup>, Qiang Du<sup>1,3</sup> and Xi Chen<sup>4</sup>

**Abstract**—This paper presents scalable traffic stability analysis for both pure autonomous vehicle (AV) traffic and mixed traffic based on continuum traffic flow models. Human vehicles are modeled by a non-equilibrium traffic flow model, i.e., Aw-Rascle-Zhang (ARZ), which is unstable. AVs are modeled by the mean field game which assumes AVs are rational agents with anticipation capacities. It is shown from linear stability analysis and numerical experiments that AVs help stabilize the traffic. Further, we quantify the impact of AV's penetration rate and controller design on the traffic stability. The results may provide insights for AV manufacturers and city planners.

## I. Introduction

Autonomous vehicles (AVs) are believed to be the foundation of the next-decade transportation system and are expected to improve the traffic flow that is presently dominated by human vehicles (HVs). Modeling AV's driving behavior and quantifying different penetration rates of AVs' impact on the traffic is of great significance.

This paper focuses on traffic stability, which is one of the most substantial traffic features. Traffic stability refers to a traffic system's asymptotic stability around uniform flows. HV traffic is observed to be an unstable system in which a small perturbation (caused by driving errors or delays) to the uniform flow will grow up with time and develop traffic congestion. By removing human errors, it is expected that AVs will help stabilize the traffic system. A field experiment [1] showed that one AV is able to stabilize the traffic with approximately twenty vehicles on a ring road.

AV's capability of stabilizing traffic is validated using microscopic models for mixed AV-HV traffic. In microscopic models, the traffic system is described by ordinary differential equations. One can carry out standard linear stability analysis to characterize such a traffic system's stability, built upon connected cruise controllers [2] or generic car-following models [3], [4]. [2], [3] considered only one AV with multiple HVs; [4] studied multiple AVs and multiple HVs but only focused on the head-to-tail stability. However, in the general case, the mixed traffic stability analysis relies on the topology of mixed vehicles and the vehicle-to-vehicle communication network, which suffers from scalability issues.

One alternative approach to address the scalability issues is the PDE approximation [5]. This approach suggests to study the stability of continuum traffic flow models which are the limits of microscopic models. The approach is well suited for the mixed traffic since one needs only to concern about the density distributions of different classes.

In continuum traffic flow models, the traffic system is described by partial differential equations (PDEs) of traffic density and velocity. For single class traffic, the Lighthill-Whitham-Richards (LWR) model [6] is the most extensively used continuum model. As a generalization, the multiclass LWR is widely used to model the interaction between two types of vehicles. [7] is among a few studies that applied multi-class LWR models to AV-HV mixed traffic and proposed networked traffic controls in the presence of AVs. Based on gas-kinetic theory, [8] proposed a multiclass macroscopic model to capture the effect of communication and information sharing on traffic flow and analyzed the model's stability with respect to the connected vehicle's penetration rate; [9], [10], [11] modeled the macroscopic traffic flow of mixed Adaptive Cruise Control (ACC) and Cooperative Adaptive Cruise Control (CACC) vehicles and analyzed how the ACC vehicle's penetration rate influences the traffic stability.

This paper models AVs using the mean field game following the authors' work [12]. In this framework, AVs are assumed to be rational, utility-optimizing agents with anticipation capabilities and play a non-cooperative game by selecting their driving speeds. AVs' utility-optimizing and anticipation behaviors are distinctive characteristics from the aforementioned continuum models. By extending [12], this paper aims to build continuum traffic flow models for both pure AV traffic and mixed AV-HV traffic based on mean field games and analyze the models' traffic stability.

The remainder of the paper is organized as follows. Section II provides an overview of the mean field game and the Aw-Rascle-Zhang model, used for modeling AVs and HVs, respectively. Section III formulates models for both pure AV traffic and mixed AV-HV traffic. Based on the proposed models, Section IV shows the linear stability analysis for the pure AV traffic and Section V demonstrates the mixed traffic's stability through numerical experiments.

## II. Preliminaries

### A. Mean Field Game

Mean field game (MFG) is a game-theoretic framework to model complex multi-agent dynamic systems [13]. In the MFG framework, a population of  $N$  rational

<sup>1</sup>Department of Applied Physics and Applied Mathematics, Columbia University

<sup>2</sup>Department of Civil Engineering and Engineering Mechanics, Columbia University

<sup>3</sup>Data Science Institute, Columbia University

<sup>4</sup>Department of Computer Science, Columbia University

utility-optimizing agents are modeled by a dynamic system. The agents interact with each other through their utilities. Assuming those agents optimize their utilities in a non-cooperative way, they form a differential game.

Exact Nash equilibria to the differential game are generally hard to solve when  $N$  is large. Alternatively, MFG considers the continuum problem as  $N \rightarrow \infty$ . By exploiting the “smoothing” effect of a large number of interacting individuals, MFG assumes that each agent only responds to and contributes to the density distribution of the whole population. Then the equilibria are characterized by a set of two PDEs: a backward Hamilton-Jacobi-Bellman (HJB) equation describing a generic agent’s optimal control provided the density distribution and a forward Fokker-Planck equation describing the population’s density evolution provided individual controls.

