UNIVERSITY OF CALIFORNIA,
IRVINE

Modified Cell Transmission Model for Bounded Acceleration

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Civil Engineering

by

Anupam Srivastava

Dissertation Committee:
Professor Wen-Long Jin, Chair
Professor Will Recker
Professor R. Jayakrishnan

2016

© 2016 Anupam Srivastava

DEDICATION

To my wife and my parents.

TABLE OF CONTENTS

Page

LIST OF FIGURES **vi**

LIST OF TABLES **viii**

ACKNOWLEDGMENTS **ix**

CURRICULUM VITAE **x**

ABSTRACT OF THE DISSERTATION **xii**

1 Introduction **1**
1.1 Research Background 1
1.2 Research Objective 4
1.3 Research Outline 6

2 Literature Review **9**
2.1 Traffic Flow Models 9
2.1.1 The Lighthill - Whitham - Richards Model 9
2.1.2 Cell Transmission Model 13
2.2 Bounded Acceleration Models 14
2.2.1 Microscopic Acceleration Models 14
2.2.2 Macroscopic Acceleration Models 18

3 Queue Discharge Features at Signalized Intersections **22**
3.1 Introduction 22
3.2 Review of queue discharge features under the Cell Transmission Model . . . 25
3.3 A modified Cell Transmission Model 27
3.3.1 A new demand function and the modified CTM 27
3.3.2 Solutions to the Riemann problem 28
3.4 Queue discharge features of the modified CTM 38
3.4.1 Discharge flow-rate and headway 38
3.4.2 Lost Times 40
3.5 Calibration and validation 42
3.5.1 Approach 42
3.5.2 Calibration 44

3.6 Conclusions 49

4 Bounded Acceleration and Lane Changing at Freeway Bottlenecks 52
4.1 Introduction 52
4.1.1 Cell Transmission Model 54
4.1.2 Lane changing models 56
4.1.3 Bounded acceleration models 56
4.2 Lane changing and bounded acceleration components 58
4.2.1 Lane changing model 58
4.2.2 Bounded acceleration and queue discharge model 60
4.2.3 Capacity at lane drop bottlenecks 61
4.3 A lane changing bounded acceleration model (LCBA-CTM) to capture capacity drop at lane-drop bottlenecks 63
4.3.1 The complete LCBA-CTM model 63
4.3.2 Stationary state with capacity drop 67
4.3.3 Need for both lane-changing and acceleration components 70
4.3.4 Downstream lane-changing region 72
4.3.5 A numerical example 73
4.4 Constant Loading Problem 75
4.4.1 Case 1: $S_0 \geq C_U$, $D_0 > C_U$: Onset of congestion, Bottleneck activation 76
4.4.2 Case 2: $S_0 \geq C_U$, $C_U \geq D_0 \geq C_C$: Dependence on initial state 79
4.4.3 Case 3: $S_0 \geq C_U$, $C_C > D_0$: Recession of congestion 81
4.4.4 Case 4: $C_C < S_0 < C_U$ 81
4.4.5 Case 5: $S_0 \leq C_C$ 82
4.5 Numerical solutions for generalized bottleneck conditions 83
4.6 Conclusions 90

5 Deriving Modified Demand Functions from Microscopic Acceleration 93
5.1 Introduction 93
5.2 Framework for obtaining demand function 94
5.2.1 Simplified approach 96
5.2.2 Shockwave approach 97
5.3 Demand functions for triangular FD 99
5.3.1 Simplified approach 100
5.3.2 Shockwave approach 104
5.4 Conclusions and discussions 106

6 Applications 108
6.1 Improved Vehicle Accelerations 108
6.1.1 Introduction 108
6.1.2 Intersection features with improved vehicular accelerations 109
6.2 Arterial Network Fundamental Diagram - Signalized Ring Road 112
6.2.1 Introduction 112
6.2.2 Simulation setup 113
6.2.3 Simulation results 114

7 Concluding Remarks **120**
7.1 Summary 120
7.2 Discussions and Future Research 122

Bibliography **127**

LIST OF FIGURES

Page

2.1 Fundamental diagrams: a) the Greenshields fundamental diagram; b) a generic concave fundamental diagram; and c) the Triangular fundamental diagram. 11
2.2 An illustration of the Cell Transmission Model 15

3.1 Discharge flow-rate and headway profiles in the traditional CTM: (3.1a, 3.1b) without lost time; (3.1c, 3.1d) with start-up lost time. 26
3.2 The modified demand function along with list of variable definitions 29
3.3 Comparison of density contour plots between the traditional and modified CTM in Case V-b. 33
3.4 Comparison of density contour plots between the traditional and modified CTM in case VII. 34
3.5 Comparison of density contour plots between the traditional and modified CTM in case I. 35
3.6 Comparison of density contour plots between the traditional and modified CTM in case II. 35
3.7 Comparison of density contour plots between the traditional and modified CTM in case III. 36
3.8 Comparison of density contour plots between the traditional and modified CTM in case IV. 36
3.9 Comparison of density contour plots between the traditional and modified CTM in case Va. 37
3.10 Comparison of density contour plots between the traditional and modified CTM in case VI. 37
3.11 Queue discharge flow-rate and headways with modified demand function . . 41
3.12 Relation between lost capacity and start-up lost time 42
3.13 Setup of the headway based calibration process 44
3.14 Schematic diagrams and aerial maps of the Peltason and Lankershim study sites 47
3.15 Headway calibration for various data sources 48

4.1 An illustration of the Cell Transmission Model showing the discretization scheme and the demand supply functions. 55
4.2 The effect of lane changing on the fundamental diagram. 59
4.3 A comparison of the demand and supply functions for the traditional CTM and the modified CTM with acceleration effects. 60

4.4 Illustration of the lane drop bottleneck and the corresponding demand and supply curves. 65
4.5 The effects of including the acceleration and lane-changing modifications in to the demand function for a lane-drop setup. 65
4.6 Stationary state condition with capacity drop in lane drop setup. 68
4.7 Illustration of the lane drop bottleneck and corresponding base (before adjusting for lane-changing impact) demand and supply curves for the numerical example. 74
4.8 Case 1: Constant loading with demand greater than uncongested downstream capacity: Onset of congestion. 78
4.9 Case 2: Constant loading with demand smaller than the uncongested downstream capacity, but greater than the congested capacity. 79
4.10 Case 3: Constant loading with demand lower than uncongested downstream capacity: Recession of congestion. 81
4.11 Case 4: Constant loading where the downstream supply is lower than the uncongested downstream capacity, but larger than the congested downstream capacity. 82
4.12 Case 5: Constant loading where the downstream supply is lower than the congested downstream capacity. 83
4.13 The demand and supply functions for the entire lane-drop location based on the calibration parameters. 84
4.14 Numerical solution showing bottleneck traffic behavior corresponding to onset and recession of congestion. 85

5.1 Demand Supply Diagram 97
5.2 The figure shows the simple and the shockwave based frameworks for deriving demand function from acceleration. 98
5.3 Constant acceleration model: Simplified approach 101
5.4 Linear decay acceleration: Simplified approach 103
5.5 Gipps' acceleration mode: Simplified approach 104
5.6 Constant acceleration model: Shockwave approach 105
5.7 Linear decay acceleration: Shockwave approach 105

6.1 Comparison of vehicle discharge properties between conventional and enhanced acceleration vehicles. 111
6.2 Simulation parameters used for the MFD study. 114
6.3 MFDs for representative cycle lengths comparing the traditional CTM, lost time CTM and modified demand CTM models. 117
6.4 Compound MFDs for multiple cycle lengths comparing the traditional CTM, lost time CTM and modified demand CTM models. 118
6.5 Average flow rate vs. cycle length for various models. The graphs show the impact of cycle length on the aggregate flow rate for varying densities of traffic. 119

LIST OF TABLES

Page

3.1 Headway fitting - Calibrated parameters for the various data sources along with the SSE obtained for each case. 46

4.1 Discharge flow-rates at the bottleneck, along with the density immediately upstream of the lane-drop location shown for various combinations of initial and boundary conditions. 87

4.2 Flow contours for various combinations of initial and boundary conditions. . 88

4.3 Density contours for various combinations of initial and boundary conditions. 89

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincerest gratitude to Professor Wenlong Jin for his guidance these last 5 years, whose academic rigor, intellect, patience and attention to detail have been deeply inspiring. I would consider myself very lucky if some of his dedication and work ethic has rubbed off on me. This thesis would not have been possible without his generous support and encouragement.

I would like to acknowledge the valuable time and advice provided on several occasions by the distinguished members on my committee, Professor Recker and Professor Jayakrishnan. Special thanks to Professor Ritchie for his wisdom and generosity, and in particular, for giving some highly enjoyable teaching experience. I am indebted to the outstanding faculty at ITS, Professor McNally, Professor Saphores, and Professor Regan for everything I learnt from them.

