Traffic-flow models: analysis, estimation and control

by

Chao Xia

B.S., University of Science and Technology of China, Hefei, China, 2012

M.Sc., Brown University, Providence, RI, 2013

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Division of Applied Mathematics at Brown University

PROVIDENCE, RHODE ISLAND

May 2017

Abstract of " Traffic-flow models: analysis, estimation and control " by Chao Xia, Ph.D., Brown University, May 2017

My research focuses on the study of traffic-flow models and their applications. Macroscopic and microscopic models are the two main approaches: macroscopic models describe the spatial quantities of traffic, such as density, velocity and flux; while microscopic models simulate the behavior of individual cars based on their interaction.

For the macroscopic model, we study the Lighthill-Whitham equation, and account for multiple traffic scenarios by modifying the original Lighthill-Whitham equation. We also study several microscopic car-following models: the optimal velocity model, the full velocity difference model, the modified GHR model and the intelligent driver model. The main research work include:

- Investigate the collision behavior of the microscopic car-following model. We theoretically prove the collision-free property of several car-following models through fast-slow system technique, and also carry out numerical simulations to provide a valid reference to the dynamics of traffic collisions.
- Apply data assimilation technique (ensemble Kalman filter and particle filter) to estimating the traffic states and uncertain parameters. An augmented approach is proposed to simultaneously assimilate the Eulerian sensor data and Lagrangian GPS data.
- Study the phenomenon of capacity discharge in the lane-drop scenario. Macroscopically, we model the lane-drop scenario with inhomogeneous Lighthill-Whitham equation, and then proposed two controlling strategies to guide vehicles smoothly through the bottleneck: (1) change driving habit through fundamental diagram; (2) merge vehicles in advance through virtual lane usage.

This dissertation by Chao Xia is accepted in its present form
by the Division of Applied Mathematics as satisfying the
dissertation requirement for the degree of Doctor of Philosophy.

Date_____ _____

Björn Sandstede, Ph.D., Advisor

Recommended to the Graduate Council

Date_____ _____

Constantine Dafermos, Ph.D., Reader

Date_____ _____

Matthew Harrison, Ph.D., Reader

Approved by the Graduate Council

Date_____ _____

Andrew G. Campbell, Dean of the Graduate School

Vitae

Education

Brown University, Providence, RI, 2012-2017

Ph.D. in Applied Mathematics, expected May 2017

M.Sc. in Applied Mathematics, 2013

University of Science and Technology of China, Hefei, China, 2008-2012

B.S. in Statistics, School for Gifted Young, 2012

Publications

C Xia and B Sandstede. Traffic control in lane-drop scenarios. In preparation, 2017.

C Xia and B Sandstede. A study of collision behavior of car-following models. Preprint, 2017.

C Xia, C Cochrane, J DeGuire, G Fan, E Holmes, M McGuirl, P Murphy, J Palmer, P Carter, L Slivinski, and B Sandstede. Assimilating Eulerian and Lagrangian data

in traffic-flow models. Physica D: Nonlinear Phenonmena. Under review, 2016.

P Wu, C Xia, F Liu, Y Wu and Y He. An integrated water strategy based on the current circumstances in China. Applied Mathematical Modeling 40 (2016) 8108-8124.

K Judd, T Stemler and C Xia. Paleoclimate variability and frequency selection by stochastic resonance in Pitchfork Bifurcations. Physical Review E. Submitted.

Selected Presentations

"Traffic flow: estimation and control." Graduate Student Seminar of the Division of Applied Mathematics, Brown University, Providence, RI, November 2016.

"Introduction to data Assimilation for traffic estimation." Oberwolfach Seminar: Data Assimilation: The Mathematics of connecting Dynamical System to Data, Oberwolfach, Germany, May 2016.

"Lagrangian data assimilation in traffic-flow models." (Poster) IPAM Workshop of Traffic Estimation in UCLA, Los Angeles, CA, Octorber 2015.

"Data assimilation for traffic states and parameters estimation." (Poster) SIAM Conference on Applications of Dynamical Systems, Snowbird, UT, May 2015.

Honors and Awards

2016 NSF Supplemental Funding (summer internship)

2012 Outstanding Undergraduate Research Project, University of Science and Technology of China

2011 Third Prize for China Undergraduate Mathematical Contest Modeling

2011 Certificate of Excellence from University of Western Australia, University of Western Australia

Professional Experience

Intern, JP. Morgan Chase & Co., 2016 Summer

Intern, Liberty Mutual Insurance, 2015 Summer

Visitor, University of Western Australia, 2011 Summer

Teaching Experience

Teaching Assistant, Brown University

APMA1650 (Statistical Inference), Fall 2013

APMA0650 (Essential Statistics), Spring 2014

Professional Associations

American Mathematical Society (AMS)

Association for Women in Mathematics (AWM)

Society for Industrial and Applied Mathematics (SIAM)

Dedication

To my loving father, who nurtured my curiosity in mathematics when I was in primary school, and supported my pursuit in mathematics in higher education.

Acknowledgements

I would first like to thank my advisor Björn Sandstede, for his mentorship in my thesis topic and his support in my professional development. He has his own philosophies of supervising students, and his responsibility, dedication and enthusiasm in research work has a positive impact on my experience as a graduate student. Aside from being an advisor in my research work, he is willing to provide resources and opportunities to develop my expertise in industrial job application. It is truly my honor to have Björn as my advisor in my graduate study.

I would also like to thank Professor C. Dafermos and Professor M. Harrison for donating their time to serve on my thesis committee. Thanks to the valuable comments and feedback for my thesis. I would like to thank the collaboration in the Summer RTG Program on "Integrating Dynamics and Stochastics", which enlightened my interest in the traffic-flow models.

I would also like to thank the team members of the research group for the informative group meetings, and helpful discussions. In particular, I would like to thank Al, Paul and Veronica for sharing a lovely office, varied opinions and study resources.

Finally, I want to express my thanks to my family: my husband Wenzhe, my parents Linyan and Qiuqin, for their continued care and love. Special thanks to the companion and encouragement from my husband in my graduate study, and his sharing of my happiness and sadness.

Contents

List of Tables

List of Figures

Chapter One

Introduction

1.1 Motivation and preliminaries

Traffic flow theory and modeling started in the 1930. This field has gained considerable attention as overall traffic demand has increased and more data as well as easy access to computing power has become available. Traffic is everywhere in our daily life. There is a growing need for accurate traffic information so that the public can efficiently schedule their trips, and the government can better provide traffic control strategies. Thus, modeling, analyzing, estimating and controlling the dynamics of traffic flow are of great importance. First of all, I would like to have a overview of traffic flow: (a) traffic-flow models (b) traffic-flow observation.

Mathematical models for traffic flow come in many different flavors. Common models range from microscopic models for individual cars to macroscopic models for car densities and possibly other quantities. As the name indicates, macroscopic models formulate the relationship among the spatial quantities such as density, velocity and flow. In contrast, microscopic models simulate individual vehicle-driver unites, so the dynamic variables of the models represent properties like the position and velocity of individual vehicles.

The macroscopic traffic models are usually described in a partial-differential equation (PDE) to build a relationship between spatial quantities [39, 54, 63]. The most well-known macroscopic model is the Lighthill-Whitham-Richard (LWR) model [54, 63], which describes the the evolution of traffic by the vehicle density according the conservation law. The LWR model can be modified to account for multiple traffic scenarios through velocity-density relation or flux-density relation. The common traffic scenarios include normal traffic, traffic lights, on-/off-ramps, construction zones, traveling bottlenecks and lane drops.

The microscopic traffic models include the acceleration and lane-changing models [61]. The acceleration models are formulated by coupled ordinary differential equations, which describe the variables of individual vehicles, such as positions and velocities. The lane-changing models involve the vehicles interaction across different lanes, and specify the rules for merging. The microscopic models are able to provide traffic simulation in a realistic road profile, and the well-known traffic simulators include AIMSUN (advanced interactive microscopic simulator for urban and nonurban networks), PARAMICS (parallel microscopic simulation) and VISSIM (Verkehr in Stadten simulation) [59]. The dynamics of the microscopic models is an interesting area so that the models can better simulate the realistic traffic.

Different aspects of traffic dynamics are captured by different measurement methods. Traffic-flow observations come from a variety of sources and are available as functions of time. Stationary sensors, such as induction loops or cameras, provide the flux, average velocity, and local density of cars that move pass the fixed sensor location. GPS data from cell phone or navigation devices, on the other hand, provide information about the positions and velocities of individual cars that move with the traffic flow. We refer to observations that come from a fixed observation location as Eulerian observations and to observations that come from parcels that move with the traffic flow as Lagrangian observations. Both Eulerian and Lagrangian observations have their own advantages and disadvantages with respect to collection cost, data accuracy, traffic coverage or even driver privacy. A lot of comparative studies are done to compare the properties of both types of observation. They are used in different circumstances for calibration, estimation and prediction based on the impacts of their advantages and disadvantages.

1.2 Thesis objectives

There are a lot of interesting traffic topics, ranging from the theoretic study of traffic-flow models, to their application in realistic traffic scenarios. There are three objectives related to traffic-flow modeling in my thesis:

- Studying the dynamics of car-following models

  Car-following models are widely used in traffic simulation because of its variability in the formula of ordinary differential equations. The traffic dynamics of these models are of great importance, and thus have received plenty of study from scholars. However, the collision behavior of car-following models remains a challenging problem. One difficulty is that the traffic system of multiple vehicles is complicated when all interactions among vehicles are taken into consideration. Previous study for making inferences about the collision behavior depended on vast numerical simulations. On one side, the study cannot provide a theoretical analysis of the car-following models from its essence. On the other side, the inferences from the simulations could be misleading in the existence of numerical error. In my thesis, we carry out a theoretical analysis of the collision behavior of four well-known car-following models. We show that the modified GHR model and the intelligent driver model are collision-free under all traffic circumstances. We also show that the numerical errors introduce collisions that the model doesn't support.

- Assimilating Eulerian and Lagrangian data in traffic-flow models

  Data assimilation of traffic flow remains a challenging problem. One difficulty is that data come from different sources ranging from stationary sensors and camera data to GPS and cell phone data from moving cars. Sensors and

cameras give information about traffic density, while GPS data provide information about the positions and velocities of individual cars. Previous methods for assimilating Lagrangian data collected from individual cars relied on specific properties of the underlying computational model or its reformulation in Lagrangian coordinates. These approaches make it hard to assimilate both Eulerian density and Lagrangian positional data simultaneously. In this thesis, we propose an alternative approach that allows us to assimilate both Eulerian and Lagrangian data. We show that the proposed algorithm is accurate and works well in different traffic scenarios and regardless of whether ensemble Kalman or particle filters are used. We also show that the algorithm is capable of estimating parameters and assimilating real traffic observations and synthetic observations obtained from microscopic models.

- Providing traffic control strategies to maximize capacity in lane-drop bottleneck

  Lane drop is a location where number of lanes provided decreases. Empirical observations at lane-drop bottleneck revealed that there is a drop in the bottleneck discharge rate when queues form in the upstream of the lane drop. Even though the research results show that lane changes are the main cause of the drop in discharge rate, the relevant traffic control to maximize the capacity is still a challenging problem. One difficulty is that the interaction between lanes are complicated to quantify, and that there are a lot of parameters to specify. Previous methods to maximize capacity either attempted to develop cooperative lane-changing models to reduce the queues in the upstream of lane drop or formulated a conservation equation for individual lane to optimize the density transfer between adjacent lanes. These approaches focus on the interaction rules between lanes, which are complicated to quantify, and not target to maximize capacity directly. In my thesis, we introduce the inhomogeneous

LWR model to account for the change of lane number, and use one continuity equation to model the total density. When the maximal capacity is guaranteed by the inhomogeneous LWR model, we attempt to smooth the velocity curve so that the embedded vehicles are comfortable to follow it. We show the traffic pattern by the inhomogeneous LWR model and its property to maximize capacity. We also show two controlling strategies to smooth and narrow the velocity gap in the solution to the inhomogeneous LWR model.

Outlines

The work is organized as follows. In Chapter 2, we provide an overview of macroscopic traffic models (LWR model) as well as the fundamental diagram, and a discussion of multiple traffic scenarios based on LWR model, which will be used to test the efficacy of our proposed assimilation method in Chapter 5. In Chapter 3, we provide an overview of microscopic traffic models, and give a detailed introduction to the well-known car-following models, along with notations and setup for the following chapter. Chapter 4 provides the analysis of the stability and collision behavior of the car-following models. The theoretical proof as well as the numerical implementation is shown to be in agreement. In Chapter 5, we have a detailed discussion of the traffic observations and data assimilation approaches. Then we proposed an alternative approach that allows us to assimilate both Eulerian and Lagrangian data, and show the accuracy and efficacy of this approach in multiple traffic scenarios. The lane-drop bottleneck is investigated in Chapter 6. The inhomogeneous LWR model is introduced to account for the multi-lane traffic scenarios. We provide two controlling strategies to maximize the capacity while providing a comfortable velocity curve for vehicles to follow. Finally, Chapter 7 includes an overall discussion, conclusion and directions for future work.

Chapter Two

Review: macroscopic traffic models

2.1 Introduction

A macroscopic traffic model is a mathematical traffic model that formulates the relationships among traffic flow characteristics like density, flow, mean speed of a traffic stream. The characteristics are locally aggregated, which vary across space and time, i.e., they correspond to dynamic fields. Thus, macroscopic models are able to describe collective phenomena such as the evolution of congested regions or the propagation of traffic wave ([73]). Macroscopic models [19, 42, 54, 63] describe traffic flow analogously to liquids or gases in motion. Hence they are sometimes called hydrodynamics models.

Since the legendary paper [54, 63] by Lighthill, Whitham and Richards, dynamical macroscopic traffic flow modeling became a central focus for both theoretical- and application-oriented research. This is a first order model based on the scalar conservation law. The first-order model was complemented by second-order model by Payne [60], in an attempt to avoid some known deficiencies of the first-order model. However, in Daganzo's note [20], it described the logical flows of the higher order continuum models. In our research, we will focus the first-order model [54, 63] and its variants.

Conventionally, the macroscopic models are derived from integrating the microscopic traffic flow models and converting the single-entity level characteristics to system level characteristics ([5, 25]). In [5], it establishes a connection between a microscopic follow-the-leader model based on ordinary differential equation and a semidiscretization of a macroscopic continuum model based on a conservation law. They also show rigorously that, at least in the homogeneous case, the macroscopic model can be viewed as the limit of the time discretization of the microscopic model

as the number of vehicles increase, with a scaling in space and time.

The foundations of macroscopic model are the hydrodynamic relation "flow equals density times speed" and the continuity equation, which describes the temporal evolution of the traffic flow characteristics. The vehicle flux-density relation is called the fundamental diagram of traffic, which is a modeling choice. Traffic flux refers to the number of passing vehicles per unit time, and traffic density refers to the number of vehicles per unit length. Then the vehicle velocity is defined such as the flow-density relation is satisfied. More details about the fundamental diagram will be provided in section §2.2.1.

The macroscopic models are particularly useful when one is interested in macroscopic quantities and the microscopic effects (lane changes, driver-vehicle type, acceleration) need not be considered. Some main applications of macroscopic model include:

- Macroscopic models can be used to study the spatiotemporal evolution of congested traffic pattern, and simulate the effects of traffic flow breakdown caused by high traffic load, bottleneck and disturbance of drivers.
- Macroscopic models can be modified to account for realistic road profiles, such as traffic lights, on-/off-ramps, change in the number of lanes, intersections, construction zone and so on.
- Flow and aggregated speed data from stationary detectors and trajectories of moving cars allow us to estimate and predict the traffic states. Since most of this information is not well known, it is typically more common to use macroscopic models for traffic estimation.
- From the analysis of the spatiotemporal dynamics on highways, we can implement model-based traffic flow optimization to increase the efficiency and

stability of traffic flow.

The traffic pattern and wave propagation are well studied in [49, 54, 63, 73]. In this chapter, we touch on the variations of macroscopic model to account for road profiles mentioned above. In chapter §5, we will apply the Lighthill-Whitham-Richards model to provide traffic evolution at later times, and focus on the estimation of traffic states and uncertain parameters. In chapter §6, we study the lane-drop bottleneck and discuss the potential optimization strategies based on the Lighthill-Whitham-Richards model.

2.2 Lighthill-Whitham-Richard model

This section reviews the theory of scalar first-order conservation law, known as the Lighthill-Whitham-Richard (LWR) partial differential model [54, 63]. The LWR model describes the evolution of traffic by the vehicle density $\rho(x,t)$ at location x and time t. For simplicity, we consider a ring road of length L but emphasize that roads with other boundary conditions can also be treated by our approach. In the absence of sinks and sources, the conservation law for the vehicle density $\rho(x,t)$ with periodic boundary conditions is given by

$$\begin{aligned} &\partial_t \rho(x,t) + \partial_x \varphi(\rho(x,t)) = 0, \qquad x \in \mathbb{T}, \\ &\rho(x,0) = \rho_0(x), \end{aligned} \tag{2.1}$$

where $\rho_0(x)$ denotes the initial data, and $\mathbb{T}$ denotes the circle of circumference L that reflects the periodic boundary conditions we employ. The equation describes the rate of change of density in terms of gradients of the flow.

The flux function $\varphi(\rho)$ expresses the dependence of the vehicle flux φ on the density ρ; this relationship is usually referred to as the fundamental diagram. We assume that φ is defined on the interval $[0, \rho_{\max}]$, where $\rho_{\max}$ is the maximal density achievable on the road.

For traffic flow, we can write $\varphi(\rho) = \rho V(\rho)$, where the function $V(\rho)$ relates velocity v and density ρ for densities $\rho \in [0, \rho_{\max}]$. The function $V(\rho)$ is a modeling choice: examples include the Daganzo–Newell velocity function [76] and the Greenshields affine velocity function [36].

2.2.1 Fundamental diagram

Fundamental diagram of traffic flow is the plot of the traffic flux φ versus the traffic density ρ. According to [64], the two-phase traffic theory divides traffic flow into free flow for low densities, and congested flow for large densities. Figure 2.1 is the fundamental diagram from Seibold's paper [64]. In the free flow regime, flux increases as density increases, while in the congested flow regime, flux decrease as density increases.

The flux-density relation can be mathematically written as $\varphi(\rho) = \rho V(\rho)$, where velocity function $V(\rho)$ relates velocity v and density ρ. The algebraic expression of the velocity function is a modeling choice, and it is typically constructed to fit experimental data. We will briefly introduce the the Daganzo–Newell velocity function [76] and the Greenshields affine velocity function [36], which are used in §5 and §6.

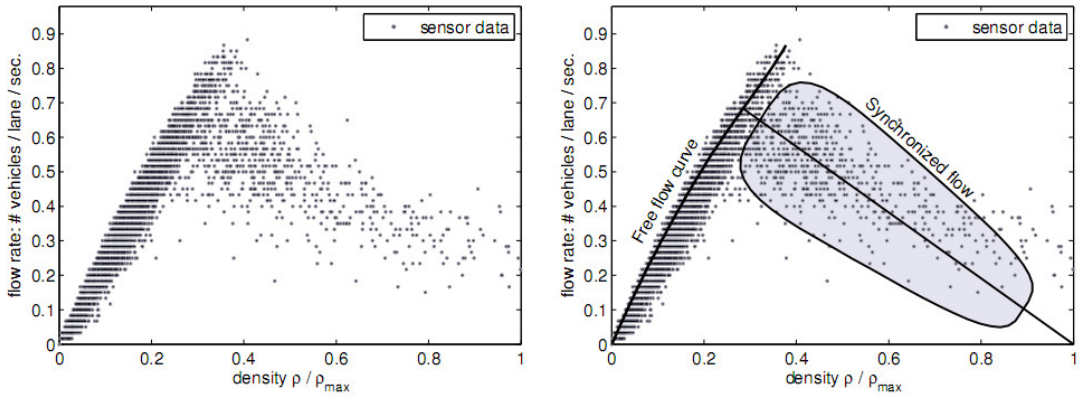


Figure 2.1: Fundamental diagram from [64]. Flow-density data obtained from sensor measurement data, aggregated over time intervals of $\Delta t = 30s$ (left). The free flow curve and a synchronized flow region of fundamental diagram (right).

Greenshields velcoity function: It is one of the earliest introduced velocity functions by Greenshields. It expresses a linear relationship between the speed and density: the velocity linearly decreases as the density increases.

$$v = V_G(\rho) = v_{\max}\left(1 - \frac{\rho}{\rho_{\max}}\right) \tag{2.2}$$

where $v_{\max}$ is the maximal free-flow velocity, and $\rho_{\max}$ is the maximal density achievable on the road. This model remains useful because of its simplicity. The velocity function and respective fundamental diagram are shown in Figure 2.2.

Daganzo-Newell velocity function: It is a widely used velocity function by assuming a constant velocity in free-flow regime, and a hyperbolic velocity in the congestion regime:

$$v = V_{DN} = \begin{cases} v_{\max}, & \text{if } \rho \leq \rho_c \\ -\omega\left(1 - \frac{\rho_{\max}}{\rho}\right), & \text{if } \rho > \rho_c \end{cases} \tag{2.3}$$

where ρ_c is called the critical density that separates the free-flow regime and congestion regime. The Daganzo-Newell velocity gives a triangular fundamental diagram. The velocity function and respective fundamental diagram are shown in Figure 2.2.

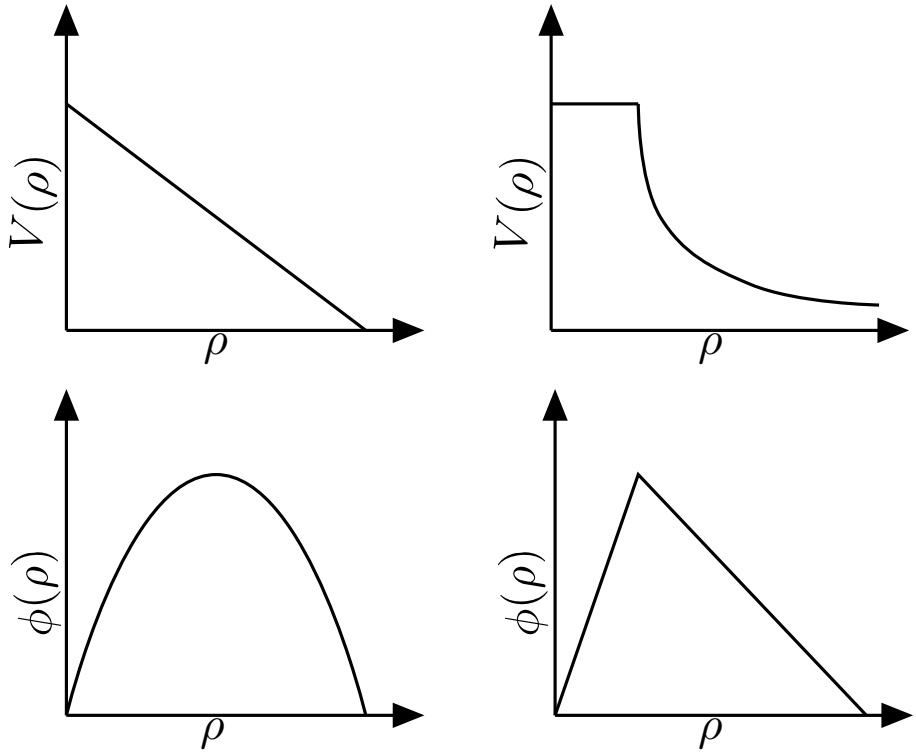


Figure 2.2: Velocity functions and respective fundamental diagrams for Greenshields (left) and Daganzo-Newell (right).

2.2.2 Characteristic curve

The LWR equation is a first-order PDE, we could use the method of characteristics to discover curves along which the PDE becomes an ODE. The LWR equation without viscous term can be written in a quasilinear form:

$$\frac{\partial \rho}{\partial t} + \varphi'(\rho)\frac{\partial \rho}{\partial x} = 0 \tag{2.4}$$

The characteristic equations for the original system with initial conditions are:

$$\begin{cases} \frac{\partial t}{\partial s} = 1, \quad t(0) = 0 \\ \frac{\partial x}{\partial s} = \varphi'(\rho), \quad x(0) = x_0 \\ \frac{\partial \rho}{\partial s} = 0, \quad \rho(0) = \rho(x_0, 0) \end{cases}$$

Because of $t = s$ from the first equation, x and ρ can be expressed as functions of t:

$$x(t) = \varphi'(\rho)t + x_0$$
$$\rho(t) = \rho(x_0, 0) = \rho_0(x_0)$$

Therefore, we get this relation $\rho(x(t), t) = \rho_0(x_0) = \rho_0(x - \varphi'(\rho)t)$. In this case, the characteristic lines are straight lines with slope $\varphi'(\rho)$, and the value of ρ remains constant along the characteristic line. Therefore, the solution of Lighthill–Whitham model comes from following characteristic curve back from (x, t) to a point when $t = 0$:

$$\rho(x, t) = \rho_0(x - \varphi'(\rho(x, t))t) \tag{2.5}$$

Riemann problem: The initial problem (2.4) with discontinuous initial condition of the form

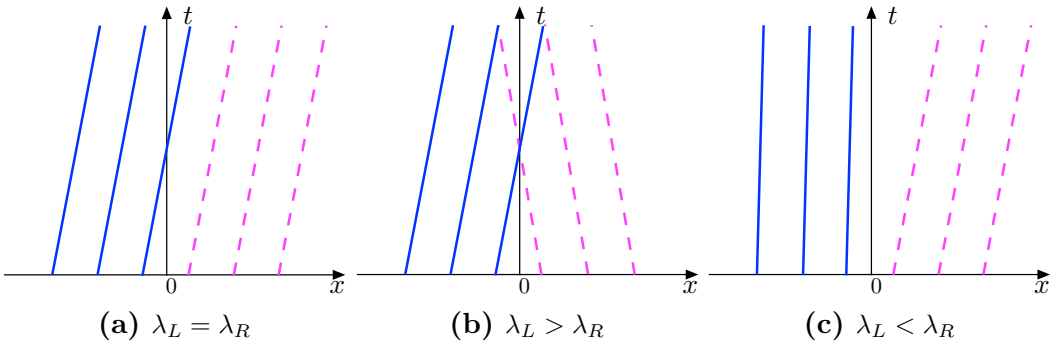
$$\rho(x, 0) = \begin{cases} \rho_L, & x \leq 0 \\ \rho_R, & x > 0 \end{cases} \tag{2.6}$$

is called a Riemann problem. By using the characteristic method presented above, we can obtain two characteristic curves:

$$x = \varphi'(\rho_L)t + x_0$$
$$x = \varphi'(\rho_R)t + x_0.$$

The slopes of the characteristic curve can be understood as wave speed. We use new notations λ_L and λ_R to represent the wave speeds in Riemann problem.

Solutions Based on the relation of λ_L and λ_R, the solution to the Riemann problem can be traveling wave, shock wave and rarefaction. The relations include $\lambda_L = \lambda_R$, $\lambda_L > \lambda_R$ and $\lambda_L < \lambda_R$, which are plotted in Figure 2.3.



(a) $\lambda_L = \lambda_R$ **(b)** $\lambda_L > \lambda_R$ **(c)** $\lambda_L < \lambda_R$

Figure 2.3: Characteristic curves of the Riemann problem when wave speeds (a) $\lambda_L = \lambda_R$, (b) $\lambda_L > \lambda_R$ and (c) $\lambda_L < \lambda_R$

We substitute the Greenshields velocity function into the flux function, and obtain $\varphi(\rho) = \rho v_{\max}(1 - \rho/\rho_{\max})$. The slope of the characteristic curve becomes $\varphi'(\rho) = v_{\max}(1 - 2\rho/\rho_{max})$, which is a linear decreasing function respect to density ρ. The Riemann solutions include:

- Traveling wave ($\lambda_L = \lambda_R = \lambda$)

  The wave doesn't change the speed from upstream to downstream, so it is a traveling wave with speed λ. The solution can be expressed as

$$\rho(x,t) = \rho_0(x - \lambda t) \tag{2.7}$$

- Shock wave ($\lambda_L > \lambda_R$)

  When the wave speed in upstream is greater than downstream, a discontinuity

will be generated and propagated. This is shock wave.

The Rankine-Hugoniot jump condition determines the position of a shock at a given time. This condition emerges when one consider the equation in integral form by integrating the LWR respect to x:

$$\frac{d}{dt}\int_{x_1}^{x_2} \rho(x,t)dx + \varphi(\rho)|_{x_1}^{x_2} = 0$$

Then we can obtain the wave speed of the shock

$$\lambda = \frac{\varphi(\rho_L) - \varphi(\rho_R)}{\rho_L - \rho_R} = \frac{\lambda_L + \lambda_R}{2} \tag{2.8}$$

The discontinuity propagating with speed λ satisfies the entropy condition because $\lambda_L > \lambda > \lambda_R$ holds.

- Rarefaction ($\lambda_L < \lambda_R$)

  The whole situation changes. Even though there is a shock wave solution satisfying the Rankine-Hugoniot condition, the shock wave doesn't satisfy the entropy condition because of $\lambda_L < \lambda < \lambda_R$.

  A weak solution satisfying the entropy condition is rarefaction wave:

$$\rho(x,t) = \begin{cases} \rho_L & x < \lambda_L t \\ w(x/t) & \lambda_L t \le x \le \lambda_R t \\ \rho_R & x > \lambda_R t \end{cases}, \tag{2.9}$$

  where $w(.)$ is a smooth function with $w(\lambda_L) = \rho_L$ and $w(\lambda_R) = \rho_R$.

So far, we are discussing the homogeneous LWR model and the solutions to the Riemann problem. In Chapter 6, we will introduce the inhomogeneous LWR model

to account for the change of lane number.

2.3 Traffic scenarios based on LWR model

As we know, the original LWR equation (2.1) models the conservation of traffic flow in absence of sources or sinks. In addition, the velocity function $V(\rho)$ is assumed not dynamic, which is unable to describes a dynamic relation between velocity and density. Therefore, the original LWR equation is limited to model more complicated road profiles.

The main road profiles include normal traffic (§ 2.3.1), on-/off-ramp (§ 2.3.2), traffic lights (§ 2.3.3), construction zone (§ 2.3.4), traveling wave (§ 2.3.5), change in the number of lanes (§ 6.2), and so on. In this section, we modify the existing LWR model to account for effects of the first five traffic scenarios. The modifications make the macroscopic model more accurate to describe the realistic traffic. The change in the number of lanes is a particular traffic scenario we will discuss in more details in § 6.2.

2.3.1 Normal traffic

In the LWR model (2.1), the adaption of speed to traffic density is instantaneous, because the relation $v(x,t) = \varphi(\rho(x,t))/\rho(x,t)$ holds if function φ and ρ are appropriately defined. This means that it takes zero time for a driver to change driving speed according to the density. It is normal that some fine structure of traffic is missing in the LWR model.

In the normal traffic, the drivers need response time to adjust the velocity, so the relation $v(x,t) = \varphi(\rho(x,t))/\rho(x,t)$ doesn't hold when density experiences big changes. In [81], the researcher exploit the relation between driver memory and viscosity to develop a viscous continuum model, which takes the drivers' response time into consideration through a diffusion term on the right-hand-side of the LWR equation. Therefore, we introduce the viscous LWR model to describe the normal traffic:

$$\rho_t + (\rho V_G(\rho))_x = \varepsilon \rho_{xx}, \tag{2.10}$$

where ε is the diffusion coefficient, which acts as a smoothing factor in the model.

The diffusion term with diffusion coefficient ε has two other advantages: this term accounts for low-level noise when updating the traffic states at later times; from the perspective of numerical solution, the introduction of small diffusion term can produce a weak solution that satisfies the entropy condition.

The viscous LWR model (2.10) will act as the base model for the following traffic scenarios: on-/off-ramps, traffic lights, construction zone and traveling bottleneck.

2.3.2 On-/off-ramps

Ramp is a road junction of the highway system: an on-ramp provides access to the specific part of a road system, while an off-ramp is one-way lane for departing a main highway. The sketch of on-/off-ramps is shown in Figure 2.4. Mathematically, on-/off-ramps imply additional in-/outflows, which have to be added to the section boundaries where the on-/off-ramps are.

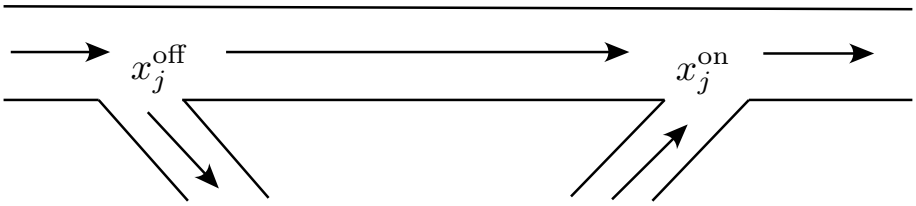


Figure 2.4: On-/off-ramps allow vehicles to enter or exit a controlled-access highway.

Let $\{x_i^{\text{on}}\}_{i\in\mathcal{I}}$ and $\{x_j^{\text{off}}\}_{j\in\mathcal{J}}$ be the positions of on-ramps and off-ramps, and $\{\varphi_i^{\text{on}}(t)\}_{i\in\mathcal{I}}$ and $\{\varphi_j^{\text{off}}(t)\}_{j\in\mathcal{J}}$ are time-dependent flows of on-ramps and off-ramps respectively. The ramp flow is positive for on-ramps, and negative for off-ramps. Instead of using a delta function $\delta(x - x_i^{\text{on}})$ and $\delta(x - x_i^{\text{off}})$ to indicate the locations of sources sinks, we introduce sech functions $f(x - x_i^{\text{on}}) = \text{sech}(x - x_i^{\text{on}})$ and $f(x-x_i^{\text{off}}) = \text{sech}(x-x_i^{\text{off}})$ to represent the spread effect of ramps. This consideration comes from the phenomenon that ramps would affect the areas nearby rather the specific points.

By addition source and sink terms to the viscous LWR equation (2.10), we get the model for on-/off-ramps scenario:

$$\rho_t + (\rho V_G(\rho))_x = \varepsilon\rho_{xx} + \sum_{i\in\mathcal{I}} \varphi_i^{\text{on}}(t) f(x - x_i^{\text{on}}) + \sum_{j\in\mathcal{J}} \varphi_j^{\text{off}}(t) f(x - x_j^{\text{off}}), \tag{2.11}$$

In this scenario, spatial shocks are generated around the on-ramp and off-ramp. Specifically, a density bump appears around the on-ramp, while a density valley appears around the off-ramp. More details about the ramp positions and ramp flows will be mentioned in § 5 when estimating traffic states in on-/off-ramps scenario.

2.3.3 Traffic lights

The traffic lights on the road force vehicles to decelerate and accelerate periodically, resulting a periodic oscillations of the vehicle density along the road. In this circumstance the velocity function is dynamic corresponding to the color of traffic lights.

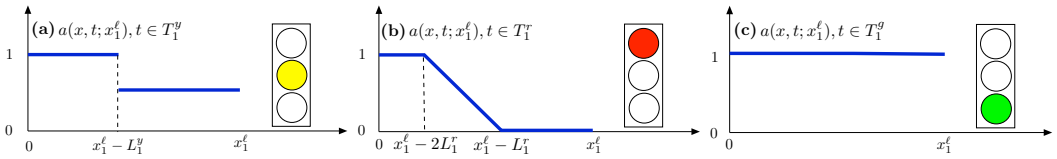
In the traffic light scenario, we place a traffic light at position x_1^ℓ on the road, and set three sub-periods for yellow light T_1^y, red light T_i^r and green light T_1^g. We introduce a deceleration factor to account for the effects from traffic light, and the velocity function becomes:

$$v(x,t) = V_G(\rho(x,t))a(x,t;x_1^\ell) \tag{2.12}$$

where $a(x,t;x_1^\ell)$ is the deceleration factor from a traffic light at position x_1^ℓ. The function $a(x,t;x_1^\ell)$ is periodic in time t with period T, which is the total length of three sub-periods $T = T_i^y + T_i^r + T_i^g$. In Figure 2.5, The traffic light cycle has a period of $T = 6$ minutes, which consists of yellow light time T_1^y, red light time T_1^r and green light time T_1^g as follows:

- During yellow lights, drivers within L_1^y miles (see Figure 2.5(a)) notice the yellow lights, and they are decelerating. The drivers out of the range will drive normally.
- During red lights, cars within L_1^r miles (see Figure 2.5(b)) are forced to stop, and those following within $2L_1^r$ are decelerating when approaching the traffic light.
- During green lights T_1^g (see Figure 2.5(c)), traffic flows normally and the ve-

locity is exactly equal to the predetermined velocity function.



(a) Yellow light period T_1^y **(b)** Red light period T_1^r **(c)** Green light period T_1^g

Figure 2.5: Deceleration factor $a(x,t;x_1^\ell)$ in traffic light scenario: (a) yellow light period; (b) Red light period; (c) Green light period.

The modified LWR model for traffic lights is mathematically given by:

$$\rho_t + \left(\rho V_G\left(\rho(x,t)\right) a(x,t;x_1^\ell)\right)_x = \varepsilon \rho_{xx}, \tag{2.13}$$

$$a(x,t;x_1^\ell) = \begin{cases} 0.5 & (x,t) \in (x_1^\ell - L_1^y, x_1^\ell) \times T_1^y \\ 0 & (x,t) \in (x_1^\ell - L_1^r, x_1^\ell) \times T_1^r \\ (x_1^\ell - L_1^r - x)/L_1^r & (x,t) \in (x_1^\ell - 2L_1^r, x_1^\ell - L_1^r) \times T_1^r \\ 1 & \text{otherwise,} \end{cases} \tag{2.14}$$

More details about the length of lights (T_1^y, T_1^r, T_1^g) and the effect region (L_1^y, L_1^r) will be mentioned in § 5 when estimating traffic states in traffic light scenario.

2.3.4 Construction zone

When approaching a construction zone, we can see signs or signals like "SLOW", "ROAD WORK AHEAD", "SINGLE LANE AHEAD" or even "STOP". Around the construction zone, typically the drivers are experiences slow speed and lane merge. In this circumstance, the velocity function is dynamic corresponding to effects of construction zone.

Let x_1^b be the location of a construction zone, which is a constant. We introduce

a bottleneck factor $a(x - x_1^b)$ to account for the slow-down effects from construction zone. Then the velocity function becomes:

$$\rho_t + \left(\rho V_G(\rho) a(x - x_1^b)\right)_x = \varepsilon \rho_{xx}, \tag{2.15}$$

where x_1^b is the location of the bottleneck. The quantity $a(x - x_1^b)$ is the bottleneck factor, which can be written as the product of severity coefficient and spread effect $a(x - x_1^b) = cf(x - x_1^b)$. The severity coefficient c reflects the degree of influence of the construction zone or traffic accident, and $f(x - x_1^b)$ represents the spread effect of the bottleneck. In this scenario, a stationary density bottleneck appears around x_1^b.