In this paper we shall formulate AV traffic as a mean field game. AVs are modeled as rational agents with predefined driving costs. Their density distribution is exactly the traffic density and the Fokker-Planck equation is the same as the continuity equation (CE) that is widely used in continuum traffic flow models.

## B. Aw-Rascle-Zhang Model

The following Aw-Rascle-Zhang (ARZ) model:

$$(CE) \quad \rho_t + (\rho u)_x = 0, \quad (1)$$

$$(ME) \quad [u + h(\rho)]_t + u[u + h(\rho)]_x = \frac{1}{\tau}[U(\rho) - u], \quad (2)$$

is a non-equilibrium continuum traffic flow model describing human driving behaviors [14], [15], where,  $\rho(x,t)$ ,  $u(x,t)$ : the traffic density and speed;  $U(\cdot)$ : the desired speed function;  $h(\cdot)$ : the hesitation function that is an increasing function of the density;  $\tau$ : the relaxation time quantifying how fast drivers adapt their current speeds to desired speeds.

Equation (1) is the continuity equation describing the flow conservation and (2) is a momentum equation (ME) prescribing human driver’s dynamic behavior.

The ARZ model is able to predict important human driving features such as stop-and-go waves and traffic instability [16]. Traffic stability is defined around uniform flows. In continuum models, uniform flows are described by constant solutions  $\rho(x,t) \equiv \bar{\rho}$ ,  $u(x,t) \equiv \bar{u}$ . The constant solutions of the ARZ model is given by  $\bar{u} = U(\bar{\rho})$ . Then the traffic stability for the ARZ model is defined as follows:

**Definition 2.1:** The ARZ model (1)(2) is stable around the uniform flow  $(\bar{\rho}, \bar{u})$  where  $\bar{u} = U(\bar{\rho})$  if for any  $\varepsilon > 0$ , there exists  $\delta > 0$  such that for any solution  $\rho(x,t), u(x,t)$  to the system:

$$\sup_{0 \leq t < \infty} \{\|\rho(\cdot, t) - \bar{\rho}\| + \|u(\cdot, t) - \bar{u}\|\} \leq \varepsilon, \quad (3)$$

whenever  $\|\rho(\cdot, 0) - \bar{\rho}\| + \|u(\cdot, 0) - \bar{u}\| \leq \delta$ . Here  $\|\cdot\|$  is a given norm. The system is linearly stable if its linearized system at  $(\bar{\rho}, \bar{u})$  is stable around the zero solution.

The ARZ model has a simple linear stability criterion [16]:

**Theorem 2.1:** The ARZ model (1)(2) is linearly stable around the uniform flow  $(\bar{\rho}, \bar{u})$  where  $\bar{u} = U(\bar{\rho})$  if and only if  $h'(\bar{\rho}) > -U'(\bar{\rho})$ .

Because of its capability of producing traffic instability, we shall use the ARZ model (1)(2) to characterize HV’s driving behavior.

## III. Model Formulation

### A. Pure AV Traffic: Mean Field Game

In this section, we will build a pure AV continuum traffic flow model based on a mean field game following [12].

Assume that a large population of homogeneous AVs are driving on a closed highway without any entrance nor exit. Those AVs anticipate others’ behaviors and the evolution of the traffic density  $\rho(x, t)$  on a predefined time horizon  $[0, T]$ . AVs control their speeds and aim to minimize their driving costs on the horizon  $[0, T]$ . Then the AVs’ optimal cost  $V(x, t)$  and optimal velocity field  $u(x, t)$  can be described by a set of HJB equations [12]:

$$(HJB) \quad V_t + uV_x + f(u, \rho) = 0, \quad (4)$$

$$u = \operatorname{argmin}_{\alpha} \{\alpha V_x + f(\alpha, \rho)\}, \quad (5)$$

where  $f(\cdot, \cdot)$  is the cost function [12].

When all AVs follow their optimal velocity controls, the system’s density evolution is described by the continuity equation:

$$(CE) \quad \rho_t + (\rho u)_x = 0. \quad (6)$$

The mean field game is described by the coupled system (4)(5)(6).

- The initial condition for the forward continuity equation (6) is given by the initial density  $\rho(x, 0) = \rho_0(x)$ .
- The terminal condition for the backward HJB equations (4)(5) is given by the terminal cost  $V(x, T) = V_T(x)$ . We will always set  $V_T(x) = 0$  meaning that the cars have no preference on their destinations.
- The choice of the spatial boundary condition depends on the traffic scenario. In this paper we assume that the highway is a ring road of fixed length  $L$  and specify the periodic boundary condition  $\rho(0, t) = \rho(L, t)$ ,  $V(0, t) = V(L, t)$ .