My sincere appreciation goes to the administrative staff at ITS and Civil Engineering: Ziggy Bates, Anne Marie DeFoe, Kathy Riley, April Heath and Lorrie Aguirre for their congeniality and efficient resolutions for all administrative issues.

I would like to acknowledge the financial support provided by University of California Transportation Center, University of California Center on Economic Competitiveness in Transportations, and the ITS-Irvine UCCONNECT fellowship, that made this research possible.

This acknowledgment would not be complete without thanking my graduate advisor, Prof Nikolas Geroliminis, who continues to provided constant reassurance and motivation every time I seek it.

I wish to thank all my friends and colleagues at ITS that made my stay in Irvine memorable. Dr. Ankoor Bhagat, Dr. Daniel Rodriguez, Dr. Suman Mitra, Dr. Sarah Hernandez, Ashley, Dr. Zhe Sun, Dr. Qijian Gan, Neda, Kate, Qinlong, Felipe, Kia, Tim, and everyone else for their friendship.

Most importantly, I would like to thank my family, whose support defines everything that I am. My mother first, for her immense patience and her unending love. My father, for always encouraging me to ask questions, and equally perhaps for putting up with the countless that he had to answer. And of course my greatest support system since I met her 8 years ago, my wife Kruthi. It was through her encouragement that I decided to pursue my studies, and it was through her numerous sacrifices and her faith in me that I was able to accomplish them. Without her unwavering support, I couldn't have dreamt of doing this.

CURRICULUM VITAE

Anupam Srivastava

EDUCATION

Doctor of Philosophy in Civil Engineering **2016**
University of California - Irvine *Irvine, CA*

Master of Science in Civil Engineering **2011**
University of Minnesota - Twin Cities *Minneapolis, MN*

Bachelor of Technology in Civil Engineering **2005**
Indian Institute of Technology - Kharagpur *Kharagpur, India*

RESEARCH AND PROFESSIONAL EXPERIENCE

Graduate Research Assistant **2007–2012**
University of California, Irvine *Irvine, CA*

Applications Engineer **2005–2008**
Oracle India Pvt. Ltd. *Hyderabad, India*

TEACHING EXPERIENCE

Teaching Assistant **2012–2015**
University of California, Irvine *Irvine, CA*

REFEREED JOURNAL PUBLICATIONS

A. Srivastava, W.-L. Jin, J.-P. Lebacque (2015). A modified Cell Transmission Model with realistic queue discharge features at signalized intersections. *Transportation Research Part B: Methodological*. Vol. 81, Part 1.

A. Srivastava, N. Geroliminis (2013). Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. *Transportation Research Part C: Emerging Technologies*. Vol. 30.

N. Geroliminis, **A. Srivastava**, P. G. Michalopoulos (2011). A Dynamic-Zone-Based Coordinated Ramp-Metering Algorithm With Queue Constraints for Minnesota's Freeways. *Intelligent Transportation Systems, IEEE Transactions on*. Vol. 12, Issue 4.

P. Parthasarathi, **A. Srivastava**, N. Geroliminis, D. Levinson (2010). The importance of being early. *Transportation*. Vol. 38, Issue 2.

REFEREED CONFERENCE PUBLICATIONS

A. Srivastava, W.-L. Jin. A Lane Changing Cell Transmission Model for Modeling Capacity Drop at Lane Drop Bottlenecks. *Transportation Research Board 95th Annual Meeting*. Washington D.C., Jan 2016.

A. Srivastava, W.-L. Jin, J.-P. Lebacque. A Modified Cell Transmission Model for Signalized Intersections. *Transportation Research Board 93rd Annual Meeting*. Washington D.C., Jan 2014.

N. Geroliminis, **A. Srivastava**, P. G. Michalopoulos. Experimental Observations of Capacity Drop Phenomena in Freeway Merges with Ramp Metering Control and Integration in a First-order Model. *Transportation Research Board 90th Annual Meeting*. Washington D.C., Jan 2011.

N. Geroliminis, **A. Srivastava**, P. G. Michalopoulos. A Coordinated ramp metering algorithm for Minnesota's freeways based on density. *13th International Conference on Intelligent Transportation Systems (ITSC)*. Madeira Island, Portugal, Sep 2010.

N. Geroliminis, **A. Srivastava**, P. G. Michalopoulos. Experimental Observations in Freeway Merges with Ramp Metering Control. *12th World Conference on Transportation Research (WCTR)*. Lisbon, Portugal, Jul 2010.

P. Parthasarathi, **A. Srivastava**, N. Geroliminis, D. Levinson. The importance of being early. *12th International Conference on Travel Behavior Research*. Jaipur, India, Dec 2009.

PROFESSIONAL AFFILIATIONS

Referee
Transportation Research Part C
Transportation Research Record
PROMET: Traffic and Transportation

Affiliations
Student Member, Transportation Research Board (TRB)

ABSTRACT OF THE DISSERTATION

Modified Cell Transmission Model for Bounded Acceleration

By

Anupam Srivastava

Doctor of Philosophy in Civil Engineering

University of California, Irvine, 2016

Professor Wen-Long Jin, Chair

Modeling capacity is an integral component towards multiple traffic engineering objectives such as design and evaluation of control strategies. Traffic dynamics at bottlenecks, both on freeways and on arterial networks, influenced by bounded acceleration and lane-changing, affect the capacity in intriguing ways. This research attempts to capture these impacts of the bounded acceleration behavior and its interplay with lane-changing, by constructing a modeling framework that accurately models traffic dynamics at bottlenecks.

Towards this goal, first a modified Cell Transmission Model (CTM) is proposed, by substituting the traditionally constant demand function with a linearly decreasing function for congested traffic. The jam-density discharge flow-rate is introduced as an additional parameter to characterize the macroscopic bounded acceleration effects. Analytically the new model is shown to reproduce observed features in the discharge flow-rate and headway at signalized intersections. Calibration with observations from existing studies, as well as new observations, further suggests that the model can reasonably capture all traffic queue discharge features.

The demand function is further modified by integrating macroscopic lane-changing effects on capacity. The Lane Changing Bounded Acceleration CTM (LCBA-CTM) thus developed, is shown to realistically model the capacity drop phenomenon at active freeway lane-drop

bottlenecks in stationary states. The capacity drop magnitude is determined by macroscopic bounded acceleration and lane-changing characteristics. Constant loading problems are analytically solved to reveal the onset and recession processes of congestion.

An addition to the framework connects microscopic acceleration profiles of vehicles to modified demand functions. This completes the framework presented by offering a mechanism to start from any acceleration model.

Finally, two applications of the modified CTM are presented illustrating the use of the framework: a) to model impacts of improved vehicle acceleration on traffic dynamics at intersections; and b) to create Macroscopic Fundamental Diagrams (MFDs) for arterial networks and compare their accuracy with traditional CTM methods.

This dissertation offers a systematic approach to incorporating bounded acceleration and lane-changing into the CTM demand functions. Such an approach is shown to capture important static and dynamic features at critical bottlenecks, including lost time and queue discharge features at signalized intersections, as well as capacity drop magnitude and the onset of capacity drop at active freeway bottlenecks. The consistency between the modified demand function and microscopic bounded acceleration models is also established.

Chapter 1

Introduction

1.1 Research Background

Traffic congestion during peak periods has been the bane for commuters experiencing the delays, as well as for traffic engineers trying to improve conditions on the network. An important measure of the roadway performance is the capacity or the maximal vehicle throughput than can be served by the system. Congestion is an outcome of the demand for the roadway network, exceeding the capacity that can be served. Thus, understanding traffic congestion often becomes synonymous to understanding capacity. Capacity can be defined either for a segment of a roadway, or at the network level. Typically, the capacity of any network is defined through the capacity of its most restricting component, individual bottlenecks. Modeling capacity at individual bottlenecks then becomes integral to traffic engineering efforts in numerous avenues such as design of control strategies, decisions about investments, and development of disaster response systems. Although the underlying mechanics is similar, bottlenecks occur with slightly different properties on freeways and on arterial intersections.

On freeway networks, bottlenecks typically exist at locations with specific geometrical fea-

tures, such as at lane-drop locations and at ramp merge locations. When traffic gets congested due to the demand at the bottleneck exceeding its capacity, the bottleneck is considered activated. An intriguing property of freeway bottlenecks is the capacity drop phenomenon. After a bottleneck gets activated, observations have shown that a fraction of the capacity is lost such that the congested capacity is lower than the uncongested capacity. The difference between the uncongested and the congested capacity is usually called the capacity drop magnitude. Of nearly equal importance as the capacity drop itself, is also the actual traffic dynamics at the bottleneck as it gets activated through the onset of congestion.