More details about the bottleneck position x_1^b, severity coefficient c and the spread effect $f(x - x_1^b)$ will be mentioned in § 5 when estimating traffic states in construction zone.

2.3.5 Traveling bottleneck

We also construct a traveling bottleneck scenario, for which we use a model similar to that of the stationary bottleneck in §2.3.4. To account for a traveling bottleneck, we allow the position of the bottleneck to be time-dependent, and the macroscopic model can be described as:

$$\rho_t + \left(\rho V_G(\rho) a(x - x_1^b(t))\right)_x = \varepsilon \rho_{xx}, \tag{2.16}$$

where $x_1^b(t)$ is the location of traveling bottleneck, which is assumed to be time-dependent.

The traveling bottleneck scenario can be used to model slow truck, moving construction zone, moving snowplows and so on. In this traffic scenario, we would like test the assimilation efficacy of our algorithm in time-dependent parameter $x_1^b(t)$.

More details about the traveling bottleneck position $x_1^b(t)$ will be mentioned in § 5 when estimating traffic states and parameters.

2. Parameter estimation for real traffic data In addition to estimating simulated traffic states and parameters, we are interested in applying the developed approach to real traffic data. In this section, we use data from the Minnesota Department of Transportation [70].

Chapter Three

Review: microscopic traffic models

3.1 Introduction to microscopic traffic models

A microscopic traffic model is an import class of model in traffic simulation. In contrast to macroscopic traffic model, microscopic model simulates individual vehicle-driver entities, so the dynamic variables of the models represent microscopic properties like the position and velocity of individual vehicles. Conventionally, the microscopic models can be used to derive the macroscopic models through micro-macro transition.

The mathematical formulations include car-following models and cellular automaton models.

- The car-following models are time-continuous models, which are defined by ordinary differential equations describing the vehicles' positions and velocities. It is assumed the accelerations of the drivers depend only on their own velocities, the velocity of the leading vehicle, and the distance to the leading vehicle. Usually the driving behavior can be affected by multiple leading vehicles. Then the acceleration function can be generalized to account for more variables. Well-known car-following models include the optimal velocity model [8], the velocity difference model [46], the GHR model [32] and the intelligent driver model [72].

- The cellular automaton models describe traffic dynamics in a completely discrete way: space is subdivided into cells, time into time steps, and speed or acceleration are integer multiples of the corresponding units. Each road section can either be occupied by a vehicle or empty. Examples of cellular automaton models include Nagel-Schreckenberg model[57], and other refined models.

Compared to the car-following models, the advantages and disadvantages of the cellular automaton models are due to the discrete scaling of time, space and state variables. Even though the cellular automaton models are easy and fast to simulate, they lack robustness because of the discrete nature. Therefore, we will focus on car-following models in our study.

Traffic simulation is an interesting and important application of car-following models. Car-following models, together with lane-changing models [61] are developed and implemented in a microscopic simulation framework. In [59], a comparative evaluation of car-following behavior in a number of well-known traffic simulators, including AIMSUN (advanced interactive microscopic simulator for urban and nonurban networks), PARAMICS (parallel microscopic simulation) and VISSIM (Verkehr in Stadten simulation).

The dynamics of car-following models is another important area for investigation. In Chapter 4, we will discuss the stability and collision behavior of car-following models.

3.2 Car-following models

Microscopic models describe traffic flow dynamics in terms of single vehicles. As the most important representatives of microscopic traffic flow models, car-following models describe traffic dynamics from the perspective of individual driver-vehicle units with interaction to the leading ones.

In continuous-time models, these car-following models are defined by coupled ordinary differential equations, and the drivers' responses are informed by their own

velocities, headways and the velocities of the leading vehicles. Many car-following models are of the form:

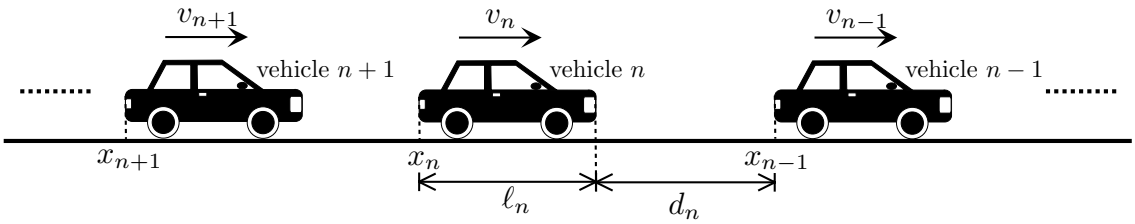


Figure 3.1: Notations for car-following models: vehicles are numbered from front to back such that vehicle n follows vehicle $n-1$. For vehicle n, x_n denotes position, v_n denotes velocity, ℓ_n denotes length and d_n denotes gap.

$$\begin{aligned} \dot{x}_n(t) &= \frac{dx_n(t)}{dt} = v_n(t), \\ \dot{v}_n(t) &= \frac{dv_n(t)}{dt} = f_{\text{accel}}\left(v_n, v_{n-1}, d_n\right), \quad d_n = x_{n-1} - x_n - \ell_n \end{aligned} \tag{3.1}$$

where

x_n : the position of vehicle n,

v_n : the speed of vehicle n,

v_{n-1} : the speed of the leading vehicle $n-1$,

ℓ_n : the length of vehicle n,

d_n : the gap between vehicle n and the leading vehicle $n-1$,

f_{accel} : the acceleration function.

In recent years, car-following models have received increasing attentions both in mathematics and engineering. Many car-following models have been developed and studied in the past half century. In [14, 17, 39], the dynamical behavior and properties of car-following models have been reviewed, including the formula, parameters, steady state, stability, trajectories, etc.

In this section, we will focus on four well-known continuous car-following models, which are frequently used in traffic simulations. These models are (i) the optimal velocity model, (ii) the full velocity difference model, (iii) the modified Gazis-Herman-Rothery model, and (iv) the intelligent driver model.

3.2.1 The optimal velocity model

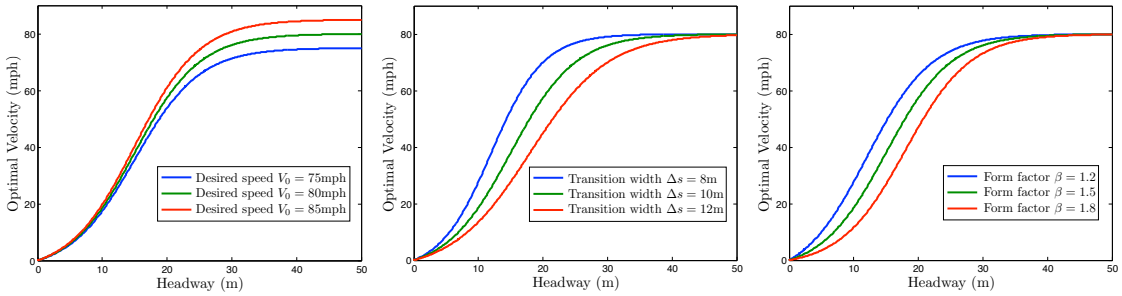
The optimal velocity model was originally proposed by Bando et al. in [8], which assumed that each vehicle has a distance-dependent optimal velocity $\mathcal{V}(\cdot)$. In this model, the driver controls the acceleration to maintain an optimal velocity according to the motion of the leading vehicle. In addition, a constant adaption time τ is introduced to reflect the driver's sensitivity to adapt to the optimal velocity. The optimal velocity model can be written as:

$$\begin{aligned} \dot{x}_n(t) &= v_n(t) \\ \dot{v}_n(t) &= \frac{\mathcal{V}(d_n) - v_n}{\tau}, \end{aligned} \tag{3.2}$$

where d_n is the gap to the leading vehicle, $\mathcal{V}(\cdot)$ is the optimal velocity function that depends on the distance to the leading vehicle, and τ is the adaption time. The acceleration equation describes the adaption of actual speed v_n to the optimal velocity $\mathcal{V}(d_n)$ on a time scale given by the adaption time τ. The smaller the adaption time τ is, the faster the vehicle adapts to the optimal velocity. The optimal velocity function $\mathcal{V}(\cdot)$ is a modeling choice: common choices are hyperbolic tangent and piecewise function. In this study, we will use the hyperbolic tangent function by Bando et al.:

$$\mathcal{V}(d_n) = V_0 \frac{\tanh(d_n/\Delta s - \beta) + \tanh\beta}{1 + \tanh\beta}$$

where V_0 is the desired speed, Δs is the transition width and β is the form factor. Figure 3.2 shows how these parameters change the shape of optimal velocity function $\mathcal{V}(\cdot)$. The bigger the desired speed V_0 is or the smaller the transition width Δs is, the greater the optimal velocity will be. For the form factor β, larger values shift the optimal velocity curve to right. Overall the optimal velocity model is easy to understand and implement, but it has a strong dependency on the parameters and simulation results can be unrealistic.



(a) Desired speed V_0 **(b)** Transition width Δs **(c)** Form factor β

Figure 3.2: Optimal velocity function in respect to parameter V_0, Δs and β

3.2.2 The velocity difference model

The main deficiency of optimal velocity model is that it does not take the speed of the leading vehicle into consideration. Thus the full velocity mode was proposed in [46] by extending the optimal velocity model with an additional linear stimulus for the speed difference. On one hand, if the speed of vehicle n is larger than that of the leading vehicle, additional deceleration will be generated to slow it down. On the other hand, if the following vehicle is slower than the leading vehicle, additional acceleration will be generated to speed it up. Thus, the full velocity difference model has better stability than the optimal velocity model. The mathematical formula is

as follows:

$$\begin{aligned} \dot{x}_n(t) &= v_n(t), \\ \dot{v}_n(t) &= \frac{\mathcal{V}(d_n) - v_n}{\tau} - \gamma(v_n - v_{n-1}) \end{aligned} \tag{3.3}$$

where $(v_n - v_{n-1})$ is the velocity difference to the leading vehicle $n-1$, and γ is the sensitivity factor for speed difference. Parameter γ measures the driver's sensitivity to speed difference: the greater γ is, the more sensitivity the driver has. When $\gamma = 0$, the full velocity difference model (3.3) reduces to the optimal velocity model (3.2). According to the simulations in [46], the full velocity difference model predicts the correct delay time of car motion and kinematic wave speed at jam density. However, the sensitivity term γ does not depend on the gap d_n, which can generate a significant and unrealistic deceleration even when the leading vehicle is far away.

3.2.3 The modified GHR model

The Gazis-Herman-Rothery (GHR) model is the most well-known model from the late fifties by Gazis et al. ([32]). This model is based on the intuition that the acceleration is proportional to the velocity difference; it also includes an explicit time delay to account for driver's reaction time. The general formula of the GHR model is a delay ODE:

$$\dot{v}_n(t+T) = \lambda(v_{n-1} - v_n)$$

where T is the reaction time, and λ is a relaxation rate. Here we modified the original GHR model based on the following considerations: firstly we replace the delay formula by the adaption to the optimal velocity with reaction time τ; then we

use the reciprocal spacing for the relaxation rate, i.e., $\lambda = \frac{\eta}{d_n}$; lastly the sensitivity $-\eta\frac{v_n - v_{n-1}}{d_n}$ to velocity difference is bounded from above in order to avoid unrealistic acceleration generated by negative speed difference $v_n - v_{n-1}$ and small gap d_n. With these sensible modifications, the stimulus is able to characterize the distance-dependent sensitivity to velocity difference as well as realistic acceleration. The modified formula can be written as:

$$\begin{aligned} \dot{x}_n(t) &= v_n(t), \\ \dot{v}_n(t) &= \frac{\mathcal{V}(d_n) - v_n}{\tau} + \min\left(-\eta\frac{v_n - v_{n-1}}{d_n}, A\right), \end{aligned} \tag{3.4}$$

where η is a constant sensitivity parameter, and A is the upper bound of acceleration. The modified GHR model is an improvement over the optimal velocity model and the full velocity difference model, since it generates distance-dependence stimulus to velocity difference.

3.2.4 The intelligent driver model

The intelligent driver model is a novel continuous model by considering the real driving behavior, which was proposed by Treiber et al. in the work [72]. In this model, the driving behaviors such as keeping a "safe distance", driving at a desired speed and preferring comfortable accelerations are taken into consideration. The main idea behind this model is to superpose (I) the acceleration tendency on a free road by comparing the current velocity to the desired velocity, and (II) the deceleration tendency in the presence of interactions by comparing the current gap

to the desired safe distance. The mathematical formula is as follows:

$$\begin{aligned} \dot{x}_n(t) &= v_n(t), \\ \dot{v}_n(t) &= \underbrace{a\left[1-\left(\frac{v_n}{V_0}\right)^{\delta}\right]}_{\text{I}} + \underbrace{\left[-a\left(\frac{d^*(v_n, v_n - v_{n-1})}{d_n}\right)^2\right]}_{\text{II}}, \end{aligned} \tag{3.5}$$

where a is the comfortable acceleration, V_0 is the desired velocity, δ is the acceleration exponent and d^* is the desired distance. In model (3.5), term I represents the acceleration tendency with exponent δ controlling the reduction of acceleration when velocity approaches the desired velocity V_0, and term II represents deceleration tendency away from the leading vehicle by comparing the current gap d_n and the desired distance d^*. The form of desired distance d^* is also a modeling choice, and we use the expression from book [73] in this paper:

$$d^*(v_n, v_n - v_{n-1}) = d_0 + \max\left(0, v_n T + \frac{v_n(v_n - v_{n-1})}{2\sqrt{ab}}\right),$$

where d_0 is the minimum gap to keep, b is the comfortable deceleration, T is the reaction time of drivers. Larger minimum gap d_0, longer time gap T, or smaller deceleration b will induce larger desired distance d^*, and thus larger gap d_n for a driver to keep. The intelligent driver model is simple, and only has a few intuitive parameters with realistic values and reproduces collective dynamics.

Chapter Four

Dynamics of car-following models

4.1 Stability analysis

Instability of traffic flow results in traffic waves, which is also called stop-and-go waves. This is mainly caused by the delays in adapting the speed to the actual traffic condition, which are consequences of finite acceleration and deceleration capability.

In this section, we would like to make a stability analysis of the car-following models without using a Taylor approximation.

4.1.1 Stability analysis of car-following model

The system of N vehicles on a ring road of length L can be describe as:

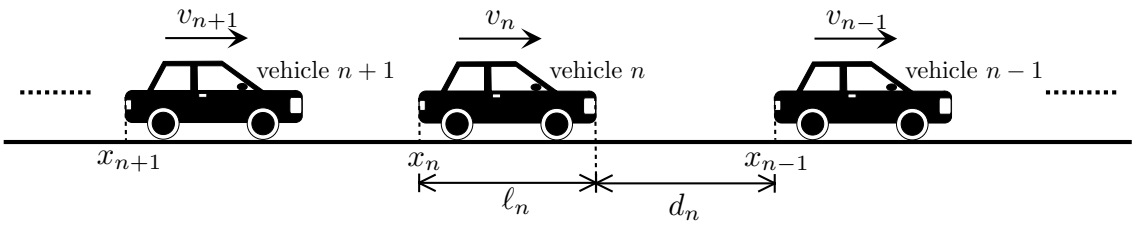


Figure 4.1: Notations for car-following models: vehicles are numbered from front to back such that vehicle n follows vehicle $n-1$. For vehicle n, x_n denotes position, v_n denotes velocity, ℓ_n denotes length and d_n denotes gap.

When studying the interactions between two adjacent vehicles, the headway rather than the vehicle length matters. Without loss of generosity, we can set the $\ell_n = 0$. By introducing the headway d_n between vehicle $n-1$ and n, the coupled ordinary differential equations become:

$$\begin{aligned} \dot{d}_n(t) &= v_{n-1}(t) - v_n(t) \\ \dot{v}_n(t) &= f_{\text{accel}}\left(v_n, v_{n-1}, d_n\right), \end{aligned} \tag{4.1}$$

The steady-state gap is $d_e = L/N$ and the steady-state velocity v_e satisfies $f_{\text{accel}}(v_e, v_e, d_e) = 0$, from which an explicit expression of v_e can be obtained as $v_e = \mathcal{V}_e(d_e)$. As in the analysis for stability, we assume small deviations y_n and u_n from steady state d_e and v_e.

$$d_n(t) = d_e + y_n(t)$$
$$v_n(t) = v_e + u_n(t)$$

(i) By substituting these in, we get odes for the deviations:

$$\begin{aligned} \frac{dy_n}{dt} &= \frac{dd_n}{dt} = u_{n-1} - u_n \\ \frac{du_n}{dt} &= \frac{dv_n}{dt} = a_v u_n + a_l u_{n-1} + a_d y_n \end{aligned} \tag{4.2}$$

where $a_v = \frac{\partial f_{\text{accel}}}{\partial v_n}|_e$, $a_l = \frac{\partial f_{\text{accel}}}{\partial v_{n-1}}|_e$ and $a_d = \frac{\partial f_{\text{accel}}}{\partial d_n}|_e$. In addition, moving along the space of steady-state solutions by simultaneously changing d_n and $v_n = v_{n-1}$must not change the acceleration f_{accel}, so we can obtain a relationship between partial derivatives in steady-state:

$$0 = a_d dd_e + (a_v + a_l)dv_e = a_d dd_e + (a_v + a_l)\mathcal{V}_e'(d_e)dd_e$$

resulting in

$$a_d = -\mathcal{V}_e'(d_e)(a_v + a_l) \tag{4.3}$$

It is assumed that $a_d > 0$, $a_v < 0$ and $a_l > 0$. Because of equation (4.3), it is easy to know that $a_v + a_l < 0$.

(ii) We can solve equation (4.2) by using Fourier-Ansatz:

$$y_n(t) = \hat{y}e^{\lambda t+ink}, \quad u_n(t) = \hat{u}e^{\lambda t+ink}$$

Inserting the ansatzs into odes (4.2) results in

$$\begin{pmatrix} \lambda & 1-e^{-ik} \\ -a_d & \lambda-(a_v+a_l e^{-ik}) \end{pmatrix} \cdot \begin{pmatrix} \hat{y} \\ \hat{u} \end{pmatrix} = 0$$

.

The resulting solvability condition assumes the form of a quadratic equation

$$\lambda^2 + p(k)\lambda + q(k) = 0 \quad \text{or} \quad \lambda_{1/2} = \frac{1}{2}\left(-p(k) \pm \sqrt{p^2(k)-4q(k)}\right)$$

where $p(k) = -a_v - a_l e^{-ik} > 0$ and $q(k) = a_d(1-e^{-ik})$.

(iii) Stability condition is equivalent to $\mathrm{Re}(\lambda_{1/2}) < 0$, i.e. $\mathrm{Re}(\sqrt{p^2(k)-4q(k)}) < \mathrm{Re}(p(k))$

Rewrite $p^2(k) - 4q(k)$ as $Re^{i\varphi} = X + iY$, then it becomes

$$\mathrm{Re}(\sqrt{p^2(k)-4q(k)}) = \sqrt{R}\cos\varphi/2 < \mathrm{Re}(p(k)) \tag{4.4}$$

with

$$\begin{aligned} &\mathrm{Re}(p(k) = -a_v - a_l\cos(-k) \\ &X = R\cos(\varphi) = (a_v + a_l\cos(-k))^2 - a_l^2\sin^2(-k) - 4a_d(1-\cos(-k)) \\ &Y = R\sin(\varphi) = 2a_l\sin(-k)(a_v + a_l\cos(-k)) + 4a_d\sin(-k) \end{aligned} \tag{4.5}$$

Take squares on inequality (4.4):

$$R\cos^2(\varphi/2) = R(1+\cos(\varphi))/2 < \mathrm{Re}(p(k))^2$$

Use $\cos\varphi = X/R$ and simplify the inequality as:

$$R < 2\mathrm{Re}(p(k))^2 - X$$

Take squares on both sides and substitute $R^2 = X^2 + Y^2$:

$$X^2 + Y^2 < (2\mathrm{Re}(p(k))^2 - X)^2$$

Remove the items X^2 on both sides:

$$Y^2 < 4\mathrm{Re}(p(k))^2 - 4X\mathrm{Re}(p(k)) \tag{4.6}$$

Next we will plug in the formula of Y and $\mathrm{Re}(p(k))$, and simplify the expression

$$\begin{aligned} &Y^2 - (4\mathrm{Re}(p(k))^2 - 4X\mathrm{Re}(p(k))) \\ &= 16a_d[a_d\sin^2(-k) + (a_v + a_l\cos(-k))(a_{v_d} - a_v)(1-\cos(-k))] < 0 \end{aligned} \tag{4.7}$$

It is assumed that $a_d > 0$, $a_v < 0$, $a_l > 0$ and $a_v + a_l < 0$. Therefore, the inequality (4.7) becomes:

$$\mathcal{V}_e'(d_e) < (a_l - a_v)\underbrace{\frac{a_v + a_l\cos(-k)}{a_v + a_l}}_{\mathrm{I}}\underbrace{\frac{1-\cos(-k)}{\sin^2(-k)}}_{\mathrm{II}}$$

Because of $a_v + a_l\cos(-k) \le a_v + a_l < 0$, the part I ≥ 1 with equality holding if $\cos(-k) = 1$. The part II is ∞ when $\cos(k) = \pm 1$, and it is $1/(1+\cos(-k))$

when $\cos(-k) \neq \pm 1$. Then we can get that II $> 1/2$ for sure.

In sum, the stability condition is

$$\mathcal{V}_e'(d_e) \leq \frac{1}{2}(a_l - a_v) \tag{4.8}$$

There are three main factors determining the stability of traffic flow: the sufficient sensitivity $\mathcal{V}_e'(d_e)$, speed sensitivity a_v and lead speed sensitivity a_l. System of greater anticipation to the lead car (bigger a_l) will favor stability, which means anticipative driver can better adjust deviations and thereby the system is stable.

(iv) the stability condition for specific models

- Optimal Velocity Model(OVM): $\mathcal{V}_e = \mathcal{V}$, $a_v = -1/\tau, a_l = 0$

$$\mathcal{V}'(d_e) < \frac{1}{2\tau}$$

- Full Velocity Difference(FVDM): $\mathcal{V}_e = \mathcal{V}$, $a_v = -1/\tau - \gamma, a_l = \gamma$

$$\mathcal{V}'(d_e) < \frac{1}{2\tau} + \gamma$$

- Modified Gazis-Herman-Potts Model(GHP): $\mathcal{V}_e = \mathcal{V}$, $a_v = -1/\tau - 1/d_e, a_l = 1/d_e$

$$\mathcal{V}'(d_e) < \frac{1}{2\tau} + \frac{1}{d_e}$$

- Intelligent Driver Model(IDM):

$$\mathcal{V}_e'(d_e) \leq \frac{(s_0 + Tv_e)(\sqrt{\frac{a}{b}}v_e + aT)}{d_e^2} + \frac{2a}{v_0}\left(\frac{v_e}{v_0}\right)^3$$

with

$$\mathcal{V}_e^{-1}(v) = \frac{s_0 + VT}{\sqrt{1-(v/v_0)^\delta}}$$
$$a_v = -\frac{4a}{v_0}\left(\frac{v_e}{v_0}\right)^3 - \frac{(s_0+Tv_e)(\sqrt{\frac{a}{b}}v_e + 2aT)}{d_e^2}$$
$$a_l = \frac{(s_0+Tv_e)\sqrt{\frac{a}{b}}v_e}{d_e^2}$$

4.2 Model calibration

In this section, the parameters of the car-following models are calibrated to the empirical traffic data. This procedure of calibration makes the car-following models not only realistic to reflect driving behaviors, but also comparable to simulate microscopic traffic.

The model parameters may fall into two types: (a) parameters that determine the macroscopic flux-density relation (fundamental diagram); (b) parameters that do not contribute the flux-density relation, but instead describe the driving habit and sensitivity: such as adaption time and comfortable acceleration.

4.2.1 Flux-density relation

Flux-density formula

Fundamental diagram of traffic flow is the plot of the traffic flux Q (the number of passing vehicles per unit time) versus the traffic density ρ (the number of vehicles

per unit length). According to [64], the two-phase traffic theory divides traffic flow into free flow for low densities, and congested flow for large densities. In the free flow regime, flux increases as density increases, while in the congested flow regime, flux decrease as density increases. Please see Figure 4.2.

Next we will induce the macroscopic flux-density relation from the microscopic car-following models through steady-state equilibrium. Firstly, the steady-state speed v_e and steady-state gap d_e satisfy the following relation:

$$f_{\text{accel}}(v_e, v_e, d_e) = 0 \Rightarrow v_e = V_e(d_e) \text{ or } d_e = D_e(v_e),$$

where D_e is the inverse function of V_e. The equilibrium formula of the car-following models are as follows: (a1) the optimal velocity model, the full velocity difference model, and the modified GHR model; (a2) the intelligent driver model

$$\text{(a1)} \quad V_e(d_e) = \mathcal{V}(d_e) = V_0 \frac{\tanh(d_n/\Delta s - \beta) + \tanh\beta}{1 + \tanh\beta}$$

$$\text{(a2)} \quad D_e(v_e) = (s_0 + v_e T)\left(1 - \left(\frac{v_e}{V_0}\right)^{\delta}\right)^{-\frac{1}{2}}$$

Secondly, a micro-macro relation between gap d_e and the macroscopic density ρ can be built as:

$$\rho = \rho(d_e) = \frac{1}{d_e + \ell},$$

where ℓ is the average length of vehicles. Thus by combining the relations above, we obtain the flux-density relation for the car-following models: (a1) the optimal velocity model, the full velocity difference model, and the modified GHR model;

(a2) the intelligent driver model

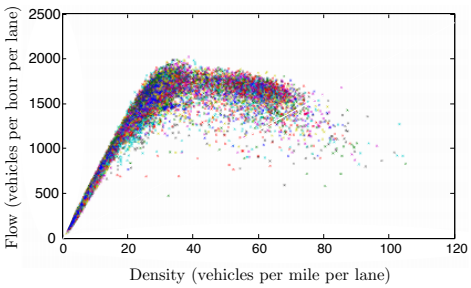
$$\begin{aligned} &\text{(a1)} \quad Q(\rho) = \rho(d_e) \times V_e(d_e) \\ &\text{(a2)} \quad Q(\rho) = \rho(d_e) \times {D_e}^{-1}(d_e) \end{aligned} \tag{4.9}$$

Flux-density data

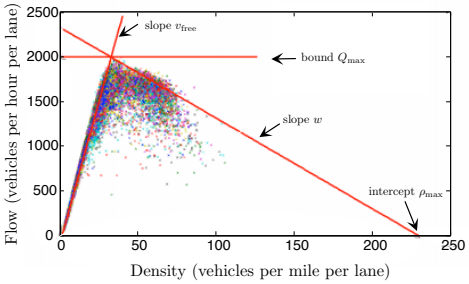
We now apply the formula above to calibrating the car-following models using data that is provided by California Department of Transportation. The traffic data is collected by the Caltrans Performance Measurement System (PeMS) [69], which provides flow, speed and occupancy data across the vehicle detector stations (vds) in the form of time series over days of operation. These sensors span the the freeway system across all major metropolitan areas of the State of California.

We recall the results from [24], which used data from a 28 mile stretch of Freeway I-880S (between 29th Avenue on-ramp in Oakland and the Auto Mall Parkway on-ramp in Fremont) [24]. The data is aggregated over time intervals of 5 minutes over 98 days that were identified as functioning over 80% between February 2007 and March 2008. The scatter plot of flux-density (density data is computed based on flow and occupancy) data is shown in Figure 4.2(a).

The paper [24] presented a method for automated and empirical calibration of freeway traffic flow characters. The parameters for flux-density relation are shown in Figure 4.2(b) and Table 4.1, including free-flow speed v_{free}, maximal flux $Q_{\max}$, congestion speed ω and jam density $\rho_{\max}$.



(a) Data for fundamental diagram



(b) Estimated parameters using fundamental diagram

Figure 4.2: Automatic calibration of the fundamental diagram from [24]. (a) The scatter of flux-density data collected by PeMS. (b) The estimation of main parameters by regression.

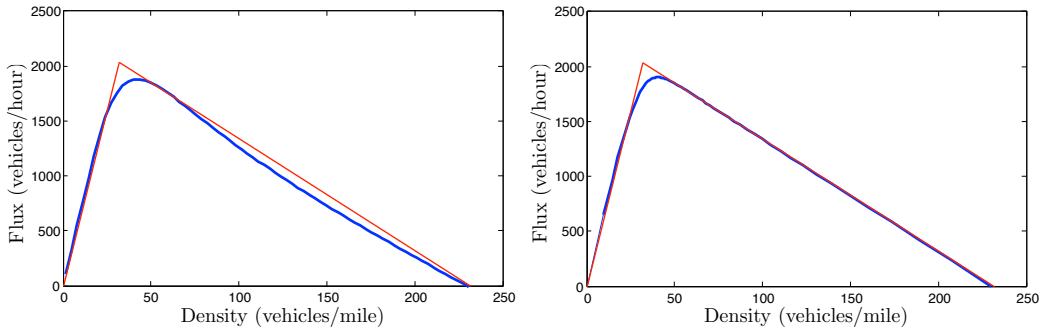
Description	Parameter	Value	Unit
Free-flow speed	v_{free}	63.3	miles/hour
Maximal flux	$Q_{\max}$	2031	vehicles/hour
Congestion speed parameter	ω	10.1	miles/hour
Jam density	$\rho_{\max}$	232	vehicles/mile

Table 4.1: Parameters from auto-calibration of the fundamental diagram

4.2.2 Model calibration

All parameters for the car-following models in § 3.2 consist of: (a) the parameters that determine the flux-density relation; (b) the parameters that describe the driving habit and sensitivity. For type (a), the parameters are estimated by tuning the flux-density curve to mimic the desired triangle shape in Figure 4.2. The flux-density relations produced by the car-following models with the calibrated parameters are shown in Figure 4.3.

For type (b), the parameters take the typical values that were used or estimated in the research work [37, 50]. In [50], the car-following behavior is studied on the basis of (publicly available) trajectory datasets recorded by a vehicle equipped with an radar sensor. The car-following models (except the modified GHR model) were calibrated by minimizing the deviation between observed dynamics and the simulated



(a) Calibration for optimal velocity model **(b)** Calibration for intelligent driver model

Figure 4.3: Calibration for flux-density relation. (a) The flux-density relation for (a1) in equation (4.9). (b) The flux-density relation for (a2) in equation (4.9). The blue solid line represents the flux-density relation for the car-following models, while the red dashed line represents the relation from PeMS traffic data.

trajectory. The remainder parameter for the modified GHR model is from [37]. The data used for calibration are the trajectories of $4,733$ vehicles provided through the FHWA Next Generation Simulation (NGSIM) project.

4.3 Collision behavior

As the main component of traffic systems, car-following models have become of increased importance in traffic simulation and safety research. In [14, 39, 59], the dynamical behaviors and properties of car-following models have been well studied and reviewed. Among these traffic dynamics, we are most interested in the collision behavior (collision-prone behavior and collision-free behavior) of car-following models, which focuses on the critical situation of potential collisions between two adjacent vehicles.

The collision behavior of traffic models are closely related to the traffic safety studies [11, 13, 75], and the collision analysis could act as a guide for microscopic simulation, especially for lane-changing models [61]. Scholars are interested in the

Description	Para	Value	Type/Ref
Optimal velocity model			
Adaption time	τ	1.5 s	(b)[50]
Desired speed	v_0	68 mph	(a)
Transition width	Δs	30 m	(a)
Form factor	β	0.5	(a)
Vehicle length	ℓ	7 m	(a)
Others			
Sensitivity factor	γ	0.65 s^{-1}	(b)[50]
Sensitivity parameter	η	12 m/s	(b)[37]
Acceleration bound	A	3 m/s^2	(b)[50]
Intelligent driver model			
Desired velocity	v_0	69 mph	(a)
Acceleration	a	1.5m/s^2	(b)[50]
Acceleration exponent	δ	8	(a)
Minimum gap	s_0	2 m	(a)
Safe time gap	T	1.5 s	(a)
Deceleration	b	3 m/s^2	(b)[50]
Vehicle length	ℓ	5 m	(a)

Table 4.2: Parameter values for car-following models. Type (a) parameters determine the flux-density relation, while type (b) parameters describe the driving habit and sensitivity terms.

drivers for collision-free behavior, which could be used to develop the sensitivity terms in new traffic models [18]. In addition, artificial intelligence, as the ultimate goal of traffic studies, would make a good use of these collision-free behavior to enhance the cooperative collision avoidance in highways [12]. However, evaluating the collision behavior of different car-following models is a complex and crucial task, which has to determine the effects of specific components and parameters. Thus, most assertions of the collision behaviors are based on the microscopic simulations rather than the theoretical analysis.

In this section, we would like to carry out a theoretical analysis of the collision behavior of four well-known continuous car-following models: the optimal velocity model [8], the full velocity difference model [46], the modified GHR model [32] and the intelligent driver model [72]. There are two main goals of the work: (1) identify whether the car-following models are collision-prone or collision-free with theoretical

analysis; (2) simulate the collision behavior of the car-following models for a comparison between theoretical and computational results. Specifically, we will show that the optimal velocity and the full velocity difference models are collision-prone regardless of the parameters, and that the modified GHR and intelligent driver models are unconditionally collision-free regardless. In addition, our study discovers that the simulation is not valid to assert the collision behavior, since the numerical errors could introduce collisions even though the theoretical results do not support them.

4.3.1 Simulation of two-vehicle system

Collision behavior can be described as collision-prone or collision-free: collision-prone means that the vehicles generate collisions in critical scenarios, while collision-free means that the vehicles can avoid collisions appropriately in any critical scenario. In the context of the car-following models in (3.1): a model is called collision-free if collision (i.e. $x_n(t) = x_{n+1}(t)$ for two adjacent vehicles n and $n+1$ at some time t) cannot occur; a model is called collision-prone if collisions could occur under some circumstances. In this section, we provide an intuition of the collision behavior through the phase portrait analysis.

Collision simulations

Consider a simple scenario: there are two vehicles on a road. The leading vehicle is moving with a constant speed c, and the other vehicle is following it according to a car-following model. This traffic system can be simplified by introducing the gap

$d_2 = x_1 - x_2$ and the speed difference $w_2 = v_1 - v_2$:

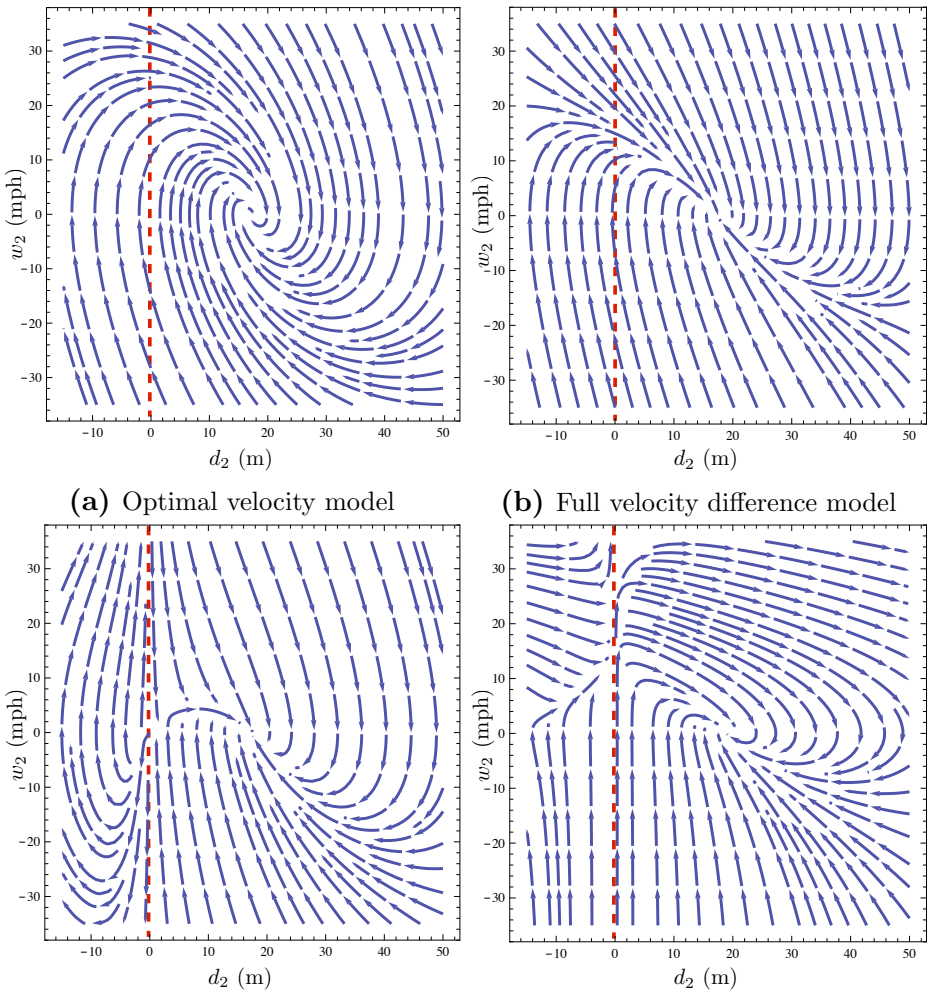
$$\begin{cases} \dot{x}_1(t) = c, \\ \dot{v}_1(t) = 0, \\ \dot{x}_2(t) = v_2(t), \\ \dot{v}_2(t) = f_{\text{accel}}(v_2, v_1, d_2), \end{cases} \quad \Longrightarrow \quad \begin{cases} \dot{d}_2(t) = w_2(t), \\ \dot{w}_2(t) = -f_{\text{accel}}(c - w_2, c, d_2), \end{cases} \tag{4.10}$$

We are interested in the behavior when the following vehicle approaches the leading one with a faster speed (i.e. $d_2(0) > 0$, $w_2(0) < 0$). If the gap $d_2(t)$ between vehicles goes negative some time, we claim that the collision-prone behavior is observed.

In our simulation, we take the parameters for the car-following models from Table 4.2, and assume that the leading vehicle moves with a constant speed $c = 10$mph. The phase portraits (d_2, w_2) of the car-following models are shown in Figure 4.4, with the collision boundary $d_2 = 0$ highlighted by the red dashed line. We can conclude:

- It is obvious to observe that the optimal velocity model and the full velocity difference model are collision-prone in this simply two-vehicle system. In the following subsection, we will justify that these two models are generally collision-prone independent of parameter settings.

- Collisions are not observed in the modified GHR model and the intelligent driver model in this simple two-vehicle system. In §4.3.2, a rigorous proof will be provided to show that these two models are collision-free in multiple-vehicle system.