The cost function represents certain driving objectives. The choice of the cost function determines AV’s driving behavior. In this paper we shall follow [12] and take the following cost function:

$$f(u, \rho) = \frac{1}{2} \left( \underbrace{\frac{u}{u_{\max}}}_{\text{kinetic energy}} \right)^2 - \underbrace{\frac{u}{u_{\max}}}_{\text{efficiency}} + \underbrace{\frac{u\rho}{u_{\max}\rho_{\text{jam}}}}_{\text{safety}}, \quad (7)$$

where,

$u_{\max}$  and  $\rho_{\text{jam}}$  are the free flow speed and the jam density;  $\frac{1}{2}(u/u_{\max})^2$  models the car’s kinetic energy;  $-u/u_{\max}$  models the car’s efficiency, minimizing this term means that the car should drive as fast as possible;  $u\rho/u_{\max}\rho_{\text{jam}}$  models the safety, it is a penalty term that restricts the car’s speed in traffic congestion;

The MFG system corresponding to the cost function (7) is [12]:

$$\left\{ \begin{array}{l} \rho_t + (\rho u)_x = 0, \\ V_t + u V_x + \frac{1}{2} \left( \frac{u}{u_{\max}} \right)^2 - \frac{u}{u_{\max}} + \frac{u \rho}{u_{\max} \rho_{\text{jam}}} = 0, \end{array} \right. \quad (8a)$$

$$\left\{ \begin{array}{l} u = g_{[0, u_{\max}]} \left( u_{\max} \left( 1 - \frac{\rho}{\rho_{\text{jam}}} - u_{\max} V_x \right) \right), \end{array} \right. \quad (8b)$$

$$\left\{ \begin{array}{l} u = g_{[0, u_{\max}]} \left( \max \{ \min \{ u, u_{\max} \}, 0 \} \right), \end{array} \right. \quad (8c)$$

where  $g_{[0, u_{\max}]}(u) = \max \{ \min \{ u, u_{\max} \}, 0 \}$  is a cut-off function which ensures the cars' speeds satisfy the constraint  $0 \leq u \leq u_{\max}$ .

[12] provides theoretical and numerical analysis on the MFG system (8a)-(8b)-(8c).

The uniform flows of the MFG system (8a)-(8b)-(8c) are given by  $\bar{u} = u_{\max} \left( 1 - \frac{\bar{\rho}}{\rho_{\text{jam}}} \right)$ . Note that Definition 2.1 does not apply to the MFG system since the system is defined and solved on a fixed time horizon  $[0, T]$ . In this case, we define traffic stability as follows:

**Definition 3.1:** The MFG system (8a)-(8b)-(8c) is stable around the uniform flow  $(\bar{\rho}, \bar{u})$  where  $\bar{u} = u_{\max} \left( 1 - \bar{\rho} / \rho_{\text{jam}} \right)$  if for any  $\varepsilon > 0$ , there exists  $\delta > 0$  such that for any  $T > 0$  and for any solution  $\rho^{(T)}(x, t), u^{(T)}(x, t)$  to the system with  $V_T(x) = 0$  on the time horizon  $[0, T]$ :

$$\sup_{0 \leq t \leq T} \left\{ \left\| \rho^{(T)}(\cdot, t) - \bar{\rho} \right\| + \left\| u^{(T)}(\cdot, t) - \bar{u} \right\| \right\} \leq \varepsilon, \quad (9)$$

whenever  $\| \rho(\cdot, 0) - \bar{\rho} \| \leq \delta$ . The system is linearly stable if its linearized system at  $(\bar{\rho}, \bar{u})$  is stable around the zero solution.

## B. Mixed Traffic: Coupled MFG-ARZ System

This section aims to develop a continuum mixed AV-HV traffic flow model. We denote  $\rho^{\text{AV}}(x, t)$  the AV density,  $\rho^{\text{HV}}(x, t)$  the HV density and

$$\rho^{\text{TOT}}(x, t) = \rho^{\text{AV}}(x, t) + \rho^{\text{HV}}(x, t), \quad (10)$$

the total density. Denote  $u^{\text{AV}}(x, t)$  and  $u^{\text{HV}}(x, t)$  the velocities of AVs and HVs, respectively.

We model HVs by the ARZ model and AVs by the MFG, respectively. The next step is to model the interactions between AVs and HVs. The interactions include the flow interaction and the dynamic interaction.

**Flow interaction.** The flow interaction relates to how the multiclass flows are computed and assigned. We follow the framework from [17] and suppose that the multiclass flows are described by the following continuity equations for both AVs and HVs:

$$(\text{CE-AV}) \quad \rho_t^{\text{AV}} + (\rho^{\text{AV}} u^{\text{AV}})_x = 0, \quad (11)$$

$$(\text{CE-HV}) \quad \rho_t^{\text{HV}} + (\rho^{\text{HV}} u^{\text{HV}})_x = 0. \quad (12)$$

**Dynamic interaction.** Each of the velocities  $u^{\text{AV}}$  and  $u^{\text{HV}}$  should depend on both AV density  $\rho^{\text{AV}}$  and HV density  $\rho^{\text{HV}}$ . The way of defining the velocities over multiclass densities characterizes the dynamic interaction. [17] summarized some possible formulations of the dynamic interaction.