On arterial roadways, bottlenecks typically form at intersection nodes. The capacity at intersections is shared by the different approaches being served by the intersection. For the case of signalized intersection, since the green time alternates between approaches, the intersection capacity can be decomposed into the green phase capacities for different phases. While there is no true capacity drop experienced at intersections, there is indeed a reduction in capacity during the start of each green phase of the signal. For vehicles stopped in a queue at the intersection, the time headways experienced as they cross the intersection stop-line at the start of a green phase have been observed to decrease with vehicle queue positions such that the first vehicle's headway is the highest. The headways are typically assumed to saturate with the fifth or the sixth vehicle in a queue. The corresponding queue discharge flow-rate at the intersection, an inverse relation of the headway, corresponding steadily increases till it saturates. The saturated rate is often called the saturated intersection capacity.

The observed headway pattern of vehicles at an intersection, is attributed to a combination of the vehicle acceleration process and the reaction time of the driver. Vehicles closer to the intersection in the initial queue cross the intersection at slower speeds compared to vehicles further in the queue that have already accelerated as they approach the intersection. Thus, the reduction in capacity during the initial seconds of the green phase can indeed

be attributed to the limited acceleration rate (bounded acceleration) of vehicles. For the fictional situation where acceleration was unbounded, all vehicles would have instantaneously accelerated to desired speeds and would have identical headways.

Past work has similarly attributed bounded acceleration as being a key contributor to the capacity drop phenomenon at freeway bottlenecks. (Hall and Agyemang-Duah, 1991) noted that capacity drop is associated with observation of queue upstream and with the acceleration process of slower vehicles as they depart from the bottleneck. However, the freeway capacity drop phenomenon has also been associated with lane-changing (Cassidy and Rudjanakanoknad, 2005).

(Lighthill and Whitham, 1955; Richards, 1956) developed what came to be known as the LWR model, a first-order macroscopic model of traffic flow that has since become the most influential framework for macroscopic traffic flow modeling. While the LWR model has countless desirable properties, being a first-order model, it also suffers from a few limitations. Modeling traffic solely under equilibrium flow conditions, vehicles within the LWR modeling framework are assumed to adjust their speeds to desired speeds instantaneously, allowing for unbounded, infinite acceleration rates. Since bounded acceleration is critical to modeling features such as capacity drop and intersection capacity reduction, the LWR model is not equipped to model such features.

Microscopic traffic flow models are capable of modeling vehicle accelerations. However, such models are often not accurate at the aggregate macroscopic level, and thus are not ideal for explaining the impacts of acceleration and lane changing on capacity. A more reasonable alternative for the objective is the use of hybrid models (Laval, 2004; Laval and Daganzo, 2006; Leclercq, 2007a). Hybrid models can capture impacts of microscopic features such as acceleration and lane-changing and study them from a macroscopic perspective. However, such models rely on tracking trajectory of a fraction of the vehicles, making them data intensive and difficult to calibrate. Further, due to the complex modeling nature combining

elements from both microscopic and macroscopic frameworks, they usually do not lend well to theoretical analysis and rely heavily on simulation tools. This limits their application towards control design problems that require optimization of a derivable objective function.

From a macroscopic perspective, various efforts have also been made to model bounded acceleration (Lebacque, 2002, 2003; Leclercq, 2002, 2007b) and lane-changing (Jin, 2010a, 2013). While they are adept at modeling either bounded acceleration or lane-changing individually, they do it such that one feature is isolated from the other. Thus, the interplay of the two features is never modeled, limiting the abilities of such models. Further, the works on bounded acceleration typically assume a constant bounded acceleration rate that is independent of the existing traffic state.

The research presented here attempts to bridge the gap in the modeling of the effects of bounded acceleration on aggregate traffic flow features and on traffic dynamics at bottlenecks. The modeling framework that is being presented would couple the lane-changing and the bounded acceleration behaviors together into a single framework such that the capacity drop is defined as an outcome of both driving behaviors. The framework presented allows for any acceleration model to be used, thus expanding on the range of behaviors that can be captured.

1.2 Research Objective

The primary objective of this research is to improve on the macroscopic modeling techniques available to traffic researchers and engineers with respect to modeling capacity flow-rates at locations of interest. More specifically, this research focuses on exploring the impacts that vehicle acceleration and, to some extent, lane-changing have on traffic dynamics at bottlenecks, and in turn on the capacity. Improvements to the Cell Transmission Modeling

framework are presented to specifically address these two traffic behaviors.

The main objectives are summarized as:

- **Propose a modified Cell Transmission Model incorporating effects of bounded acceleration.**

  Using a signalized intersection as testing grounds, the modified CTM model should be able to reproduce observed patterns in intersection discharge traffic properties. The model should accurately represent traffic dynamics at the intersection, focusing not only on the aggregate intersection capacity, but also other features such as observed headway distributions of vehicles by queue position, and correspondingly, the evolution of the discharge flow rates over the green phase.

- **Develop the Lane Changing Bounded Acceleration CTM (LCBA-CTM) model by further integrating a lane-changing component into the modified CTM.**

  The lane-changing and the bounded acceleration components of the models should both be compatible together with the CTM framework. The LCBA-CTM thus developed, should support the existence of capacity drop under stationary state conditions, as well as be able to model traffic dynamics details such as the onset and recession of congestion at an active lane-drop bottleneck. The model should further be able to predict the capacity drop magnitude for a given setup from calibrated lane-changing and bounded acceleration parameters.

- **Propose a mechanism for deriving demand functions from known acceleration profiles.**

  The objective is to bridge the gap between microscopic acceleration behaviors and macroscopic modeling. The framework would be able to convert any well known unrestricted acceleration model such as the constant acceleration model or the linear

decay acceleration model, into the corresponding macroscopic demand function that represents the acceleration behavior.

1.3 Research Outline

This dissertation research is organized into seven chapters. Chapter 1 introduced the background and set up the motivation for the research. Chapter 2 provides a brief literature review pertinent to the research. The Lighthill Whitham Richards model and the Cell Transmission Model (CTM) are reviewed. These models form the foundational traffic modeling framework that is built upon in the research.

Chapter 3 introduces a modified CTM by redefining the CTM demand function to model the effect of vehicle acceleration on flow-rate. The new demand is defined as a linearly decreasing function of density under congestion, replacing the traditional demand that is constant under congestion. The new definition introduces a new variable in the form of either the slope of the congested demand, the demand corresponding to jam density, or the intercept of the demand function with the density axis, in addition to the conventional parameters of the fundamental diagram. Using signalized intersections as the setup, the new model is shown analytically to reproduce observed discharge features such as the discharge flow-rate and headways, improving the modeling of traffic dynamics. The model is calibrated against observations from existing studies on intersection headways and shown to model discharge features accurately.

Chapter 4 presents the Lane-Changing Bounded Acceleration CTM (LCBA-CTM). The LCBA-CTM combines the modified demand function based acceleration model presented in Chapter 3 with a lane-changing model based on ideas presented in (Jin, 2010a,b). The integrated model is shown to capture the capacity drop phenomenon at active freeway lane-drop

bottlenecks. The lane-drop location is broken down into an upstream region that is unaffected by lane-changing, a region that experiences the entirety of the lane-changing process, and a downstream acceleration region. The demand at the bottleneck is first modified for the acceleration process by introducing the congested demand slope as in Chapter 3, and then adjusted for the impact of lane-changing. The chapter further shows how the capacity drop magnitude can be predicted from calibrating the acceleration and lane-changing components. The model setup is analyzed under the constant loading problem scenarios to show how stationary states can be computed for any combination of initial and boundary conditions.

In Chapter 5, a framework to derive the modified demand function from known acceleration profiles (speed-acceleration functions) of vehicles is presented. The demand functions corresponding to the triangular fundamental diagram coupled with some popular acceleration models are derived. This work helps bridge the final gap between the microscopic acceleration process of individual vehicles and the resulting macroscopic traffic dynamics and macroscopic capacity.

Chapter 6 briefly explores two applications of the modified CTM. The first section investigates how traffic dynamics at signalized intersections may be modeled for technological improvements such as autonomous and electric vehicles. The second section uses the modified demand function model to create Macroscopic / Network Fundamental Diagarms (MFD / NFD) for simple arterial networks. MFDs are generated, through simulation, for a single signalized ring road under three models for contrast: a traditional CTM, a lost time based CTM, and the modified demand CTM. The simulation results first highlight how traditional CTM can greatly overestimate intersection performances due to the instantaneous acceleration / zero lost time assumption intrinsic to the model. The lost-time based model is able to refine the MFD prediction greatly. Results also illustrate that while the lost time based model is reasonably accurate for scenarios with adequately long green times, the modified

demand CTM outperforms the lost time based model for low green time scenarios. This is intuitively expected due to the modified demand CTM's ability to accurately model the dynamic queue discharge process while the lost-time based model relies only on the aggregate features.