(a) Optimal velocity model

(b) Full velocity difference model

(c) Modified velocity model

(d) Intelligent driver model

Figure 4.4: Phase portraits of car-following models. The x-axis represents the gap d_2 between vehicles, and the y-axis represents the speed difference w_2. The red dashed line $d_2 = 0$ represents the critical situation of collisions.

Collisions in the optimal velocity and full velocity difference models

The two-vehicle system (4.10) is specified for the optimal velocity model (with $\gamma = 0$) and the full velocity difference model:

$$\begin{cases} \dot{d}_2 = w_2, \\ \dot{w}_2 = -\frac{1}{\tau}\left(\mathcal{V}(d_2) + w_2 - c\right) - \gamma w_2. \end{cases} \tag{4.11}$$

The phase portrait of the system is sketched in Figure 4.5. The trajectory through the origin is critically collision-free ($d_2 \geq 0$, but $d_2(t) = 0$ for some time t). For the trajectories to its left, the gap d_2 will definitely go negative, and thus generating collisions. The region (d_2, w_2) that could lead to negative gap is called the collision region, which is the area to the left of the critical trajectory in Quadrant IV.

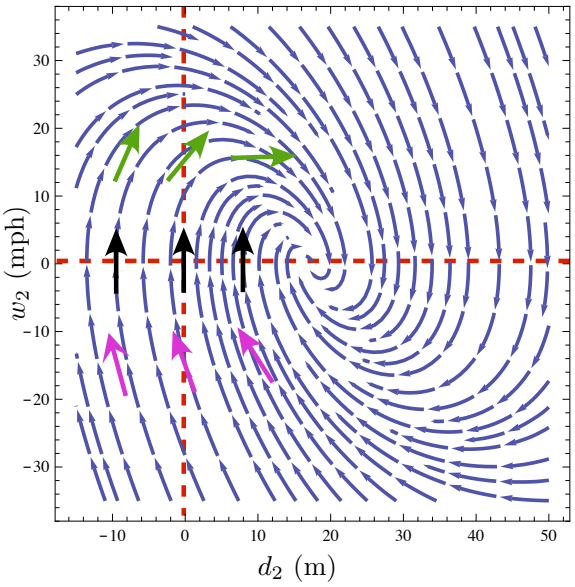


Figure 4.5: Analysis of collision-prone behavior in optimal velocity model and full velocity difference model.

Let us investigate the field direction along the y axis, i.e. the collision boundary $d_2 = 0$:

- Origin ($d_2 = 0, w_2 = 0$): the filed direction is up because of $(\dot{d}_2, \dot{w}_2) = (0, \frac{c}{\tau} > 0)$. Along the x axis with small gap $d_2 < \mathcal{V}^{-1}(c)$, the field direction is also upright. Please see the black arrows in Figure 4.5.

- Negative y axis ($d_2 = 0, w_2 < 0$): the filed direction is up-left because of $(\dot{d}_2, \dot{w}_2) = \big(w_2 < 0, \frac{c}{\tau} - (\frac{1}{\tau} + \gamma)w_2 > 0\big)$. Here the constants c, τ, γ are all positive. Thus, the gap d_2 will go negative soon and collisions occur. Please see the magenta arrows in Figure 4.5.

- Positive y axis near the origin ($d_2 = 0, w_2 > 0$): the filed direction is always to the right because of $\dot{d}_t = w_2 > 0$. Thus the two-vehicle system goes back to the normal situation from collisions. Please see the green arrows in Figure 4.5.

Based on the analysis above, the optimal velocity model and full velocity difference model are collision-prone, independently of parameter settings. However, the collision region can be reduced through appropriate parameter settings, such as the sensitivity parameter γ. By adding the speed sensitivity $-\gamma w_2$, the full velocity difference model has a much smaller collision region compared to the optimal velocity model (see Figure 4.4). This is why the full velocity model is often claimed to be collision-free in simulation [73] even though it is collision-prone in theory.

4.3.2 Proof of collision-free behavior

In this section, we aim to prove that the modified GHR model and the intelligent driver model are collision-free. The analysis approach developed in this paper can be applied to other car-following models to study their collision behavior. There are two traffic scenarios studied here: the collision-free proof for traffic on a straight road is presented first, while the last part focuses on traffic on a ring road.

The collision behavior of the car-following models can be concluded as:

Theorem 4.1. *The optimal velocity model* (3.2) *and the full velocity difference model* (3.3) *are collision-prone, which is independent of parameter settings.*

Theorem 4.2. *The modified GHR model* (3.4) *and the intelligent driver model* (3.5) *are collision-free under all traffic scenarios no matter it is a straight road or a ring road.*

Traffic on a straight road

There are n vehicles on a straight road: vehicles are labeled $1, 2, \ldots, n$ from front to back, and the leading vehicle 1 is moving at a constant speed c. As below, we define

x_i : the position of vehicle i, $\quad i = 1, \ldots, n$

v_i : the speed of vehicle i, $\quad i = 1, \ldots, n$

$d_i = x_{i-1} - x_i$: the headway of vehicle i, $\quad i = 2, \ldots, n$

$w_i = v_{i-1} - v_i$: the speed difference of vehicles i and $i-1$, $\quad i = 2, \ldots, n$

The full system of n vehicles can be expressed in terms of x_i and v_i, which can be also written in terms of d_i and w_i:

$$\begin{cases} \dot{x}_1 = c \\ \dot{v}_1 = 0 \\ \dot{x}_2 = v_2 \\ \dot{v}_2 = f_{\text{accel}}(v_2, c, d_2) \\ \cdots \end{cases} \quad \Longrightarrow \quad \begin{cases} \dot{d}_2 = w_2 \\ \dot{w}_2 = -f_{\text{accel}}(c - w_2, c, d_2) \\ \dot{d}_3 = w_3 \\ \dot{w}_3 = f_{\text{accel}}(v_2, v_1, d_2) - f_{\text{accel}}(v_3, v_2, d_3) \\ \cdots \end{cases} \tag{4.12}$$

The idea behind the proof is mathematical induction. Firstly, we show that the system of 2 vehicles is collision-free. We then show that system of j vehicles is also collision-free given the system of $j-1$ is known to be collision-free, which is the inductive step. In the proof, we only consider the critical scenario that the following vehicle is approaching the leading vehicle with a faster speed but a small gap ahead. Otherwise collisions never happen.

Modified GHR model As mentioned earlier, we focus on the critical scenario when $\underline{\text{(1a) } w_j < 0}$ and the gaps are small. Then the full system for the modified GHR model can be simplified as:

$$\begin{cases} \dot{d}_2 = w_2 \\ \dot{w}_2 = -\eta \frac{w_2}{d_2} - \frac{1}{\tau}(\mathcal{V}(d_2) + w_2 - c) \\ \dot{d}_3 = w_3 \\ \dot{w}_3 = -\eta \frac{w_3}{d_3} + \eta \frac{w_2}{d_2} + \frac{1}{\tau}(\mathcal{V}(d_2) - \mathcal{V}(d_3) - w_3) \\ \cdots \end{cases} \tag{4.13}$$

where the stimulus $-\eta \frac{w_i}{d_i}$ can provide strong repulsions when the gaps d_i are very small, and thus these terms will dominate in a critical scenario. We use the technique of fast system to analyze the modified GHR model by substituting $d_j = \varepsilon q_j$ and

$t = \varepsilon T$:

$$\begin{cases} q_2' = w_2 \\ w_2' = -\eta \frac{w_2}{q_2} - \frac{\varepsilon}{\tau}(\mathcal{V}(\varepsilon q_2) + w_2 - c) \\ q_3' = w_3 \\ w_3' = -\eta \frac{w_3}{q_3} + \eta \frac{w_2}{q_2} + \frac{\varepsilon}{\tau}(\mathcal{V}(\varepsilon q_2) - \mathcal{V}(\varepsilon q_3) - w_3) \\ \cdots \end{cases} \tag{4.14}$$

where the terms contain ε can be treated as small perturbations. When $\varepsilon = 0$, the system (4.14) becomes the fast system:

$$\begin{cases} q_2' = w_2 \\ w_2' = -\eta \frac{w_2}{q_2} \\ q_3' = w_3 \\ w_3' = -\eta \frac{w_3}{q_3} + \eta \frac{w_2}{q_2} \\ \cdots \end{cases} \tag{4.15}$$

Next we will divide the proof into two parts: firstly we would like to show that the fast system (4.15) is collision-free; then we go back to the original system (4.14), and show that the perturbations do not impact the collision-free property.

(i) The fast system could be generally written as

$$\begin{cases} Q_j' = W_j, \\ W_j' = \alpha_j \frac{W_{j-1}}{Q_{j-1}} - \beta_j \frac{W_j}{Q_j}, \quad \alpha_j \geq 0, \beta_j > 0 \end{cases} \tag{\#}$$

We start with a general form of fast system with non-negative parameters α_j and positive parameter β_j. These parameters can be specified based on the car-following models. The solution of system (#) satisfies:

$$\begin{aligned}\beta_j \frac{Q_j'}{Q_j} + W_j' &\overset{(\#)}{=} \beta_j \frac{W_j}{Q_j} + \alpha_j \frac{W_{j-1}}{Q_{j-1}} - \beta_j \frac{W_j}{Q_j} \\ &= \alpha_j \frac{W_{j-1}}{Q_{j-1}} \\ &= \alpha_j \frac{Q_{j-1}'}{Q_{j-1}}\end{aligned}$$

Taking the integral on both sides gives:

$$\frac{d}{dt}\left(\beta_j \ln Q_j + W_j\right) = \frac{d}{dt}(\alpha_j \ln Q_{j-1}), \tag{4.16}$$

Thus during the period of $\underline{\text{(1a) } w_j < 0}$, we conclude the following:

$$\begin{aligned}Q_j(T) &= Q_j(0) \left(\frac{Q_{j-1}(T)}{Q_{j-1}(0)}\right)^{\frac{\alpha_j}{\beta_j}} \exp\left\{\frac{W_j(0) - W_j(T)}{\beta_j}\right\} \\ &\geq Q_j(0) \left(\frac{Q_{j-1}(T)}{Q_{j-1}(0)}\right)^{\frac{\alpha_j}{\beta_j}} \exp\left\{\frac{W_j(0)}{\beta_j}\right\}\end{aligned} \tag{4.17}$$

Next we will show that the gaps $\{Q_j(T)\}_{j=2,3,\cdots,n}$ never go negative during the critical traffic scenario given the positive initial conditions $\{Q_j(0) > 0\}_{j=2,3,\cdots,n}$:

- For the second vehicle with $\alpha_2 = 0, \beta_2 = \eta$, the gap Q_2 satisfies:

$$Q_2(T) \geq Q_2(0) \exp\left\{\frac{W_2(0)}{\eta}\right\} = B_2$$

  which guarantees that the second vehicle never collides with the leading vehicle, and the positive lower bond is B_2.

- For the third vehicle with $\alpha_3 = \beta_3 = \eta$, the gap Q_3 satisfies:

$$Q_3(T) \geq Q_3(0)\frac{Q_2(T)}{Q_2(0)}\exp\left\{\frac{W_3(0)}{\eta}\right\} = B_3$$

  where the gap for the second vehicle $Q_2(t)$ has been proved positive for any time T no-negative. Thus, the third vehicle never collides with the second vehicle in a critical scenario.

- $\cdots$

- For the last vehicle n, we can similarly prove that it will not collide with its leading vehicle $n-1$. Thus, by using this induction method, the fast system (4.15) is surely collision-free.

(ii) Next we analyze the perturbation terms in equation (4.14) and evaluate their impact in the collision-free property. As mentioned earlier, collisions are only possible when the gaps are very small $0 < \{q_j\}_{j=2,3,\cdots,n} \ll 1$. Thus, we study the perturbation terms only when the gaps are less than the predetermined thresholds.

- For the second vehicle, the perturbation term is in bold as follows:

$$\begin{cases} q_2' = w_2 \\ w_2' = -\eta\frac{w_2}{q_2} - \frac{\varepsilon}{\tau}(\boldsymbol{\mathcal{V}(\varepsilon q_2) + w_2 - c}), \quad \boldsymbol{c \geq 0} \end{cases}$$

  The threshold for q_2 is $\underline{\text{(1b) } 0 < \varepsilon q_2 \leq \mathcal{V}^{-1}(w_2)}$, which defines a corner area in Quadrant IV enclosed by the curve $q_2 = 0, w_2 \leq 0$ and the curve $\mathcal{V}(\varepsilon q_2) + w_2 = 0$. Then the perturbation term satisfies:

$$-\frac{\varepsilon}{\tau}(\mathcal{V}(\varepsilon q_2) + w_2 - c) \overset{(1b)}{\geq} -\frac{\varepsilon}{\tau}(0 - c) \geq 0$$

Take the differential of trajectory $w_2(q_2)$:

$$\frac{dw_2}{dq_2} = \frac{w_2'}{q_2'} = -\frac{\eta}{q_2} + \frac{1}{w_2}\left(-\frac{\varepsilon}{\tau}(\mathcal{V}(\varepsilon q_2) + w_2 - c)\right) \overset{(1a)(1b)}{\leq} -\frac{\eta}{q_2} \tag{4.18}$$

The slope along the trajectory $w_2(q_2)$ is negative, which is steeper than that of the fast system (4.15). The comparison of the slopes is demonstrated in Figure 4.6. Once a trajectory $w_2(q_2)$ enters the threshold area (defined by $0 < \varepsilon q_2 \leq \mathcal{V}^{-1}(w_2)$), it will cross the trajectories of the fast system nearby to the upright direction. As proved earlier, the fast system is collision-free, and its trajectories never go across the collision boundary $d_j = 0$. Therefore, it is easy to observe that the original system is more conservative than the fast system under the critical circumstance.

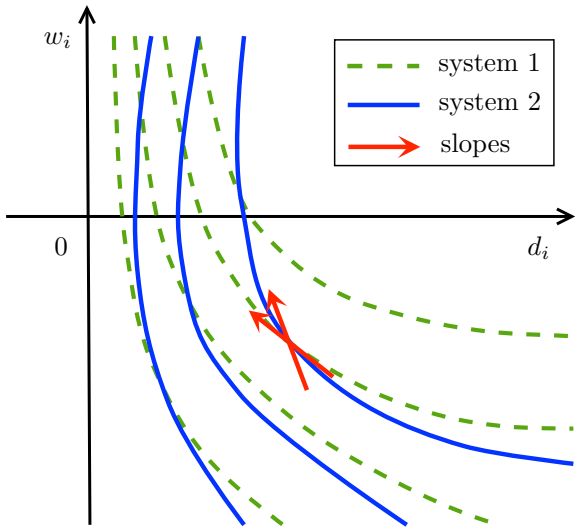


Figure 4.6: Slope comparison. System 1 (dashed green line) represents the fast system (4.15) and system 2 (solid blue) represents the original car-following system (4.14).

- For the third vehicle, the perturbation term is in bold as follows:

$$\begin{cases} q_3' = w_3 \\ w_3' = \eta\frac{w_2}{q_2} - \eta\frac{w_3}{q_3} + \frac{\varepsilon}{\tau}(\boldsymbol{\mathcal{V}(\varepsilon q_2) - \mathcal{V}(\varepsilon q_3) + w_3}) \end{cases}$$

It is known that the second vehicle never collide, i.e. $q_2(T) > 0$, so we set the threshold for gap q_3 as $\underline{(1c)\ 0 < q_3 \leq q_2}$. Otherwise, collisions can not occur. Then the perturbation term satisfies:

$$\frac{\varepsilon}{\tau}(\mathcal{V}(\varepsilon q_2) - \mathcal{V}(\varepsilon q_3) - w_3) \overset{(1c)}{\geq} \frac{\varepsilon}{\tau} w_3 \overset{(1a)}{\geq} 0$$

Take the differential of trajectory $w_3(q_3)$:

$$\begin{aligned} \frac{dw_3}{dq_3} = \frac{w_3'}{q_3'} &= \frac{\eta w_2}{w_3 q_2} - \frac{\eta}{q_3} + \frac{1}{w_3}\left(\frac{\varepsilon}{\tau}(\mathcal{V}(\varepsilon q_2) - \mathcal{V}(\varepsilon q_3) - w_3)\right) \\ &\overset{(1a)(1b)}{\leq} \frac{\eta w_2}{w_3 q_2} - \frac{\eta}{q_3} \end{aligned} \tag{4.19}$$

The slope along the trajectory $w_3(q_3)$ is negative and steeper than that of the fast system. Thus, we could infer that the third vehicle never collide based on the similar analysis and comparison of slopes in Figure 4.6.

- $\cdots$
- For the last vehicle n, we can prove the same results. Thus, the original system (4.14) of the modified GHR model is collision-free.

In summary, the modified GHR model has been proved collision-free on a straight road.

Intelligent driver model The analysis for the intelligent driver model is similar to that in the previous section. We reform the formula of the intelligent driver model under the critical traffic scenario $\underline{(1a)\ w_j < 0}$, and compare it with the proved collision-free fast system.

The full system for the intelligent driver model can be simplified as:

$$\begin{cases} \dot{d}_2 = w_2 \\ \dot{w}_2 = a\left[-1 + \left(\frac{v_2}{V_0}\right)^{\delta} + \left(\frac{s^*(v_2,-w_2)}{d_2}\right)^2\right] \\ \dot{d}_3 = w_3 \\ \dot{w}_3 = a\left[\left(\frac{v_3}{V_0}\right)^{\delta} - \left(\frac{v_2}{V_0}\right)^{\delta} + \left(\frac{s^*(v_3,-w_3)}{d_3}\right)^2 - \left(\frac{s^*(v_2,-w_2)}{d_2}\right)^2\right] \\ \dots \end{cases} \tag{4.20}$$

As mentioned earlier, collision are only possible when the gaps are very small $0 < d_j \ll 1$, so we study the system only when the gaps are less than the predetermined thresholds.

- For the second vehicle, we only consider the circumstance of small gap with (2b) $d_2 < \min(1, d_0/2)$. The acceleration of the speed difference can be bounded from below:

$$\begin{aligned} \dot{w}_2 &\overset{(4.20)}{=} a\left[-1 + \left(\frac{v_2}{V_0}\right)^{\delta} + \left(\frac{d_0 + v_2T - v_2w_2/2\sqrt{ab}}{d_2}\right)^2\right] \\ &\overset{(2b)}{\geq} a\left[-1 + \left(\frac{s_0}{2d_2} + \frac{s_0}{2d_2} + \frac{v_2T}{d_2} + \frac{v_2w_2/2\sqrt{ab}}{d_2}\right)^2\right] \\ &\overset{(2b)}{\geq} a\left[\left(\frac{s_0}{2d_2}\right)^2 + \left(\frac{v_2w_2/2\sqrt{ab}}{d_2}\right)^2\right] \\ &\overset{(2b)}{\geq} a\min\left(\frac{s_0^2}{4}, \frac{v_2^2}{4ab}\right)\frac{1+w_2^2}{d_2} \\ &\geq a\min\left(\frac{s_0^2}{4}, \frac{v_2^2}{4ab}\right)\frac{-2w_2}{d_2} \\ &= -\beta_2\frac{w_2}{d_2}, \end{aligned} \tag{4.21}$$

where the right-hand-side is a formula of fast system as in (#) by introducing

$\alpha_2 = 0$ and $\beta_2 = 2a \min\left(\frac{s_0^2}{4}, \frac{v_2^2}{4ab}\right)$. Take the differential of trajectory $w_2(d_2)$:

$$\frac{dw_2}{dd_2} = \frac{\dot{w}_2}{\dot{d}_2} \overset{(4.21)(2a)}{\leq} -\frac{\beta_2}{d_2} \tag{4.22}$$

The slope along the trajectory $w_2(d_2)$ is negative and steeper than that of the fast system. Thus the second vehicle never collides based on the analysis and comparison of slopes in Figure 4.6.

- For the third vehicle, we only consider the circumstance of small gap with (2c) $d_3 < \min(d_2, 1, d_0/2)$. The acceleration of the speed difference can be bounded from below:

$$\begin{aligned}
\dot{w}_3 &\overset{(4.20)}{=} a\left[\left(\frac{v_3}{V_0}\right)^{\delta} - \left(\frac{v_2}{V_0}\right)^{\delta} + \left(\frac{d_0 + v_3 - v_3 w_3/2\sqrt{ab}}{d_3}\right)^2 - \left(\frac{d_0 + v_2 - v_2 w_2/2\sqrt{ab}}{d_2}\right)^2\right] \\
&\overset{(2a)}{\geq} a\left[\left(\frac{d_0 + v_3 - v_3 w_3/2\sqrt{ab}}{d_3}\right)^2 - \left(\frac{d_0 + v_2 - v_2 w_2/2\sqrt{ab}}{d_2}\right)^2\right] \\
&\overset{(2c)}{\geq} 4a\left[\left(\frac{d_0 + v_3 - v_3 w_3/2\sqrt{ab}}{d_3}\right) - \left(\frac{d_0 + v_2 - v_2 w_2/2\sqrt{ab}}{d_2}\right)\right] \\
&\overset{(2c)}{\geq} 4a\left(\frac{d_0 + v_3}{d_3} - \frac{d_0 + v_2}{d_3} + \frac{-v_3 w_3/2\sqrt{ab}}{d_3} - \frac{-v_2 w_2/2\sqrt{ab}}{d_2}\right) \\
&\overset{(2a)}{\geq} 4a\left(\frac{-v_3 w_3/2\sqrt{ab}}{d_3} - \frac{-v_2 w_2/2\sqrt{ab}}{d_2}\right) \\
&= \alpha_3 \frac{w_2}{d_2} - \beta_3 \frac{w_3}{d_3}
\end{aligned} \tag{4.23}$$

where the right-hand-side is also a formula of the fast system as in (4.15) by introducing $\alpha = \frac{2v_2\sqrt{a}}{\sqrt{b}}$ and $\beta_3 = \frac{2v_3\sqrt{a}}{\sqrt{b}}$. Take the differential of trajectory

$w_3(d_3)$:

$$\frac{dw_3}{dd_3} = \frac{\dot{w}_3}{\dot{d}_3} \overset{(4.23)(2a)}{\leq} \alpha_3 \frac{w_2}{w_3 d_2} - \beta_3 \frac{1}{d_3} \tag{4.24}$$

The slope along the trajectory $w_3(d_3)$ is negative and steeper than that of the fast system. Thus the third vehicle never collide based on the analysis and comparison of slopes in Figure 4.6.

- $\cdots$
- For the last vehicle n, we can prove the same results. Thus, the original system (4.20) of the intelligent driver model is collision-free.

Traffic on a ring road

There are n vehicles on a ring road with the same notations as on a straight road. Unlike the straight road scenario, there is no leading vehicle moving at a constant c, and each vehicle can be influenced by the rest of vehicles rather than only the leading vehicles. Despite the increased complexity of the ring road scenario, similar mathematical analysis can be applied to prove the collision-free property.

The full system of n vehicles can be expressed in terms of x_i and v_i, which can

be also written in terms of d_i and w_i:

$$\begin{cases} \dot{x}_1 = v_1 \\ \dot{v}_1 = f_{\text{accel}}(v_1, v_n, d_1) \\ \dot{x}_2 = v_2 \\ \dot{v}_2 = f_{\text{accel}}(v_2, v_1, d_2) \\ \dots \end{cases} \Longrightarrow \begin{cases} \dot{d}_1 = w_1 \\ \dot{w}_1 = f_{\text{accel}}(v_n, v_{n-1}, d_n) - f_{\text{accel}}(v_1, v_n, d_1) \\ \dot{d}_2 = w_2 \\ \dot{w}_2 = f_{\text{accel}}(v_1, v_n, d_1) - f_{\text{accel}}(v_2, v_1, d_2) \\ \dots \end{cases} \tag{4.25}$$

The idea behind the proof is mathematical induction. Specifically, the proof can be divided into three steps:

(i) The vehicles are on a ring road, so the sum of the gaps is equal to the length of road L all the time:

$$d_1(t) + d_2(t) + \cdots + d_n(t) = L$$

The averaged gap is L/n. At time t_0, there is at least one vehicle k_0, whose headway is above the averaged gap L/n (Pigeonhole Principle). Since the vehicle speed is bounded by the maximal speed V_0, there must exist a time period $\Delta = L/nV_0$ such that the vehicle k_0 will not collide with its leader. Thus, vehicle k_0 is collision-free during time $[t_0, t_1]$, where $t_1 = t_0 + \Delta$.

(ii) Given a collision-free vehicle k_0 is collision-free, we intend to show that its following vehicles are all collision-free during the period $[t_0, t_1]$. Firstly, we consider its following vehicle $k_0 + 1$. For any time $t \in [t_0, t_1]$, the following vehicle $k_0 + 1$ must stay in one of the following circumstances:

(a) $w_{k_0+1} \geq 0$ or $d_{k_0+1} \geq \min(d_{k_0}, d_0/2, 1)$: there is no need to worry about collisions between vehicle k_0 and $k_0 + 1$. There is because positive speed difference w_{k_0+1} is increasing the gap d_{k_0+1}, and that collisions cannot occur with the gap d_{k_0+1} larger than the predetermined threshold.

(b) $w_{k_0+1} < 0$ and $0 < d_{k_0+1} < \min(d_{k_0}, d_0/2, 1)$: according the proof in §4.3.2, the slope of trajectory $w_{k_0+1}(d_{k_0+1})$ is negative and steeper than that of the fast system (#), which has already shown collision-free. Thus, the vehicle $k_0 + 1$ never collides based on the comparison of slopes in Figure 4.6.

By mathematical induction, we can repeat this process and show that all vehicles are collision-free during the period $[t_0, t_1]$.

(iii) By repeating step 1 and 2, we can show that the system is collision-free during $[t_1, t_2], [t_2, t_3], \cdots$. The length of each period is $\Delta = L/nV_0$

(iv) It is obvious $\Delta + \Delta + \Delta + \cdots = \infty$, so the traffic system is theoretically collision-free forever.

4.3.3 Numerical results

The main goal is to study the collision-prone and collision-free behaviors of car-following models: the optimal velocity model and the full velocity difference model are shown collision-prone independent of parameter settings through the phase portrait analysis in §4.3.1; while the modified GHR model and the intelligent driver model are shown collision-free through mathematical analysis in §4.3.2.

In this section, we carry out the traffic simulations to test the consistency of

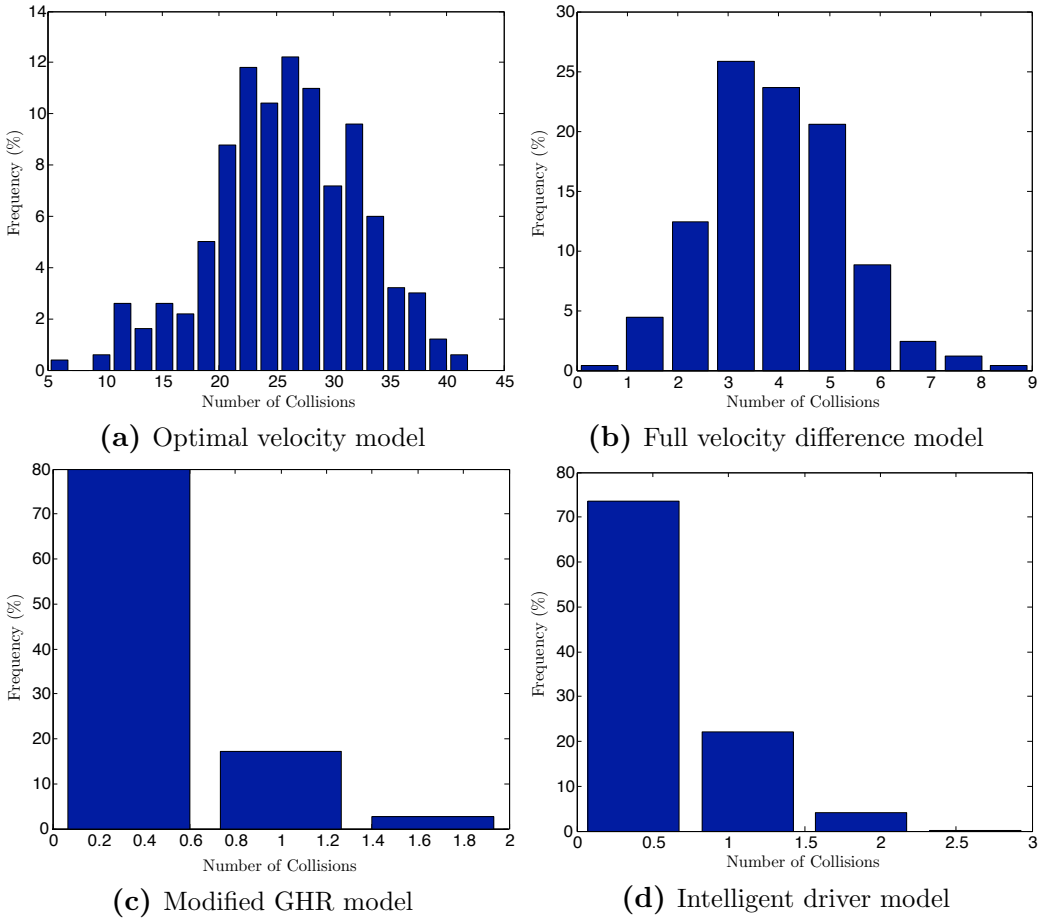
collision behaviors to the theoretical analysis. The parameter settings are based on the calibration results in §4.2. In addition, we investigate how the numerical errors potentially influence the collision behaviors, and discover that the numerical errors can introduce collisions even though the model does not support them.

Monte Carlo simulation: counts of collisions

We put 28 vehicles on a ring road of 350m length: the initial positions and velocities are randomly picked according to the uniform distribution; with the same initial conditions, the vehicles will move 20 seconds based on the four car-following models. In order to make the results representative of the collision behaviors, we implement the Monte Carlo simulation with 500 iterations. We count the number of collisions by analyzing the vehicle trajectories, and the resulting distribution is in Figure 4.7.

- The averaged number of collisions in the optimal velocity model reaches as many as 25.85, while it is only 3.87 for the full velocity difference model. This is consistent to our phase portrait analysis in §4.3.1. For the full velocity difference model, it is uncommon to observe collisions in simulation if the initial conditions are not extremely critical. This is why this model is often claimed to be collision-free in simulation [73].

- There are a few collisions in the modified GHR model and the intelligent driver model, even though they are collision-free theoretically. This is because the numerical errors bring collisions that are not supposed to exist. Thus, the collision behaviors in simulation cannot used to conclude whether a car-following model is collision-prone or collision-free.



(a) Optimal velocity model

(b) Full velocity difference model

(c) Modified GHR model

(d) Intelligent driver model

Figure 4.7: Trajectories of car-following models. The x-axis represents the bins of collisions observed, and the y-axis represents the counts of collision in percentage.

Why can collisions occur in simulations of collision-free models?

It is an interesting phenomena that collisions still numerically exist in the theoretically collision-free car-following models. An intuition for this comes from the phase portraits in Figure 4.4: when the trajectory gets close to the y-axis ($d_i = 0$), it is easy to be pushed across the y-axis by numerical errors. Therefore, we are interested in studying whether the collisions in the simulations of collision-free car-following models could diminish or even disappear with finer meshes of time.

Monte Carlo simulation The basic setting is the same: there are28 vehicles on a ring road of 350m length with randomly picked initial conditions. We run the numerical scheme for different time step h, and repeat this process with 4000 iterations for different initial conditions. The meshes of time are getting finer and finer according to the following table:

Total time	T	20s	20s	20s	$\cdots$	20s
Mesh points	n	400	800	1200	$\cdots$	6000
Time step	$h = T/n$	0.0500s	0.0250s	0.0167s	$\cdots$	0.0033s

Table 4.3: Setting of time steps in Monte Carlo simulation

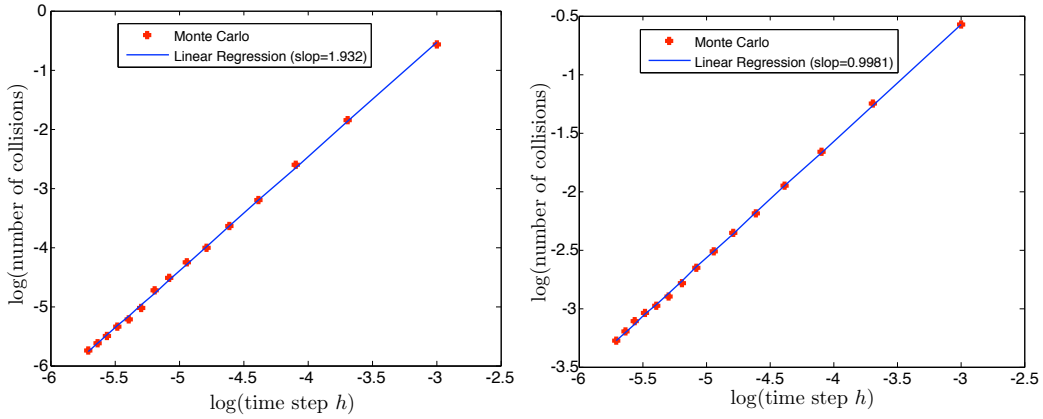
There are two ways to record the trajectories of vehicles, and thus there are two ways to count collisions respectively:

(i) Update the positions of vehicles every h time, but only record the trajectories at the fixed time points $[0.05s, 0.10s, 0.15s, \cdots, 20s]$, i.e. the recording time interval is $\Delta T = 0.5s$. Then the number of collisions are counted based on the trajectories that are linearly interpolated among these 400 fixed time points.

(ii) Update the positions of vehicles every h time, and record the trajectories every

time step until 20s $[h, 2h, 3h, \cdots, nh = 20s]$, i.e. the recording time interval is $\Delta T = h$. Then the number of collisions are counted based on the trajectories that are linearly interpolated among these $n = 20/h$ time points.

The main difference between the two methods above is the time interval ΔT for recording trajectories and measuring collisions. The intuition tells that the more frequent we check vehicles' positions, the more collisions we observe. The Monte Carlo simulations are implemented in the modified GHR model with Euler explicit scheme, and the results are shown in Figure 4.8.



(a) log-log plot with $\Delta T = 0.05$ **(b)** log-log plot with $\Delta T = h$

Figure 4.8: The relationship of collision counts over time step h. The red dots represent the simulation results, and the blue line is a fitted line from linear regression.

The simulation results are valid to show that the collisions caused by the numerical errors will diminish with smaller time step h. Thus, by choosing a sufficiently small time step h in simulations will effectively avoid these unrealistic collisions from the collision-free car-following models. In Figure 4.8, we also observe that the number of collisions is approximately proportional to the exponentiation of time step h: for the first counting method with $\Delta T = 0.05s$, Counts $\propto \mathcal{O}(h^2)$, and for the second counting method with $\Delta T = h$, Counts $\propto \mathcal{O}(h^1)$.

Even thought it is hard to figure out how the collisions are related to numerical

errors, we guess that they are proportional to the truncation error of the Euler explicit scheme. The local truncation error of Euler explicit scheme is h^2, which coincides with the first counting method, and the global truncation error is h^1, which also coincides with the second counting method. An descriptive explanation might be: for a numerical scheme with accuracy $\mathcal{O}(h^p)$, the local truncation error is proportional to $\mathcal{O}(h^{p+1})$), which directly affects the collision behaviors in one time step h. Then by checking M points in the vehicle trajectories, we could observe $M\mathcal{O}(h^{p+1})$ collisions in the simulations. For the first counting method with time interval $\Delta T = 0.05$, we check $M = 400$ equally distributed points along the trajectories, and thus observe around $400\mathcal{O}(h^{p+1}) = \mathcal{O}(h^{p+1})$ in simulations. For the second counting method with time interval $\Delta T = h$, we check $n = 20/h$ equally distributed points along the trajectories, and thus observe around $20/h\mathcal{O}(h^{p+1}) = \mathcal{O}(h^p)$.

Investigation for Euler explicit scheme In the previous part, we come up with a guess that the number of observed collisions is related to the accuracy of the numerical scheme. The general proof should be complicated and beyond the scope of this paper. In this part, we would like to provide a brief discussion about Euler explicit scheme in a simplified traffic situation. Here we only consider the second counting method with a time interval $\Delta T = h$, and show that the number of collisions is proportional to $\mathcal{O}(h^1)$.

We only put two vehicles on a straight road with the leading vehicle moves at a constant speed c, and study what (d, w) values can generate collisions. In order to have a collision within a time step h, the following relation $d(t) + w(t)h < 0$ must hold. This relation sketches a triangle region, and a collision will definitely occur within one time step h if current (d, w)) falls into it (red checks in Figure 4.9). In addition, any initial conditions $(d(0), w(0))$ whose trajectory interacts with this

triangle area will definitely generate a collision, too. Thus, the (d, w) values that generates collisions are in the triangle region plus the supplementary region (red diamonds in Figure 4.9).

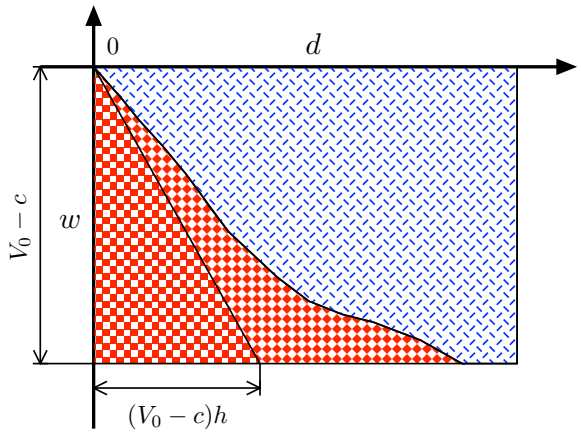


Figure 4.9: Region partition of initial condition (d, w): the red checks and diamonds area is collision-prone; the blue lines area is collision-free.

It is observed that the area of this (d, w) region depends on the picked time step h, so we could study how the collisions diminishes by studying how the region area shrink when time step h decreases. The regions are sketched in red for different time step h in Figure 4.10. By linear regression, the region area is proportional to $\mathcal{O}(h^1)$.