In this paper we model an asymmetric dynamic interaction between AVs and HVs by introducing multiclass

densities into the HJB equations and the momentum equation of the system. For HVs, we assume that HVs only observe the total density  $\rho^{\text{TOT}}$  to adapt their speeds. The momentum equation (2) in the ARZ model then becomes:

$$\begin{aligned} [u^{\text{HV}} + h(\rho^{\text{TOT}})]_t + u^{\text{HV}} [u^{\text{HV}} + h(\rho^{\text{TOT}})]_x = \\ \frac{1}{\tau} [U(\rho^{\text{TOT}}) - u^{\text{HV}}]. \end{aligned} \quad (13)$$

We take the Greenshields desired speed function  $U(\rho) = u_{\max} (1 - \rho / \rho_{\text{jam}})$ .

For AVs, we assume that AVs observe both AV and HV densities. We model AVs' reaction to multiclass densities by introducing an extra term into the AV's cost function. The AV's modified cost function for mixed traffic is:

$$\begin{aligned} f(u^{\text{AV}}, \rho^{\text{AV}}, \rho^{\text{HV}}) = & \underbrace{\frac{1}{2} \left( \frac{u^{\text{AV}}}{u_{\max}} \right)^2}_{\text{kinetic energy}} - \underbrace{\frac{u^{\text{AV}}}{u_{\max}}}_{\text{efficiency}} \\ & + \underbrace{\frac{u^{\text{AV}} \rho^{\text{TOT}}}{u_{\max} \rho_{\text{jam}}} + \beta \frac{\rho^{\text{HV}}}{\rho_{\text{jam}}}}_{\text{safety}}, \end{aligned} \quad (14)$$

where the safety is modeled by two penalty terms: one is similar to the penalty term in (7) but the congestion is modeled by the total density  $\rho^{\text{TOT}}$ , the other quantifies HV's impact on AV's speed selection and the parameter  $\beta$  represents AV's sensitivity to HV's density. From (14) we can derive the corresponding HJB equations.

Summarizing all above, we obtain the following coupled MFG-ARZ system:

$$\left\{ \begin{array}{l} \rho_t^{\text{AV}} + (\rho^{\text{AV}} u^{\text{AV}})_x = 0, \\ V_t + u^{\text{AV}} V_x + \frac{1}{2} \left( \frac{u^{\text{AV}}}{u_{\max}} \right)^2 - \frac{u^{\text{AV}}}{u_{\max}} + \frac{u^{\text{AV}} \rho^{\text{TOT}}}{u_{\max} \rho_{\text{jam}}} \end{array} \right. \quad (15a)$$

$$\left. \begin{array}{l} + \beta \frac{\rho^{\text{HV}}}{\rho_{\text{jam}}} = 0, \\ u^{\text{AV}} = g_{[0, u_{\max}]} \left( u_{\max} \left( 1 - \frac{\rho^{\text{TOT}}}{\rho_{\text{jam}}} - u_{\max} V_x \right) \right), \end{array} \right. \quad (15c)$$

$$\rho_t^{\text{HV}} + (\rho^{\text{HV}} u^{\text{HV}})_x = 0, \quad (15d)$$

$$\begin{aligned} [u^{\text{HV}} + h(\rho^{\text{TOT}})]_t + u^{\text{HV}} [u^{\text{HV}} + h(\rho^{\text{TOT}})]_x = \\ \frac{1}{\tau} \left[ u_{\max} \left( 1 - \frac{\rho^{\text{TOT}}}{\rho_{\text{jam}}} \right) - u^{\text{HV}} \right], \end{aligned} \quad (15e)$$

$$\rho^{\text{TOT}} = \rho^{\text{AV}} + \rho^{\text{HV}}. \quad (15f)$$

- The initial conditions are given by the initial densities  $\rho^{\text{AV}}(x, 0) = \rho_0^{\text{AV}}(x)$ ,  $\rho^{\text{HV}}(x, 0) = \rho_0^{\text{HV}}(x)$  and the initial velocity  $u^{\text{HV}}(x, 0) = u_0^{\text{HV}}(x)$ .
- The terminal condition is given by the terminal cost  $V(x, T) = V_T(x)$ . We will always set  $V_T(x) = 0$ .
- We specify the periodic boundary conditions for all of  $\rho^{\text{AV}}$ ,  $\rho^{\text{HV}}$ ,  $u^{\text{HV}}$  and  $V$ .