Finally Chapter 7 concludes with a summary of the work, a discussion of the implications of the research, and comments on potential future work that the research inspires.

Chapter 2

Literature Review

In this chapter, first, a review of the LWR model and of the CTM model is presented, followed by discussion on some of the most prominent work in macroscopic modeling of bounded acceleration.

2.1 Traffic Flow Models

2.1.1 The Lighthill - Whitham - Richards Model

The LWR model (Lighthill and Whitham, 1955; Richards, 1956) laid the foundation for the majority of work on macroscopic modeling of traffic that followed since. The modeling framework was developed based on an analogy of comparing traffic flow to flow of compressible fluids.

The LWR model is a first order macroscopic model. In the model, traffic is represented through three state variables: flow q, density k, and speed v, all of which are functions of location s and time t.

The basic equations that show the relationships between traffic states in the model are given as:

R1 Law of conservation of vehicle mass,

$$\frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0 \tag{2.1}$$

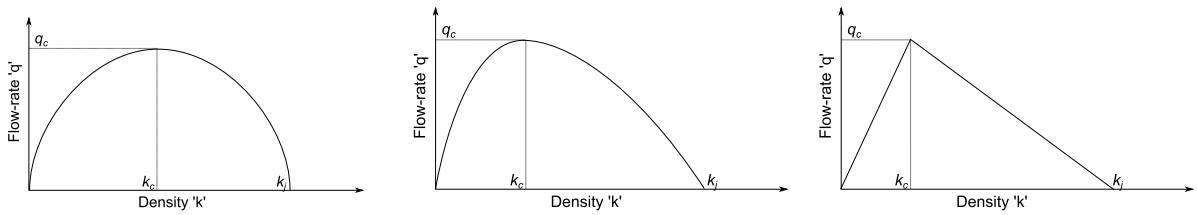
R2 Law of continuous media,

$$q = kv \tag{2.2}$$

R3 Equilibrium speed-density relationship,

$$v = V_e\left(k\right) \tag{2.3}$$

The equilibrium speed-density relationship in Rule R3 is defined as the speed-density, or flow-density relation under steady state traffic conditions. This relationship is typically referred to as the fundamental diagram of traffic flow. Various studies have proposed models that fit into this space, such as the Greenshields model (Greenshields et al., 1935), Greenberg model (Greenberg, 1959), Underwood model (Underwood, 2008), and Pipes-Munjal model (Pipes, 1967). The Greenshields' fundamental diagram (Greenshields et al., 1935) suggested that speed is a linear function of density, decreasing as density increases, and was the first study to present fundamental diagram. The Greenshields' model is especially recognized for its simplicity and for having mathematically desirable features such as being continuous and continuously differentiable in the entire domain. Another model that has become popular more recently is the triangular fundamental diagram, initially derived from car-following models but later verified often against observations (Drake et al., 1967; Munjal et al., 1971).



(a) Greenshields (b) General (c) Triangular

Figure 2.1: Fundamental diagrams: a) the Greenshields fundamental diagram; b) a generic concave fundamental diagram; and c) the Triangular fundamental diagram.

Combining R2 (Eq. 2.2) and R3 (Eq. 2.3), gives:

$$q = kv = kV_e(k) = Q_e(k) \tag{2.4}$$

And finally, plugging (Eq. 2.4) into R1 (Eq. 2.1), the conservation law can be re-written as:

$$\frac{\partial k}{\partial t} + \frac{\partial Q_e(k)}{\partial x} = 0. \tag{2.5}$$

This final conservation equation is a non-strict hyperbolic conservation law, with density k as the conserved value, and flow q as the flux.

In the LWR model, the analytical solutions for the above conservation equation, are shown to be derivable through the introduction of characteristic waves and shock waves. Characteristic waves (kinematic waves) are responsible for propagation of traffic states corresponding to slight changes in flow or density, while shock waves correspond to the propagation of larger discontinuities in traffic states across the wave boundaries. The speeds of the two types of waves are given by:

$$\text{Characteristic wave speed} = \frac{\partial q}{\partial k}, \tag{2.6}$$

and,

$$\text{Shock wave speed} = \frac{\Delta q}{\Delta k}(\text{Rankine - Hugoniot}). \tag{2.7}$$

In addition to the three rules listed above that define the conservation law for traffic modeling, the solution of hyperbolic conservation laws typically requires two further rules:

R4 Generation of weak solutions

R5 Entropy conditions

The wave solutions as presented in the LWR model provide conditions to generate weak solutions from a given set of initial and boundary conditions on a roadway, thus fulfing Rule R4. However, they do not guarantee uniqueness of the solution. In the original work (Lighthill and Whitham, 1955) entropy conditions (Rule R5) such as rarefaction fan solution to the intersection queue discharge problem, are implied but not explicitly cited (see (Ansorge, 1990; Lax, 1954)). Thus, a complete traffic modeling technique typically requires the introduction of entropy conditions for the LWR model (Lebacque, 1996; Godunov, 1959).

Various modeling techniques have been developed based on the LWR framework as given by (Eq. 2.5) in order to model traffic over roadway networks at various scales. These can broadly be classified as methods that seek to find explicit solutions (Stephanopoulos et al., 1979; Michalopoulos et al., 1981), or use discretization schemes (Daganzo, 1995a; Lebacque, 1984). One particular discretization scheme, the Cell Transmission Model (CTM) (Daganzo, 1995a), is of great relevance to this work, and is introduced in the following section.

2.1.2 Cell Transmission Model

In 1995, Daganzo (Daganzo, 1995a) proposed a discretized framework, the Cell Transmission Model (CTM), for solving the LWR model, which has since become one of the most accepted tools used towards macroscopic modeling of traffic. Roadway space is discretized into cells (length Δx) as shown in Figure 2.2a, while time is discretized into time steps (length Δt).

CTM introduces the concept of demand (or sending flow), and supply (or receiving flow) as functions of the densities in each cell. The boundary flux (flow-rate across the shared boundary of any two cells) is determined as the minimum of the demand in the upstream cell and supply in the downstream cell.

If the roadway is divided into N cells, we number the cells in an increasing order in the direction of flow of traffic such that the most upstream cell is numbered 1, and the most downstream cell is numbered N. Further, the variables k_i, Φ_i, and k_c are used to represent the density in cell i, flux across the boundary between cells $i-1$ and i, and the critical density respectively. For a generic fundamental diagram (given by $q = Q_e(k)$), the formulations for deriving the demand $D(k_i)$ and supply $S(k_i)$ in each cell, and the flux Φ_i across each cell boundary is given as:

$$\text{Demand} = D(k_i) = Q_e(min\{k_i, k_c\}), \tag{2.8}$$

$$\text{Supply} = S(k_i) = Q_e(max\{k_i, k_c\}), \tag{2.9}$$

$$\text{Boundary Flux} = \Phi_i = min\{D(k_{i-1}), S(k_i)\}, \tag{2.10}$$

Through the rest of this work, unless otherwise explicitly declared, the triangular flow-density relationship (Drake et al., 1967; Munjal et al., 1971) would be used as the fundamental relationship for traffic. The triangular fundamental diagram is represented by the following

flow-density relationship:

$$Q(k) = min\{v_f k, w(k_j - k)\}. \tag{2.11}$$

The corresponding supply and demand relations (Figure 2.2b) and the boundary flux are:

$$\text{Demand} = D(k_i) = min\{v_f k_i, q_c\}, \tag{2.12}$$

$$\text{Supply} = S(k_i) = min\{q_c, w(k_j - k_i)\}, \tag{2.13}$$

$$\text{Boundary Flux} = \Phi_i = min\{D(k_{i-1}), S(k_i)\}, \tag{2.14}$$

where, the free flow speed v_f, the capacity flow-rate q_c, the jam density k_j, and the speed of the congested shock wave $-w$ are properties of the triangular fundamental diagram.