4.3.4 Discussion

We aim to study the collision behavior of the car-following models. Specifically, we studied four well-known continuous car-following models: the optimal velocity model, the full velocity difference model, the modified GHR model, and the intelligent driver model. The parameter setting is based on the model calibration using historical traffic data. The optimal velocity model and the full velocity difference model has been shown collision-prone independent of parameter settings through the phase portrait analysis in §4.3.1, while the modified GHR and the intelligent driver model

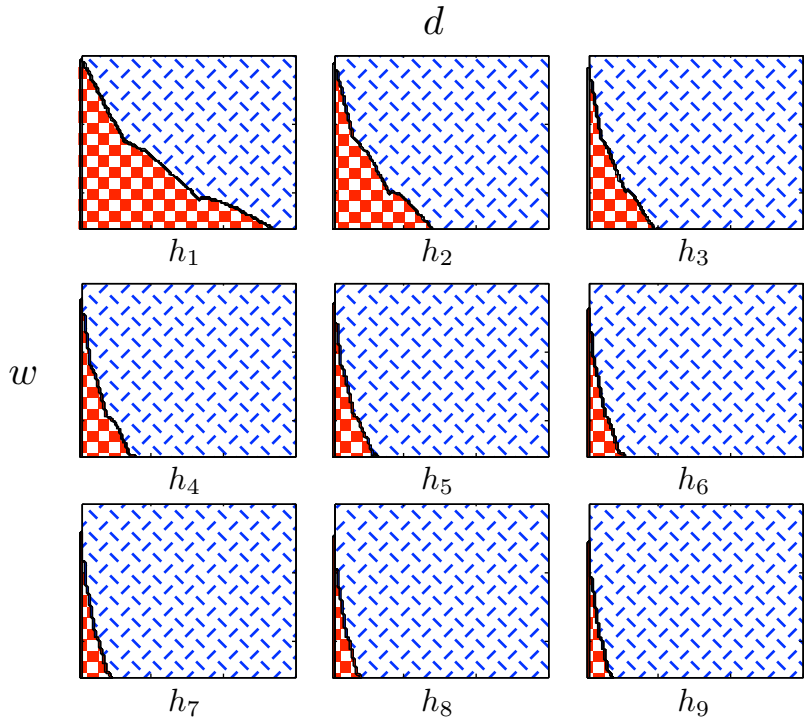


Figure 4.10: (d, w) collision region for different time step h ($h_1 : h_2 : \cdots : h_8 : h_9 = 1 : 2 : \cdots : 8 : 9$).

are proved to be collision-free through mathematical analysis in §4.3.2.

In addition to the theoretical analysis of the collision property, we carry out experiments to simulate the trajectories of the car-following models. The simulation results are consistent with our theoretical analysis except that numerical errors introduce collisions that the model does not support. More simulations are implemented to show that the number of the observed collisions is proportional to the truncation error of the numerical scheme.

The study of collision behavior provides a theoretical reference for the car-following models. The collision behavior will be taken into consideration when designing new models or studying artificial intelligence. In the future work, we will combine these collision-free models with lane-changing model to construct a complete traffic system.

Chapter Five

Data assimilation for traffic flow estimation

5.1 Introduction

Estimating the traffic state on a network of highway is a challenging problem. With the availability of stationary sensor data, traffic cameras, and GPS cell phone data, more and more data can now be used to predict car density, velocity and travel time in real time. Predicting traffic involves (i) a model that provides the traffic at later times based on given initial condition and (ii) techniques to incorporate and assimilate actual traffic-state observations into the initial conditions to improve the model forecast. To formulate a framework for estimating traffic flow, we therefore need a mathematical model of traffic flow, a sense of what types of observations are available, and a scheme for assimilating these observations into the model. We now discuss these three ingredients in turn.

Mathematical models for traffic flow come in many different flavors. Common models range from cellular automata or microscopic car-following models for individual cars to macroscopic discrete or partial-differential equation (PDE) models for car densities, velocities, and possibly other quantities [39, 54, 63]. Knowledge of the positions and velocities of all cars on a road or road network is required to successfully employ a microscopic model. Since most of this information is not well known, it is typically not feasible to use such models for traffic flow prediction. Thus, macroscopic models are used more frequently for traffic estimation. A commonly used minimal macroscopic PDE model is the Lighthill–Whitham equation [54]. This model is discussed in more detail in §2.

Traffic-flow observations come from a variety of sources and are available as functions of time. Stationary sensors, such as induction loops or cameras, provide the flux, average velocity, and local density of cars that move past the fixed sensor

location. GPS data from cell phones or navigation devices, on the other hand, provide information about the positions and velocities of individual cars that move with the traffic flow. We refer to observations that come from a fixed observation location as Eulerian observations and to observations that come from parcels (cars) that move with the traffic flow as Lagrangian observations.

There are a number of different data-assimilation techniques available that help incorporate traffic observations at discrete times into an underlying model. Most of these are based on a Bayesian statistics analysis that treats the forecast from the model as the prior distribution and then calculates a posterior distribution based on the available observations. Commonly, traffic estimators employ ensemble Kalman filters for this calculation. The main limitation of the existing data-assimilation techniques stems from the fact that sensor and GPS data require very different assimilation approaches: it is therefore technically involved to use existing techniques to assimilate Eulerian and Lagrangian observations simultaneously.

The main contribution of our research is to propose a method that provides an efficient alternative approach to the assimilation of Lagrangian GPS data, whilst keeping the ability to simultaneously assimilate Eulerian sensor data. In this part, we will discuss the efficacy and accuracy of the proposed method for assimilating sensor and GPS data in traffic flow. Specifically, we will demonstrate that the proposed algorithm is robust with respect to different traffic scenarios and its implementation using ensemble Kalman filter or particle filter. Thus, our method provides a viable alternative to existing data assimilation techniques in their ability to assimilate both Eulerian and Lagrangian data.

5.2 Underlying macroscopic traffic model

We use the viscous Lighthill-Whitham equation as the underlying core macroscopic mathematical model of traffic flow. The velocity v and density ρ is linearly related by the Greenshields function.

$$\rho_t + (\rho V_G(\rho))_x = \varepsilon \rho_{xx}, \tag{5.1}$$

$$V_G(\rho) = v_{\max}\left(1 - \frac{\rho}{\rho_{\max}}\right), \tag{5.2}$$

where $v_{\max}$ is the maximal free-flow velocity. Traffic jams manifest themselves as shock waves in (5.1), see Figure 5.1, which travel with characteristic speed c, where positive speeds correspond to propagation of traffic jams against the traffic flow.

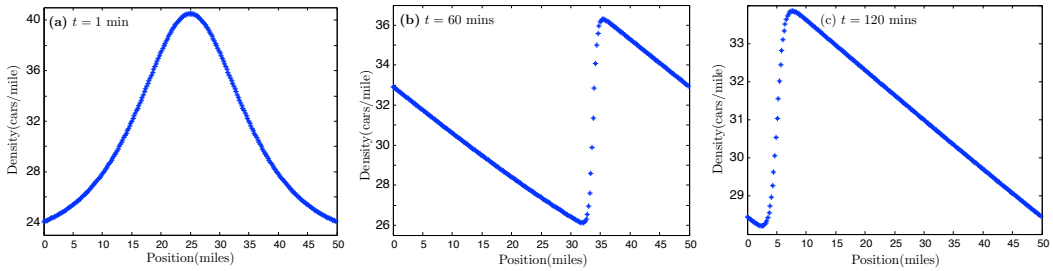


Figure 5.1: Evolution into a shock in the viscous Lighthill–Whitham equation (5.1). The shock is moving backward with gradually decreasing amplitude.

Given a solution $\rho(x,t)$ of (5.1), we can recover the movement of individual cars. Indeed, if $(p_i(t))_{i=1,2,\cdots N_c}$ denote the positions of N_c individual cars as functions of time, then these functions satisfy the ordinary differential equations

$$\frac{dp_i}{dt} = V(\rho(p_i,t)), \qquad i = 1,2,\cdots,N_c, \tag{5.3}$$

which can be solved to provide car positions for all times given their initial locations at $t = 0$.

The macroscopic model (5.1) can be used to describe normal traffic flow in the absence of ramps, traffic lights, and construction zones. In §2, we have added additional terms to the model (5.1) to capture the effects of on- and off-ramps, traffic lights, road construction sites, and traveling bottlenecks. The modifications allow us to simulate common traffic scenarios and therefore enable us to better test and validate the data assimilation strategies proposed here. In the next section, we will focus on (5.1) but stress that the techniques outlined in this section will also apply to the modifications of Lighthill-Whitham model.

5.3 Eulerian and Lagrangian observations

There are two main categories of observation models in traffic state estimation: one is Eulerian observation in which data is collected from stationary sensors on the road such as cameras, induction loops, and radar; the other type is Lagrangian observation, which includes the position and velocity data from moving vehicles equipped with GPS devices or cell phones.

As explained earlier, it is technically involved to use existing techniques to assimilate Eulerian and Lagrangian observation simultaneously. If we use a partial differential equation for the car density $\rho(x,t)$ at position x and time t as our underlying model, then sensor data produce estimates for the car density $\rho(x_0,t)$ at the sensor location x_0, which can be compared directly to the prediction made by the model. GPS data, on the other hand, are Lagrangian in nature, that is, they provide the location $x(t)$ and velocity $v(t)$ of an individual car as it moves with the traffic but cannot produce an estimate of the car density at that moving location. In particular, such data cannot be assimilated directly into a density-based model.

Two approaches have been considered to accomplish the assimilation of Lagrangian data. First, the underlying model can be written in terms of Lagrangian coordinates [79, 80] that move with the speed of the underlying cars, which makes it possible to assimilate Lagrangian observations. However, it is then much more difficult to assimilate Eulerian observations (such as sensor data); moreover, the approach outlined in [79, 80] was applied only to certain types of Eulerian observations (namely the velocities of cars measured at fixed sensor locations, but not the densities or fluxes recorded at these fixed locations). Second, one can convert the partial differential equation for the density to an equation for the velocity by inverting the density-velocity function $V(\rho)$ and then assimilate velocity observations [76, 77]; however, this inversion may not be possible for other models.

We now introduce our proposed approach that can efficiently assimilate Eulerian and Lagrangian data simultaneously.

5.3.1 Continuous observation models

Eulerian observation data: Sensors deployed at fixed positions along roads are capable of counting the number of vehicles passing by as well as reporting the flux at their location as a function of time. Thus, if there are N_s sensors positioned at the fixed locations $(q_i)_{i=1,2,\cdots,N_s}$ on a ring road $\mathbb{T}$, then the complete system of traffic model (M) and Eulerian observation operators (D) is given by

$$\begin{aligned} &\text{(M)} \quad \rho_t(x,t) + \partial_x \varphi(\rho(x,t)) = \varepsilon \rho_{xx}(x,t), \qquad x \in \mathbb{T} \\ &\text{(D)} \quad h^{\varphi}(\rho) := (\varphi(\rho(q_i,t)))_{i=1,2,\cdots,N_s}. \end{aligned} \tag{5.4}$$

An advantage of Eulerian sensor data is their high accuracy: modern traffic sensors are so sensitive that the probability of missing or over-counting a vehicle is below 0.1% ([1]). However, the Eulerian data is expensive to collect, and fixed locations of sensors limit the observation coverage. Note also that the Eulerian observation function $h^{\varphi}(\cdot)$ maps the vehicle density ρ to the vehicle flux $\varphi(\rho)$ and is therefore typically nonlinear: the nonlinearity of the observation function will influence the efficacy of data assimilation, especially for the ensemble Kalman filter, and requires special treatment which will be discussed in §5.4.2.

Lagrangian observation data: Lagrangian observations consist typically of the positions and velocities of moving vehicles collected by GPS devices or cell phones. If there are N_c cars equipped with these devices, and their positions and velocities are denoted by $(p, v) := (p_j, v_j)_{j=1,2,\cdots,N_c}$, then the complete system of the traffic model (M), capable of tracking the macroscopic traffic density and the positions of the N_c individual cars, and the Lagrangian observation operators (D) is given by

$$\begin{aligned} &\text{(M)} \quad \begin{cases} \rho_t(x,t) + \partial_x \varphi(\rho(x,t)) = \varepsilon \rho_{xx}(x,t), \qquad x \in \mathbb{T} \\ \dot{p_j}(t) = v_j(t) = V(\rho(p_j(t),t)), \quad j = 1, 2, \cdots, N_c \end{cases} \\ &\text{(D)} \quad \begin{cases} h^p(p,v) = p \\ h^v(p,v) = v. \end{cases} \end{aligned} \tag{5.5}$$

Thus, in contrast to previous methods for assimilating Lagrangian data that relied on its reformulation in Lagrangian coordinates or reshaping the macroscopic traffic model in terms of velocity or spacing [79, 80], our approach adds the governing equations for the positions of individual cars to the macroscopic model by exploiting that our constitutive law relates density and velocity.

The main advantage of Lagrangian data is the broad coverage of traffic states as the tracked vehicles travel along the road. The disadvantage is that density information is not collected directly. Furthermore, the observation noise of Lagrangian data is relatively large (see below for details) and depends, for instance, on the carrier's signal and weather conditions.

Eulerian and Lagrangian observation data: We now outline how Eulerian and Lagrangian data can be assimilated simultaneously by concatenating the systems (5.4) and (5.5) that we introduced above. The resulting system for the traffic flow model (M) and the observation operator (D) is given by

$$\begin{aligned}
\text{(M)} \quad & \begin{bmatrix} \rho_t(x,t) \\ \dot{p_j}(t) \\ v_j(t) \end{bmatrix} = \begin{bmatrix} \varepsilon\rho_{xx}(x,t) - \partial_x\varphi(\rho(x,t)) \\ V(\rho(p_j(t),t)) \\ V(\rho(p_j(t),t)) \end{bmatrix} \quad \begin{matrix} \\ j = 1,2,\cdots,N_c \\ j = 1,2,\cdots,N_c \end{matrix} \\
\text{(D)} \quad & \begin{bmatrix} h^{\varphi}(\rho) \\ h^{p}(p,v) \\ h^{v}(p,v) \end{bmatrix} = \begin{bmatrix} (\varphi(\rho(q_i,t)))_{i=1,2,\cdots,N_s} \\ p \\ v \end{bmatrix}.
\end{aligned} \tag{5.6}$$

The framework proposed in (5.6) is capable of assimilating both Eulerian sensor and Lagrangian GPS data as it is a closed system for the density ρ and the locations and velocities of the N_c vehicles from which GPS data are gathered. We note that we will consider N_c to be fixed but emphasize that our framework allows this quantity to depend on function of time as long as it changes slowly compared to the time scale over which GPS data are collected.

Observation errors: We assume that the observation errors for sensor, GPS position, and GPS velocity data are independent and normally distributed with zero

mean. As mentioned above, the observation noise η_i^{φ} for sensor data has a variance of $0.1\%\,\varphi(q_i,t)$. Based on the current accuracy of GPS data, we assume that position and velocity measurements have errors with standard deviation of 5.12m ([35]) and 0.0707m/s ([58, 62]), respectively. Using these assumptions, the covariance matrix of the observation errors can be written as

$$\mathbf{R}(t) = \begin{bmatrix} 0.1\% \left(\varphi(q_i,t)\right)_{i=1,\dots,N_s} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (5.12\text{m})^2\, \mathbb{I}_{N_c} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (0.0707\text{m/s})^2\, \mathbb{I}_{N_c} \end{bmatrix}, \tag{5.7}$$

where $\mathbb{I}_N$ denotes the $N \times N$ identity matrix.

5.3.2 Discretized observation models

We now discuss the numerical implementation of the continuous scheme (5.6), which is obtained by discretization in space and time. Recall that we consider a ring road $\mathbb{T}$ of length L.

First, we consider the discretization of the traffic model (M) given by

$$\begin{bmatrix} \rho_t(x,t) \\ \dot{p_j}(t) \\ v_j(t) \end{bmatrix} = \begin{bmatrix} \varepsilon\rho_{xx}(x,t) - \partial_x\varphi(\rho(x,t)) \\ V(\rho(p_j(t),t)) \\ V(\rho(p_j(t),t)) \end{bmatrix} \quad \begin{matrix} \\ j = 1,2,\cdots,N_c \\ j = 1,2,\cdots,N_c. \end{matrix} \tag{5.8}$$

For the spatial discretization, we pick a large integer M, define a spatial step size $\Delta x := L/M$, and choose M consecutive mesh points $x_m \in \mathbb{T}$ that are equally spaced along the ring road with distance Δx. Time is discretized by choosing a small positive time step ΔT and evaluating solutions at times $t = n\Delta T$ for integers $n \in \mathbb{N}$.

The traffic state at time $t = n\Delta T$ consists of the densities at each mesh point $x_m \in \mathbb{T}$ as well as the GPS positions and velocities of tracked vehicles and is therefore given by

$$\mathbf{X}(t) := [\underbrace{\rho(x_1,t),\ldots,\rho(x_M,t)}_{\text{densities}},\underbrace{p_1(t),\ldots,p_{N_c}(t)}_{\text{GPS positions}},\underbrace{v_1(t),\ldots,v_{N_c}(t)}_{\text{GPS velocities}}]^T. \tag{5.9}$$

The discretized system obtained from solving (5.8) numerically can then be written as

$$\mathbf{X}(t) = \boldsymbol{f}(\mathbf{X}(t-\Delta T);\boldsymbol{\theta}(t-\Delta T)), \tag{5.10}$$

where $\boldsymbol{\theta}(t)$ denotes system parameters, which may be constant or vary in time. The function $\boldsymbol{f}$ is obtained using the following numerical schemes. First, we discuss how we discretize the viscous Lighthill–Whitham equation for the density $\rho(x,t)$ that appears in (5.8): for periodic boundary conditions, we approximate the spatial derivatives in the viscous Lighthill–Whitham equation for the density $\rho(x,t)$ using spectral differentiation matrices in physical space as outlined in [71, (3.10) in §3]; for non-periodic boundary conditions, we discretize the spatial derivatives using finite differences. Next, we describe how we evaluate the equations for the positions and velocities (p_j, v_j) that appear in (5.8): to determine the values $\rho(p_j,t)$ of the density at arbitrary positions p_j, we interpolate ρ from its known values $\rho(x_m,t)$ at the mesh points x_m using trigonometric interpolation ([40]). Finally, we integrate the resulting system of differential equations from time $t-\Delta T$ to time t using a third-order Runge–Kutta method.

We now move to the observations. Using the observation errors introduced above,

Eulerian and Lagrangian observations can be written as

$$\mathbf{Y}(t) = \boldsymbol{h}(\mathbf{X}(t)) + \boldsymbol{\eta}(t), \qquad \boldsymbol{\eta}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}(t)), \tag{5.11}$$

where $\boldsymbol{h}(\mathbf{X}(t))$ with $\mathbf{X}(t)$ given in (5.9) is defined via

$$\boldsymbol{h}(\mathbf{X}(t)) := (\varphi(\rho(q_1, t)), \ldots, \varphi(\rho(q_{N_s}, t)), p_1(t), \ldots, p_{N_c}(t), v_1(t), \ldots, v_{N_c}(t))^T. \tag{5.12}$$

In order to evaluate $\rho(q_i, t)$ at the location q_i of a sensor, we can either arrange that q_i coincides with a mesh point x_m or again use trigonometric interpolation ([40]). We remark that the expression in (5.12) can be modified easily to account for the situation where we have observations from sensors only or from GPS data only available for data assimilation.

5.4 Ensemble methods

The goal of data assimilation is to accurately predict the states and parameters of a numerical model by incorporating available information obtained from observations into the model. In our work, this technique is carried out to estimate the traffic states and uncertain parameters in the traffic flow models described above. A general introduction to data assimilation and filters is provided in this section.

We first review two common sequential data assimilation methods, the particle filter (PF) and the ensemble Kalman filter (EnKF). The goal of each of these methods is to estimate the Bayesian prior and posterior probability distributions using an ensemble of possible states, as they change in time. First, define the state of interest to be $\mathbf{X}$, and the observation $\mathbf{Y}$ to be a noisy function of the state: $\mathbf{Y} = h(\mathbf{X}) + \eta$,

where the observation error is normally distributed $\eta \sim \mathcal{N}(0, \mathbf{R})$ with $\mathcal{N}(\cdot, \cdot)$ denoting the Gaussian distribution. Then, we are interested in estimating the posterior probability distribution $p(\mathbf{X}|\mathbf{Y})$ given a prior distribution $p(\mathbf{X})$ based on our current knowledge of the state, and a likelihood distribution $p(\mathbf{Y}|\mathbf{X})$ of the observations $\mathbf{Y}$ given the state $\mathbf{X}$, based on our knowledge of the noise in the observations. Bayes' rule gives the true posterior distribution in this case:

$$p(\mathbf{X_t}|\mathbf{Y_{1:t}}) \propto p(\mathbf{Y_t}|\mathbf{X_t}) \int p(\mathbf{X_t}|\mathbf{X_{t-1}}) p(\mathbf{X_{t-1}}|\mathbf{Y_{1:t-1}}) \tag{5.13}$$

where $p(\mathbf{Y}|\mathbf{X})$ is Gaussian likelihood with $\mathbf{Y}|\mathbf{X} \sim \mathcal{N}(h(\mathbf{X}), \mathbf{R})$. In both the PF and EnKF methods, these distributions are approximated by a weighted ensemble of the state $\mathbf{X}$ given by $\{\mathbf{X}_i, w_i\}_{i=1}^{N_e}$ which implies the distribution

$$\sum_{i=1}^{N_e} w_i \delta(\mathbf{X} - \mathbf{X}_i) \qquad \text{with} \qquad \sum_{i=1}^{N_e} w_i = 1\,, \tag{5.14}$$

where $\delta(\mathbf{X} - \mathbf{X}_i)$ is the Dirac delta centered at $\mathbf{X}_i$. Generally for the EnKF, $w_i = 1/N_e$. Between the observation times, the weights are kept fixed and the state variables are evolved according to the dynamics of the system. The main difference between the two methods comes at the time when observations are available. This is described in the next two subsections. We will use the notation that $\{\mathbf{X}_i^f, w_i^f\}_{i=1}^{N_e}$ is the ensemble from the prior distribution $p(\mathbf{X})$ whereas $\{\mathbf{X}_i^a, w_i^a\}_{i=1}^{N_e}$ is from the posterior distribution $p(\mathbf{X}|\mathbf{Y})$.

5.4.1 Particle filter

The posterior is approximated by updating the weights but leaving the particle positions fixed, i.e., $\mathbf{X}_i^a = \mathbf{X}_i^f := \mathbf{X}_i$. The updated weights are obtained by applying

Bayes' rule to the weights as follows:

$$w_i^a = \frac{p(\mathbf{Y}|\mathbf{X}_i)w_i^f}{\sum_{j=1}^{N_e} p(\mathbf{Y}|\mathbf{X}_j)w_j^f}, \tag{5.15}$$

that is, the updated weights are found by multiplying the likelihood of that particle by the previous weight and normalizing to sum to 1. This is the simplest implementation of the PF, also known as sequential importance sampling ([27, 34]). For a more complete derivation of the particle filter and the proposal distributions, see [26, 66, 68].

Due to the finite nature of the approximation and the recursive updating of the weights, sequentially applying this algorithm eventually leads to one particle with very high weight, while the rest of the particles have almost zero weight (so-called "filter divergence" or "weight collapse".) To prevent the weights of the ensemble concentrating on a single state, various resampling methods for the ensemble states may be used ([74]). The basic idea behind each of these methods is to monitor when a predetermined threshold is hit (for example, when the "effective sample size" becomes small ([51])), and to then resample the particles from the discrete approximation of the posterior distribution and reset all weights to $1/N_e$. We approximate the effective sample size to be

$$N_{\text{eff}} \approx \frac{1}{\sum_{i=1}^{N_e} w_i^2} = \begin{cases} N_e & \text{for } w_j = 1/N_e \\ 1 & \text{for } w_j = \delta_{ij} \end{cases} \tag{5.16}$$

as in [51]. We will apply resampling when $N_{\text{eff}} < N_{\text{eff}}^{\text{thresh}}$ for a predetermined threshold $N_{\text{eff}}^{\text{thresh}}$ that will typically be a small fraction of the total number of particles N_e.

One major drawback of the PF is that it has been shown to fail in high dimensions ([67]). Thus, in the case where the state of interest is a spatially-discretized function over some domain, the PF is intractable. However, when the state dimension is small enough and the number of particles is large enough, the PF can provide an accurate approximation to the exact Bayesian posterior distribution. This is especially useful if this distribution is skewed or multimodal, which is often the case when the dynamical system governing the state is nonlinear.

5.4.2 Ensemble Kalman filter

Like the PF, the EnKF [28, 29] employs an ensemble of state vectors $\{\mathbf{X}_i\}_{i=1\ldots N_e}$ to represent the posterior distribution; however, unlike the PF, the ensemble members are equally weighted for the entire assimilation window. Instead of updating the weights at analysis times, the members themselves are updated according to an ensemble approximation of the traditional Kalman filter update step, given here by the so-called perturbed observation EnKF [15, 29, 43]

$$\mathbf{X}_i^a = \mathbf{X}_i^f + \mathbf{K}(\mathbf{Y} - \mathbf{H}\mathbf{X}_i^f + \boldsymbol{\eta}_i), \tag{5.17}$$

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}\right)^{-1}, \tag{5.18}$$

where $\mathbf{X}_i^f$ is the forecast of the i^{th} ensemble member, $\mathbf{X}_i^a$ is the i^{th} updated (analysis) ensemble member, $\mathbf{H}$ is the observation operator, $\mathbf{K}$ is the Kalman gain matrix, $\mathbf{R}$ is the covariance of the observation error, and $\eta_i \sim \mathcal{N}(0, \mathbf{R})$ are the observation perturbations. $\mathbf{P}^f$ is the forecast ensemble covariance given by

$$\mathbf{P}^f = \frac{1}{N_e - 1}\sum_{i=1}^{N_e}(\mathbf{X}_i^f - \bar{\mathbf{X}}^f)(\mathbf{X}_i^f - \bar{\mathbf{X}}^f)^T, \quad \bar{\mathbf{X}}^f = \frac{1}{N_e}\sum_{i=1}^{N_e}\mathbf{X}_i^f. \tag{5.19}$$

Nonlinear observation function: We note that the update step requires that the observation operator $\mathbf{H}$ is linear. However, the observation function given in (§5.12 equation 5.12) is nonlinear, and we will therefore outline how we adjust the update step to accommodate the nonlinear observation function. Instead of calculating the operator $\mathbf{H}$ and the covariance matrix $\mathbf{P}^f$ separately when estimating the state $\mathbf{X}_i^a$ in (5.17) and (5.18), we follow [56] and treat $\mathbf{P}^f\mathbf{H}^T$ and $\mathbf{H}\mathbf{P}^f\mathbf{H}^T$ as follows as single entities:

$$\mathbf{H}\mathbf{P}^f\mathbf{H}^T = \frac{1}{N_e - 1}\sum_{i=1}^{N_e}(\mathbf{Z}_i^f - \bar{\mathbf{Z}}^f)(\mathbf{Z}_i^f - \bar{\mathbf{Z}}^f)^T, \qquad \bar{\mathbf{Z}}^f := \frac{1}{N_e}\sum_{i=1}^{N_e}\mathbf{Z}_i^f, \tag{5.20}$$

$$\mathbf{P}^f\mathbf{H}^T = \frac{1}{N_e - 1}\sum_{i=1}^{N_e}(\mathbf{X}_i^f - \bar{\mathbf{X}}^f)(\mathbf{Z}_i^f - \bar{\mathbf{Z}}^f)^T, \qquad \bar{\mathbf{X}}^f := \frac{1}{N_e}\sum_{i=1}^{N_e}\mathbf{X}_i^f, \tag{5.21}$$

where $\mathbf{Z}_i^f := h(\mathbf{X}_i^f)$ denotes the ensembles projected into observation space. Similarly, the term $\mathbf{H}\mathbf{X}_i^f$ in (5.17) is replaced by $\mathbf{Z}_i^f$.

Localization: There has been a lot of work towards improving the EnKF to make it feasible for very high dimensional systems; in particular, covariance inflation and localization have provided significant improvement of the performance of the EnKF (see [4, 38, 43, 44]). Since the ensemble size is often much smaller than the dimension of the state, spurious correlation will arise in the sample covariance matrices. Localization is a method in which these spatially long-range correlations are diminished: in particular, this is often implemented via a Schur product of the sample covariance matrix and some sort of cutoff matrix.

One common localization method applies an exponential decay function, such as the Gaspari–Cohn function in [31], to the forecast covariance matrix $\mathbf{P}^f$. In our formulation of the Kalman gain matrix, the forecast covariance matrix $\mathbf{P}^f$ is not

directly accessible as we provide only $\mathbf{H}\mathbf{P}^f\mathbf{H}^T$ and $\mathbf{P}^f\mathbf{H}^T$ through (5.20) and (5.21), respectively. As in [44], we could now apply elementwise localization separately to each of these two matrices: localization of $\mathbf{P}^f\mathbf{H}^T$ would account for the assumption that observations at a given location would only affect the density and car positions near that location, while localization of $\mathbf{H}\mathbf{P}^f\mathbf{H}^T$ would correspond to localizing the effect observations at different positions have on each other.

Alternatively, if we interpret the operator $\mathbf{K} = (\mathbf{K}_{ij})$ loosely as a weight matrix where the (i, j)th entry $\mathbf{K}_{ij}$ determines how much the jth entry of the observation vector affects the ith entry of the state vector, then we can apply localization directly to the matrix $\mathbf{K}$ to ensure that observations at a certain location affect only nearby states. Specifically, we assume that only mesh points within 0.5 miles of the location of a sensor are affected by Eulerian observations. We then define an exponential decay localization function for the jth column of the Kalman gain matrix $\mathbf{K}$ via

$$\mathrm{e}^{-d|x_m-(q_j+s)|} \quad \text{with} \quad |x_m - q_j| < 0.5 \text{ miles}, \tag{5.22}$$

where x_m is the coordinate of the mth mesh point, q_j is the coordinate of the jth sensor, and d is a decay coefficient that describes the decay rate of the localization function. Since cars detected by a sensor will have travelled a certain distance s during time step ΔT, we center the localization at $q_j + s$. Throughout, we use the parameters $d = 0.5$ and $s = 0.35$ miles. We use the same localization function also for Lagrangian observations: we set $d = 1.2$ and choose $s = 0$ as there will be no time delay for positions and velocities collected from GPS data.

Figure 5.2 shows that applying localization directly to the Kalman gain matrix turns out to be very similar to applying localization to the forecast covariance matrix. Indeed, our numerical results show that the observation covariance $\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}$ is

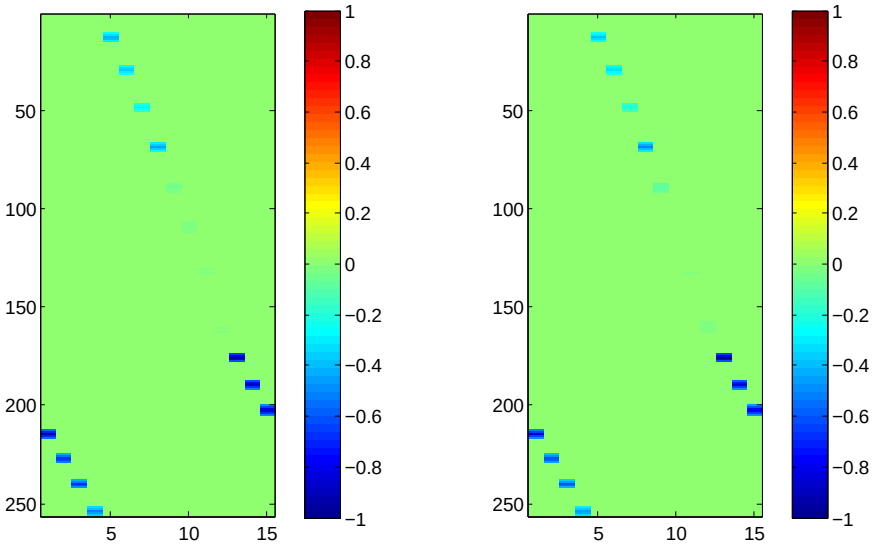


Figure 5.2: Comparison of the scaled Kalman gain matrices with localization applied to the forecast covariance matrix (left) and directly to the Kalman gain matrix (right).

dominated by its diagonal elements: using the notation $\mathbf{I}$ for the identity matrix, we then see that the Kalman gain matrix, with localization applied to the forecast covariance matrix, can be approximated by the Kalman gain matrix with direct localization:

$$\begin{aligned}\left(\mathbf{loc} \circ (\mathbf{P}^f \mathbf{H}^T)\right) \left(\mathbf{I} \circ (\mathbf{H}\mathbf{P}^f \mathbf{H}^T) + \mathbf{R}\right)^{-1} &= \mathbf{loc} \circ \left(\mathbf{P}^f \mathbf{H}^T \left(\mathbf{I} \circ (\mathbf{H}\mathbf{P}^f \mathbf{H}^T) + \mathbf{R}\right)^{-1}\right) \\ &\approx \mathbf{loc} \circ \left(\mathbf{P}^f \mathbf{H}^T \left(\mathbf{H}\mathbf{P}^f \mathbf{H}^T + \mathbf{R}\right)^{-1}\right).\end{aligned}$$

Inflation: The EnKF is not immune to filter divergence, where the prior ensemble spans a smaller and smaller space until the observations no longer have any effect in the analysis step ([4]). The EnKF is prone to artificial collapse of the covariance matrix. To address this issue, the covariance may be "inflated" before the Kalman update step is performed. We use two different inflation algorithms for the EnKF. First, we employ a constant scalar inflation for the simulations of multiple traffic

scenarios as the localization algorithm described above does not significantly tighten the covariance matrix. Second, for the computations with realistic traffic data, we use adaptive inflation as described in [2, 3].

5.4.3 Parameter estimation

The macroscopic model we utilize may contain parameters whose values are not known: for instance, the expression (5.2) for the Greenshields velocity function contains the parameters $v_{\max}$ and $\rho_{\max}$, which will generally be unknown.

We briefly describe three approaches for how data assimilation may be used to estimate parameters in addition to the state of interest. One approach is the augmented-vector method, in which the parameters to be estimated are appended to the state vector ([6, 55]) and are updated concurrently. Another approach is the "separated" method [9], in which the state is updated first, followed by updating the parameters, using localization on the state and non-global parameters. The third approach applies a two-stage filter [65] that consists of applying the EnKF to the state estimate and the PF to the parameters.

There are then several methods that may be used to model the evolution of the parameters between assimilation steps. The parameters may be kept fixed between time steps (persistence model, see [23]), but this can lead to degeneracy (sample attrition) when using PF methods. In [55], the authors suggest an artificial evolution of parameters between observations, in which the parameter evolution model is given

by

$$\theta_{t+1} = \theta_t + \xi_{t+1}, \tag{5.23}$$

$$\xi_{t+1} \sim \mathcal{N}(0, \mathsf{W}_{t+1}), \tag{5.24}$$

where W_{t+1} is a specified covariance matrix. In [78], multiplicative parameters were estimated by evolving the parameters under a temporally smoothed version of the persistence model. It was shown that this results in smooth temporal variation of the parameters, which in turn prevents blow-up of the model.

The augmented-vector approach is used to estimate parameters, and artificial evolution is introduced to time-varying parameters.

5.5 Numerical results

5.5.1 Performance criteria

In this paper, we use both absolute and relative root-mean-square-errors (RMSE) as performance indicators to assess different scenarios. The absolute performance index at time t_n is defined as

$$\mathrm{RMSE_A} = \sqrt{\frac{\sum_{m=1}^{M}(\rho_m(t_n) - \hat{\rho}_m(t_n))^2}{M}}, \tag{5.25}$$

while the corresponding relative performance index is given by

$$\mathrm{RMSE_R} = \frac{\mathrm{RMSE_A}}{\rho_{\max}}, \tag{5.26}$$

where $\rho_{\max}$ is the maximal density in the macroscopic model. Here, M is the number of mesh points used to discretize a given highway stretch, $\rho_m(t_n)$ denotes the true density at the mth mesh point x_m at time t_n, and $\hat{\rho}_m(t_n)$ denotes the corresponding estimation from data assimilation. For those traffic scenarios with unknown maximal density $\rho_{\max}$, the denominator in the relative RMSE is replaced by the average $\sum_{m=1}^{M} \rho_m(t_n)/M$ of the true density.

5.5.2 Traffic scenarios with true underlying model

The main goal of this part is to apply data assimilation to estimate traffic states and parameters using the theory described in § 5.3 and § 5.4. Specifically, we will investigate

(i) assimilation of Eulerian and Lagrangian data in dynamic traffic flow;

(ii) efficacy of Lagrangian data assimilation;

(iii) impact of sensor location on data assimilation;

(iv) parameter estimation of traffic flow models.

We note that the first three topics are investigated under the assumption that the estimator has the exact values of model parameters, and the emphasis is placed on the capability of traffic state estimators. However, in the last topic, the parameter estimation is activated with the intent of investigating the capability of parameter estimators.

Basic setup: In the simulations outlined in § 5.5.2, the underlying model for the data assimilation is the same as the model generating the truth; in each case, the viscous Lighthill-Whitham equation (5.1) is used with additional terms to simulate various traffic scenarios, to be outlined individually below. We will show that data assimilation can be used to reduce the error in simulating these scenarios when imperfect initial data is used. The initial data for density is set as the truth plus 10% noise through Fourier coefficients of the density curve.

For each traffic scenario and associated macroscopic system, we consider traffic flow on a ring road of length $L = 50$ miles, which is equally divided into 256 mesh points for the numerical computation. The truth is generated by initializing the system with the profile

$$\rho_0(x) = 0.5\rho_{\max} + 0.4\rho_{\max}\text{sech}(x - L/2), \tag{5.27}$$

and the parameters

$$\rho_{\max} = 45 \text{ cars/mile}, \quad v_{\max} = 75 \text{ miles/hour}, \quad \varepsilon = 0.1. \tag{5.28}$$

For the data assimilation algorithm, the underlying model is taken to be the same as that generating the synthetic truth; however, the initial conditions for the ensemble are the noisy density curves after adding noise directly through the Fourier coefficient. Aside from the fourth investigation, where parameter estimation is employed, we use the same systems parameters in both the generation of the truth and the data assimilation.

The Eulerian and/or Lagrangian observations are collected and assimilated every one minute, and the entire simulation has 180 total updates. For the Eulerian

observation system, we place $N_s = 8$ stationary sensors equidistantly along the ring road. For the Lagrangian observation system, we pick $N_c = 15$ cars with GPS devices that are initially equally spaced along the ring road. The ensemble size is 30 for the EnKF and 300 for the PF because of the high dimension of the state vector.

In each of the following investigation, we will outline the traffic scenarios to be studied and the associated macroscopic models, followed by the results of data assimilation in estimating the traffic states which emerge.

1. Data assimilation of Eulerian and Lagrangian data First, we demonstrate that the proposed algorithm can be used to assimilate Eulerian, Lagrangian, and combined Eulerian-Lagrangian observations into a macroscopic model to accurately estimate traffic states. To test the algorithm, we consider a ring road with a traffic light for our simulations and assimilate sensor and GPS data. In order to better evaluate the performance of the algorithm, we will compare the results also to those obtained by simulating the macroscopic model with the same initial data and parameter values but without assimilating observations.