The mixed traffic's uniform flows are defined as the system's constant solutions  $\rho^{\text{AV}}(x, t) \equiv \bar{\rho}^{\text{AV}}$ ,  $\rho^{\text{HV}}(x, t) \equiv \bar{\rho}^{\text{HV}}$ ,

$$\rho^{\text{TOT}}(x, t) \equiv \bar{\rho}^{\text{TOT}} = \bar{\rho}^{\text{AV}} + \bar{\rho}^{\text{HV}}, \quad (16)$$

and

$$u^{\text{AV}}(x,t) \equiv u^{\text{HV}}(x,t) \equiv \bar{u} = u_{\max} \left( 1 - \frac{\bar{\rho}^{\text{TOT}}}{\rho_{\text{jam}}} \right). \quad (17)$$

Since AVs are modeled by a mean field game, the mixed traffic system (15a-15f) is defined and solved on a predefined time horizon  $[0, T]$ . Similar to Definition 3.1, the mixed traffic system's stability is defined as:

**Definition 3.2:** The system (15a-15f) is stable around the uniform flow  $(\bar{\rho}^{\text{AV}}, \bar{\rho}^{\text{HV}}, \bar{u})$  which satisfies (17) if for any  $\varepsilon > 0$ , there exists  $\delta > 0$  such that for any  $T > 0$  and for any solution  $\rho^{\text{AV},(T)}(x,t), u^{\text{AV},(T)}(x,t), \rho^{\text{HV},(T)}(x,t), u^{\text{HV},(T)}(x,t)$  to the system with  $V_T(x) = 0$  on the time horizon  $[0, T]$ :

$$\sup_{0 \leq t \leq T} \sum_{i=\text{AV,HV}} \left\| \rho^{i,(T)}(\cdot, t) - \bar{\rho}^i \right\| + \left\| u^{i,(T)}(\cdot, t) - \bar{u} \right\| \leq \varepsilon, \quad (18)$$

whenever

$$\sum_{i=\text{AV,HV}} \left\| \rho^{i,(T)}(\cdot, 0) - \bar{\rho}^i \right\| + \left\| u^{\text{HV},(T)}(\cdot, 0) - \bar{u} \right\| \leq \delta. \quad (19)$$

The system is linearly stable if its linearized system at  $(\bar{\rho}^{\text{AV}}, \bar{\rho}^{\text{HV}}, \bar{u})$  is stable around the zero solution.

#### IV. Pure AV Traffic: Linear Stability Analysis

In this section we will carry out the standard linear stability analysis for the MFG system (8a)-(8b)-(8c).

By scaling to dimensionless quantities we assume  $u_{\max} = 1$  and  $\rho_{\text{jam}} = 1$ . In addition we remove the speed constraint  $0 \leq u \leq u_{\max}$  since the existence of the constraint does not change the system's stability when  $0 < \bar{u} < u_{\max}$ . Then we eliminate  $V$  from the system (8a)-(8b)-(8c) and obtain a simpler system of  $\rho$  and  $u$ :

$$\begin{cases} \rho_t + (\rho u)_x = 0, \\ u_t + uu_x - (\rho u)_x = 0. \end{cases} \quad (20)$$

Fix a uniform flow  $(\bar{\rho}, \bar{u})$  where  $\bar{u} = 1 - \bar{\rho}$ . Suppose that the system (20) has the initial condition  $\rho(x, 0) = \bar{\rho} + \tilde{\rho}_0(x)$  and the terminal condition  $V_T(x) = 0$ . Here  $\tilde{\rho}_0(x)$  is any small perturbation.

Then we linearize the system (20) near the uniform flow  $(\bar{\rho}, \bar{u})$ . Suppose  $\rho(x, t) = \bar{\rho} + \tilde{\rho}(x, t)$ ,  $u(x, t) = \bar{u} + \tilde{u}(x, t)$ . Note that  $\bar{u} = 1 - \bar{\rho}$ , we get the following linearized system:

$$\begin{cases} \tilde{\rho}_t + (1 - \bar{\rho})\tilde{\rho}_x + \bar{\rho}\tilde{u}_x = 0, \\ \tilde{u}_t + (\bar{\rho} - 1)\tilde{\rho}_x + (1 - 2\bar{\rho})\tilde{u}_x = 0. \end{cases} \quad (21)$$

(21) is also a forward-backward system with the initial condition  $\tilde{\rho}(x, 0) = \tilde{\rho}_0(x)$  and the terminal condition  $\tilde{\rho}(x, T) + \tilde{u}(x, T) = 0$ .

**Proposition 4.1:** The linearized system (21) is stable near the zero solution for all  $0 < \bar{\rho} < 1$ .

We provide a computer-assisted proof for Proposition 4.1 in the Appendix. The analytical proof is left for future research. As a corollary of Proposition 4.1 we have the following results on the MFG system's stability:

**Corollary 4.2:** The MFG system (20) is linearly stable around the uniform flow  $(\bar{\rho}, \bar{u})$  where  $\bar{u} = 1 - \bar{\rho}$  for all  $0 < \bar{\rho} < 1$ .

Our analysis shows that the proposed MFG system for AVs is always stable even if each AV only aims to optimize his own utility. Then we turn our attention to the mixed traffic and study whether the existence of AVs can stabilize the unstable HV traffic.

#### V. Mixed Traffic: Numerical Experiments

In this section, we will demonstrate the stability of the mixed traffic system (15a-15f) by numerical experiments. We will run numerical simulations in different scenarios and check the stability in those simulations automatically with a stability criterion. Then we discuss how AVs' different penetration rates and different controller designs influence the stabilizing effect.