The density in each cell is then updated for the next time step according to:

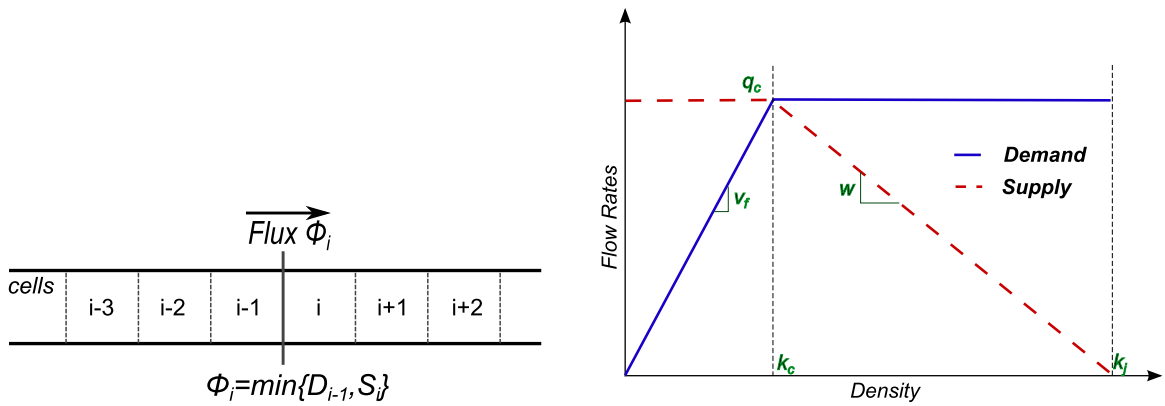
$$k_i(t + \Delta t) = k_i(t) + \frac{\Delta t}{\Delta x} \left(\Phi_i \left(t \right) - \Phi_{i+1} \left(t \right) \right), \tag{2.15}$$

and Δx and Δt follow the Courant-Friedrich-Lewy condition (Courant et al., 1967 [orig.: 1928]).

2.2 Bounded Acceleration Models

2.2.1 Microscopic Acceleration Models

The acceleration component of a microscopic car-following model typically refers to elements corresponding to both the acceleration and the braking process of vehicles. The acceleration process itself can further be broken down into the car-following acceleration, cruise control acceleration, and free speed acceleration. For the purpose of this study, the free speed



(a) Cells and Boundary Flux

(b) Triangular FD demand, supply

Figure 2.2: An illustration of the Cell Transmission Model. The figure shows a). the space discretization of the roadway link into cells, and b). a typical pair of demand and supply functions corresponding to the triangular fundamental diagram.

acceleration process where the acceleration is not limited by downstream traffic conditions or by speed limits, is most important. Models that capture this acceleration can also be called bounded acceleration models, where the boundedness typically is induced by physical constraints such as engine performance. There have been many acceleration models proposed over time with increasing levels of detail. Such models can be classified broadly into two family of models, the first dealing with only kinematics of the vehicle (such as speed), and the second incorporating other features of the vehicle-environment system (such as engine power, drag, and friction) [Rakha et al. (2004)].

Kinematics family of acceleration models

The most commonly used models of vehicle acceleration use a relationship between the current speed and location of the vehicle, and the acceleration rate applied. The following sections cover a representation of the most common models used that fall under the kinematics family.

Constant Acceleration

The constant acceleration model (eq 2.16) provides the most basic acceleration estimation.

The model uses constant acceleration rate regardless of the current speed of the vehicle [Akçelik and Biggs (1987)]. The constant acceleration model uses a simplification of the acceleration process which might not necessarily represent realistic acceleration variations [Searle (1999)], but even in its simplicity, can provide realistic modeling estimates for aggregate and macroscopic features.

$$a_i(t) \leq a_1 \tag{2.16}$$

Dual Regime Acceleration

The dual regime acceleration model (eq 2.17) builds upon the constant acceleration model and results from studies that suggested that acceleration of vehicles changes with gear, with lower gears allowing a higher acceleration rate than higher gears. The dual regime model uses a speed threshold as a proxy for the gear and sets one of two constant acceleration rates based on whether the current speed of the vehicle is higher or lower than the speed threshold [Bham and Benekohal (2004)].

$$a_i(t) \leq \left\{ \begin{array}{ll} a_{high} & for\ 0 \leq v < u_{thresh} \\ a_{low} & for\ v \geq u_{thresh} \end{array} \right\} \tag{2.17}$$

Linear Decay Acceleration

The linear decay acceleration model takes the idea of the acceleration reducing with speed, like the dual regime model, and achieves this through a linear relationship between the speed and acceleration (eq 2.18). The acceleration rate is maximum at zero speed just as the vehicle begins the acceleration process, and reduced gradually as the speed increases. This model provides for a way of overcoming the most visible shortcoming of the constant acceleration models (lower acceleration rates are observed at higher speeds) without adding unnecessary

complexity.

$$a_i(t) \leq a_2 \left(1 - \phi_i(t)\right) \tag{2.18}$$

$$\phi_i(t) = \frac{v_i(t)}{v_f} \tag{2.19}$$

Non Linear Models

Various non linear models of higher complexity have been suggested over time. Akçelik and Biggs (1987) mentions the two-term and three-term sinusoidal, and polynomial models of acceleration, and These models were developed over the simple models explained earlier motivated by observations that acceleration rates typically peak at a speed that is close to, but not equal to zero. Thus, they mimic a behavior where acceleration first quickly increases with speed, reaches a peak, and then decreases with speed over a longer range. Equation 2.20 presents a representative non-linear model that is based on Gipps' complete car following model [Gipps (1981)]:

$$a_i(t) \leq a_3 \left(1 - \phi_i(t)\right) \sqrt{0.025 + \phi_i(t)}, \tag{2.20}$$

$$\phi_i(t) = \frac{v_i(t)}{v_f}. \tag{2.21}$$

Dynamics family of acceleration models

The second family of models incorporates additional parameters involved in the vehicle acceleration process such as the vehicle engine power, frictional losses, mechanical losses, traction on road, grade, air drag etc [Rakha et al. (2004), Searle (1999)]. In doing so, the models attempt to model acceleration based on the factors that actually directly contribute to it, thus steering away from simpler mathematical models that are calibrated purely empirically. These models are also better equipped at intrinsically accommodating for differences in vehicle types and roadway conditions. However, by definition such models tend to involve a

larger set of parameters, and are often too complex to be easily incorporated into aggregate level modeling frameworks, and thus would not be explored as part of the current study.

2.2.2 Macroscopic Acceleration Models

First order macroscopic models of traffic flow such as the LWR model usually ignore the transitional acceleration states of traffic and assume that vehicles can instantaneously accelerate to their desired speeds. The acceleration process is assumed to be completed within an infinitesimally small time and distance.

Higher order models, containing additional components to describe the spatio-temporal evolution of speed (Payne, 1971; Whitham, 2011; Helbing, 2001) are able to include the transitional states and thus are able to bound acceleration. However, there has been much debate about the validity and physical correctness of such models (Daganzo, 1995b; Papageorgiou, 1998; Aw and Rascle, 2000; Zhang, 2003; Helbing and Johansson, 2009).

In 1997, (Lebacque, 1997) first proposed the importance of introducing a bounded acceleration feature into the LWR model through the MBA model by introducing a moving boundary condition whenever a rarefaction wave is formed, thus emulating the realistic trajectory of the leading vehicle in the platoon. (Leclercq, 2002, 2007a) proposed an alternate FBA model based on field of constraints defining the maximum allowed speeds locally. The most prominent contribution perhaps was by Lebacque (Lebacque, 2003) when he introduced the two-phase bounded acceleration model. The intent was to develop an LWR based model, that retained the desirable features of the LWR model (such as analytical calculability), while being able to overcome the underlying infinite acceleration assumption of existing model.

Lebacque's Two Phase Model

(Lebacque, 2003) introduced a two phase model of traffic to capture bounded acceleration of vehicles. In this model, traffic is allowed to exist in one of two phases: a regular LWR phase which models traffic in equilibrium state; and a bounded acceleration (BA) phase with traffic in transitional second-order state with constant acceleration rate equal to the bounded acceleration value. Vehicles either operate under equilibrium conditions with constant speed and no acceleration, or, accelerate at a constant rate up to the desired equilibrium speeds. Lebacque's model is given as:

$$V = V_e(K), \frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} \leq A, \tag{2.22}$$

$$V \leq V_e(K), \frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} = A, \tag{2.23}$$

where A is the maximum achievable acceleration or the bounded acceleration limit and is considered a constant, independent of current traffic state.

The vehicle trajectories in the BA phase, which are parabolic due to constant acceleration, are shown to translate into straight line trajectories on the flow-density graph. This feature makes the generation of the analytical solution feasible for the BA phase. The BA-LWR transitions take the form of shockwaves not dissimilar to the LWR shockwaves, with the difference being that the BA state might be a state inside the fundamental diagram instead of being on the fundamental diagram. The LWR-BA transitions are assumed to propagate from vehicle to vehicle with a delay determined by a reaction time through a phenomenological model.

A simplified model, using the notion of an internal recovery state is also presented, based on the argument that the BA phases exist for short lengths of the roadway, such that the spatial extent of the BA phase can be altogether ignored. The BA-LWR transition in this

case is broken down between the LWR to recovery state boundary and the recovery state to BA phase boundary, where shockwaves exist between the LWR and the recovery states, and rarefaction waves are formed between the recovery and the BA phases.