The traffic light scenario is described in § 2.3.3 by the following equation

$$\rho_t + \left(\rho V_G\left(\rho(x,t)\right) a(x,t;x_1^\ell)\right)_x = \varepsilon \rho_{xx},$$

$$a(x,t;x_1^\ell) = \begin{cases} 0.5 & (x,t) \in (x_1^\ell - L_1^y, x_1^\ell) \times T_1^y \\ 0 & (x,t) \in (x_1^\ell - L_1^r, x_1^\ell) \times T_1^r \\ (x_1^\ell - L_1^r - x)/L_1^r & (x,t) \in (x_1^\ell - 2L_1^r, x_1^\ell - L_1^r) \times T_1^r \\ 1 & \text{otherwise,} \end{cases}$$

The traffic light is placed in the middle of the road with $x_1^\ell = 25$. We set the lights

length $(|T_1^y|, |T_1^r|, |T_1^y|) = (10, 190, 400)$ seconds and response distances $(L_1^y, L_1^r) = (1, 0.8)$ miles in order to generate significantly oscillating traffic flow.

We use the above model and the initial conditions and parameters from the basic setup to generate the truth. The same system and parameters are used for the underlying data assimilation model, with perturbed initial conditions.

Results: We carried out simulations using the algorithm described in § 5.3 with the EnKF and PF for (i) Eulerian observations only, (ii) Lagrangian observations only, (iii) both Eulerian and Lagrangian observations, and (iv) no observations. The simulation results based on the traffic lights scenario are shown in Figure 5.3: panels (a) and (b) contain the results for the EnKF and PF, respectively. The relative RMSE for the cases (i)-(iii) ranges between 1% and 2% after 3 hours, while the relative RMSE for case (iv) (no observations) is above 7%. The results suggest that our algorithm is capable of assimilating any combination of Eulerian and Lagrangian data and performs well independently of whether we use the EnKF or PF.

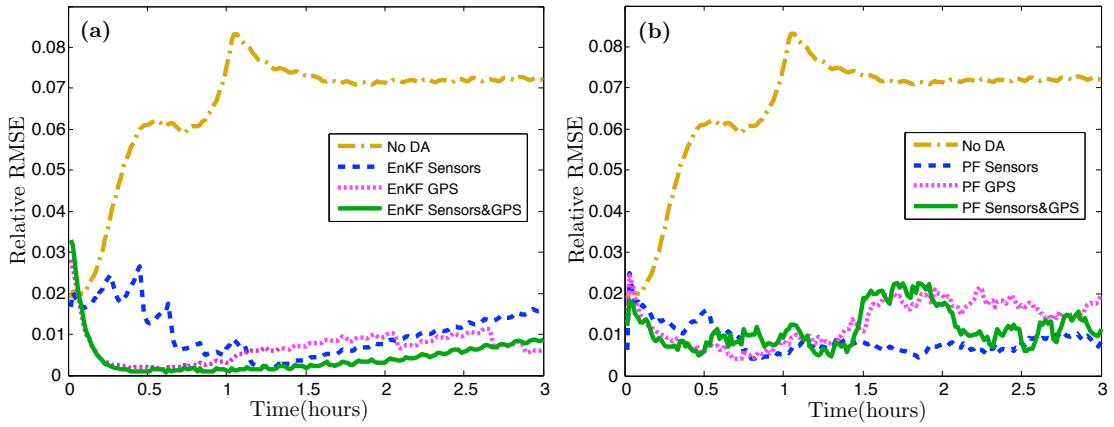


Figure 5.3: Traffic state estimation in the traffic light scenario. Shown is the relative RMSE for assimilating no data (yellow dashed dot), Eulerian data (blue dashed), Lagrangian data (magenta dotted), and both Eulerian and Lagrangian data (green solid) using the (a) EnKF and (b) PF.

2. Efficacy of Lagrangian data assimilation In practice, either position or velocity data can be collected through GPS devices. Thus, we are interested in the efficacy and accuracy of different Lagrangian observations. According to the statistical data of GPS devices on the internet, the error of position data has a standard deviation around 5.12 meters [35], while the error of velocity data has a standard deviation around 0.0707 m/s [58, 62]. In this section, both the EnKF and PF are used to assimilate three groups of different Lagrangian observations in a normal traffic scenarios: (1) positions of vehicles, (2) velocities of vehicles, (3) combined positions and velocities of vehicles.

we simulate a scenario of unimpeded traffic flow, which we model via the viscous Lighthill-Whitham system (2.10)

$$\rho_t + (\rho V_G(\rho))_x = \varepsilon \rho_{xx}$$

with no additional terms added. The initial conditions and parameters are as outlined in basic setup.

Results: The simulation results are shown in Figure 5.4. Figure 5.4(a) shows the relative RMSE of Lagrangian data assimilation for different observation data using the EnKF. The error for assimilating position data (solid blue) approaches 3.5%, the error for assimilating velocity data (dashed blue) approaches 0.4% and the error for assimilating both data (dotted blue) approaches 0.1%. In contrast, Figure 5.4(b) shows the comparison based on the PF: the error for assimilating velocity data (dashed magenta) approaches 0.25%, and the errors for assimilating position data and both data (solid and dotted magenta lines) approach 0.4%. The phenomenon that assimilating both is less accurate initially is related to the "curse of dimension-

ality" of the PF: the dimension of observations is doubled when assimilating both data, but we are unable to correspondingly increase the ensemble size to compensate.

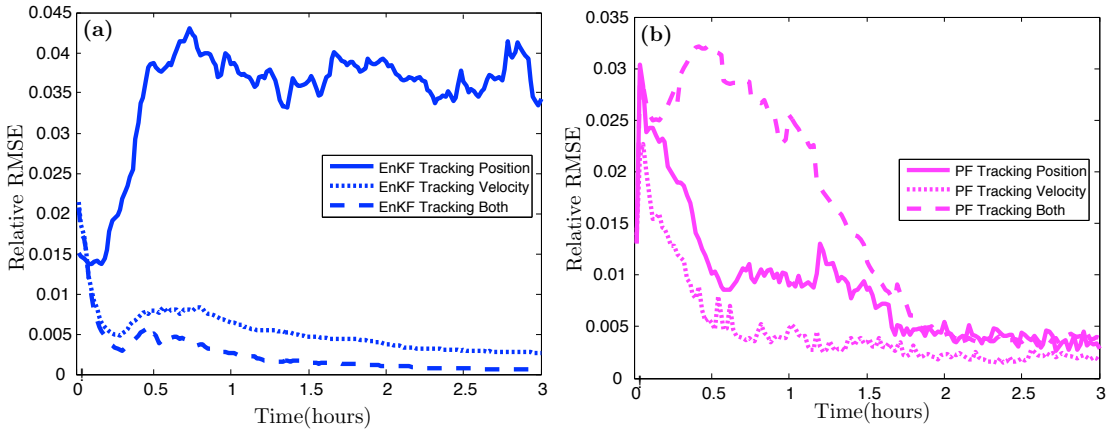


Figure 5.4: Traffic state estimation comparison for Lagrangian data assimilation based on different observation data from normal traffic flow modeled by (2.10). Shown is the relative RMSE for the (a) EnKF and (b) PF, where we assimilate position data only (solid), velocity data only (dotted), and combined position and velocity data (dashed).

For both the EnKF and PF, we can see that assimilating velocity data is more accurate compared to assimilating position data, which is consistent with the fact that velocity data has smaller observation noise than position data. However, the PF has a stronger tolerance of observation noise since the difference of error between assimilating velocity data and position data is much smaller compared to those by the EnKF. In addition, the performance of the EnKF can be improved with combined observation data, while the performance of the PF is limited by the number of particles with combined observation data. Therefore, we will use combined Lagrangian data for EnKF, while only use velocity data for PF in other simulations.

3. Impact of sensor location on data assimilation When the Department of Transportation installs sensors on the road, an important consideration is where to install them in order to optimize the amount of traffic information collected by them. In this section, we are interested in studying the impact of the configuration of sensors

in the presence of on-/off-ramps or bottlenecks in order to efficiently place sensors near ramps and bottlenecks. We will use two traffic scenarios for investigation: an on-/off-ramps scenario and a stationary bottleneck scenario.

The on-/off-ramps scenario is described in § 2.3.2 with the following formula:

$$\rho_t + (\rho V_G(\rho))_x = \varepsilon \rho_{xx} + \sum_{i \in \mathcal{I}} \varphi_i^{\text{on}}(t) f(x - x_i^{\text{on}}) + \sum_{j \in \mathcal{J}} \varphi_j^{\text{off}}(t) f(x - x_j^{\text{off}}),.$$

we place one on-ramp at $x_1^{\text{on}} = 37.5$ miles and one off-ramp at $x_1^{\text{off}} = 12.5$ miles on the road, and set constant flow $\varphi_i^{\text{on}}(t) = 2\rho_{\max}$ for the on-ramp and $\varphi_i^{\text{off}}(t) = 4\rho_{\max}$ for the off-ramp. Three comparative experiments are designed where a sensor is installed (1) at the ramp, (2) 0.5 miles upstream to the ramp, and (3) 0.5 miles downstream to the ramp.

The stationary bottleneck scenario is describe in § 2.3.4 with the following formula:

$$\rho_t + \left(\rho V_G(\rho) a(x - x_1^b)\right)_x = \varepsilon \rho_{xx}.$$

We set the bottleneck at $x_1^b = 25$ miles, severity coefficient $c = 1$ and spread effect $f(x) = 1 - 0.5\text{sech}(x)$. In this scenario, a stationary density bottleneck appears around x_1^b. Three comparative experiments are designed where a sensor is installed (1) at the bottleneck, (2) one mile upstream to the bottleneck, and (3) one mile downstream to the bottleneck.

Results: The simulation results for ramps are shown separately in Figure 5.5(a) and 5.5(b), and those for the bottleneck are shown in Figure 5.5(c). In Figure 5.5(a), it is observed that the performance of data assimilation is improved by moving a

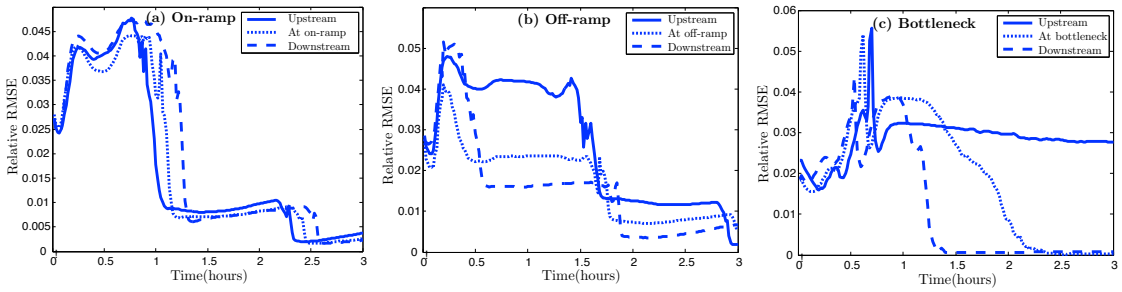


Figure 5.5: Evaluation of various sensors configurations using the EnKF. Results for different sensor locations near (a) on-ramps, (b) off-ramps, and (c) bottlenecks are shown, where the solid curves correspond to sensor locations upstream to the target location, the dotted curves represent sensors at the target location, and the dashed curves represent locations downstream to the target location.

sensor from downstream to upstream of the on-ramp. While in Figure 5.5(b), unlike the on-ramp case, the performance is improved by moving a sensor from upstream to downstream of the off-ramp. In Figure 5.5(c), it shows that the performance is improved by moving a sensor from upstream to downstream of the bottleneck. By studying the shape of density on the ring road, we find that a density bump is generated gradually upstream to the on-ramp. Therefore, installing a sensor slightly upstream can provide more traffic information and thereby improve the performance of the EnKF. In contrast, a density valley is generated gradually downstream to the off-ramp or the bottleneck, so installing a sensor slightly downstream can provide more traffic information.

In summary, the locations of sensors do impact the performance of data assimilation when using the EnKF: the simulation results suggest installing sensors upstream of on-ramps and downstream of off-ramps and bottlenecks. In contrast, the PF performs well regardless of the locations of sensors (simulation results not shown).

4. Parameter estimation of traffic-flow models In previous parts, we showed that our algorithm successfully assimilates traffic states when the parameters in the

underlying traffic model are known. However, in practical applications, these parameters are not known and need to be estimated when traffic data are assimilated. Hence, we are interested in the efficacy of data assimilation when traffic parameter estimation is activated. In this part, we consider two traffic scenarios for investigation: an on-/off-ramps scenario and traveling bottleneck scenario. For each scenario, the basic setup for generating the truth and the underlying model for data assimilation is the same as in the setup; however, the parameters in the underlying traffic model for data assimilation are now taken as unknown.

The model for the on-/off-ramps scenario is as described in the 3rd case for on/off-ramps, and the unknown parameters include maximal density $\rho_{\max}$, maximal velocity $v_{\max}$, the fluxes of on-ramp ρ_1^{on} and the fluxes of off-ramp ρ_1^{off}. These parameters are taken to be constant throughout the simulation.

The traveling bottleneck scenario is described in § 2.3.5 with the formula:

$$\rho_t + \left(\rho V_G(\rho) a(x - x_1^b(t))\right)_x = \varepsilon \rho_{xx}.$$

The position of the bottleneck is assumed to move back and forth periodically according to the formula $x_1^b(t) = L/2 + L/4\cos(2\pi t/3)$. The unknown parameters include maximal density $\rho_{\max}$, maximal velocity $v_{\max}$, taken to be constant, and the bottleneck location $x_1^b(t)$, which is time-dependent.

Results: As shown in Figure 5.6(a), the relative RMSE by using the EnKF (blue dashed line) and the PF (magenta dashed line) hovers around 3% and 2%, respectively, after 3 hours of assimilating Eulerian observations. In Figure 5.6(b), the relative RMSE by using the EnKF (blue dash line) and the PF (magenta dash line) is around 10% and 20%, respectively, after 3 hours of assimilating Lagrangian ob-

servations. Observations of the flux, which is the product of density and velocity, can balance the estimation of $\rho_{\max}$ and $v_{\max}$ and thereby provide accurate estimation with an error below 5%. However, Lagrangian observations, especially of velocity data, cannot provide as good a balance between density and velocity as Eulerian observations, and we indeed find that Lagrangian observations are less efficient when the parameters are unknown. We also see in Figure 5.6(c) and 5.6(d) that the PF estimator (magenta dashed) is more sensitive when considering the estimated fluxes as these oscillate around the true flux value near the ramps.

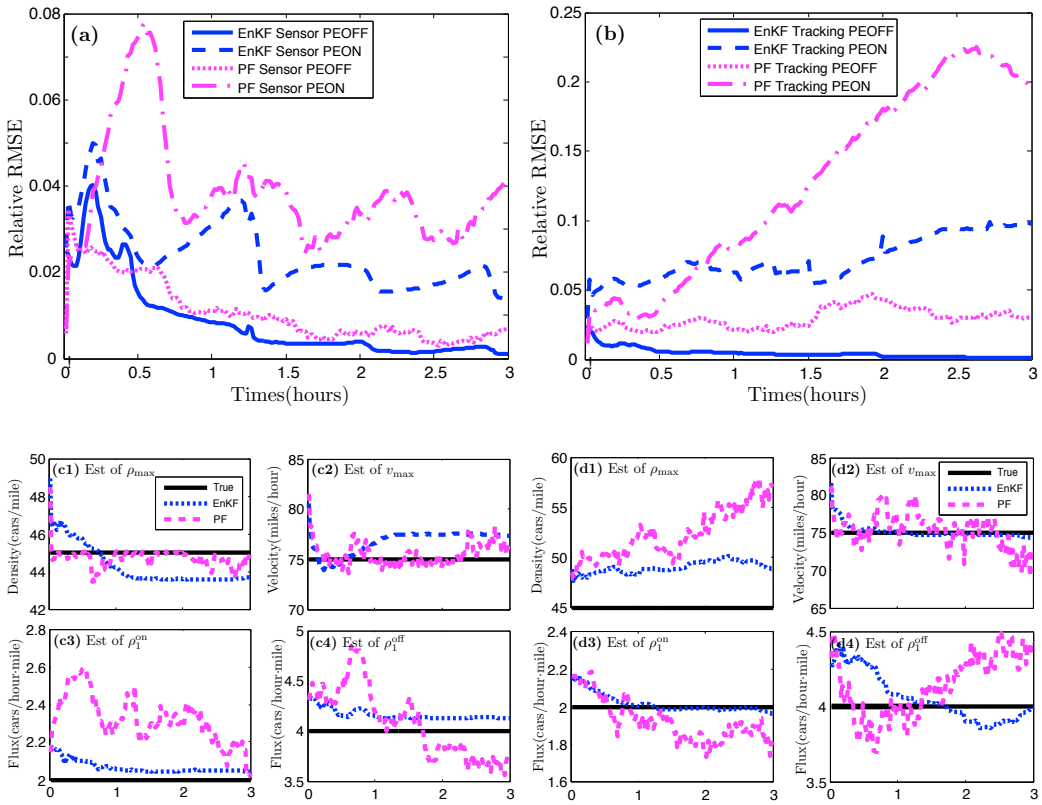


Figure 5.6: Estimation of traffic states and parameters in the ramps scenario: shown are the relative RMSE for assimilating (a) Eulerian and (b) Lagrangian observations as well as the results of parameter estimation for (c) Eulerian and (d) Lagrangian observations. Estimated are the maximal density $\rho_{\max}$, the maximal velocity $v_{\max}$, the flux of the on-ramp ρ_1^{on}, and the flux of the off-ramp ρ_1^{off}. PEOFF represents that the parameters are known, while PEON represents that the parameter estimation is considered.

In the bottleneck scenario, the conclusions are similar as those in the ramps scenario. The relative RMSEs are below 5% in Figure 5.7(a) and below 9% in

Figure 5.7(b), which means our approach has provided a very good state estimator. In addition, Figure 5.7(c) and 5.7(d) demonstrate the dynamical tracking capability of our approach for the time-varying parameter $x_1^b(t)$.

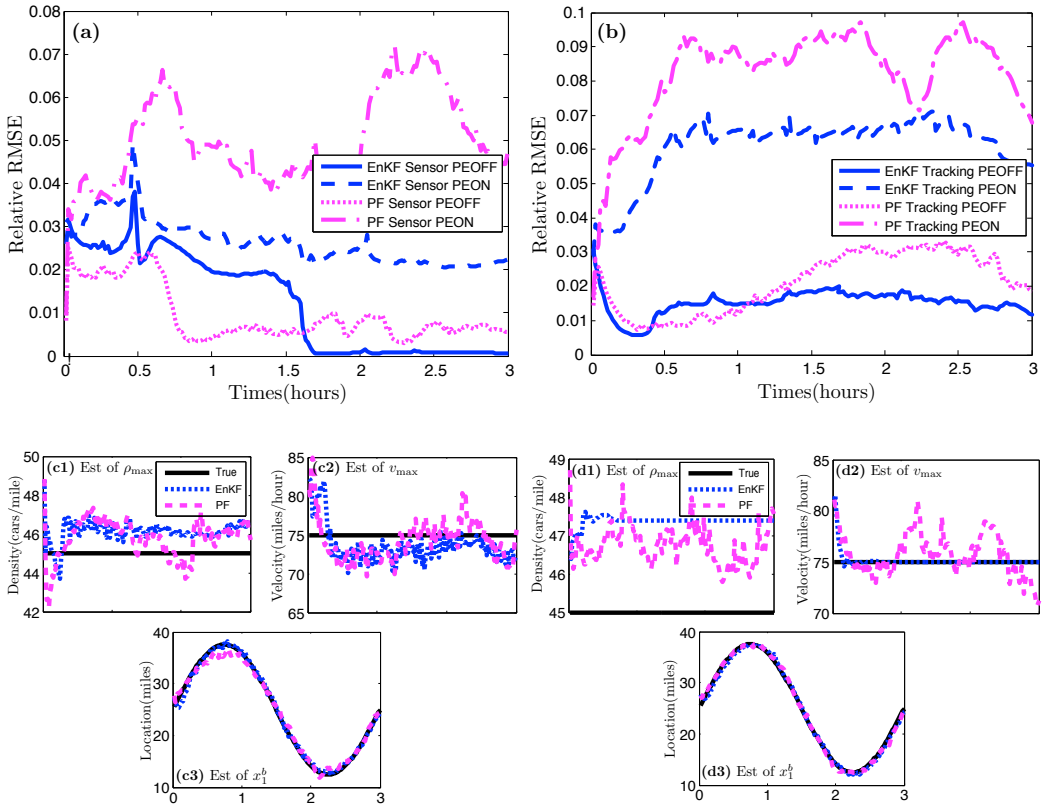


Figure 5.7: Estimation of traffic states and parameters in the traveling bottleneck traffic scenario: shown are the relative RMSE for assimilating (a) Eulerian and (b) Lagrangian observations and the results of parameter estimation for (c) Eulerian and (d) Lagrangian observations, where we estimate the maximal density $\rho_{\max}$, the maximal velocity $v_{\max}$, and the time-dependent bottleneck location x_1^b.

In summary, our approach is capable of providing accurate estimations for both constant and time-varying parameters. Generally Lagrangian observations are more efficient when estimating traffic states, but Eulerian observations appear to be slightly more efficient when parameter estimation is activated.

5.5.3 Traffic scenarios with unknown underlying model

The simulations in §5.5.2 used the same traffic model for (i) generating the truth and (ii) assimilating data in our algorithm. The main goal of this section is to test the implementation of data assimilation to estimate traffic states when the truth is not generated by the underlying model for the data assimilation scheme.

More specifically, in case 1, we will use data assimilation to estimate traffic states using observations generated from a microscopic traffic flow model. Then, in case 2, we will apply parameter estimation to real traffic data obtained from the Minnesota Department of Transportation.

5. Truth generated from microscopic models We are interested in testing the sensitivity with respect to the model used in the data assimilation algorithm in a controlled environment. In this section, we generate synthetic observations using an extended optimal-velocity type microscopic model and assimilate the data using the macroscopic LWR equation under two different scenarios of highway traffic. These scenarios, to be described in more detail below, are obtained by different choices of the parameters and initial conditions in the microscopic model.

We begin with an overview of the microscopic model, followed by the results when using data assimilation to estimate traffic states in each of the scenarios. In our simulations below, we will study each scenario separately using the EnKF to assimilate Lagrangian observations (vehicle positions) from the microscopic model to estimate the macroscopic traffic density. To compare the predicted traffic densities with the truth, we transform the microscopic traffic data into continuous densities on the ring road.

The microscopic optimal-velocity model attempt to model traffic through differential equations governing the evolution of the individual cars' positions and velocities. In general, these models allow for each car to adjust its velocity and acceleration based on the position and/or velocity of its neighbors. We focus on an extended optimal-velocity type model introduced in [16] in which the drivers adjust their velocity according to not only the headway but also the relative velocity to the car in front. In this model, the position $p_n(t)$ and the individual target headway $s_n(t)$ of the nth vehicle evolve according to

$$\begin{aligned} \tau\ddot{p}_n &= V_g \tanh\left(\frac{p_{n+1}-p_n-s_n}{\ell_0}\right) + V_0 - \dot{p}_n \\ \alpha\dot{s}_n &= \frac{V_g}{\ell_0}\left(\bar{s}-s_n\right) - \beta\left(\dot{p}_{n+1}-\dot{p}_n\right), \end{aligned} \tag{5.29}$$

where the dot symbol denotes d/dt. Here τ is the reaction time, and the right hand side of the equation is the so-called optimal velocity function. The adjustment $\mathcal{V}$ depends on the difference of the headway $p_{n+1}-p_n$ from the target headway s_n that encompasses the car length plus a safety distance. A common choice for this function is $\mathcal{V}(u)=\tanh(u)$, which we will use in this paper. Since $\mathcal{V}(0)=0$, the quantity V_0 represents an optimal velocity at which the car drives when the headway is equal to the optimal headway s_n. The quantity V_g represents a velocity gain, and by considering an infinite headway, we have $\mathcal{V}(\infty)=1$, so that the quantity V_g+V_0 gives an effective speed limit for the drivers. The parameter ℓ_0 is a characteristic length scale which describes the pace at which the driver attempts to achieve the optimal velocity.

The second equation describes the evolution of the individual target headway $s_n(t)$: the term $\bar{s}-s_n$ describes the relaxation of $s_n(t)$ to an optimal headway $\bar{s}$, while the term $\beta(\dot{p}_{n+1}-\dot{p}_n)$ takes into account that drivers may increase speed

if the car in front does so. The parameters α and β are dimensionless: α is a measure of the overall adjustment time of the individual headways, while β measures how proactively drivers react to changes of the relative velocity. Note that setting $(\alpha, \beta) = 0$ recovers the standard optimal-velocity model [8].

We consider the situation of N cars driving on a circular road of fixed length L under the generalized OV model described above which is achieved by adding the periodicity conditions

$$p_{n+N} = p_n + L, \quad s_{n+N} = s_n \ . \tag{5.30}$$

It was shown in [16] that the model (5.29) admits both free-flow and traffic jam (traveling wave) solutions. In this paper, we take the following choice of the parameters $V_g, V_0, \ell_0, \bar{s}, \tau$ which were obtained from Japanese highway data ([7, 30]):

$$V_g = 37.57 \text{ mi/hr}, \quad V_0 = 34.30 \text{ mi/hr}, \quad \ell_0 = 38.15 \text{ ft}, \quad \tau = 0.5 \text{ s}, \quad \bar{s} = 82.00 \text{ ft} \ , \tag{5.31}$$

and we consider two different scenarios of highway traffic, to be described below, obtained by different choices of the parameters α, β, N, L and initial conditions. In the first scenario, a square wave evolves into a shock with gradually decreasing amplitude (a solution that can also qualitatively be found in the macroscopic LWR model), while the second scenario describes the emergence of a traveling wave with fixed amplitude (which cannot be generated by the LWR model).

Results (square wave scenario): The first scenario is that of a long circular road consisting of a large number of cars engaged in free flow interacting with a

small traffic jam region. We set $(\alpha, \beta) = (1, 4)$ and the length $L = 22$ miles. To initialize the positions of the cars, we place cars equally spaced along the road with a fixed density of 40 cars per mile except for a two mile stretch in which the density is increased to 130 cars per mile for a total of $N = 1060$ cars. This gives that the initial headways are $s_0^f = 131.95$ ft in the free flow region and $s_0^j = 40.61$ ft in the traffic jam region. The initial velocities of the cars are taken to be

$$v_0^f = V_g \tanh\left(\frac{s_0^f - \bar{s}}{\ell_0}\right) + V_0, \quad v_0^j = V_g \tanh\left(\frac{s_0^j - \bar{s}}{\ell_0}\right) + V_0 \ , \tag{5.32}$$

for the free flow and jam regions, respectively. We initialize the system with the above conditions and compute the local density at each time step using a circular kernel density estimator in Matlab. In terms of density, the initial condition resembles a square wave which evolves into a shock with decreasing amplitude over time (Figure 5.8).

The simulation results for the first scenario are shown in Figure 5.8. We observe that the estimated density (green circled line) gradually converges to the true value (blue starred line). This is plausible as the viscous Lighthill–Whitham model can generate a traveling shock with decreasing amplitude as demonstrated in §5.2. With the aid of data assimilation, the macroscopic model provides an accurate estimation for the traveling shock wave transferred from microscopic traffic data.

6. Truth generated from microscopic models

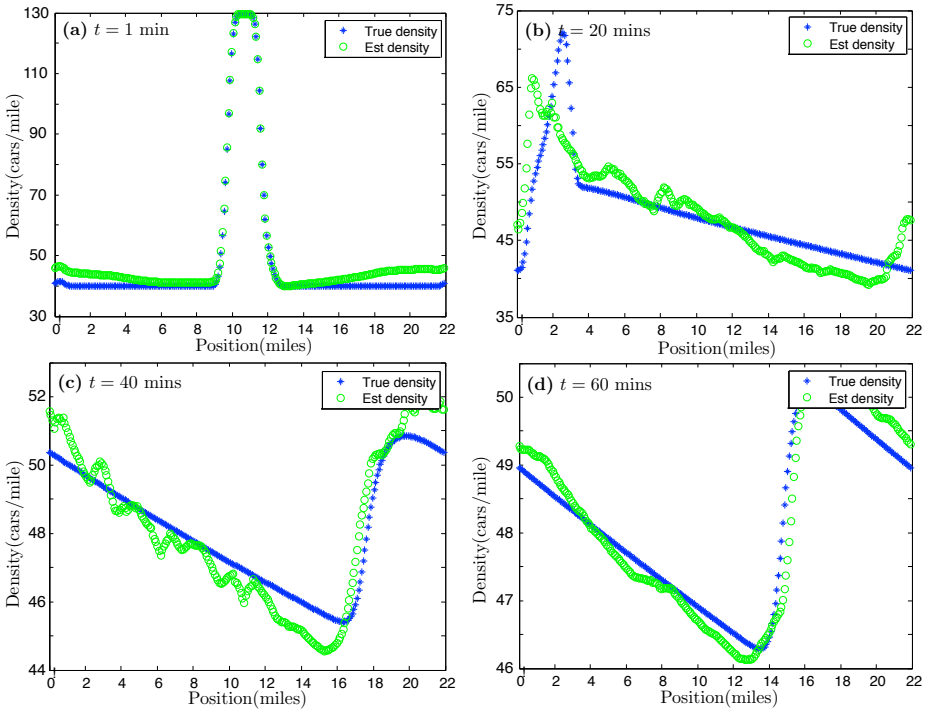


Figure 5.8: Estimating microscopic data in scenario 1 using the EnKF. True and estimated density values are represented by blue pluses and green circles, respectively.

Results (traveling wave scenario): The second scenario describes the propagation of a traffic jam or traveling wave solution. We set $(\alpha, \beta) = (4, 0.01)$ and the length $L = 2.5$ miles. We place $N = 150$ cars offset from an equal spacing of length $(n-1)L/N$ on the road with initial positions and velocities

$$p_n(0) = \frac{(n-1)L}{N} + 10\sin\left(\frac{2\pi(n-1)}{N}\right), \quad v_n(0) = V_g \tanh\left(\frac{s_0 - \bar{s}}{\ell_0}\right) + V_0, \tag{5.33}$$

where $s_0 = 87.97$ ft is the initial average headway. Figure 5.9 shows the evolution of the local density of this system. After a short time, a traffic jam, or traveling wave, emerges that travels backwards against the flow of traffic.

Also shown in Figure 5.9 are the results of the simulation using data assimilation. We observe that the estimated density (green circles) is smoothed out rapidly and does not converge to the true density value (blue pluses). Unlike the previous case, the macroscopic model cannot provide an accurate estimation even though data assimilation is used. This is not unexpected as the viscous Lighthill–Whitham model cannot generate a traveling square wave shape with stable amplitude and is therefore not capable of reproducing the correct density profile.

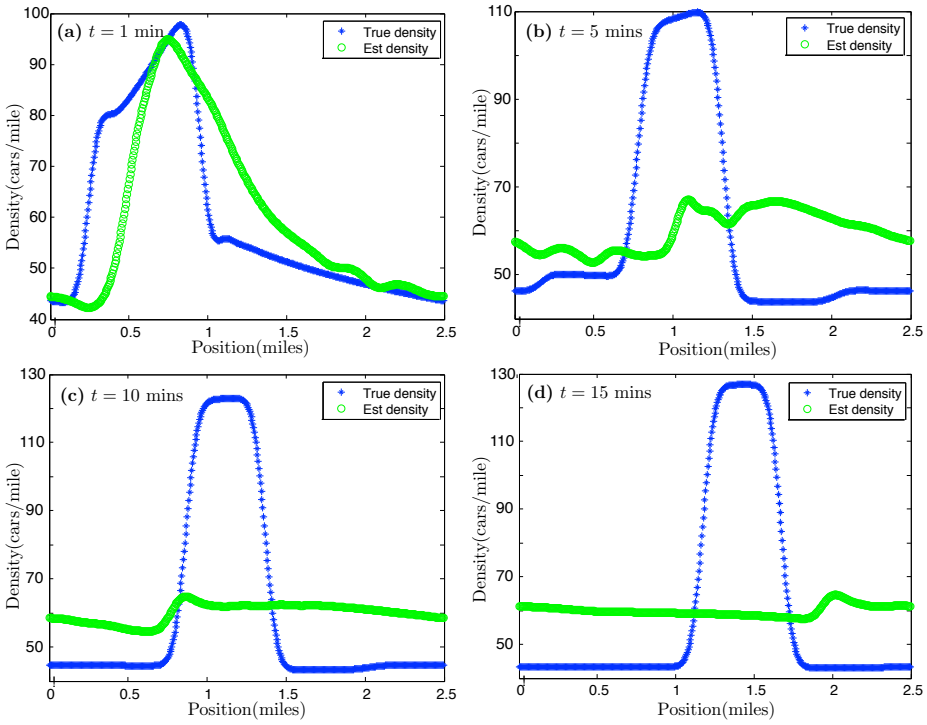


Figure 5.9: Estimating microscopic data in scenario 2 using the EnKF. True and estimated density values are represented by blue pluses and green circles, respectively.

In summary, a macroscopic model may produce inaccurate results if the model is not capable of generating dynamical features present in the traffic flow data that need to be assimilated.

2. Parameter estimation for real traffic data In addition to estimating simulated traffic states and parameters, we are interested in applying the developed

approach to real traffic data. In this section, we use data from the Minnesota Department of Transportation [70].

A freeway strip of I-35E with length 7.175 miles between Main Street and Hwy 96E is considered. There are 14 consecutive detectors S1535-S1548 along this stretch, as well as three on-ramps and one off-ramp (see Figure 5.10). We picked two representative time periods from a day for simulation: one is late night (00:01am-00:30am) when free flow is expected, and the other is rush hour (05:31pm-06:00pm) when traffic jams are expected. The Minnesota Department of Transportation provides traffic density, velocity and flow data collected from the sensors, and the density data of ramps. All traffic data are collected each minute.

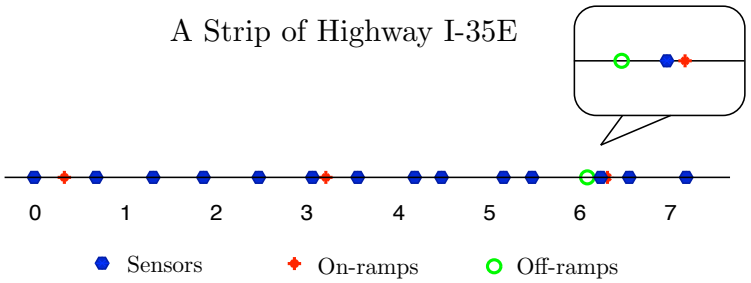


Figure 5.10: A strip of highway I-35E in Minnesota. Cars are moving from left to right. Sensors, on-ramps, and off-ramps are labelled using blue stars, red pluses, and green circles, respectively.

We use a macroscopic model that takes the on- and off-ramps into account, and assimilate the flux data collected from the sensors using the EnKF. We also estimate the unknown maximal density $\rho_{\max}$ and the maximal velocity $v_{\max}$. The density profiles of the ramps are artificially taken as unknown and are replaced by the average density value during a period. In contrast to the simulation presented in previous sections, we use an adaptive inflation algorithm [2] here.

Results: The simulation results for late night traffic are shown in Figure 5.11. The relative RMSE by assimilating all observation approaches 15%, while the rela-

tive RMSE by not assimilating observations approaches 40%. The parameter estimation in 5.11(b) shows that estimated maximal velocity $\hat{v}_{\max}$ is approximately 85 miles/hour, and the maximal density $\hat{\rho}_{\max}$ decreases to below 30 cars/mile. This is a reasonable description of late night traffic.

The simulation results for rush hour traffic are shown in Figure 5.12. The approach results in a decrease of error from 15% to 8%. Since the flow of ramps is quite stable during rush hour, the average density value for ramps are close to the true density value which explains why our approach results in a less drastic improvement when compared to the previous case. In Figure 5.12(b), the estimated maximal velocity $\hat{v}_{\max}$ is still approximately 85 miles/hour, but the maximal density $\hat{\rho}_{\max}$ increases to 150 cars/mile, which is characteristic of traffic flow during rush hour.

In summary, the efficacy of our approach has been validated by the real traffic data from Minnesota Department of Transportation. Estimation for traffic states is significantly improved in both late night and rush hour conditions, and the estimation for parameter is consistent with characteristics of traffic scenarios.

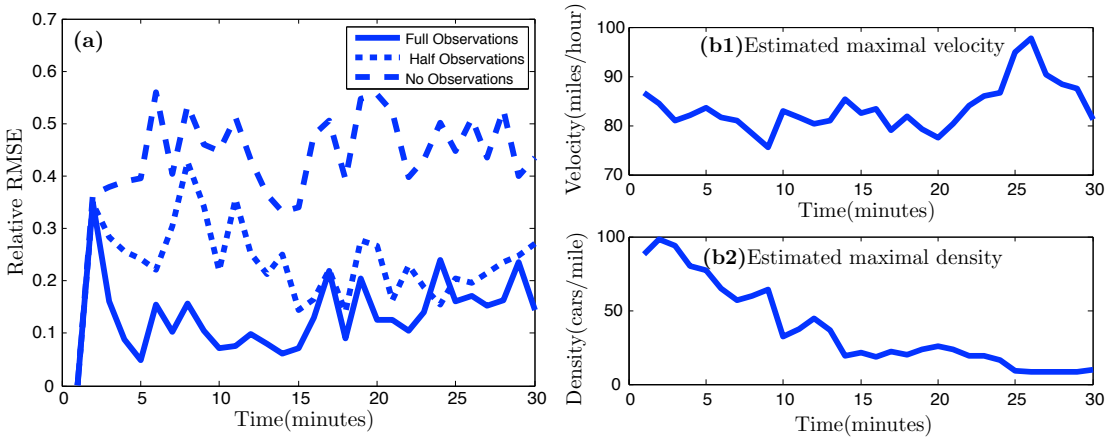


Figure 5.11: Relative RSME (a) and parameter estimate (b) are shown for data assimilation of real traffic data taken during late night. Observations are taken from all sensors (solid), half the sensors (dotted), or no sensors (dashed).