##### A. Experimental Settings

Take vehicles' free flow speed  $u_{\max} = 30 \text{ m/s}$  and the jam density  $\rho_{\text{jam}} = 1/7.5 \text{ m}$ . Choose the hesitation function  $h(\rho)$  in the ARZ model to be:

$$h(\rho) = 9 \text{ m/s} \cdot \left( \frac{\rho/\rho_{\text{jam}}}{1 - \rho/\rho_{\text{jam}}} \right)^{1/2}, \quad (22)$$

which has the same form as the one used in [16]. For all of the numerical experiments, the length of the ring road  $L = 1 \text{ km}$  and the length of the time horizon  $T = 2L/u_{\max}$ .

For the system (15a-15f) and its arbitrary uniform flow solution  $(\bar{\rho}^{\text{AV}}, \bar{\rho}^{\text{HV}}, \bar{u})$ , the initial densities are set to be:

$$\rho_0^i(x) = \bar{\rho}^i + 0.1 \times \bar{\rho}^i \sin(2\pi x/L), \quad (23)$$

for  $i = \text{AV, HV}$  so that the initial perturbations on both AV and HV densities are sine waves whose magnitudes are 10% of the respective uniform states. The HV's initial velocity is set to be:

$$u_0^{\text{HV}}(x) \equiv \bar{u} = u_{\max} \left( 1 - \frac{\bar{\rho}^{\text{TOT}}}{\rho_{\text{jam}}} \right), \quad (24)$$

where  $\bar{\rho}^{\text{TOT}} = \bar{\rho}^{\text{AV}} + \bar{\rho}^{\text{HV}}$  so that there is no initial perturbation on HV's velocity. The AV's terminal cost is always set to be  $V_T(x) = 0$ .

It is not easy to check the conditions (18)-(19) directly. Alternatively we shall use a simplified stability criterion. Suppose  $\rho^{\text{AV},(T)}(x, t), u^{\text{AV},(T)}(x, t), \rho^{\text{HV},(T)}(x, t), u^{\text{HV},(T)}(x, t)$  is any solution to the system, we define an error function:

$$E(t) = \sum_{i=\text{AV,HV}} \left\| \rho^{i,(T)}(\cdot, t) - \bar{\rho}^i \right\| + \left\| u^{i,(T)}(\cdot, t) - \bar{u} \right\|, \quad (25)$$

for  $0 \leq t \leq T$  and the system is said to be unstable if:

$$\max_{0 \leq t \leq T} E(t) \geq 2E(0), \quad (26)$$

otherwise it is said to be stable. The stability criterion (26) is checked automatically in the numerical experiments. It is validated in the experiments with no presence of AVs that the criterion (26) predicts the same stability as the ARZ model's analytical stability criterion.

## B. Numerical Method

To solve the coupled MFG-ARZ system (15a-15f) numerically, we apply a finite difference method (FDM) on the spatial-temporal grids. We discretize the continuity equations (15a)(15d) by the Lax-Friedrichs scheme. We discretize the HJB equations (15b)(15c) of the MFG by an upwind scheme [12]. The momentum equation (15e) of the ARZ model is transformed into its conservative form with a relaxation term. Then we apply a hybrid scheme with an explicit Lax-Friedrichs scheme for the conservation part and an implicit Euler scheme for the relaxation part. Finally we compress all equations into a large nonlinear system and solve the system by Newton's method [12].

## C. Numerical Results

In the first group of experiments we fix  $\beta = 0$  and try different pairs of  $\bar{\rho}^{\text{AV}}$  and  $\bar{\rho}^{\text{HV}}$ . We restrict the values to be under  $\bar{\rho}^{\text{AV}} + \bar{\rho}^{\text{HV}} \leq 0.75\rho_{\text{jam}}$  to avoid the total density exceeding the jam density. We check the system's stability from each numerical experiment and plot the results in the phase diagram between the normalized AV and HV density, see Figure 1a. We observe that when the HV density is fixed, adding AVs can stabilize the traffic. when the AV density is large enough, the mixed traffic is always stable.

In the second group of experiments we still keep  $\beta = 0$  but try different total densities  $\bar{\rho}^{\text{TOT}}$  and different AV's penetration rates. Then we plot the results in the phase diagram between the AV's penetration rate and the normalized total density, see Figure 1b. We observe that when the total density is fixed, traffic becomes more stable with a higher portion of AVs. In addition, the minimal AV's penetration rate to make the traffic stable increases as the total density increases. We also observe that when the AV's penetration rate is large enough, the mixed traffic is always stable.

Figure 2 compares the total density evolution between a stable example and an unstable example. When the total density is  $\bar{\rho}^{\text{TOT}} = 0.4\rho_{\text{jam}}$ , the pure HV traffic is unstable while 30% AVs can stabilize the mixed traffic. In the former case, the initial perturbation on the total density grows up and develops a shock; In the latter case, the same initial perturbation decays and the total density converges to a uniform flow.