Field of Constraints / Maximal Velocity Model

(Leclercq, 2007a) proposed a 'field of constraints' model where a maximal allowed speed $V_m(x,t)$ is introduced for each point in space and time. This maximal speed is computed by restricting acceleration to a constant bounded acceleration value 'A' similar to that in the two-phase model. Thus, once again we have:

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} = A \tag{2.24}$$

The allowed speed in the model for each space-time segment is applied as a field of constraints such that acceleration can never exceed the bounded acceleration. The field of constraints is applied intrinsically to the LWR model by modifying the fundamental diagram. The non-linear solution to this relation was shown to have a flux function $\text{flux}(V) = V^2/2$. The solution to the system introduces characteristic waves with speed equal to the derivative of the flux function, or V. Further, all rarefaction wave regions obtained through the solution, are replaced with empty regions of no activity, as part of the entropy conditions. This is akin to the physical observation that no vehicle should have a speed higher than the first exiting vehicle in the platoon in a dispersion scenario.

Leclercq examined the application of the model to the specific case of bounded acceleration close to fixed bottlenecks. The bounded acceleration zone is introduced as a finite length transition area where vehicles accelerate. The acceleration states are summarized through:

$$V(x)\frac{\partial}{\partial x}V(x) = AV(x) = \sqrt{V_u^2 + 2Ax} \tag{2.25}$$

According to the model, the traffic states in the transitional area should be inside the fundamental diagram curve and on a straight line with flow equal to the downstream supply and velocity increasing progressively according to the equation above. The fundamental diagram thus takes the form of a trapezoidal transformation of the upstream fundamental diagram restricted to the downstream capacity flow as the plateau of the trapezoidal curve. Further, the acceleration rate and the length of the downstream acceleration zone L are related:

$$V_d = \sqrt{V_u^2 + 2AL}. \tag{2.26}$$

where V_d is the downstream speed.

Chapter 3

Modified Cell Transmission Model for Realistic Queue Discharge Features at Signalized Intersections

3.1 Introduction

On a road connecting two signalized intersections, a traffic queue grows when vehicles join it at the upstream queue tail, and dissipates when vehicles leave it from the downstream queue head. The two processes are usually referred to as the queue build-up and discharge processes, respectively. In particular, the queue build-up process occurs when traffic lights at the upstream intersection are green or yellow for some vehicles entering the road, and the queue discharge process occurs when traffic lights at the downstream intersection are green or yellow for vehicles on the road. Therefore the interplay between the two processes determines the formation and dissipation of the traffic queue.

Any complete arterial model of traffic thus needs to model both the build-up and the dis-

charge processes as well as incorporate network node models, merge-diverge models, and models for any traffic control systems. Queue build-up models typically study the arrival pattern of traffic through the modeling of upstream controls causing vehicles to platoon, and the dispersion of such platoons as they approach the tail of the queue. Platoon dispersion modeling was first studied and modeled in (Pacey, 1956; Robertson, 1969), and has stayed an actively studied field. More recently, various studies such as (Geroliminis and Skabardonis, 2005) have tried to improve on the modeling of the dispersion and thus of travel time of vehicles on links.

Queue discharge features such as the discharge headways and the discharge flow-rate are other critical components to design and analysis of traffic signals at signalized intersections, as they are used to calculate lost times, effective green times, and capacities of different movements. There have been many empirical studies on headway distribution, which can be influenced by the number of lanes, vehicle types, measurement location, and stochasticity (Gerlough and Wagner, 1967; King and Wilkinson, 1977; Kunzman, 1978; Lee and Chen, 1986; Moussavi and Tarawneh, 1990; Niittymäki and Pursula, 1997; Al-Ghamdi, 1999; Hung et al., 2002; Li and Wang, 2006). The earliest queue discharge model developed for traffic at signalized intersections was perhaps in (Webster, 1958), which used existing queue analysis techniques coupled with traffic observations to estimate queue lengths and delays at intersections. Phenomenological models such as (Akçelik and Besley, 2002), and simulation models such as (Tong and Hung, 2002), have been able to correctly predict various queue discharge features. However these models are not easily applied to study traffic dynamics.

Microscopic car-following models have been applied to study the queueing process on a signalized road network and the resulting discharge features (Tian et al., 2002; Bloomberg and Dale, 2000). At the macroscopic level, the Lighthill-Whitham-Richards (LWR) model (Lighthill and Whitham, 1955; Richards, 1956) has been widely applied to calculate the queue length at a signalized intersection and analyze traffic dynamics in a signalized road

network, usually without considering lane channelization. For example, the signal control problem is solved based on the LWR model (Michalopoulos et al., 1981); and periodic traffic patterns and corresponding macroscopic fundamental diagram were studied with the network kinematic wave model for a double-ring network in (Jin et al., 2013).

Since (Lighthill and Whitham, 1955), it has been shown with the characteristics method that, when the downstream traffic light turns green, the discharge flow-rate at the stop-line is always at capacity until the queue disappears. This suggests that the discharge headways at the stop-line are also constant. However, observed headways decrease with the number of vehicles and approach a saturation value with the 5th or 6th vehicle (Greenshields et al., 1946). Further in (Akçelik and Besley, 2002), models for the distributions of both discharge headways and flow-rates were proposed based on observations. In literature, when applying the LWR model to analyze the queueing process at a signalized intersection, one usually assumes that the discharge process only occurs during an "effective" green time, not the actual green time (Stephanopoulos et al., 1979). Here the yellow signal and the impact of the dilemma zone are also approximately included in the effective green time. Note that the effective green time is defined at the microscopic level, and is equal to the phase time minus the start-up and clearance lost times (Roess et al., 2010). But it is not clear whether such an approximation is reasonable or not. In summary, the LWR model cannot capture the realistic queue discharge features.

By nature, the unrealistic queue discharge property of the LWR model, in which the discharge flow-rate immediately increases to the capacity when the green period starts, stems from the implicit assumption of an infinite acceleration rate for vehicles leaving an initial standing queue. In literature, some efforts have been made to introduce bounded acceleration to the LWR model on a homogeneous road or at a bottleneck (Lebacque, 1997, 2003; Leclercq, 2002). But these models turn out to be of higher orders and are challenging to analyze, calibrate, and validate.

In this chapter, I propose to modify the Cell Transmission Model (CTM) (Daganzo, 1995a), which is the Godunov discrete version of the LWR model, to capture queue discharge features after traffic light turns green. I further focus specifically on the through movements and ignore impacts of all turning movements. CTM has been used towards a variety of tasks such as network modeling (Jin et al., 2013), dynamic traffic assignment (Lo, 1999), and urban intersections (Chen et al., 2008). However, the traditional CTM cannot capture the queue discharge features. Here I apply a new demand function, which linearly decreases in density under over-saturated conditions. Since such a modified demand function was first introduced in (Lebacque, 2003), where the bounded nature of acceleration in discharging vehicles was studied within the framework of higher-order kinematic wave models, it is not surprising the modified CTM can replicate realistic discharge headway and flow-rates. One advantage of the presented model is that only one new jam demand parameter is added, and the model can be analyzed and calibrated with minimal efforts.

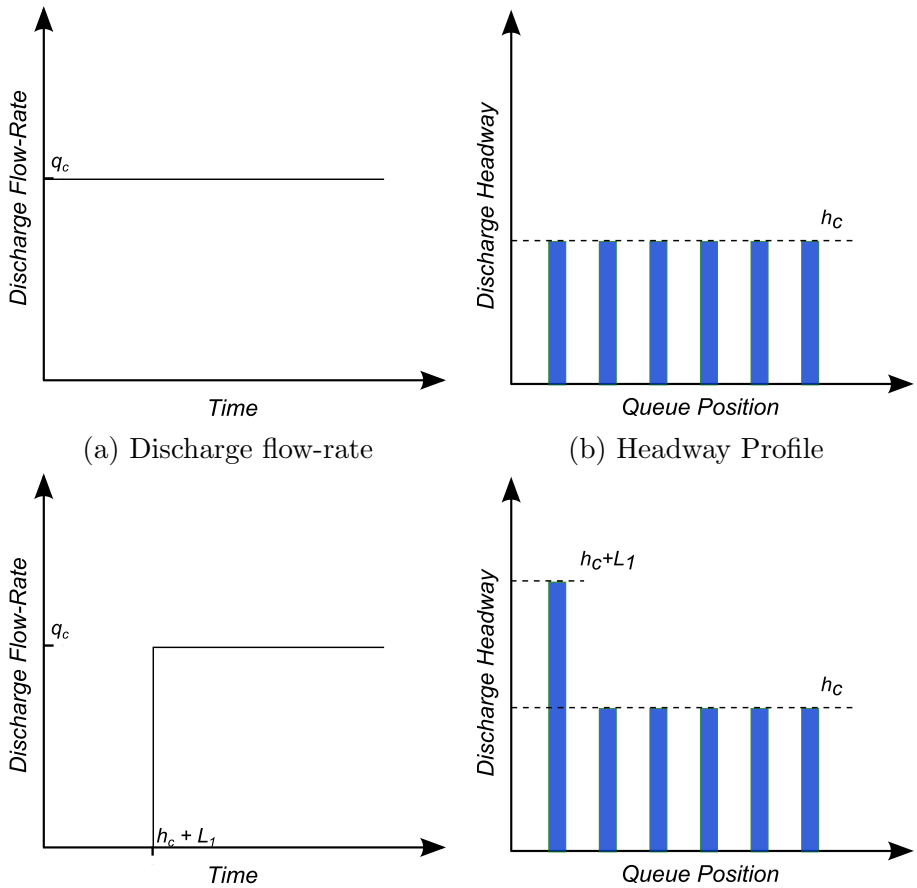
3.2 Review of queue discharge features under the Cell Transmission Model