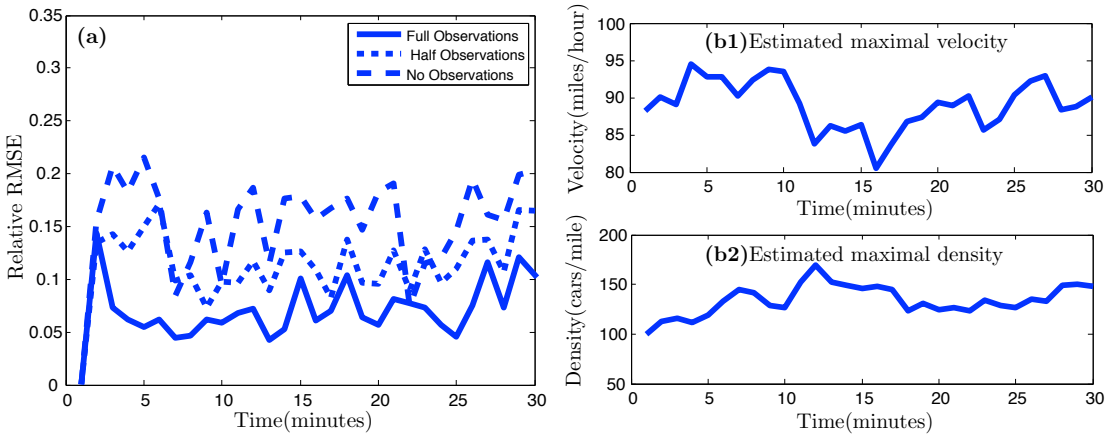


Figure 5.12: Relative RSME (a) and parameter estimate (b) are shown for data assimilation of real traffic data collected during rush hour. Observations are taken from all sensors (solid), half the sensors (dotted), or no sensors (dashed).

5.6 Discussion

We studied several aspects of traffic flow prediction: mathematical traffic models, observation models, and data assimilation techniques. Specifically, we reviewed the Lighthill-Whitham macroscopic model that is used as the basic underlying traffic model in data assimilation (§5.2). We then discussed two types of observations: Eulerian sensor data and Lagrangian GPS data, and developed a formulation that is capable of assimilating the Eulerian and Lagrangian observations simultaneously (§5.3). We also investigated how our approach could be used to estimate traffic states as well as parameters using the EnKF and PF (§5.4).

The initial motivation of this paper was to propose an efficient approach that could assimilate both Eulerian sensor data and Lagrangian GPS data simultaneously. The idea behind our approach is to append the differential equations for the positions and velocities of the vehicles to the macroscopic traffic model in order to solve them simultaneously. Unlike the previous methods, our approach is capable of handling flux data, position data and velocity data efficiently without reformulation

in Lagrangian coordinates or reshaping the macroscopic traffic model. This approach has been shown to accurately estimate traffic states in different traffic scenarios with true underlying traffic model (Case 1, Case 4). We also studied how the choice of filters and observations affect data assimilation: compared to the EnKF, the PF is less sensitive to observation noise and sensor locations, but its computation cost is higher (Case 2); the positions of sensor locations impact the accuracy of the EnKF, and relevant suggestions for installing sensors are provided (Case 3). However, this approach is less accurate when the true traffic model is unknown and an estimated underlying traffic model is used (Case 5, Case 6).

One limitation we reported on in §5.5.3 is that the accuracy of data assimilation is reduced significantly if the underlying traffic model cannot reproduce the actual traffic flow from which observations are collected: in such a situation, the model error is very large and will therefore dominate the overall error. Hence, to increase accuracy in such cases requires the development of better, more accurate traffic flow models that are capable of reproducing a wider range of traffic flows.

Another limitation is that the proposed approach for assimilating Lagrangian observations requires knowledge of the velocity of a vehicle at a given position: in our macroscopic model for the density, the velocity was given explicitly as a function of the density, which allowed us to access the velocity in our algorithm. If the underlying traffic-flow model does not provide information about the velocity of vehicles at any specified position, our method for assimilating Lagrangian observations will not work.

There are various extensions that could be implemented to improve the framework presented here. First, the proposed data assimilation approach was only applied to simplified traffic scenarios such as on and off ramps, traffic lights, and bottlenecks that were implemented in the LWR equation. In practice, the road network is typ-

ically more complicated and may include, for instance, intersections and multiple lanes. An exciting extension would be to expand the above framework to more complex and realistic traffic scenarios. Second, it would be interesting to see whether the algorithm could be extended to use parameter estimation to relay traffic information to drivers. For instance, the period of traffic lights could be estimated by data assimilation, and thus the waiting time could be provided to drivers. Data assimilation could also be used to predict the location and duration of congestion caused by slow trucks or road constructions.

Chapter Six

Traffic control in lane-drop scenarios

6.1 Introduction

A lane drop is defined as a location on a highway where the number of lanes provided decreases. The reasons for a lane drop can be varied: the road construction may close the usage of some lanes during road work; an unexpected traffic accident also restricts the usage of the shoulder lane; the highway may reduce the number of lanes when space is limited, and so on.

The merge bottleneck from lane drop is an interesting traffic phenomena. An increase in lane-changing maneuvers will form queues in the upstream of the lane drop, and the queues are propagating backward. Traffic was studied upstream and downstream of a bottleneck in the lane drops. Recent empirical observations [10] at freeway merge bottlenecks revealed: a drop in the bottleneck discharge rate when queues form upstream, which is also called capacity drop. Further, it is shown that the bottleneck's discharge flow is about 10% lower than the prevailing flow observed prior to queue formation. Upon bottleneck activation, flow reduction occurs sequentially in time and space. In [53], numerical simulation results were in agreement with the findings of the empirical studies at merge bottlenecks. In addition, it strongly suggested that lane changes are the main cause of the drop in discharge rate.

Two types of approaches have been considered to reduce the capacity drop of merge bottleneck. One type of approach attempts to develop cooperative lane-changing models so that the lane-changing behavior is more efficient. Like in [41, 48], cooperative lane-changing systems were developed microscopically in order to smooth the merges. However, this microscopic approach is complicated to specify and there are lot of parameters to quantify. The other type of approach [52] embeds the vehicles as discrete particles in a multi-lane stream using the kinematic wave model.

It attempts to maximize capacity through efficiently controlling the lane-changing ratio between adjacent lanes. However, this macroscopic approach formulates a conservation equation for individual lane.

In this chapter, we would like to introduce a macroscopic approach to model and analyze the multi-lane traffic, by using only one continuity equation. Then we will develop controlling strategies based on this continuity equation to maximize the road capacity in lane drop. The chapter is organized as follows. In § 6.2, we describe the inhomogeneous LWR equation that used to model multi-lane traffic, and analyze the solutions to Riemann problems. In § 6.3, we develop two controlling strategies in order to maximize road capacity in lane drop. Then we conclude with a discussion of our results and open problems in § 6.4.

6.2 Inhomogeneous LWR model

In this section we will first briefly review the homogeneous LWR model, and then introduce the inhomogeneous LWR model to describe the multi-lane traffic by taking the number of lanes into consideration(§ 6.2.1). In addition, we analyze the solutions to the Riemann problems of inhomogeneous LWR equation by applying the supply–demand theory (§ 6.2.2). In § 6.2.3, we explore the traffic pattern on a ring road and the efficiency of LWR equation to organize the road capacity.

The LWR model studies the traffic dynamics based on the conservation law. We have described the LWR model in § 2.2. In this model, the traffic is described by

the vehicle density ρ and vehicle flux φ:

$$\frac{\partial \rho}{\partial t} + \frac{\partial \varphi}{\partial x} = 0, \tag{6.1}$$

where flux is a function of density via $\varphi(\rho)$. The solutions to the Riemann problems of homogeneous LWR model are already discussed in § 2.2 using the method of characteristics, which include traveling wave, shock wave and rarefaction. In addition, we generated multiple traffic scenarios by adding additional terms to capture the respective effects.

The vehicle flow is defined as a function of traffic density ρ: $\varphi = \varphi(\rho)$, which is independent of the location x. Therefore, the LWR model is only for modeling a homogeneous road, where flow function is uniform with respect to location x. That is, the LWR model has limited modeling capacity when the traffic scenario becomes complicated.

In this section, we will focus on the traffic scenarios with multiple lanes, where the number of lanes are location dependent, and introduce the inhomogeneous LWR model to account for the changes of the number of lanes.

6.2.1 Model description

We are interested in modeling the traffic scenarios with multiple lanes macroscopically. One well-known method to model multi-lanes was developed by Daganzo in [21, 22, 53]. It formulated a conservation equation for each lane, and the right hand side of each conservation equation is the net lane-changing rate onto the current lane from adjacent lanes. The lane-changing rate consists of the sources from neighbor

lanes and sink to neighbor lanes, which depends on the nearby traffic at a specific location and time.

However, the interaction between lanes can be very complicated to quantify in lane-drop scenarios if we use a conservation equation for each lane. In addition, we are only interested in the dynamics of the aggregated traffic characteristics instead of the inner interaction between lanes. Therefore, we use another approach mentioned in [73] (chapter 7.2), which formulates the averaged density and flux with only one continuity equation:

$$\frac{\partial(I\rho)}{\partial t} + \frac{\partial(I\varphi)}{\partial x} = 0 \tag{6.2}$$

where ρ is the averaged density per lane and φ is the averaged flux per lane. The function $I(x)$ is the number of lanes at location x, which models the the lane changes before a lane closure or after a new lane open. The number of lanes is integer in reality, but $I(x)$ can be non-integer to account for partial lane usage. For example, a value of $I = 1.5$ of a realistic 2-lane road indicates the second/shoulder lane is less frequently used, and the flow on this lane is only half of the averaged flow on the other lane.

By introducing the notations for total density and total flux

$$\rho_{\text{total}}(x,t) = I(x)\rho(x,t),$$
$$\varphi_{\text{total}}(x,t) = I(x)\varphi(x,t),$$

the conservation law (6.2) for multi-lane traffic can be rewritten in a form that is similar to the homogeneous LWR model:

$$\frac{\partial \rho_{\text{total}}}{\partial t} + \frac{\partial \varphi_{\text{total}}}{\partial x} = 0. \tag{6.3}$$

Unlike the homogeneous LWR model, the total vehicle density ρ_{total} and total vehicle flux φ_{total} are not directly related with respected to a fixed flux-density relation. Instead, they are related corresponding to a lane-dependent flux-density relation:

$$\begin{aligned}\varphi_{\text{total}}(x,t) &= \varphi_{\text{total}}(\rho_{\text{total}}, I(x)) \\ &= I(x)\varphi\left(\frac{\rho_{\text{total}}}{I(x)}\right). \end{aligned} \tag{6.4}$$

A lane-drop traffic scenario

In the whole chapter, we will focus on a simplified lane-drop traffic scenario where it changes from two lanes to one lane. The analysis of solutions and traffic controlling strategies in the following sections are also based on this specific lane-drop scenario. Extension to generalized multi-lane traffic scenarios will be discussed in final section § 6.4. The lane-drop scenario is shown in the figure below:

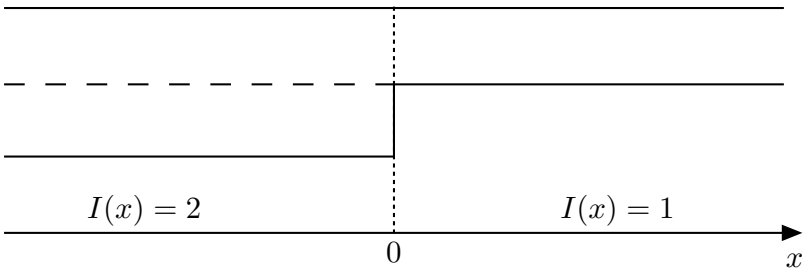


Figure 6.1: Lane-drop scenario from two lanes to one lane.

The lane drop is at the origin with $x = 0$, and there are two lanes in the upstream $x < 0$ and one lane in the downstream $x > 0$.

The fundamental diagram is a modeling choice, which is the plot of traffic flux φ versus traffic density ρ. According to [64], the two-phase traffic theory divides traffic flow into free flow for low densities, and congested flow for large densities,

which together form a reversed λ shape. The fundamental diagram we use to model multi-lane traffic is shown in Figure 6.2. The shape of the fundamental diagram is determined by the parameters from [24]. (see model calibration in § 4.2)

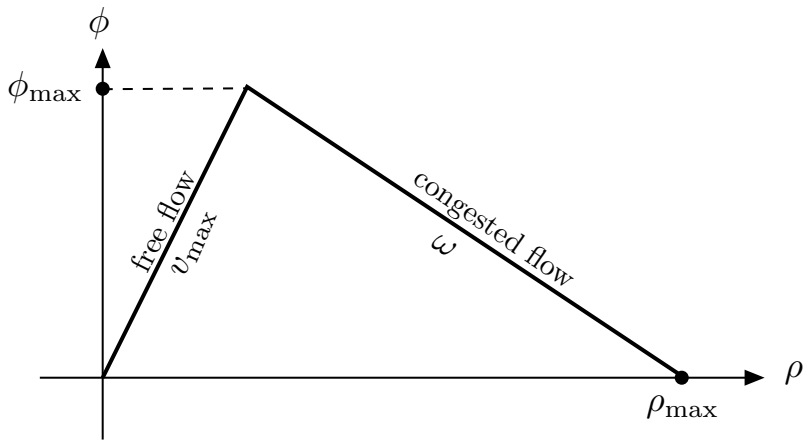


Figure 6.2: Fundamental diagram $\varphi(\rho)$. The slope of the free flow region is the maximal velocity $v_{\max}$, and the slope of the congested flow is the congestion parameter ω. The maximal density is $\rho_{\max}$, and the maximal flux is $\varphi_{\max}$.

Description	Parameter	Value	Unit
Free-flow speed	$v_{\max}$	63.3	miles/hour
Maximal flux	$\varphi_{\max}$	2031	vehicles/hour
Congestion speed parameter	ω	-10.1	miles/hour
Jam density	$\rho_{\max}$	232	vehicles/mile

Table 6.1: Parameters from auto-calibration of the fundamental diagram

The critical density ρ_{crit} is the divider of free flow and congested flow in the fundamental diagram. It is easy to calculate the critical density via

$$\rho_{\text{crit}} = \mathcal{Q}_{\max}/v_{\max} \approx 32.09 \text{vehicles/mile} \tag{6.5}$$

The critical density is a very important bifurcation point in multi-lane traffics. We will talk about it with more details when analyzing the solution to multi-lane traffic in the following subsection.

By using the Equation (6.4), we obtain the fundamental diagram for multiple-lane traffic. It is of the same reversed λ shape, but scaled in both x-axis and y-axis according to the lane number $I(x)$. The Figure 6.3 shows the fundamental diagram for the simplified lane-drop scenario in Figure 6.1.

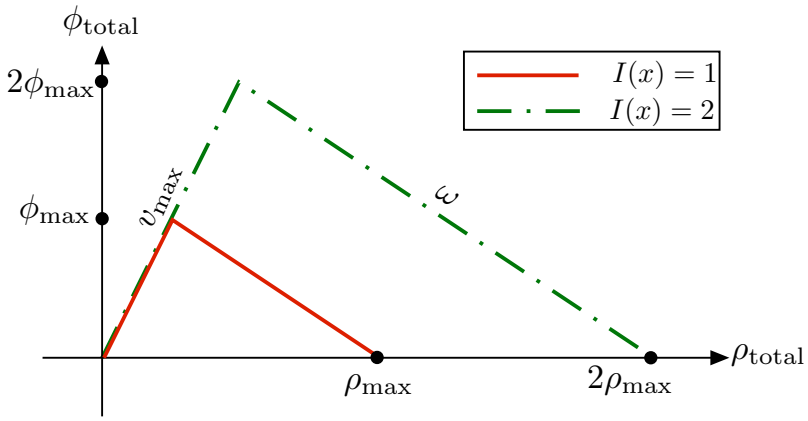


Figure 6.3: The flux-density relation of $\varphi_{\text{total}}(\rho_{\text{total}}, I(x))$ in lane-drop scenario. The red solid line represents the fundamental diagram for road of one lane, while the green dashed dotted line is for road of two lanes.

6.2.2 Solutions to Riemann problems

Next, we will investigate the solutions to the inhomogeneous LWR equation. The problem with discontinuous initial traffic states and lane number of the following form:

$$\rho_{\text{total}}(x,0) = \begin{cases} \rho_L, & x \le 0 \\ \rho_R, & x > 0 \end{cases}, \quad I(x) = \begin{cases} 2, & x \le 0 \\ 1, & x > 0 \end{cases}. \tag{6.6}$$

is called a Riemann problem to the inhomogeneous LWR model. In § 2.2, we have reviewed the characteristic method, and used it to solve the homogeneous LWR model. The solutions include traveling wave, shock wave and rarefaction. For the inhomogeneous LWR model (6.2), there are some analytical and numerical studies [42, 45, 47] to investigate the solutions to Riemann problems. In our study, we adopt

the method called wave separation, which was introduced in article [42]. Specifically, the separation method includes:

(i) first, we introduce a stationary wave between the road branches of different number of lanes. The stationary wave is subject to entropy condition. More details about the entropy condition between road branches are explained as follows.

(ii) then, we use the characteristic method to solve the homogeneous LWR model for each road branch with constant number of lanes. The solution for each road branch is also subject to the entropy condition. The entropy condition for homogeneous LWR is listed earlier in § 2.2.

According to the article [42], the solutions to Riemann problems of the inhomogeneous LWR model exist and are unique if the following two entropy conditions are satisfied:

Entropy 1: the waves from left (upstream) to right (downstream) should increase their wave speeds so that they don't cross each other.

Entropy 2: the standing wave can not cross the transition curve $\rho_{\text{total}}/I(x) = \rho_{\text{crit}}$. The left side of the transition curve is $\rho_{\text{total}}/I(x) < \rho_{\text{crit}}$, and the right side of the transition curve is $\rho_{\text{total}}/I(x) \geq \rho_{\text{crit}}$.

According to [47], the entropy condition can be transferred into a supply-demand framework. Supply–demand framework is relatively easy to understand the traffic flow. Another advantage about supply–demand framework is that it can guide you to find standing wave between branches through analyzing the boundary flux.

Therefore, we will use it to construct the standing wave that satisfies the entropy conditions, and then get the unique solution to the Riemann problem. Next let us briefly review the supply–demand theory in traffic.

Supply and demand function

We start with showing the supply and demand function in multi-lane traffic scenarios. The traffic supply function computes the traffic flow to supply for the upstream cell: the supply of the free flow is the maximal flow, while the supply of the congested flow is only itself. The mathematical expression is as follows:

$$\begin{aligned} S(\rho_{\text{total}}) &= I(x)\varphi(\max(\rho_{\text{crit}}, \rho_{\text{total}}/I(x))) \\ &= \begin{cases} I(x)\varphi(\rho_{\text{crit}}), & \rho_{\text{total}}/I(x) < \rho_{\text{crit}} \\ I(x)\varphi(\rho_{\text{total}}/I(x)), & \rho_{\text{total}}/I(x) \geq \rho_{\text{crit}} \end{cases} \end{aligned} \tag{6.7}$$

The traffic demand function computes the traffic flow to demand from the downstream cell: the demand of the free flow is itself, while the demand of the congested flow is the maximal flow. The mathematical expression is as follows:

$$\begin{aligned} D(\rho_{\text{total}}) &= I(x)\varphi(\min(\rho_{\text{crit}}, \rho_{\text{total}}/I(x))) \\ &= \begin{cases} I(x)\varphi(\rho_{\text{total}}/I(x)), & \rho_{\text{total}}/I(x) \leq \rho_{\text{crit}} \\ I(x)\varphi(\rho_{\text{crit}}), & \rho_{\text{total}}/I(x) \geq \rho_{\text{crit}} \end{cases} \end{aligned} \tag{6.8}$$

Let two adjacent cells be i and $i+1$ with total vehicle density $\rho_{\text{total},i}$ and $\rho_{\text{total},i+1}$. According to the supply–demand theory, the boundary flux is determined by the supply from downstream and the demand from upstream. Specifically, the boundary flux from cell i to cell $i+1$ is the smaller of the supply of the cell $i+1$ and the

demand of the cell i. This is in agreement to our intuition.

$$\varphi^*_{i,i+1} = \min(S(\rho_{\text{total},i+1}), D(\rho_{\text{total},i})) \tag{6.9}$$

The Godunov method is used to solve the LWR model, which is a conservative numerical scheme suggested by S. K. Godunov [33]. It is a conservative finite-volume method which solves exact, or approximate Riemann problems at each inter-cell boundary. The Godunov method is equivalent to the supply–demand theory when calculating the boundary flux. The formula of Godunov method can be rewritten as:

$$\rho^{j+1}_{\text{total},i} = \rho^{j}_{\text{total},i} - \frac{\Delta t}{\Delta x}\left(\varphi^*_{i,i+1} - \varphi^*_{i-1,i}\right) \tag{6.10}$$

Analytic solution to Riemann problem

It is convenient to find the unique solution to the Riemann problem by using the idea of boundary flux. Given two initial vehicle density ρ_L and ρ_R in the Riemann problem, we can first calculate the boundary flux at the origin $x = 0$. That is $\varphi^*_{L,R} = \min(S(\rho_R), D(\rho_L))$ by the supply–demand theory. As the standing wave has speed 0 at the origin, the flux close to the lane drop must equals to the boundary flux.

Next we have to figure out the two transitional vehicle densities in the upstream ρ_A and downstream ρ_B of the standing wave. At least one of the transitional densities

can be determined by the calculated boundary flux as follows:

$$\text{if } \varphi^*_{L,R} = D(\rho_L) \qquad \Longrightarrow \qquad \rho_A = \min(\rho_{\text{crit}}, \rho_L/I(x)) \tag{6.11}$$

$$\text{if } \varphi^*_{L,R} = S(\rho_R) \qquad \Longrightarrow \qquad \rho_B = \max(\rho_{\text{crit}}, \rho_R/I(x)) \tag{6.12}$$

For instance, if $\varphi^*_{L,R} = D(\rho_L) < S(\rho_R)$, then we can figure out ρ_A by using the formula above. Next, we have to specify the other transitional density ρ_B. First, the vehicle flow corresponding to ρ_B should be the same as the boundary flow $\varphi^*_{L,R}$ because of the standing wave at origin. As the fundamental diagram has a reversed λ shape, there are at most two admissible densities for ρ_B. Then we can apply the two entropy conditions **Entropy 1** and **Entropy 2** to pick the unique transitional density for ρ_B.

The last step is to solve two Riemann problem for homogeneous LWR equations in the upstream and downstream branches. The characteristic method can be applied here. The two Riemann problems are:

$$\rho(x,0) = \begin{cases} \rho_L, & x \leq 0 \\ \rho_A, & x = 0- \end{cases} \tag{6.13}$$

$$\rho(x,0) = \begin{cases} \rho_B, & x = 0+ \\ \rho_R, & x > 0 \end{cases} \tag{6.14}$$

The typical density solution to the Riemann problem is sketched in Figure 6.4. There is a standing wave in the origin that is subject to entropy condition. Two wave solutions are in upstream and downstream based on the characteristic method.

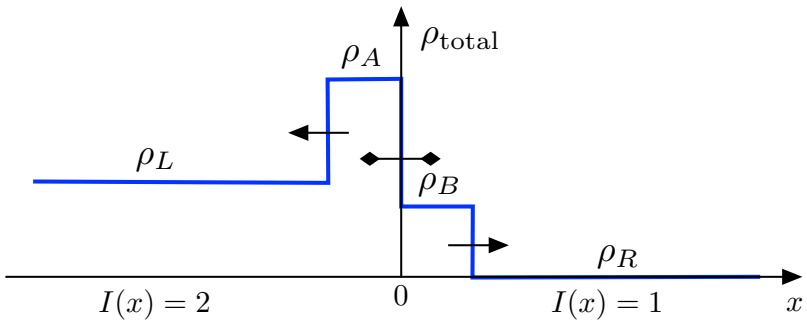


Figure 6.4: The wave solution to Riemann problem of inhomogeneous LWR. ρ_L and ρ_R are the initial conditions. ρ_A and ρ_B are the transitional traffic states in the upstream and downstream of the standing wave.

Examples

We will show how to apply the wave separation method to solve the following Riemann problem:

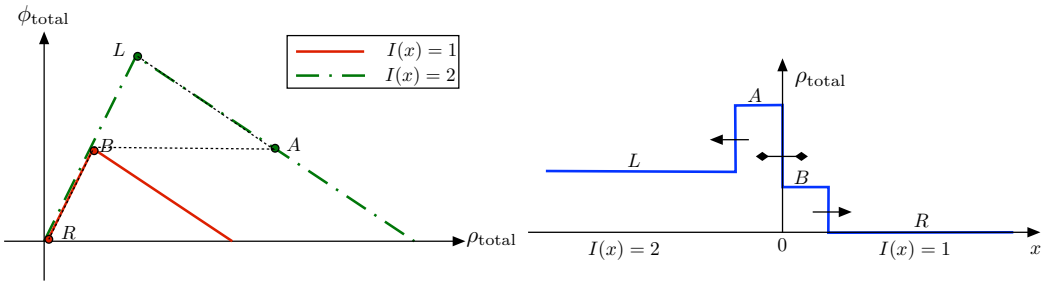
$$\rho_{\text{total}}(x,0) = \begin{cases} 2\rho_{\text{crit}}, & x < 0 \\ 0, & x > 0 \end{cases}, \qquad I(x) = \begin{cases} 2, & x < 0 \\ 1, & x > 0 \end{cases}. \tag{6.15}$$

First, we have to compute the boundary flux by using the supply–demand theory:

$$D(\rho_L) = 2\varphi_{\max}$$
$$S(\rho_R) = \varphi_{\max}$$
$$\varphi^*_{L,R} = \min(S(\rho_R), D(\rho_L)) = \varphi_{\max}$$

The computations show that the standing wave has boundary flux of $\varphi_{\max}$. The downstream transitional vehicle density is $\rho_B = \max(\rho_R, \rho_{\text{crit}}) = \rho_{\text{crit}}$ because of $\varphi^*_{L,R} = S(\rho_R)$. The upstream transitional vehicle density can be ρ_{crit} or $\rho_{\text{crit}} + \rho_{\max}$ because their respective vehicle fluxes equal the boundary flux $\varphi^*_{L,R}$. By checking the two entropy conditions, we find that only $\rho_A = \rho_{\text{crit}} + \rho_{\max}$ satisfies the entropy conditions.

The initial traffic conditions as well as the transitional traffic states are indicated in the fundamental diagram in Figure 6.5(a). The wave solution by using the separation method is illustrated in Figure 6.5(b). It is a shock wave between state L and A, which is moving backward with speed $\omega = \frac{\varphi_{\text{total}}(\rho_L,2)-\varphi_{\text{total}}(\rho_A,2)}{\rho_L-\rho_A}$. The standing wave is stationary at the origin. It is a traveling wave between state B and R, which is moving forward with speed $v_{\max}$. We will justify that this weak solution satisfies both entropy conditions.



(a) Traffic states in fundamental diagram. **(b)** Wave solution in density ρ_{total}.

Figure 6.5: The weak solution to Riemann problem (6.15) of inhomogeneous LWR model.

Entropy 1: the waves from left (upstream) to right (downstream) should increase their wave speeds so that they don't cross each other. The waves' speeds from upstream to downstream are $\omega, 0$ and $v_{\max}$. It is obvious that the relation $\omega < 0 < v_{\max}$ holds.

Entropy 2: the standing wave can not cross the transition curve $\rho_{\text{total}}/I(x) = \rho_{\text{crit}}$. The two transitional states ρ_A and ρ_B around the standing wave both satisfy $\rho_{\text{total}}/I(x) \geq \rho_{\text{crit}}$. That means they do not across the transition curve.

6.2.3 Traffic pattern on a ring road

With the ability to solve the Riemann problems of inhomogeneous LWR model, we can explore the traffic in more complicated road structure. We are most interested in the traffic pattern by the inhomogeneous LWR on a ring road with two branches: one branch is of one lane, and the other branch is of two lanes. On one side, we could observe how the inhomogeneous LWR model generates congestion pattern. On the other side, we would like to understand how LWR model efficiently organizes and distributes traffic.

The ring road profile is shown in Figure 6.16. There are two road branches: the one-lane branch is of length L_1 and the two-lane branch is of length L_2. Here, we just set the length ratio $L_2 : L_1 = 1 : 1$ for investigating the traffic pattern, and then make a conclusion for a general case in the end of this subsection.

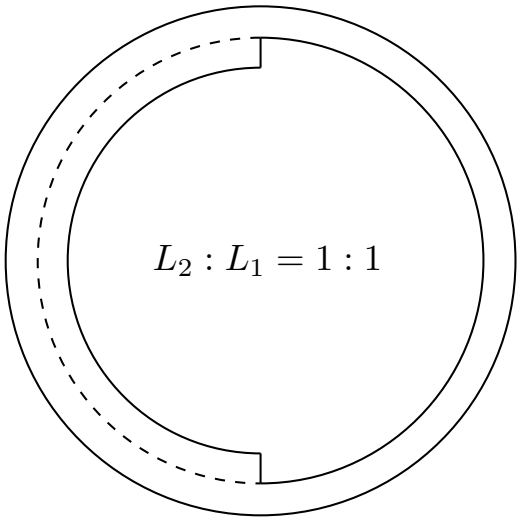


Figure 6.6: A ring road with two branches: one-lane branch with length L_1 and two-lane branch with length L_2. The ratio is $L_2 : L_1 = 1 : 1$.

The initial densities on the ring road are described in equation (6.16) and Figure 6.7. This Riemann problem has periodic boundary condition at $x = L_1/-L_2$

(we will use $x = L_1$ to represent the boundary in the following discussion).

$$\rho_{\text{total}}(x,0) = \begin{cases} \rho_2, & -L_2 < x < 0 \\ \rho_1, & 0 < x < L_1 \end{cases}, \qquad I(x) = \begin{cases} 2, & -L_2 < x < 0 \\ 1, & 0 < x < L_1 \end{cases}. \tag{6.16}$$

Road capacity refers the maximal traffic flow (vehicles per hour) obtainable on a

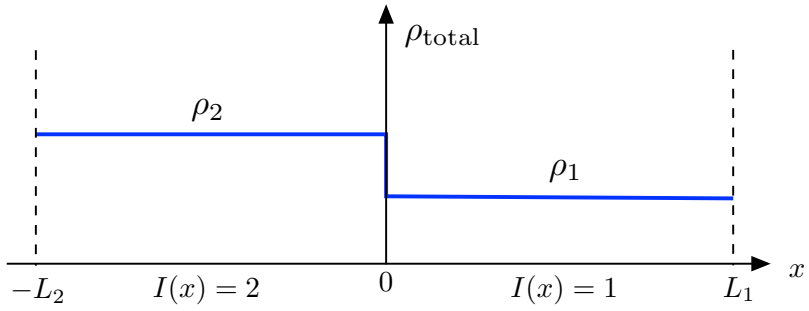


Figure 6.7: Initial conditions of the Riemann problem on a ring road with two branches. The first branch is of two lanes from $-L_2$ to 0, and the second branch is of one lane from 0 to L_1. The boundary is periodic.

given road using all available lanes. Therefore, the two-lane branch has double road capacity of the one-lane branch. The road capacity of this ring road is restricted by the one-lane branch, so the overall road capacity is only $\varphi_{\max}$. In the fundamental diagram of the ring road, there are two points that can provide the exact road capacity $\varphi_{\max}$. Please see Figure 6.8 for positions of these two critical points C_1 and C_2. The analysis shows that the averaged density $(\rho_1 + \rho_2)/2$ determines the traffic pattern of the ring road, and there are three regimes controlled by the points C_1 and C_2. The three regimes are:

$$\text{Regime 1: } \frac{\rho_1 + \rho_2}{2} \le \rho_{C_1} \tag{6.17}$$

$$\text{Regime 2: } \rho_{C_1} < \frac{\rho_1 + \rho_2}{2} \le \frac{\rho_{C_1} + \rho_{C_2}}{2} \tag{6.18}$$

$$\text{Regime 3: } \frac{\rho_1 + \rho_2}{2} > \frac{\rho_{C_1} + \rho_{C_2}}{2} \tag{6.19}$$

$$\text{with} \quad \rho_{C_1} = \rho_{\text{crit}}, \quad \rho_{C_2} = \rho_{\text{crit}} + \rho_{\max}$$

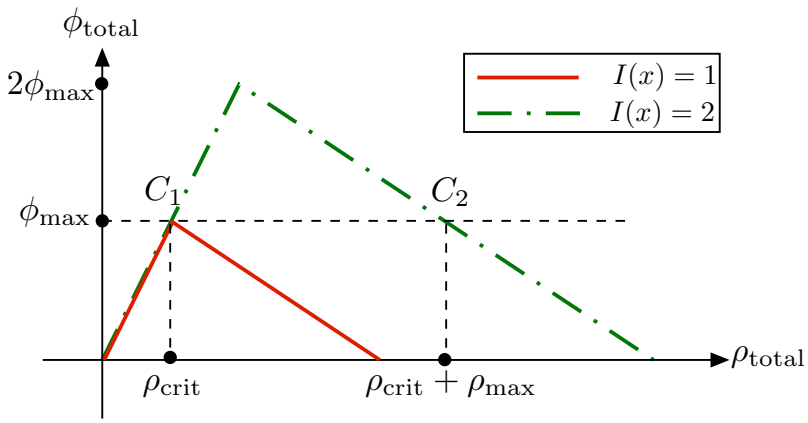


Figure 6.8: The fundamental diagram of the ring road. There are two points C_1 and C_2 which can provide the road capacity $\varphi_{\max}$.

Next we will show the traffic pattern in each regime as well as provide the analytic solution to the Riemann problems:

- **Regime 1:** $\frac{\rho_1+\rho_2}{2} \leq \rho_{C_1}$

  ρ_{C_1} is the critical density in the fundamental diagram, where maximal traffic flux is achieved. In a one-lane road, when the vehicle density is less than ρ_{C_1}, it is free flow; when the vehicle density is above ρ_{C_1}, it is congested flow. Therefore, the intuition tells us that it should be free flow everywhere if the initial average density is not more than the critical density ρ_{C_1}.

  For the Regime $\frac{\rho_1+\rho_2}{2} \leq \rho_{C_1}$, there is at least one density (ρ_1 or ρ_2) should be less than or equal to the critical density ρ_{C_1}. There are two sub cases for this Regime 1: (a) both ρ_1 and ρ_2 are less than or equal to the critical density; (b) only one density of ρ_1 and ρ_2 is less than or equal to the critical density.

  - Case 1: $\rho_1 \leq \rho_{C_1}$ and $\rho_2 \leq \rho_{C_1}$

    This is a trivial case. When the initial densities on both road branch are less than or equal to the critical density ρ_{C_1}, it is a traveling wave with

maximal speed $v_{\max}$. The traveling wave consists of two traffic states: ρ_1 and ρ_2. The distance between the two discontinuities are exactly the length L_1 of road branch.

In Figure 6.9, we show the traveling wave solution to the following Riemann problem (6.20) in which both ρ_1 and ρ_2 are less than or equal to the critical density ρ_{C_1}. The traveling wave in Figure 6.9 has two density states: $0.5\rho_{C_1}$ and $0.8\rho_{C_1}$. Even though we switch the initial density for two branches, the stationary wave solution is the same.

$$\rho_{\text{total}}(x,0) = \begin{cases} 0.5\rho_{C_1} \\ 0.8\rho_{C_1} \end{cases} \quad \text{or} \quad \begin{cases} 0.8\rho_{C_1}, & -L_2 < x < 0 \\ 0.5\rho_{C_1}, & 0 < x < L1 \end{cases}. \tag{6.20}$$

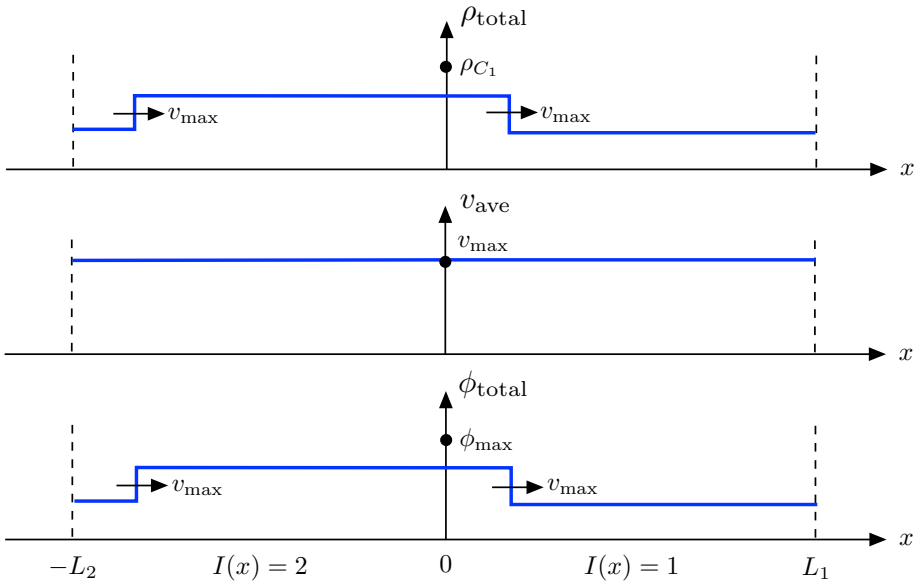


Figure 6.9: The weak solution to Riemann problem (6.20) on a ring road. The plots are vehicle density, velocity and flux from top to bottom.

– Case 2: $\rho_1 > \rho_{C_1}$ or $\rho_2 > \rho_{C_1}$

In this case, one of the branches has an initial density above the critical density ρ_{C_1}. Even though the analytic solution is more complicated in the

beginning, the stationary solution is still a traveling wave with maximal speed $v_{\max}$. The traveling wave consists of two traffic states: $\min(\rho_1, \rho_2)$ and ρ_{C_1}. The distance between the two discontinuities is not necessarily the length of roach branch, but is uniquely determined by the conservation of vehicles. Let D be the length with density ρ_{C_1}, then it must satisfy the relation $\rho_1 L_1 + \rho_2 L_2 = \min(\rho_1, \rho_2)(L_1 + L_2 - D) + \rho_{C_1} D$.