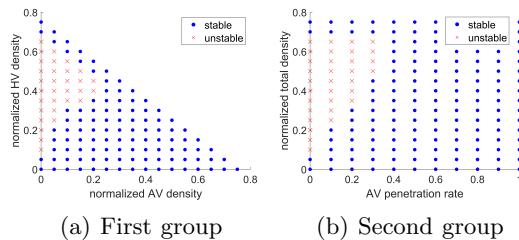


Fig. 1: Stability regions for the first and second groups of experiments

In the third group of experiments we fix the total density  $\bar{\rho}^{\text{TOT}} = 0.5\rho_{\text{jam}}$  and vary the AV's penetration rate and the parameter  $\beta$ . Then we plot the results in the phase diagram

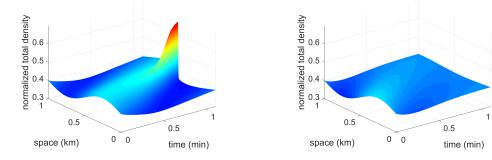


Fig. 2: Evolution of normalized total density when  $\beta = 0$ ,  $\bar{\rho}^{\text{TOT}} = 0.4\rho_{\text{jam}}$ , 0% AV (left) and 30% AVs (right)

between  $\beta$  and the AV's penetration rate, see Figure 3. We observe that for any fixed  $\beta$ , increasing AV's penetration rate makes the traffic more stable. When the AV's penetration rate is fixed but higher than 20%, increasing  $\beta$  makes the traffic more stable. This means that when AV is more sensitive to HV, the traffic becomes more stable.

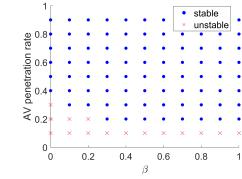


Fig. 3: Stability region for the third group of experiments

## VI. Conclusion

This paper presents continuum traffic flow models for both pure AV traffic and mixed AV-HV traffic. The pure AV traffic is modeled by a mean field game and the linear stability analysis shows the traffic is always stable. The mixed AV-HV traffic is modeled by a coupled MFG-ARZ system. To demonstrate the mixed traffic stability analysis, three groups of numerical experiments are performed. In particular, we characterize the stability regions over AV density and HV density as well as over total density and AV's penetration rate in the mixed traffic. We also quantify the impact of the AV controller parameter on traffic stability. In future work, we plan to develop analytical stability analysis for mixed traffic and discuss the relation between more general AV controller designs and stability under different types of AV-HV interactions.

## APPENDIX

Apply Fourier analysis to (21), denote  $\hat{\rho}(\xi, t)$  and  $\hat{u}(\xi, t)$  the Fourier modes of  $\tilde{\rho}(x, t)$  and  $\tilde{u}(x, t)$ ,  $\xi = \frac{2k\pi x}{L}$  ( $k \in \mathbb{Z}$ ). For any frequency  $\xi$ :

$$\begin{cases} \hat{\rho}_t + i\xi(1 - \bar{\rho})\hat{\rho} + i\xi\bar{\rho}\hat{u} = 0, \\ \hat{u}_t + i\xi(\bar{\rho} - 1)\hat{\rho} + i\xi(1 - 2\bar{\rho})\hat{u} = 0. \end{cases} \quad (27)$$

It is an ODE system with the initial condition  $\hat{\rho}(\xi, 0) = \hat{\rho}_0(\xi)$  where  $\hat{\rho}_0(\xi)$  is the Fourier transform of  $\tilde{\rho}_0(x)$  and the terminal condition  $\hat{\rho}(\xi, T) + \hat{u}(\xi, T) = 0$ . The linear PDE system (21) is stable in  $L^2$  norm if and only if there exists a universal constant  $C > 0$  such that for any  $T > 0$  and  $\xi$ , the solution of the ODE system (27) on  $[0, T]$  satisfies:

$$|\hat{\rho}(\xi, t)|^2 + |\hat{u}(\xi, t)|^2 \leq C|\hat{\rho}_0(\xi)|^2, \forall t \in [0, T]. \quad (28)$$

The ODE system (27) is homogeneous. We can assume without loss of generality that  $\hat{\rho}_0(\xi) = 1$ . To check the condition (28) we directly solve this boundary value problem of the ODE system (27). Denote  $r = \sqrt{\bar{\rho}(5\bar{\rho}-4)}$ ,  $\eta = \xi t$ ,  $\lambda = \xi T$  and  $S = \exp(-\frac{1}{2}i\eta(r-3\bar{\rho}+2))$ , the solution is:

$$\hat{\rho}(\xi, t) = S \frac{(r+\bar{\rho})e^{ir\eta} + (r-\bar{\rho})e^{ir\lambda}}{r+\bar{\rho} + (r-\bar{\rho})e^{ir\lambda}}, \quad (29)$$

$$\hat{u}(\xi, t) = -S \frac{(r+3\bar{\rho}-2)e^{ir\eta} + (r-3\bar{\rho}+2)e^{ir\lambda}}{r+\bar{\rho} + (r-\bar{\rho})e^{ir\lambda}}, \quad (30)$$

when  $\bar{\rho} \neq \frac{4}{5}$  or  $\hat{\rho}(\xi, t) = e^{\frac{1}{5}i\eta \frac{5i-2\eta+2\lambda}{5i+2\lambda}}$  and  $\hat{u}(\xi, t) = -e^{\frac{1}{5}i\eta \frac{5i-\eta+\lambda}{5i+2\lambda}}$  when  $\bar{\rho} = \frac{4}{5}$ .