For a through movement, we can assume that the traffic light turns green at $t = 0$ and the stop-line is at $x = 0$. In a queue discharge scenario, the initial traffic density is:

$$k(x, 0) = \begin{cases} k_j, & x < 0 \\ 0, & x > 0 \end{cases} \tag{3.1}$$

If in CTM we let the queue discharge process immediately start at $t = 0$, then the upstream demand and the downstream supply at $x = 0$ are always q_c, and the discharge flow-rate equals the saturation flow-rate; i.e., $q(0, t) = q_c$ for $t > 0$ till the queue disappears or the traffic light turns red. Correspondingly the queue discharge headway is the saturation headway

$h = h_s = 1/q_c$. The solutions are shown in Figure 3.1a and Figure 3.1b. Such a model would clearly over-estimate the intersection capacity as the product of green time and capacity flow-rate (Dion et al., 2004) and may lead to suboptimal intersection design.



(a) Discharge flow-rate (b) Headway Profile

(c) Discharge flow-rate (d) Headway Profile

Figure 3.1: Discharge flow-rate and headway profiles in the traditional CTM: (3.1a, 3.1b) without lost time; (3.1c, 3.1d) with start-up lost time.

To accommodate the capacity lost due to start-up acceleration of vehicles, a start-up lost time (Webster, 1958) can be introduced in CTM. Traditionally, the start-up lost time is measured as total time difference between actual observed headways of vehicles and the saturation headway. Denoting the lost time as L_1. Thus the discharge flow-rate in CTM

becomes

$$q(0,t) = \begin{cases} 0, & t < L_1 \\ q_c, & t \geq L_1 \end{cases} \tag{3.2}$$

Correspondingly the first vehicle's lost time is increased by L_1. These solutions are shown in Figure 3.1c and Figure 3.1d. This model is more realistic, but even if the start-up lost time is accurate, it is not clear whether the cumulative flow is accurate. In addition, the discharge headways still do not match the observed pattern, and the corresponding dynamics are not correct.

3.3 A modified Cell Transmission Model

3.3.1 A new demand function and the modified CTM

In (Lebacque, 1984), where a network traffic simulation model was developed based on a Godunov discretization of the LWR model, it was argued that the queue discharge flow-rate should gradually increase from 0 to the saturation flow-rate after the traffic light turns green. But the implementation of this discharge feature was not discussed.

In this study, CTM is modified by applying a new demand function, which was introduced in (Lebacque, 2003) for a bounded acceleration extension of the LWR model. In the demand function, traffic demand in congested traffic is less than capacity and decreases with density. Note that the congested part of the demand function can be convex or concave in (Lebacque, 2003). (Monamy et al., 2012) applied the idea of a modified demand function at a node to capture the recovery flow under congested conditions. The demand used in the study is dependent on the number of vehicles currently 'stored' in the node. The resultant was an overall reduced recovery demand that was able to model features similar to observed capacity

drop at test locations. In this study, a demand function that is linearly decreasing in density for congested traffic is used. That is,

$$D'(k) = \min\{v_f k,\ q_j + c^*(k_j - k)\} = \min\{v_f k,\ c^*(k_j^* - k)\}. \tag{3.3}$$

Figure 3.2 shows the new demand function along with definitions of parameters. In addition to the free-flow speed, jam density, and critical density, which are used in the definition of the traditional demand function, a new variable, called jam demand, q_j, is introduced as the demand at the jam density.

From the definitions of the variables, the following relations follow immediately:

$$q_c = c^*(k_j^* - k_c) = w(k_j - k_c), \tag{3.4}$$

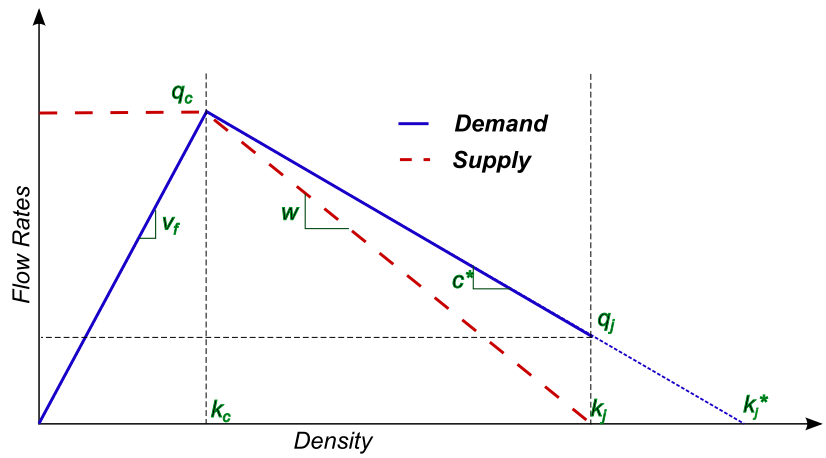
$$q_j = c^*(k_j^* - k_j). \tag{3.5}$$

The slope of the demand curve c^* (not to be confused with shock wave speed), and the extrapolated jam density for demand k_j^* can be derived from the four basic parameters and are also shown in the figure.

Therefore, the modified CTM is only different from the traditional CTM in the definition of the demand in congested traffic and has an additional parameter, the jam demand, but the definition of supply, the flux function, and the density updating equation are still the same as those in traditional CTM (Eq. 2.13)-(Eq. 2.15).

3.3.2 Solutions to the Riemann problem

The new demand function and, therefore, the modified CTM are primarily proposed to obtain more realistic solutions under the queue discharge initial condition in (Eq. 3.1). In



(a) Demand-Supply

v_f	*Free Flow Speed*
q_c	*Capacity Flow Rate*
k_c	*Critical Density*
k_j	*Jam Density*
k_j^*	*Projected Jam Density for Demand*
q_j	*Jam Demand*
w	*Slope of Supply Curve when* $k > k_c$
c^*	*Slope of Demand Curve when* $k > k_c$

(b) Variables

Figure 3.2: The modified demand function along with list of variable definitions

this subsection, however, it is demonstrated that the new model is also well defined under more general Riemann initial conditions:

$$k(x,0) = \begin{cases} k_u, & x < 0 \\ k_d, & x > 0 \end{cases} \tag{3.6}$$

where k_u is the upstream density and k_d the downstream density. The corresponding initial flow-rates are denoted by $q_u = Q(k_u)$ and $q_d = Q(k_d)$.

Following (Lebacque, 1996), the numerical solutions to the modified CTM in the seven scenarios are considered, where traffic is (strictly) under-critical if density is (strictly) lower than the critical density, and (strictly) over-critical if density is (strictly) higher than the critical density.

Note that the new demand function is different from the traditional function only when $k > k_c$. Thus the solutions to the Riemann problem are the same for most of the cases, and the focus is only on the scenarios when the modified CTM has different solutions with possible additional sub-scenarios. The seven resulting scenarios are presented below.

Case I, II: Both upstream and downstream under-critical: $k_u > k_d$ or $k_u \leq k_d$.

Since the new demand function is the same as the traditional one in these cases, the solutions to the Riemann problem remain the same with the modified CTM. Case I (Figure 3.5) results in the formation of a rarefaction wave, while a shock wave is generated in Case II (Figure 3.6).

Case III, IV: Upstream under-critical, downstream over-critical, $q_u < q_d$ or $q_u \geq q_d$.

Similar to cases I and II, these cases lead to the same solutions between the modified and traditional CTM. A shock wave is formed in the solution for both Case III (Figure 3.7) and Case IV (Figure 3.8). In Case III, the shock wave has a positive speed, while in Case IV, the shock wave has a negative speed, as can be deduced from the flow-density diagram.