In Figure 6.10, we show that the traveling wave solution to the following Riemann problem (6.21) in which only one of ρ_1 and ρ_2 is less than or equal to the critical density ρ_{C_1}. The traveling wave in Figure 6.9 has two density states: $0.5\rho_{C_1}$ and $1\rho_{C_1}$. The length with density ρ_{C_1} is $1.4L_1$. Even though we switch the initial density for two branches, the stationary wave solution is the same.

$$\rho_{\text{total}}(x,0) = \begin{cases} 1.2\rho_{C_1} \\ 0.5\rho_{C_1} \end{cases} \quad \text{or} \quad \begin{cases} 0.5\rho_{C_1}, & -L_2 < x < 0 \\ 1.2\rho_{C_1}, & 0 < x < L1 \end{cases}. \tag{6.21}$$

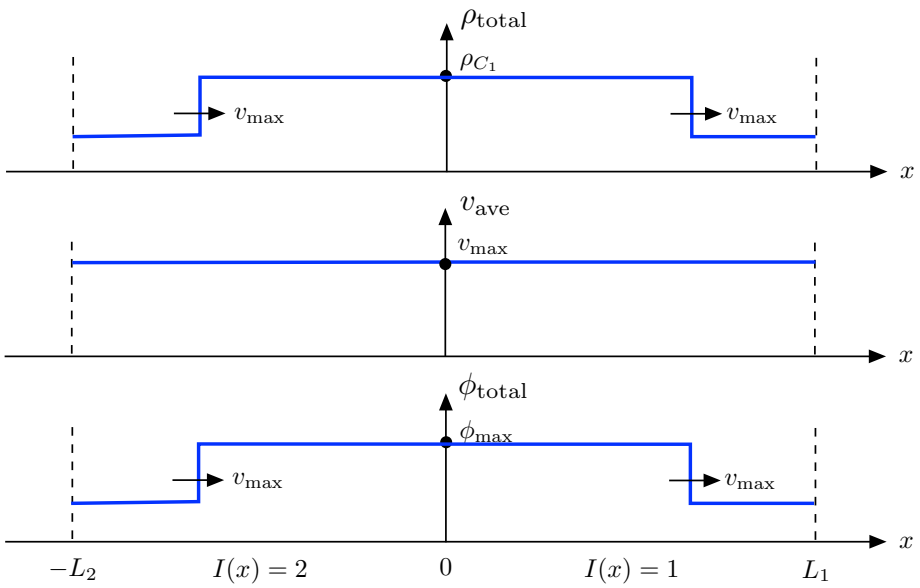


Figure 6.10: The weak solution to Riemann problem (6.21) on a ring road. The plots are vehicle density, velocity and flux curves from top to bottom.

- **Regime 2:** $\rho_{C_1} < \frac{\rho_1+\rho_2}{2} \leq \frac{\rho_{C_1}+\rho_{C_2}}{2}$

  ρ_{C_1} is the critical density in the fundamental diagram of one-lane road where maximal traffic flux is achieved. In the fundamental diagram of two-lane road, there are two densities ρ_{C_1} and ρ_{C_2} whose respective fluxes are exactly the maximal traffic flux $\varphi_{\max}$ of the one-lane road. When the average road density is above the critical density ρ_{C_1}, it is impossible to see free flow everywhere. Congestion exists somewhere.

  For the condition $\frac{\rho_1+\rho_2}{2} \leq \frac{\rho_{C_1}+\rho_{C_2}}{2}$, there is at least one inequality of $\rho_1 \leq \rho_{C_1}$ and $\rho_2 \leq \rho_{C_2}$ hold. When $\rho_1 \leq \rho_{C_1}$, the capacity supply of the one-lane road branch is $\varphi_{\max}$. When $\rho_2 \leq \rho_{C_2}$, the capacity supply of the two-lane road branch is above $\varphi_{\max}$. Therefore, there is at least one road branch that can provide capacity $\varphi_{\max}$, and generate free flow with density ρ_{C_1}. Then the free flow stays on the road branch of one lane after traveling. In contrast, congestion occurs on the road branch of two lanes in the upstream of the lane drop location (origin $x = 0$). Because the average density is bounded by $\frac{\rho_{C_1}+\rho_{C_2}}{2}$, the congestion will only occur on the road branch of two lanes, and not propagate backward to the branch boundary $x = L_1$. The free flow on the road branch of one lane is not influences by the congestion.

  The stationary weak solution is a standing wave, which consists of two traffic states: ρ_{C_1} and ρ_{C_2}. The density ρ_{C_2} only occurs on the road branch of two lanes in the upstream of the lane drop location. The two discontinuities are both standing waves. The length of the density ρ_{C_2} is uniquely determined by the conservation of vehicles. Let D be the length with density ρ_{C_2}, then it must satisfy the relation $\rho_1 L_1 + \rho_2 L_2 = \rho_{C_1}(L_1 + L_2 - D) + \rho_{C_2} D$. Even though it is not free flow everywhere, we can observe the maximal vehicle flux $\rho_{\max}$ everywhere.

The parameter setting of the fundamental diagram indicates $\rho_{C_2} \approx 8.23\rho_{C_1}$. Then the Regime 2 becomes $\rho_{C_1} < \frac{\rho_1+\rho_2}{2} \leq 4.615\rho_{C_1}$. In Figure 6.11, we show the stationary wave solution to the following Riemann problem (6.22) in which $\rho_1 + \rho_2 = 7\rho_{C_1}$. The stationary wave solution in Figure 6.11 has two density states: ρ_{C_1} and ρ_{C_2}. Even though we switch the initial density for two branches, the stationary wave solution is the same.

$$\rho_{\text{total}}(x,0) = \begin{cases} 3\rho_{C_1} \\ 4\rho_{C_1} \end{cases} \quad \text{or} \quad \begin{cases} 4\rho_{C_1}, & -L_2 < x < 0 \\ 3\rho_{C_1}, & 0 < x < L1 \end{cases}. \tag{6.22}$$

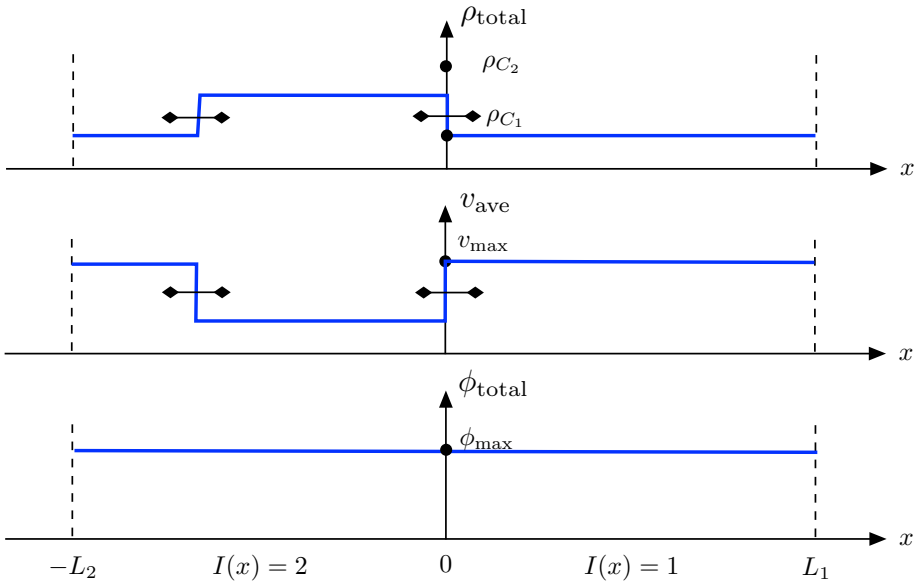


Figure 6.11: The weak solution to Riemann problem (6.22) on a ring road. The plots are vehicle density, velocity and flux curves from top to bottom.

The extreme situation is shown in Figure 6.12 with initial condition $\frac{\rho_1+\rho_2}{2} = 4.615\rho_{C_1}$. The stationary wave has two density states: ρ_{C_1} and ρ_{C_2}. The discontinuities are standing with zero speed. The congestion makes up the whole road branch of two lanes.

- **Regime 3:** $\frac{\rho_1+\rho_2}{2} > \frac{\rho_{C_1}+\rho_{C_2}}{2}$

As mentioned earlier, in the fundamental diagram of two-lane road, there are

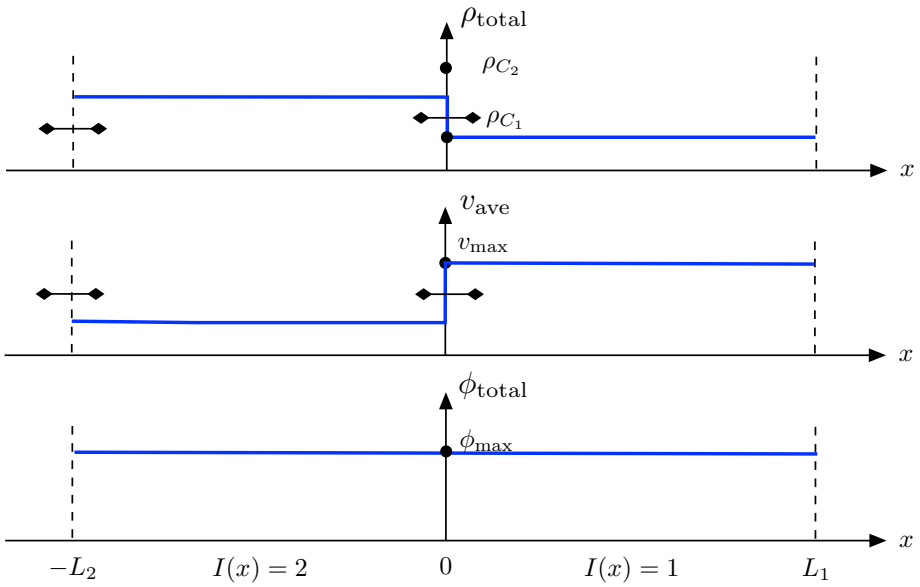


Figure 6.12: The weak solution to Riemann problem ($\frac{\rho_1+\rho_2}{2} = 4.615\rho_{C_1}$) on a ring road. The plots are vehicle density, velocity and flux curves from top to bottom.

two densities ρ_{C_1} and ρ_{C_2} whose respective fluxes are exactly the maximal traffic flux $\varphi_{\max}$ of the one-lane road. When the average road density is above $\frac{\rho_{C_1}+\rho_{C_2}}{2}$, it is impossible to see a sustainable free flow on the road branch of one lane. That is, congestion could exist everywhere on the ring road.

The stationary solution consists of two shock waves, which are moving backward with speed ω on each road branch. The shock wave evolves into another shock wave when it hits the origin ($x = 0$) or the boundary ($x = L_1$). In total, there are four types of shock waves even though we can observe only two shock waves at a time. For each road branch, the shock discontinuity consist of two fixed density states, which are uniquely determined by the initial conditions. Next, we will introduce the shock waves case by case.

For the condition $\frac{\rho_1+\rho_2}{2} > \frac{\rho_{C_1}+\rho_{C_2}}{2}$, there is at least one inequality of $\rho_1 \geq \rho_{C_1}$ and $\rho_2 \geq \rho_{C_2}$ hold. The two cases for Regime 3 are: (a) both inequalities hold for ρ_1 and ρ_2; (b) only one inequality holds for ρ_1 or ρ_2.

– Case 1: $\rho_1 \geq \rho_{C_1}$ and $\rho \geq \rho_{C_2}$

The road capacity of both branches are restricted by the supply $\varphi_{\text{total}}(\rho_1, 1)$ and $\varphi_{\text{total}}(\rho_2, 2)$. Then the stationary solution has a capacity switch between $\varphi_{\text{total}}(\rho_1, 1)$ and $\varphi_{\text{total}}(\rho_2, 2)$ on each road branch. There are two corresponding densities called $\tilde{\rho}_1$ and $\tilde{\rho}_2$ which satisfy

$$\varphi_{\text{total}}(\rho_1, 1) = \varphi_{\text{total}}(\tilde{\rho}_2, 2),$$
$$\varphi_{\text{total}}(\tilde{\rho}_1, 1) = \varphi_{\text{total}}(\rho_2, 2).$$

Please see Figure 6.13 for the positions of these densities in fundamental diagram. The density $\tilde{\rho}_2$ has equal flux to density ρ_1, and the density $\tilde{\rho}_1$ has equal flux to density ρ_2. The shock wave on the road branch of one lane is made up with ρ_1 and $\tilde{\rho}_1$, and the shock wave on the road branch of two lanes is made up with ρ_2 and $\tilde{\rho}_2$. In total, there are four possible types of shock waves on the ring road, and they will disappear and regenerate when hitting the origin or boundary. The evolution rules are: shock$(\rho_2, \tilde{\rho}_2) \leftrightarrow$ shock$(\rho_1, \tilde{\rho}_1)$ and shock$(\tilde{\rho}_2, \rho_2) \leftrightarrow$ shock$(\tilde{\rho}_1, \rho_1)$.

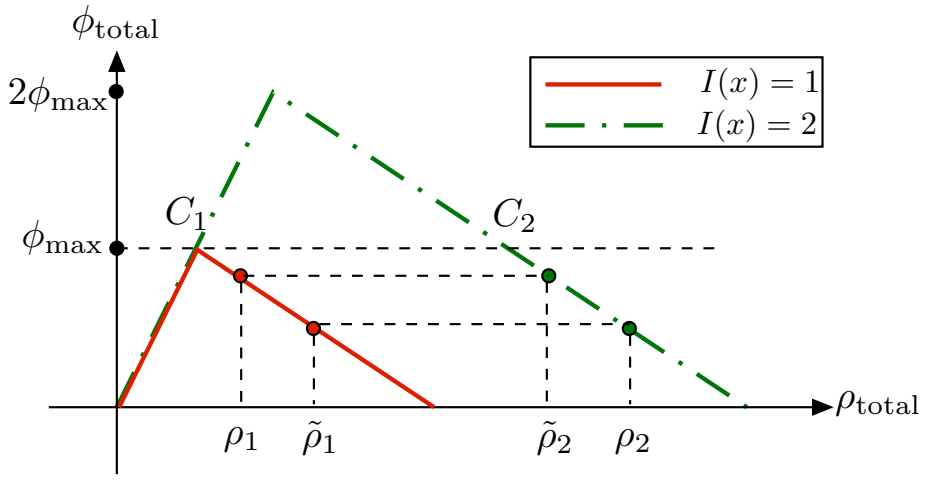


Figure 6.13: The fundamental diagram of the ring road. The density $\tilde{\rho}_2$ has equal vehicle flux to the given initial density ρ_1. The density $\tilde{\rho}_1$ has equal vehicle flux to the given initial density ρ_2.

In Figure 6.14, we show the shock wave solution to the following Riemann problem (6.23) in which $\rho_1 > \rho_{C_1}$ and $\rho_2 > \rho_{C_2}$. Both shock waves are

moving with speed ω. When the shock wave hits the origin ($x = 0$) or the boundary ($x = L_1$), it disappears on the current road branch, but evolves into another shock on the other road branch. In this case, the distance between the two observed shocks are exactly the length L_1 of roach branch because the two shocks waves are generated simultaneously.

$$\rho_{\text{total}}(x,0) = \begin{cases} \rho_{C_2} + 2\rho_{C_1}, & -L_2 < x < 0 \\ 2\rho_{C_1}, & 0 < x < L1 \end{cases}. \tag{6.23}$$

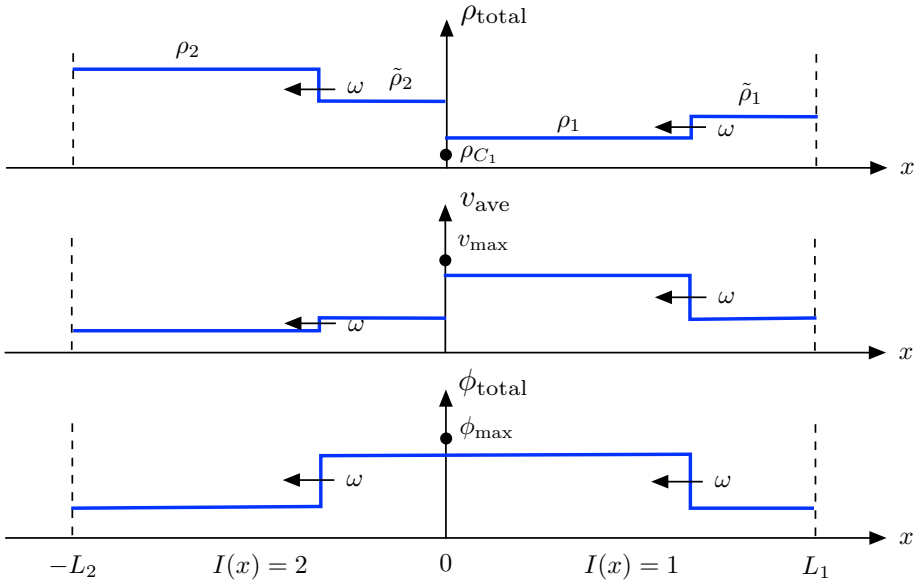


Figure 6.14: The weak solution to Riemann problem (6.22) on a ring road. The plots are vehicle density, velocity and flux curves from top to bottom.

– Case 2: $\rho_1 < \rho_{C_1}$ or $\rho < \rho_{C_2}$

In this case, one of the inequality $\rho_1 < \rho_{C_1}$ or $\rho < \rho_{C_2}$ holds. Even though the analytic solution is more complicated at the beginning, the stationary solution still consists of two shocks wave. Because of the inequality $\rho_1 < \rho_{C_1}$ or $\rho < \rho_{C_2}$, the capacity supply could somehow reach the maximal flux $\varphi_{\max}$. Without loss of generosity, we just assume $\rho_1 < \rho_{C_1}$ fo the following analysis. Then the stationary solution has a capacity switch

between $\varphi_{\max}$ and $\varphi_{\text{total}}(\rho_2, 2)$. The four densities states in the stationary solution are ρ_{C_1}, ρ_{C_2}, ρ_2 and $\tilde{\rho}_1$:

$$\varphi_{\text{total}}(\rho_{C_1}, 1) = \varphi_{\text{total}}(\rho_{C_2}, 2),$$
$$\varphi_{\text{total}}(\tilde{\rho}_1, 1) = \varphi_{\text{total}}(\rho_2, 2).$$

In Figure 6.15, we show the piecewise shock wave solution to the following Riemann problem (6.24) in which $\rho_1 \leq \rho_{C_1}$ and $\rho_2 > \rho_{C_2}$. Both shock waves are moving with speed ω. When the shock wave hits the origin ($x = 0$) or the boundary ($x = L_1$), it disappears on the current road branch, but evolves into another shock on the other road branch. The evolution rules are: shock$(\rho_2, \rho_{C_2})\rightarrow$ shock$(\rho_{C_1}, \tilde{\rho}_1)$ and shock$(\rho_{C_2}, \rho_2)\rightarrow$ shock$(\tilde{\rho}_1, \rho_{C_1})$. In this case, the distance between the two observed shocks are not necessary the length L_1 of roach branch, but is uniquely determined by the conservation of vehicles. Let D be the length with capacity $\varphi_{\max}$, then it must satisfy the relation $\rho_1 L_1 + \rho_2 L_2 = \frac{\rho_{C_1}+\rho_{C_2}}{2} D + \frac{\rho_2+\tilde{\rho}_1}{2}(L_1 + L_2 - D)$.

$$\rho_{\text{total}}(x, 0) = \begin{cases} \rho_{C_2} + 2\rho_{C_1}, & -L_2 < x < 0 \\ 0.5\rho_{C_1}, & 0 < x < L1 \end{cases}. \tag{6.24}$$

Generalization of traffic pattern on a ring road

The discovery of traffic pattern on a ring road can be generalized to any ring road profile. The general road profile is shown in Figure 6.16. There are two road branches: the k_1-lane branch is of length L_{k_1}, and the k_2-lane branch is of length L_{k_2} with the

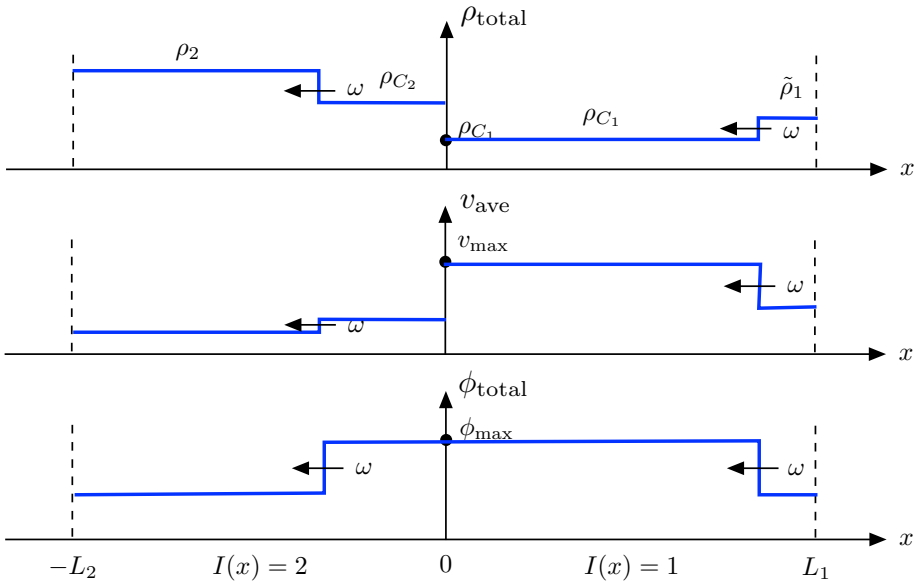


Figure 6.15: The weak solution to Riemann problem (6.24) on a ring road. The plots are vehicle density, velocity and flux curves from top to bottom.

length ratio $L_{k_2} : L_{k_1} = \lambda_2 : \lambda_1$. The Riemann problem is described in (6.25) with periodic boundary condition at $x = L_{k_1}$ or $-L_{k_2}$.

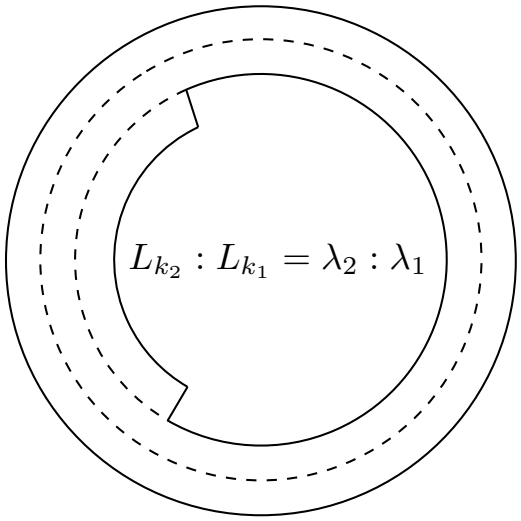


Figure 6.16: A ring road with two branches: k_1-lane branch with length L_{k_1} and k_2-lane branch with length L_{k_2}. The ratio is $L_{k_2} : L_{k_1} = \lambda_2 : \lambda_1$.

$$\rho_{\text{total}}(x,0) = \begin{cases} \rho_2, & -L_{k_2} < x < 0 \\ \rho_1, & 0 < x < L_{k_1} \end{cases}, \qquad I(x) = \begin{cases} k_2, & -L_{k_2} < x < 0 \\ k_1, & 0 < x < L_{k_1} \end{cases}. \tag{6.25}$$

Without loss of generosity, we just assume $k_2 > k_1$. The road capacity of the k_1-lane branch is $k_1\rho_{\max}$. In the fundamental diagram of the ring road, there are two points that can provide exact $k_1\rho_{\max}$ capacity. These two points are called C_1 and C_2. The averaged density $\frac{\lambda_1\rho_1+\lambda_2\rho_2}{\lambda_1+\lambda_2}$ determines the traffic pattern of the ring road. There are three regimes controlled by the points C_1 and C_2.

$$\text{Regime 1: } \frac{\lambda_1\rho_1 + \lambda_2\rho_2}{\lambda_1 + \lambda_2} \leq \rho_{C_1} \tag{6.26}$$

$$\text{Regime 2: } \rho_{C_1} < \frac{\lambda_1\rho_1 + \lambda_2\rho_2}{\lambda_1 + \lambda_2} \leq \frac{\lambda_1\rho_{C_1} + \lambda_2\rho_{C_2}}{\lambda_1 + \lambda_2} \tag{6.27}$$

$$\text{Regime 3: } \frac{\lambda_1\rho_1 + \lambda_2\rho_2}{\lambda_1 + \lambda_2} > \frac{\lambda_1\rho_{C_1} + \lambda_2\rho_{C_2}}{\lambda_1 + \lambda_2} \tag{6.28}$$

$$\text{with} \quad \rho_{C_1} = k_1\rho_{\text{crit}}, \quad \rho_{C_2} = k_1\rho_{\text{crit}} + (k_2 - k_1)\rho_{\max}$$

6.3 Controlling strategies

In this section, we will develop two controlling strategies based on the multi-lane LWR model such that the wave solution can provide smooth velocity solution without losing the road capacity. The vehicles are embedded in this macroscopic stream and their movements obey the velocity solution.

The study of traffic pattern on a ring road indicates that the multi-lane LWR model is efficient in organizing traffic density, and thus capable to maximize the capacity. In order to better show our proposed controlling strategies, we use a traffic scenario to explain: there is a straight road with a lane drop at the origin ($x = 0$). A wave of vehicles (density $2\rho_{\text{crit}}$) is approaching the lane drop. The scenario can be

mathematically formulated as:

$$\rho_{\text{total}}(x,0) = \rho_0(x) = \begin{cases} 2\rho_{\text{crit}}, & x << 0 \\ 0, & \text{otherwise} \end{cases}, \qquad I(x) = \begin{cases} 2, & x < 0 \\ 1, & x > 0 \end{cases}. \tag{6.29}$$

Then complete LWR model as well as the initial conditions and boundary conditions is as follows:

$$\frac{\partial \rho_{\text{total}}}{\partial t} + \frac{\partial \varphi_{\text{total}}}{\partial x} = 0 \tag{6.30}$$

$$\text{IC: } \rho_{\text{total}}(x,0) = \rho_0(x) \tag{6.31}$$

$$\text{BC: } \rho_{\text{total}}(-\infty,t) = 2\rho_{\text{crit}}, \quad \rho_{\text{total}}(\infty,t) = 0 \tag{6.32}$$

Let us analyze the wave solution to the equation above. First, it is equivalent to a Riemann problem of homogeneous LWR equation in the upstream branch ($x < 0$). The solution is a traveling wave with speed $v_{\max}$. When the traffic density $2\rho_{\text{crit}}$ reaches the lane drop ($x = 0$), it becomes a Riemann problem of inhomogeneous LWR equation with initial conditions: $\rho_L = 2\rho_{\text{crit}}$ and $\rho_R = 0$. We have figured out the unique solution that satisfy the entropy conditions. (Please see § 6.2.2 for the solution analysis.) The evolution of the solution is sketched in the Figure 6.17.

The stationary solution consists of three waves: the shock wave $\rho_L - \rho_A$ moving backward with speed ω, the standing wave $\rho_A - \rho_B$ at the lane drop, and the traveling wave $\rho_B - \rho_R$ moving forward with speed $v_{\max}$.

With the density solution to the multi-lane LWR equation, we can compute the respective velocity/flux solution by using the velocity/flux function. By taking the

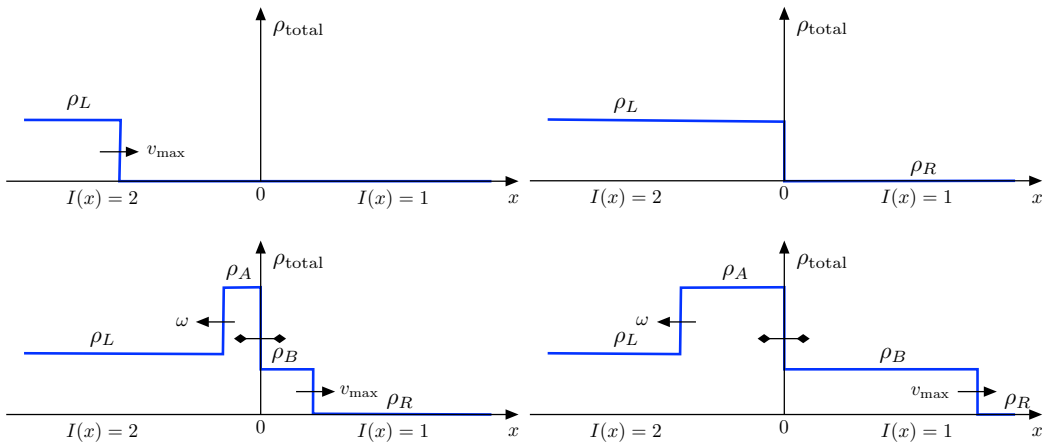


Figure 6.17: The evolution of the solution to Riemann problem (6.29). In the top figures, a traveling wave moving forward with speed $v_{\max}$. In the bottom figures, the solution consists of a shock wave moving backward, a standing wave at the origin and a traveling wave moving forward.

lane number $I(x)$ into consideration, the velocity and flux functions become:

$$\varphi_{\text{total}}(\rho_{\text{total}}, I(x)) = I(x)\varphi\left(\frac{\rho_{\text{total}}}{I(x)}\right)$$

$$v_{\text{ave}}(\rho_{\text{total}}, I(x)) = \begin{cases} \frac{\varphi_{\text{total}}}{\rho_{\text{total}}}, & \rho_{\text{total}} \neq 0 \\ v_{\max}, & \rho_{\text{total}} = 0 \end{cases}$$

We plot the velocity and flux curves together with the density solution to the multi-lane LWR equation in Figure 6.18. From the flux curve in the bottom, we observe that the total flux φ_{total} equals to the road capacity $\rho_{\max}$ in the downstream even though there is a queue in the upstream. In other words, the phenomena of capacity drop is not observed in the multi-lane LWR model.

This is an exciting discovery. If the vehicles on the road can adjust the speed based on the velocity solution from multi-lane LWR model and use zipper merging, then there is no so-called capacity drop. This maximizes the capacity usage of highways, and efficiently avoids delays. It is not hard to realize if all the vehicles are self-driving cars or they can communicate with a central control system. The vehicle at position x and time t moves with speed $v_{\text{ave}}(x,t)$ of the velocity solution. Because

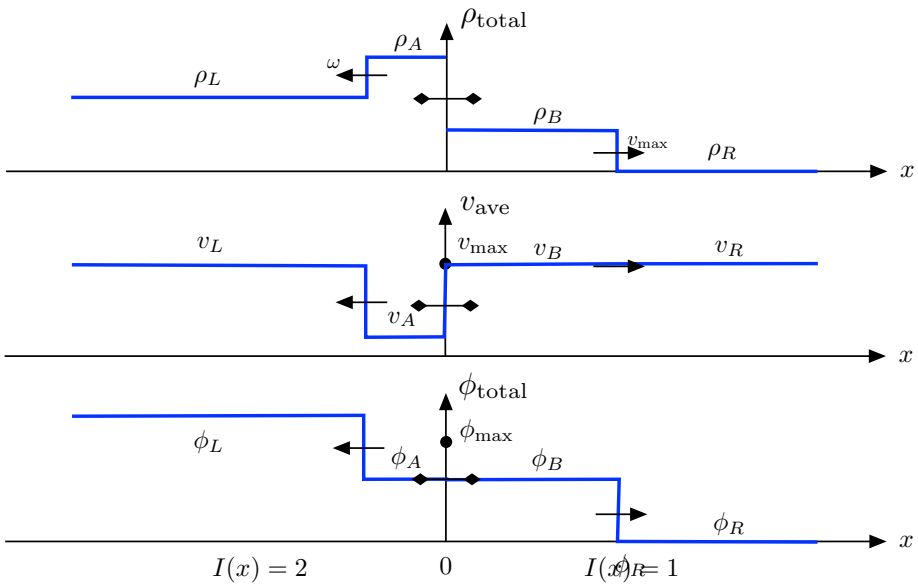


Figure 6.18: The wave solution to (6.29). The plots are density curve, velocity curve and flux curve from top to bottom.

the density–velocity relation follows the fundamental diagram, the vehicles always stay at a safe velocity based on current headways. Therefore, we are interested in guiding vehicles to go through the lane-drop bottleneck by obeying the velocity solution from the multi-lane LWR.

However, the velocity curve generated by the multi-lane LWR system is not applicable for a vehicle to follow because of the huge velocity jumps. Even though, the self-driving cars can adjust speed fast enough to follow the velocity curve, it definitely sacrifices the comfort of the passengers by using extreme acceleration or deceleration.

Therefore, our goal is to build a multi-lane LWR system, which can generate efficient and comfortable average velocity $v_{\text{ave}}(x,t)$ for vehicles to follow while maximizing the capacity. The desired average velocity $v_{\text{ave}}(x,t)$ is a smoothly changing curve that avoids extreme acceleration and deceleration. Here, we still use the fundamental diagram to restrict the velocity–density relation in order to guarantee the

safety. There are two potential controlling strategies to generate a comfortable average velocity: (a) lower the velocity jumps v_L/v_A and v_B/v_A; (b) smooth the velocity jumps v_L/v_A and v_B/v_A.

In the next two subsections, we will show the controlling strategies through providing the traffic explanation and mathematical formula.

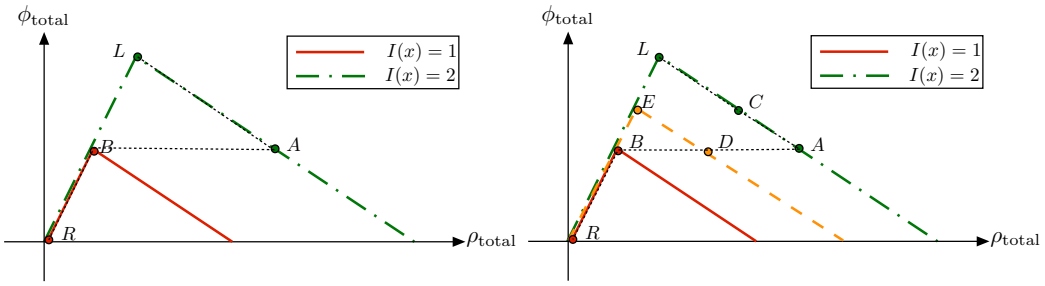
6.3.1 Virtual lane usage

When the vehicles approach the lane drop, the vehicles on the dropping lane will merge into other existing lanes gradually. The common senses tell us that drivers should reduce the usage of the dropping lanes and merge into other lanes earlier. Otherwise, the vehicles have to cut in by forcing the vehicles on the target lane to stop. The rushed cut-in can disrupt the ongoing traffic flow and bring stop-and-go wave.

The idea of early merge when approaching the lane drop can be mathematically understood as virtual lane usage. In the original multi-lane LWR model, the number of lanes $I(x)$ jumps from 2 to 1 immediately at the origin ($x = 0$). It means drivers are still taking full use of the shoulder lane until the lane drop. As explained in § 6.2, $I(x)$ can be non-integer to account for partial lane usage. For example, a value of $I(x) = 1.5$ represents the shoulder lane is less frequently used. In order to demonstrate the idea of virtual lane usage, the lane number $I(x)$ should gradually decreases from 2 to 1.

We pick a road strip of length L in the upstream of the lane drop, and set the lane number $I(x)$ as 1.5 instead of 2. The idea is shown in Figure 6.19. As shown in

Figure 6.19(a), the density gap between state L and A is determined by the difference of lane number $I(x)$. When we pick a road strip of 1.5 lanes in Figure 6.19(b), there are more transitional states between the gap L/A. The fundamental diagram of the virtual 1.5 lanes is in the orange dashed line.



(a) Original fundamental diagram (green dashed dotted line) for one-lane and two-lanes traffic in multi-lane LWR model.

(b) Introduced fundamental diagram (orange dashed line) for 1.5-lane traffic in the upstream of the lane drop.

Figure 6.19: Comparison of fundamental diagram: (a) lane number $I(x)$ decreases from 2 to 1; (b) lane number $I(x)$ decreases from 2 to 1.5, then to 1.

Next let us analyze the wave solution to the multi-lane LWR system with virtual lane usage. There are two lane drops in the road profile now. First, the solution is a traveling wave with speed $v_{\max}$ in the upstream branch ($x < -L$). Then, it becomes a Riemann problem of inhomogeneous LWR model with initial conditions: $2\rho_{\text{crit}}$ and 0 in the branch $(-\infty, 0)$. By using the supply–demand theory, it is not hard to compute the boundary flux and figure out the two transitional states C and E. The density E becomes a new traveling wave with speed $v_{\max}$ in the branch $(-L, 0)$. When the state E arrives the second lane drop $x = 0$, another Riemann problem of inhomogeneous LWR model is generated with initial conditions: ρ_E and 0. By using the supply–demand theory, we figure out the two transitional states D and B. The shock wave E/D move backwards with speed ω, and evolves into shock wave C/A across the first lane drop ($x = -L$). The wave solution, including density, velocity and flux, is plotted in Figure 6.20.

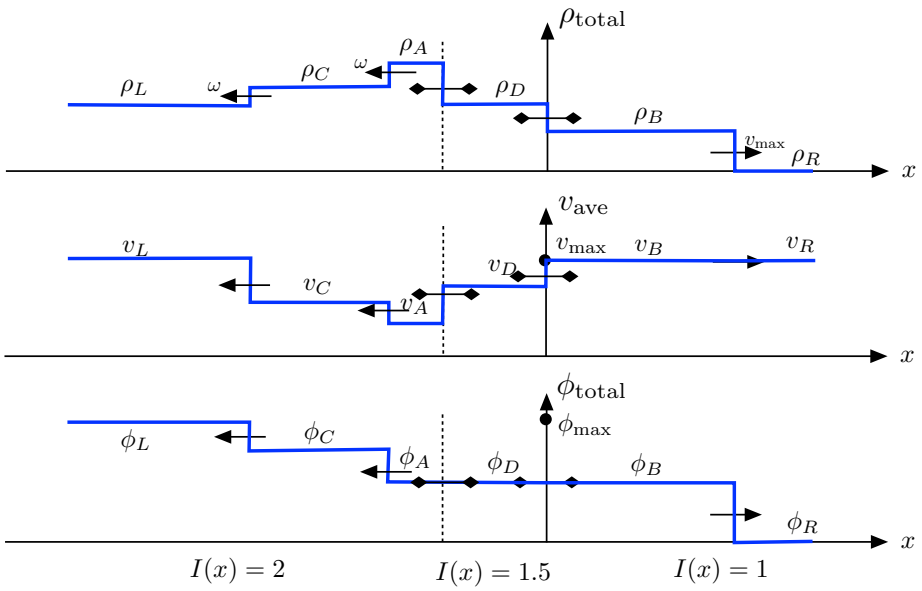


Figure 6.20: The wave solution to (6.29) after introduce a virtual lane $I(x) = 1.5$. The plots are density curve, velocity curve and flux curve from top to bottom.

There are two shock waves L/C and C/A moving backward with the same speed ω. There are two standing wave at the lane drop location $x = 0$ and $x = -L'$. It is a traveling wave between state B and R in the downstream, which is moving forward with speed $v_{\max}$. We will justify that this weak solution satisfies both entropy conditions.

Entropy 1: the waves from left (upstream) to right (downstream) should increase their wave speeds so that they don't cross each other. The waves' speeds from upstream to downstream are still $\omega, 0$ and $v_{\max}$. It is obvious that the relation $\omega < 0 < v_{\max}$ holds.

Entropy 2: the standing wave can not cross the transition curve $\rho_{\text{total}}/I(x) = \rho_{\text{crit}}$. The two transitional states ρ_A and ρ_D at lane drop $x = 0$ satisfy $\rho_{\text{total}}/I(x) \geq \rho_{\text{crit}}$. The two transitional states ρ_D and ρ_B at lane drop $x = 0$ satisfy $\rho_{\text{total}}/I(x) \geq \rho_{\text{crit}}$. That means that they do not across the transition curve.