Define:

$$E_{\bar{\rho}}(\lambda) = \max_{0 \leq \eta \leq \lambda \text{ or } \lambda \leq \eta \leq 0} [|\hat{\rho}(\xi, t)|^2 + |\hat{u}(\xi, t)|^2]. \quad (31)$$

Then to check (28) it suffices to check the boundedness of the function  $E_{\bar{\rho}}(\lambda)$  for all  $0 < \bar{\rho} < 1$ . We do this by computing the values of  $E_{\bar{\rho}}(\lambda)$  from discrete values of  $\bar{\rho}$  and  $\lambda$ . The computation shows that for any  $\bar{\rho}$ ,  $E_{\bar{\rho}}(\lambda)$  is bounded when  $|\lambda| \rightarrow \infty$ .

## ACKNOWLEDGMENT

The authors would like to thank Data Science Institute from Columbia University for providing a seed grant for this research.

## References

- [1] R. E. Stern, S. Cui, M. L. Delle Monache, R. Bhadani, M. Bunting, M. Churchill, N. Hamilton, H. Pohlmann, F. Wu, B. Piccoli et al., "Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments," *Transportation Research Part C: Emerging Technologies*, vol. 89, pp. 205–221, 2018.
- [2] I. G. Jin and G. Orosz, "Connected cruise control among human-driven vehicles: Experiment-based parameter estimation and optimal control design," *Transportation research part C: emerging technologies*, vol. 95, pp. 445–459, 2018.
- [3] S. Cui, B. Seibold, R. Stern, and D. B. Work, "Stabilizing traffic flow via a single autonomous vehicle: Possibilities and limitations," in *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2017, pp. 1336–1341.
- [4] C. Wu, A. M. Bayen, and A. Mehta, "Stabilizing traffic with autonomous vehicles," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1–7.
- [5] Y. Zheng, S. E. Li, K. Li, and L.-Y. Wang, "Stability margin improvement of vehicular platoon considering undirected topology and asymmetric control," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 4, pp. 1253–1265, 2016.
- [6] M. J. Lighthill and G. B. Whitham, "On kinematic waves II. A theory of traffic flow on long crowded roads," *Proc. R. Soc. Lond. A*, vol. 229, no. 1178, pp. 317–345, 1955.
- [7] M. W. Levin and S. D. Boyles, "A multiclass cell transmission model for shared human and autonomous vehicle roads," *Transportation Research Part C: Emerging Technologies*, vol. 62, pp. 103–116, 2016.
- [8] D. Ngoduy, S. Hoogendoorn, and R. Liu, "Continuum modeling of cooperative traffic flow dynamics," *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 13, pp. 2705–2716, 2009.
- [9] K. Porfyri, I. Nikolos, and M. Papageorgiou, "Stability analysis of a macroscopic traffic flow model for adaptive cruise control systems," in *ASME 2015 International Mechanical Engineering Congress and Exposition*. American Society of Mechanical Engineers, 2015, pp. V012T15A002–V012T15A002.
- [10] A. I. Delis, I. K. Nikolos, and M. Papageorgiou, "Simulation of the penetration rate effects of ACC and CACC on macroscopic traffic dynamics," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2016, pp. 336–341.
- [11] —, "A macroscopic multi-lane traffic flow model for ACC/CACC traffic dynamics," *Transportation Research Record*, vol. 2672, no. 20, pp. 178–192, 2018.
- [12] K. Huang, X. Di, Q. Du, and X. Chen, "A game-theoretic framework for autonomous vehicles velocity control: Bridging microscopic differential games and macroscopic mean field games," *arXiv preprint arXiv:1903.06053*, 2019.
- [13] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.
- [14] A. Aw and M. Rascle, "Resurrection of "second order" models of traffic flow," *SIAM Journal on Applied Mathematics*, vol. 60, no. 3, pp. 916–938, 2000.
- [15] H. M. Zhang, "A non-equilibrium traffic model devoid of gas-like behavior," *Transportation Research Part B: Methodological*, vol. 36, no. 3, pp. 275–290, 2002.
- [16] B. Seibold, M. R. Flynn, A. R. Kasimov, and R. R. Rosales, "Constructing set-valued fundamental diagrams from jamiton solutions in second order traffic models," *arXiv preprint arXiv:1204.5510*, 2012.
- [17] S. Fan and D. B. Work, "A heterogeneous multiclass traffic flow model with creeping," *SIAM Journal on Applied Mathematics*, vol. 75, no. 2, pp. 813–835, 2015.