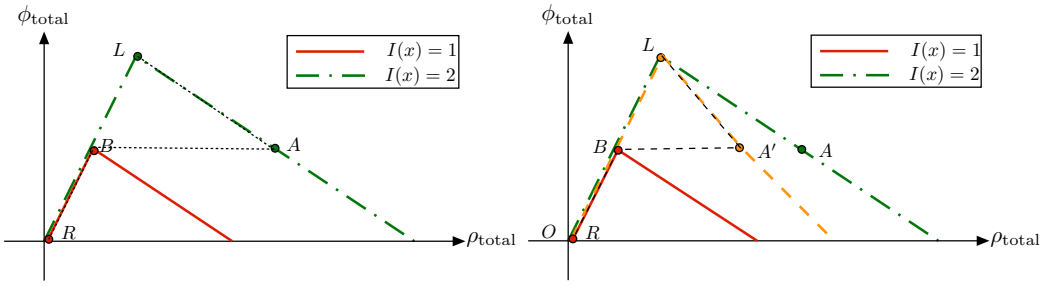
6.3.2 Reshaped fundamental diagram

When the vehicles approach the lane drop, congestions occur because of non-uniform lane-changing behaviors. The common senses tell us that drivers should slow down when approaching the lane drop in attempt to leave enough headway for the merging purpose. For example, drivers usually drive 50mph when the headway is 20m. Then the drivers should drive only 40mph with the same headway when there is a lane drop ahead. This action can avoid over-congestion so that vehicles are relatively further away from each other.

The change of driving habit when approaching the lane drop can be mathematically understood as the change of the velocity–density relation in the regime of congested flow in the fundamental diagram. The change of velocity–density relation is equivalent to the change of the flux–density relation, so we can reshape the flux-density relation around the lane drop area. Therefore, an alternative fundamental diagram is introduced as the controlling strategy in order to narrow the density gap ρ_L/ρ_A, and thus narrow the velocity gap v_L/v_A.

The comparison is shown in Figure 6.21. As shown in Figure 6.21(a), the density gap between state L and A is determined by the shape of the flux-density relation in the congested flow. When we propose an alternative flux-density relation in Figure 6.21(b) under the original relation, the density gap is narrowed. The alternative flux-density relation is in the orange dashed line.

Next let us analyze the wave solution to the multi-lane LWR system with the modified flux-density relation. First, the solution is a traveling wave with speed $v_{\max}$ in the upstream branch ($x < 0$). Then, it becomes a Riemann problem of inhomogeneous LWR model with initial conditions: $\rho_L = 2\rho_{\text{crit}}$ and $\rho_R = 0$. By



(a) Original fundamental diagram (green dashed dotted line) for two-lane traffic in multi-lane LWR model.

(b) Alternative fundamental diagram (orange dashed line) for two-lane traffic near lane drop.

Figure 6.21: Comparison of fundamental diagram: (a) original shape with transitional state A; (b) alternative shape with transitional state A'.

using the supply–demand theory, it is not hard to compute the boundary flux and figure out the two transitional states A' and B. The wave solution, including density, velocity and flux, is plotted in Figure 6.22.

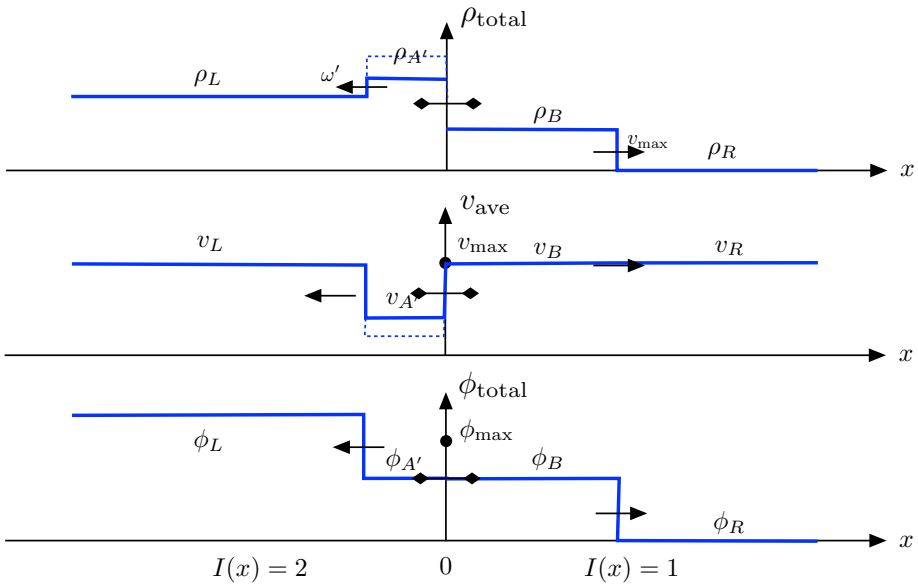


Figure 6.22: The wave solution to (6.29) after reshaping the fundamental diagram. The plots are density curve, velocity curve and flux curve from top to bottom. The original wave solution is in blue dashed line.

It is a shock wave between state L and A', which is moving backward with speed $\omega' = \frac{\varphi_L - \varphi_{A'}}{\rho_L - \rho_{A'}}$. The standing wave is stationary at the origin. It is a traveling wave between state B and R, which is moving forward with speed $v_{\max}$. We will justify

that this weak solution satisfies both entropy conditions.

Entropy 1: the waves from left (upstream) to right (downstream) should increase their wave speeds so that they don't cross each other. The waves' speeds from upstream to downstream are $\omega', 0$ and $v_{\max}$. It is obvious that the relation $\omega' < 0 < v_{\max}$ holds.

Entropy 2: the standing wave can not cross the transition curve $\rho_{\text{total}}/I(x) = \rho_{\text{crit}}$. The two transitional states $\rho_{A'}$ and ρ_B around the standing wave both satisfy $\rho_{\text{total}}/I(x) \geq \rho_{\text{crit}}$. That means that they do not across the transition curve.

6.3.3 Optimization in controlling strategies

In the last two parts, we introduced idea of traffic control via reshaped fundamental diagram and virtual lane usage. In this part, we would like to provide more insights for the parameter setting in the two controlling strategies.

virtual lane usage

The goal of virtual lane usage is to smooth the velocity jump so that the vehicles are comfortable to follow. Consider the velocity jump that consists the higher velocity v_L and the lower velocity v_A. The jump v_L/v_A is moving backward with speed ω, and the jump v_A/v_L is standing with speed zero. Please see Figure 6.23 for lane number $I(x)$ and average velocity $v_{\text{ave}}(x)$.

First, we investigate the comfort requirement on the velocity solution v_{ave}, then

we attempt to obtain the comfort requirement on the virtual lane $I(x)$ through the relation between $v_{\text{ave}}(x)$ and $I(x)$.

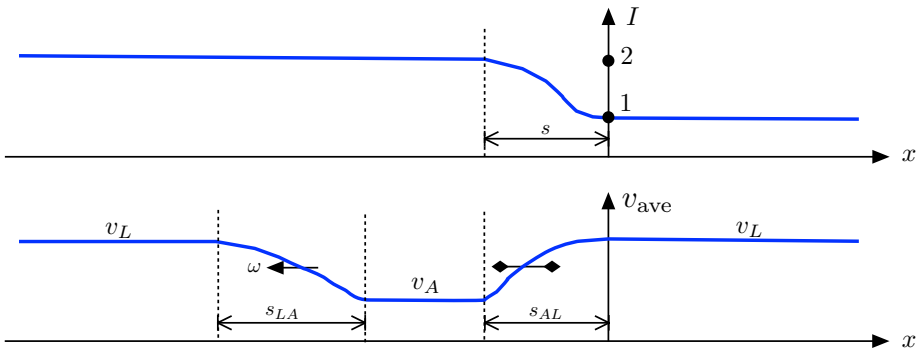


Figure 6.23: Lane number $I(x)$ (top) and average velocity $v_{\text{ave}}(x)$ (bottom) using virtual lane usage.

The comfortable acceleration and deceleration are denoted by a and b (see § 4.2). The comfort requirements on velocity v_{ave} include: the deceleration from v_L to v_A should not be greater than $|b|$; the acceleration from v_A to v_L should not be greater than a. The deceleration can be computed as:

$$\begin{aligned} &\lim_{\Delta t \to 0} \frac{v_{\text{ave}}(x) - v_{\text{ave}}(x + (v_{\text{ave}}(x) + \omega)\Delta t)}{\Delta t} \\ =&\lim_{\Delta t \to 0} \frac{v_{\text{ave}}(x) - v_{\text{ave}}(x) - (v_{\text{ave}}(x) + \omega)v'_{\text{ave}}(x)\Delta t}{\Delta t} \\ =&-(v_{\text{ave}}(x) + \omega)v'_{\text{ave}}(x) \end{aligned} \tag{6.33}$$

The acceleration can be computed as:

$$\begin{aligned} &\lim_{\Delta t \to 0} \frac{v_{\text{ave}}(x + v_{\text{ave}}(x)\Delta t) - v_{\text{ave}}(x)}{\Delta t} \\ =&\lim_{\Delta t \to 0} \frac{v_{\text{ave}}(x) + v_{\text{ave}}(x)v'_{\text{ave}}(x)\Delta t - v_{\text{ave}}(x)}{\Delta t} \\ =&v_{\text{ave}}(x)v'_{\text{ave}}(x) \end{aligned} \tag{6.34}$$

Therefore, the requirements on the velocity v_{ave} can be written as:

$$(v_{\text{ave}}(x) + \omega)v'_{\text{ave}}(x) \leq |b| \tag{6.35}$$

$$v_{\text{ave}}(x)v'_{\text{ave}}(x) \leq a \tag{6.36}$$

In Figure 6.23, let s represent the span function $I(x)$ takes from 2 to 1, s_{LA} represent the span from velocity v_L to v_A, and s_{AL} represent the span from velocity v_A to v_L. The length of the spans satisfy the following relation:

$$s_{AL} = 1s \tag{6.37}$$

$$s_{LA} = \frac{v_{\max} + |w|}{v_{\max}} s \tag{6.38}$$

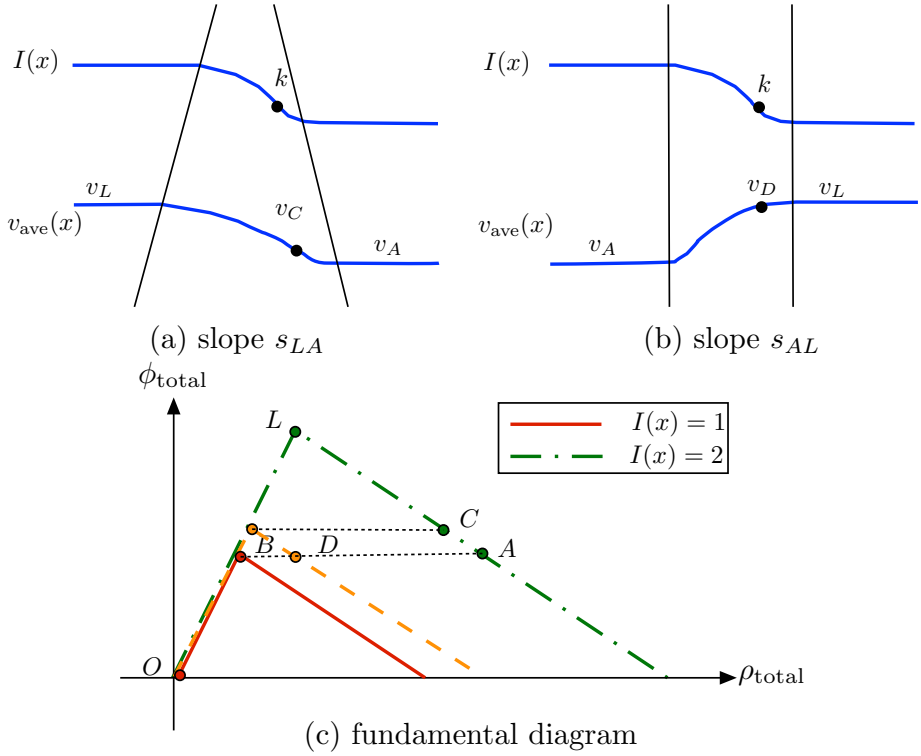
The shapes of s_{AL} and s_{LA} in velocity $v_{\text{ave}}(x)$ are determined by the lane number $I(x)$ via the fundamental diagram. Then we would like to derive the relation between lane number $I(x)$ and the shapes of s_{AL} and s_{LA} by using the fundamental diagram. The relation is shown in Figure 6.24.

The virtual lane $I(x) = k$ will bring two transitional states: state C in slope s_{LA} and state D in slope s_{AL}. The density and flux of these states can be inferred form the fundamental diagram, and then we can compute the velocity by using $v_{\text{ave}} = \varphi_{\text{total}}/\rho_{\text{total}}$:

$$\varphi_C = k\varphi_{\max}, \rho_C = k\rho_{\text{crit}} + (2-k)\rho_{\max} \quad \Longrightarrow \quad v_C = \frac{k\varphi_{\max}}{k\rho_{\text{crit}} + (2-k)\rho_{\max}}$$

$$\varphi_D = \varphi_{\max}, \rho_D = \rho_{\text{crit}} + (k-1)\rho_{\max} \quad \Longrightarrow \quad v_D = \frac{\varphi_{\max}}{\rho_{\text{crit}} + (k-1)\rho_{\max}}$$

By incorporating the scaling factor in Equation (6.37) and (6.38) into consideration, the relation between velocity lane number $I(x)$ and velocity $v_{\text{ave}}(x)$ can be explicitly expressed as follows. We are only interested in the shapes of s_{AL} and s_{LA}, so neither



(a) slope s_{LA} (b) slope s_{AL}

(c) fundamental diagram

Figure 6.24: The relation between lane number $I(x)$ and velocity $v_{\text{ave}}(x)$. (a)&(b) The slopes s_{LA} and s_{AL} corresponding to lane number $I(x)$; (c) fundamental diagram with virtual lane.

the time term nor the position shift is included in the expression.

$$v_{\text{ave}}(x) = \frac{I(x)\varphi_{\max}}{I(x)\rho_{\text{crit}} + (2 - I(x))\rho_{\max}} \tag{6.39}$$

$$v_{\text{ave}}\left(\frac{v_{\max} + |w|}{v_{\max}}x\right) = \frac{\varphi_{\max}}{\rho_{\text{crit}} + (I(x) - 1)\rho_{\max}} \tag{6.40}$$

By substituting the relation into the requirements on the velocity v_{ave}, we can get the requirements on the lane number $I(x)$. From the perspective of comfort, there is a variety of functions $I(x)$ that could provide smooth velocity changes that satisfy the comfort requirements in Equations (6.35) and (6.36). From the perspective of efficiency, we would like to find a function $I(x)$ that has the shortest span s. Thus,

the road design becomes an optimization question:

$$\min s(I(x)) \quad \text{given} \quad \begin{cases} (v_{\text{ave}}(x)+\omega)v'_{\text{ave}}(x) \le |b| \\ v_{\text{ave}}(x)v'_{\text{ave}}(x) \le a \end{cases} \tag{6.41}$$

This method can be generalized to other velocity gaps. The key consideration of comfort is enforced through the comfortable acceleration a and deceleration b.

reshaped fundamental diagram

The goal of reshaped fundamental diagram is to narrow the velocity gap v_L/v_A. However, the speed of the shock wave will increase when the gap is narrowed. In Figure 6.25, we show the contrary relation between $v_{A'}$ and ω':

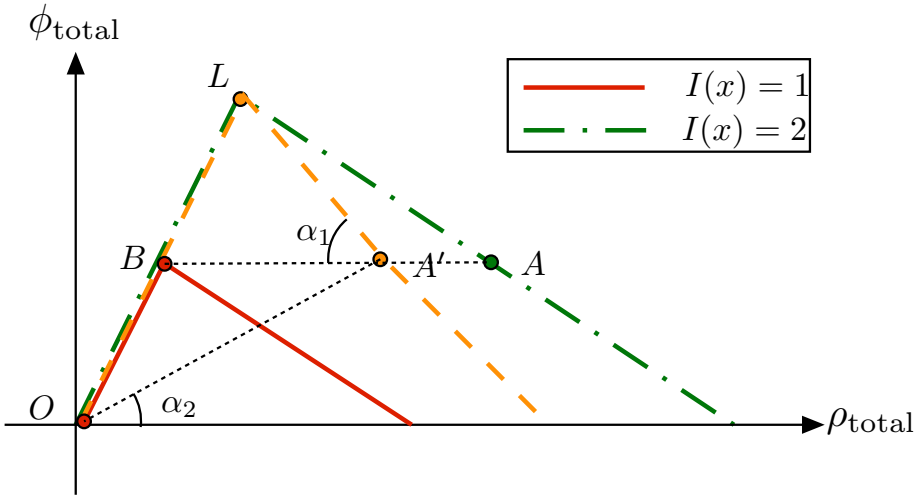


Figure 6.25: The velocity gap and shock wave speed in reshaped fundamental diagram. A' is the intersection of the alternative fundamental diagram with line BA.

The velocity at state A' is equivalent to the slope of line OA, and the speed of shock wave is equivalent to the slope of line LA'. It can be observed that: the velocity gap $v_L - v_{A'}$ is narrowed when A' moves aways from state A; in contrast, the

shock wave speed ω' increases when A' moves aways from state A. How to choose an alternative fundamental diagram in this controlling strategy depends on the optimization question we are interested. Here I only provide two potential objective functions:

- Usually drivers prefer not to vary velocity dramatically, so the smaller gap $v_L - v_{A'}$ is preferred.
- As mentioned in the virtual lane usage, another objective could be the comfortable experience is acceleration and deceleration.

From the practical perspective, the objective function could be a combination of both velocity gap and the span. The weights w_1 and w_2 are used to balance the two sub-objectives. Like the virtual lane usage, we prefer to have smaller span s. Unlike the analysis earlier, we should use the reshaped fundamental diagram to derive the relation between v_{ave} and $I(x)$. The optimization question becomes:

$$\min w_1(v_L - v_{A'}) + w_2 s(I(x)) \quad \text{given} \quad \begin{cases} (v_{\text{ave}}(x) + \omega')v'_{\text{ave}}(x) \leq |b| \\ v_{\text{ave}}(x)v'_{\text{ave}}(x) \leq a \end{cases} \tag{6.42}$$

6.4 Discussion

In this chapter, we aim to study the lane-drop bottleneck, and provide relevant controlling strategies to maximize capacity. In the introduction, we described the lane-drop bottleneck, and the capacity loss phenomena. Observations show that the queues formed by the lane-changing behavior will bring capacity loss in the downstream of lane drop.

In § 6.2, we introduced a macroscopic approach to model the multi-lane traffic, which uses only one continuity equation (multi-lane LWR model). Then we analyzed the solutions to the multi-lane LWR model by applying the supply–demand theory. The study of traffic pattern demonstrated that the multi-lane LWR model is efficient in organizing the traffic density and maximizing the road capacity. Therefore, we attempted to modify the multi-lane LWR equation, such that the vehicles can comfortably follow the velocity solution to the modified equation.

We proposed two controlling strategies in § 6.3 to overcome the velocity gap in the original velocity solution. The strategy of virtual lane usage represents the idea of early merge when approaching the lane drop. It aims to smooth the velocity gap in order to generate comfortable acceleration and deceleration for the vehicles. The lane usage I should gradually decrease in the upstream of the lane drop under the comfort requirements. The strategy of reshaped fundamental diagram represents the idea of slow-down driving when approaching the lane drop. It attempts to narrow the velocity gap by sacrificing the speed of the backward shock wave. The slope of the congested flow in the fundamental diagram should be appropriately reshaped to narrow the velocity gap.

The biggest advantage of the proposed traffic control is that it efficiently maximizes the capacity through controlling the macroscopic flow using the multi-lane LWR model. The two control strategies represent the common senses in lane drop, which is easy to understand and quantify. The design of the target multi-lane LWR system can be concluded as an optimization question, and the objective function is a modeling choice.

The proposed controlling strategies can applied to artificial intelligence of traffic. In the future, the vehicles will be self-driving or they could communicate with a

central control system. Under this circumstance, their movements can obey the velocity solution to the multi-lane LWR system. That is, the control in traffic flow is reflected on the movement of individual vehicle.

Chapter Seven

Conclusion

In the project of traffic-flow models, I have a systematic study of both macroscopic and microscopic traffic models as well as their dynamics and applications. In this thesis, we mainly investigate three interesting traffic-flow problems: the collision behavior of car-following models, traffic estimation using data assimilation techniques, and the traffic control in lane-drop bottleneck. The discoveries and results are summarized as follows:

Collision behavior of car-following models

First, we studied the collision behavior of four well-known car-following models: the optimal velocity model, the full velocity difference model, the modified GHR model and the intelligent driver model. The parameter setting is based on the model calibration using historical traffic data.

- We applied the phase portrait techniques to showing that the optimal velocity model and the full velocity difference model are collision-prone. This collision-prone behavior is independent of parameter setting.
- We also applied the fast-slow system techniques and mathematical induction to proving that the modified GHR model and the intelligent driver model are collision-free in any traffic scenarios.
- The simulations results are in agreement with the theoretical analysis. However, collisions can be occasionally observed from the trajectories generated by the collision-free models: the modified GHR model and the intelligent driver model.
- More simulations are implemented to show that the numerical errors introduce

collisions that the model does not support. In addition, the number of observed collisions is proportional to the truncation error of the numerical scheme.

In the study of the collision behavior of car-following model, we identified the driver to prevent collisions in critical situation, which can be taken into consideration when designing new acceleration models in the future. In addition, the collision-free model can be used in artificial intelligence, in which the self-driving vehicles can move according to the collision-free acceleration models.

In the future work, we would like to study the collision behavior of lane-changing models so that we can combine these safe car-following model and lane-changing to construct a complete traffic system.

Traffic estimation using data assimilation

Second, we investigated the data assimilation techniques which can be used to estimate the traffic states and uncertain parameters, including ensemble Kalman filter and particle filter. The initial motivation was to propose an efficient approach that could assimilate both Eulerian and Lagrangian GPS data simultaneously.

- We proposed an alternative approach that allows us to assimilate both Eulerian and Lagrangian GPS data simultaneously. The idea behind it is to append the differential equations for the positions and velocities of the vehicles to the macroscopic traffic model in order to solve them simultaneously.
- For ensemble Kalman filter and particle filter, we used custom localization, inflation, weight collision diagnose and resampling methods, which are suitable for traffic flow estimation.

- We added additional terms to the LWR model to capture the effects of on-/off-ramps, traffic lights, road construction and traveling bottlenecks. The proposed approach is accurate and works well in different traffic scenarios. In addition, the proposed approach can improve the accuracy when applied to realistic traffic data from Minnesota Department of Transportation.
- We compared the performance of ensemble Kalman filter and particle filter. The proposed approach is accurate and works well regardless of which ensemble Kalman filter or particle filter is used. Compared to ensemble Kalman filter, the particle filter is less sensitive to observation noise and sensor locations, but its computation cost is higher.

One limitation we reported is that the accuracy of data assimilation can be reduced significantly if the underlying traffic model cannot reproduce the actual traffic flow from which observation are collected. Hence, more accurate traffic models and calibration are needed to improve the estimation accuracy.

There are various extension that could be implemented in the future work. An exciting extension would be to apply the proposed data assimilation approach to more complicated and realistic road network. Second extension would be to use parameter estimation to relay traffic information to drivers. For instance, data assimilation can be used to predict the location and duration of congestion caused by slow trucks or road constructions.

Traffic control in lane-drop bottleneck

We also studied the lane-drop bottleneck, in which there is capacity drop when queues form in the upstream of the lane drop. Capacity drop brings unnecessary

congestion, and reduce the usages of roads. Therefore, we would like to propose some controlling strategies in order to maximize the capacity.

- We introduced only one continuity equation (inhomogeneous LWR model) to model the multi-lane traffic. This macroscopic approach focuses on the aggregated traffic characteristics instead of the interactions between lanes.
- The multi-lane LWR model was shown efficient in organizing the traffic density, and thus maximize the traffic flow in the lane drop. Therefore, we would like to have vehicles adjust their speeds based on the velocity solution to the multi-lane LWR model.
- We applied two controlling strategies to adjust the velocity curve such that it is comfortable for vehicles to follow. The first strategy targeted to smooth the velocity gaps in the velocity solution through virtual lane usage. The second strategy attempted to narrow the velocity gaps in the velocity solution through reshaped fundamental diagram.

In the future work, we would like to combine the proposed controlling strategies with zipper merging to build an intelligent control system. The acceleration and lane-changing behavior of individual vehicles are communicated through this system in order to reduce traffic delay in lane-drop bottleneck. In addition, we look forward to applying this approach to other traffic bottlenecks, for instance ramp bottleneck and accident bottleneck.

Bibliography

[1] Administration, F.H.. Over-roadway sensor technologies. http://www.fhwa.dot.gov/policyinformation/pubs/vdstits2007/05pt2.cfm; 2017.

[2] Anderson, J.L.. An adaptive covariance inflation error correction algorithm for ensemble filters. Tellus 2007;59A:210–224.

[3] Anderson, J.L.. Spatially and temporally varying adaptive covariance inflation for ensemble filters. Tellus A 2009;61(1):72–83.

[4] Anderson, J.L., Anderson, S.L.. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Mon Wea Rev 1999;127(12):2741–2758.

[5] Aw, A., Klar, A., Rascle, M., Materne, T.. Derivation of continuum traffic flow models from microscopic follow-the-leader models. SIAM Journal on Applied Mathematics 2002;63(1):259–278.

[6] Baek, S.J., Hunt, B.R., Kalnay, E., Ott, E.. Local ensemble Kalman filtering in the presence of model bias. Tellus 2006;58A(3):293–306.

[7] Bando, M., Hasebe, K., Nakanishi, K., Nakayama, A.. Analysis of optimal velocity model with explicit delay. Phys Rev E 1998;58(5):5429.

[8] Bando, M., Hasebe, K., Nakayama, A., Shibata, A., Sugiyama, Y.. Dynamical model of traffic congestion and numerical simulation. Physical Review E 1995;51(2):1035.

[9] Bellsky, T., Berwald, J., Mitchell, L.. Nonglobal parameter estimation using local ensemble Kalman filtering. Mon Wea Rev 2014;142(6):2150–2164.

[10] Bertini, R.L., Leal, M.T.. Empirical study of traffic features at a freeway lane drop. Journal of Transportation Engineering 2005;131(6):397–407.

[11] Bevrani, K., Chung, E., Miska, M.. Evaluation of the ghr car following model for traffic safety studies. In: Proceedings of the 25th ARRB Conference. ARRB Group Ltd; 2012. p. 1–11.

[12] Biswas, S., Tatchikou, R., Dion, F.. Vehicle-to-vehicle wireless communication protocols for enhancing highway traffic safety. IEEE communications magazine 2006;44(1):74–82.

[13] Bonsall, P., Liu, R., Young, W.. Modelling safety-related driving behaviourimpact of parameter values. Transportation Research Part A: Policy and Practice 2005;39(5):425–444.

[14] Brackstone, M., McDonald, M.. Car-following: a historical review. Transportation Research Part F: Traffic Psychology and Behaviour 1999;2(4):181–196.

[15] Burgers, G., van Leeuwen, P.J., Evensen, G.. Analysis scheme in the ensemble Kalman filter. Mon Wea Rev 1998;126:1719–1724.

[16] Carter, P., Christiansen, P., Gaididei, Y., Gorria, C., Sandstede, B., Sorensen, M., Starke, J.. Multi-jam solutions in traffic models with velocity-dependent driver strategies. SIAM J Appl Math 2014;74(6):1895–1918.

[17] Chandler, R.E., Herman, R., Montroll, E.W.. Traffic dynamics: studies in car following. Operations research 1958;6(2):165–184.

[18] CHUNG, S.B., SONG, K.H., HONG, S.Y., KHO, S.Y.. Development of sensitivity term in car-following model considering practical driving behavior of preventing rear end collision. Journal of the Eastern Asia Society for Transportation Studies 2005;6:1354–1367.

[19] Daganzo, C.F.. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transportation Research Part B: Methodological 1994;28(4):269–287.

[20] Daganzo, C.F.. Requiem for second-order fluid approximations of traffic flow. Transportation Research Part B: Methodological 1995;29(4):277–286.

[21] Daganzo, C.F.. A behavioral theory of multi-lane traffic flow. part i: Long homogeneous freeway sections. Transportation Research Part B: Methodological 2002;36(2):131–158.

[22] Daganzo, C.F.. A behavioral theory of multi-lane traffic flow. part ii: Merges and the onset of congestion. Transportation Research Part B: Methodological 2002;36(2):159–169.

[23] Dee, D.P., Da Silva, A.M.. Data assimilation in the presence of forecast bias. Q J Roy Meteor Soc 1998;124(545):269–295.

[24] Dervisoglu, G., Gomes, G., Kwon, J., Horowitz, R., Varaiya, P.. Automatic calibration of the fundamental diagram and empirical observations on capacity. In: Transportation Research Board 88th Annual Meeting. volume 15; 2009. .

[25] Di Francesco, M., Rosini, M.D.. Rigorous derivation of nonlinear scalar conservation laws from follow-the-leader type models via many particle limit. Archive for Rational Mechanics and Analysis 2015;217(3):831–871.

[26] Doucet, A., Freitas, N.D., Gordon, N.. An introduction to sequential Monte Carlo methods. In: Sequential Monte Carlo methods in practice. Springer; 2001. p. 3–14.

[27] Doucet, A., Godsill, S., Andrieu, C.. On sequential Monte Carlo sampling methods for Bayesian filtering. Stat Comput 2000;10:197–208.

[28] Evensen, G.. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J Geophys Res: Oceans 1994;99(C5):10143–10162.

[29] Evensen, G.. The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean dynamics 2003;53(4):343–367.

[30] Gaididei, Y.B., Gorria, C., Berkemer, R., Kawamoto, A., Shiga, T., Christiansen, P.L., Sørensen, M.P., Starke, J.. Controlling traffic jams by time-modulating the safety distance. Phys Rev E 2013;.

[31] Gaspari, G., Cohn, S.E.. Construction of correlation functions in two and three dimensions. Q J Roy Meteor Soc 1999;125(554):723–757.

[32] Gazis, D.C., Herman, R., Rothery, R.W.. Nonlinear follow-the-leader models of traffic flow. Operations research 1961;9(4):545–567.

[33] Godunov, S.K.. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. Matematicheskii Sbornik 1959;89(3):271–306.

[34] Gordon, N.J., Salmond, D.J., Smith, A.F.M.. Novel approach to nonlinear-non-Gaussian Bayesian state estimation. IEE Proc-F 1993;140(2):107–113.

[35] Gov, G.. GPS accuracy. http://www.gps.gov/systems/gps/performance/accuracy/.

[36] Greenshields, B.D., Channing, W., Miller, H., et al. A study of traffic capacity. In: Highway research board proceedings. National Research Council (USA), Highway Research Board; volume 1935; 1935. .

[37] Hamdar, S., Mahmassani, H.. From existing accident-free car-following models to colliding vehicles: exploration and assessment. Transportation Research Record: Journal of the Transportation Research Board 2008;(2088):45–56.

[38] Hamill, T.M., Whitaker, J.S., Snyder, C.. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. Mon Wea Rev 2001;129(11):2776–2790.

[39] Helbing, D.. Traffic and related self-driven many-particle systems. Rev of Mod Phys 2001;73(4):1067.

[40] Henrici, P.. Applied and computational complex analysis, discrete Fourier analysis, Cauchy integrals, construction of conformal maps, univalent functions. volume 3. John Wiley & Sons, 1993.

[41] Hidas, P.. Modelling lane changing and merging in microscopic traffic simulation. Transportation Research Part C: Emerging Technologies 2002;10(5):351–371.

[42] Holden, H., Risebro, N.H.. A mathematical model of traffic flow on a network of unidirectional roads. SIAM Journal on Mathematical Analysis 1995;26(4):999–1017.

[43] Houtekamer, P.L., Mitchell, H.L.. Data assimilation using an ensemble Kalman filter technique. Mon Wea Rev 1998;126:796–811.

[44] Houtekamer, P.L., Mitchell, H.L.. A sequential ensemble Kalman filter for atmospheric data assimilation. Mon Wea Rev 2001;129(1):123–137.

[45] Isaacson, E., Temple, B.. Nonlinear resonance in systems of conservation laws. SIAM Journal on Applied Mathematics 1992;52(5):1260–1278.

[46] Jiang, R., Wu, Q., Zhu, Z.. Full velocity difference model for a car-following theory. Physical Review E 2001;64(1):017101.

[47] Jin, W.L., Chen, L., Puckett, E.G.. Supply-demand diagrams and a new framework for analyzing the inhomogeneous lighthill-whitham-richards model. In: Transportation and Traffic Theory 2009: Golden Jubilee. Springer; 2009. p. 603–635.

[48] Kato, S., Tsugawa, S., Tokuda, K., Matsui, T., Fujii, H.. Vehicle control algorithms for cooperative driving with automated vehicles and intervehicle communications. IEEE Transactions on Intelligent Transportation Systems 2002;3(3):155–161.

[49] Kerner, B.S.. The physics of traffic: empirical freeway pattern features, engineering applications, and theory. Springer, 2012.

[50] Kesting, A., Treiber, M.. Calibrating car-following models by using trajectory data: Methodological study. Transportation Research Record: Journal of the Transportation Research Board 2008;(2088):148–156.

[51] Kong, A., Liu, J., Wong, W.. Sequential imputations and Bayesian missing data problems. J Am Stat Assoc 1994;89:278–288.

[52] Laval, J., Cassidy, M., Daganzo, C.. Impacts of lane changes at merge bottlenecks: a theory and strategies to maximize capacity. In: Traffic and Granular Flow05. Springer; 2007. p. 577–586.

[53] Laval, J.A., Daganzo, C.F.. Lane-changing in traffic streams. Transportation Research Part B: Methodological 2006;40(3):251–264.

[54] Lighthill, M.J., Whitham, G.B.. On kinematic waves. ii. a theory of traffic flow on long crowded roads. Proceedings of the Royal Society of London Series A Mathematical and Physical Sciences 1955;229(1178):317–345.

[55] Liu, J., West, M.. Combined parameter and state estimation in simulation-based filtering. In: Sequential Monte Carlo methods in practice. Springer; 2001. p. 197–223.

[56] Mandel., J.. Efficient implementation of the ensemble Kalman filter. University of Colorado at Denver and Health Sciences Center, Center for Computational Mathematics, 2006.

[57] Nagel, K., Schreckenberg, M.. A cellular automaton model for freeway traffic. Journal de physique I 1992;2(12):2221–2229.

[58] Net, G.I.. GPS speed. http://www.gpsinformation.net/main/gpsspeed.htm.

[59] Panwai, S., Dia, H.. Comparative evaluation of microscopic car-following behavior. Intelligent Transportation Systems, IEEE Transactions on 2005;6(3):314–325.

[60] Payne, H.J.. Models of freeway traffic and control. Mathematical models of public systems 1971;.

[61] Rahman, M., Chowdhury, M., Xie, Y., He, Y.. Review of microscopic lane-changing models and future research opportunities. IEEE transactions on intelligent transportation systems 2013;14(4):1942–1956.

[62] Replay, G.A.. GPS and speed measurement. http://gpsactionreplay.free.fr/index.php?menu=6&choice=5.

[63] Richards, P.I.. Shock waves on the highway. Oper Res 1956;4(1):42–51.

[64] Rosales, R.R., Kasimov, A.R., Flynn, M.R., Seibold, B.. Constructing set-valued fundamental diagrams from jamiton solutions in second order traffic models 2013;.

[65] Santitissadeekorn, N., Jones, C.K.R.T.. Two-state filtering for joint state-parameter estimation. arXiv preprint arXiv:14035989 2014;.

[66] Snyder, C.. In: Proceedings of the ECMWF Seminar on Data Assimilation for Atmosphere and Ocean. 2011. .

[67] Snyder, C., Bengtsson, T., Bickel, P., Anderson, J.. Obstacles to high-dimensional particle filtering. Mon Wea Rev 2008;136(12):4629–4640.

[68] Snyder, C., Bengtsson, T., Morzfeld, M.. Performance bounds for particle filters using the optimal proposal. Mon Wea Rev 2015;143(11):4750.

[69] of Transportation, C.D.. The freeway performance measurement (pems). http://pems.dot.ca.gov/; .

[70] of Transportation, M.D.. Traffic forecasting & analysis. http://www.dot.state.mn.us/traffic/data/; .

[71] Trefethen, L.N.. Spectral methods in MATLAB. SIAM, 2000.

[72] Treiber, M., Hennecke, A., Helbing, D.. Congested traffic states in empirical observations and microscopic simulations. Physical Review E 2000;62(2):1805.

[73] Treiber, M., Kesting, A.. Traffic flow dynamics. Traffic Flow Dynamics: Data, Models and Simulation, Springer-Verlag Berlin Heidelberg 2013;.

[74] van Leeuwen, P.J.. Particle filtering in geophysical systems. Mon Wea Rev 2009;137:4089–4114.

[75] Vogel, K.. A comparison of headway and time to collision as safety indicators. Accident analysis & prevention 2003;35(3):427–433.

[76] Work, D.B., Blandin, S., Tossavainen, O.P., Piccoli, B., Bayen, A.M.. A traffic model for velocity data assimilation. Appl Math Res Express 2010;2010(1):1–35.

[77] Work, D.B., Tossavainen, O.P., Blandin, S., Bayen, A.M., Iwuchukwu, T., Tracton, K.. An ensemble Kalman filtering approach to highway traffic estimation using GPS enabled mobile devices. In: 2008 47th IEEE Decis. Contr. P. IEEE; 2008. p. 5062–5068.

[78] Yang, X., Delsole, T.. Using the ensemble Kalman filter to estimate multiplicative model parameters. Tellus 2009;61A(5):601–609.

[79] Yuan, Y., Van Lint, H., Van Wageningen-Kessels, F., Hoogendoorn, S.. Network-wide traffic state estimation using loop detector and floating car data. J Intell Transport S: Technology, Planning, and Operations 2014;18(1).

[80] Yuan, Y., Van Lint, J.W.C., Wilson, R.E., Van Wageningen-Kessels, F., Hoogendoorn, S.P.. Real-time Lagrangian traffic state estimator for freeways. IEEE Trans Intell Transp Syst 2012;13(1):59–70.

[81] Zhang, H.M.. Driver memory, traffic viscosity and a viscous vehicular traffic flow model. Transport Res B-Meth 2003;37(1):27–41.