Case V: Upstream over-critical, downstream over-critical, $k_u > k_d$.

When the upstream density is higher, in the traditional model, a rarefaction wave arises along with accelerating vehicles, and the discharging flow-rate $q(0,t) = S(k_d)$ since $D(k_u) = q_c > S(k_d)$. But in the modified CTM, since $D'(k_u)$ can be smaller than $S(k_d)$, this case should be further separated into two sub-cases.

- **Case V-a:** $D'(k_u) \geq S(k_d)$**.**

  In this case, both the traditional demand $D(k_u)$ as well as the modified demand $D'(k_u)$ is greater than or equal to the downstream supply. Thus, the downstream supply determines the boundary flux in either of the models, and the solutions for both the traditional and modified CTM are identical, resulting in a rarefaction fan (Figure 3.9).

- **Case V-b:** $D'(k_u) < S(k_d)$**.**

  This condition is same as

$$k_c < k_d < k_c + \frac{w^*}{w}(k_u - k_c). \tag{3.7}$$

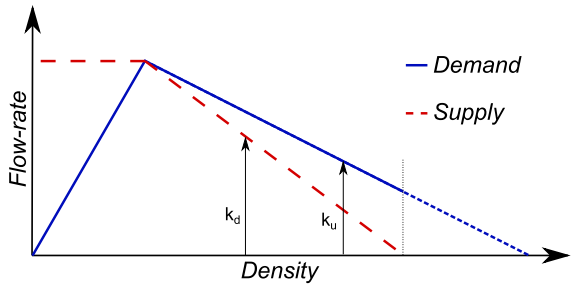
  As an example, the solutions to the traditional and modified CTM are shown in Figure 3.3, where $v_f = 35.19$, $q_c = 1900$, $w = 12.18$, $q_j = 750$, $k_u = 80$ vpm, and $k_d = 60$ vpm. Clearly (Eq. 3.7) is satisfied. It can be seen that, in the modified CTM, the solutions show a discharge like behavior: The initial speed of a discharging vehicle is lower than the speed of downstream block of vehicles. This creates a 'vacuum' condition (zero density) between the first discharged vehicle and the downstream platoon. This region increases in length as the discharged vehicle accelerates up to the downstream platoon's speed. The discharged vehicle is able to accelerate beyond the speed of the downstream platoon due to this 'vacuum' area and eventually catches up with the downstream block creating a shockwave.

Case VI: Upstream over-critical, downstream over-critical, $k_u \leq k_d$**.**

In this case, the upstream demand is always higher than the downstream supply, and the discharging flow-rate is controlled by the downstream supply. Thus in this case the solution is once again unchanged from the traditional model (Figure 3.10).

Case VII: Upstream strictly over-critical, downstream strictly under-critical.

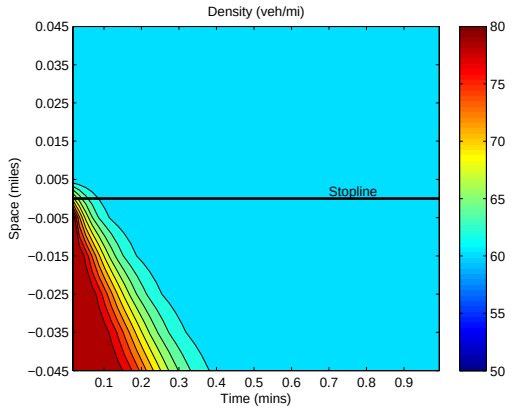
This case is a more generalized form of the queue discharge scenario in (Eq. 3.1), and vehicles discharge from congested platoons to an unrestrictive downstream link. The qualitative results of such a case are discussed in Section 3.4 (see Figure 3.11) and further from data validation results in Figure 3.15. The upstream demand stays in the modified over-critical region, while the downstream supply stays at capacity. The scenario results in formation of rarefaction waves as shown in Figure 3.4.



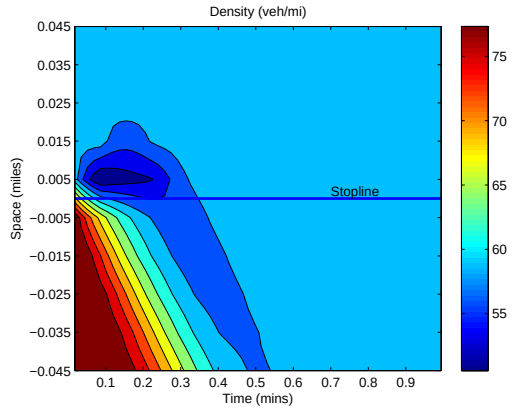
(a) Initial densities

Capacity	1900 veh/hr
Critical Density	54 veh/mi
Jam Density	210 veh/mi
Jam Demand	750 veh/hr
Cell Length	0.01 mi
Initial k_u	80 veh/mi
Initial k_d	60 veh/mi

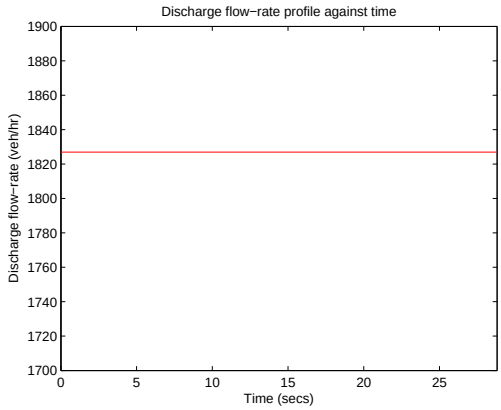
(b) Parameter values



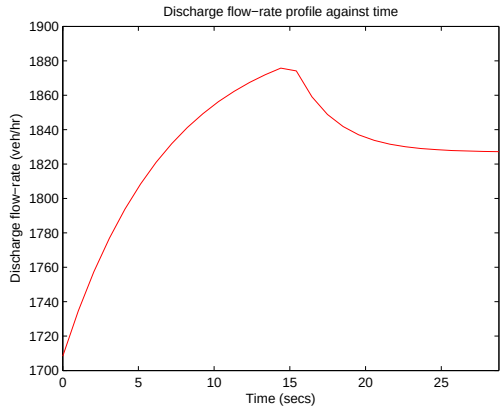
(c) Trad. CTM densities



(d) Mod. CTM densities

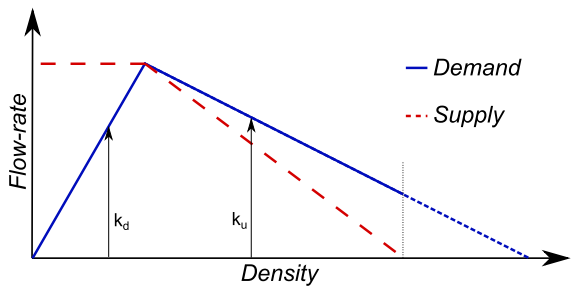


(e) Trad. CTM flow-rate



(f) Mod. CTM flow-rate

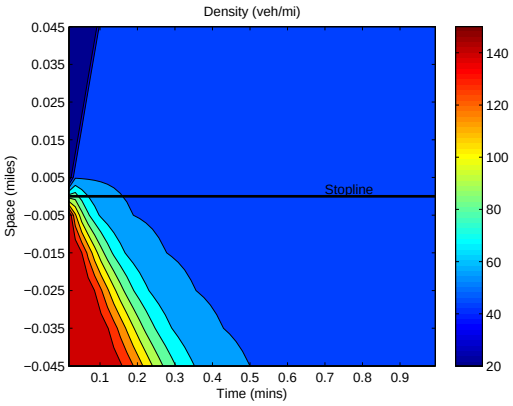
Figure 3.3: Comparison of density contour plots between the traditional and modified CTM in Case V-b.



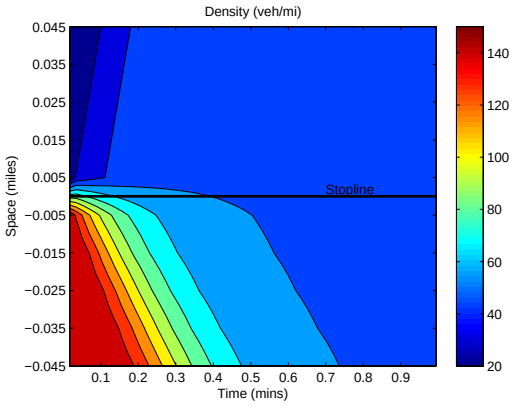
(a) Initial densities

Capacity	1900 veh/hr
Critical Density	54 veh/mi
Jam Density	210 veh/mi
Jam Demand	750 veh/hr
Cell Length	0.01 mi
Initial k_u	150 veh/mi
Initial k_d	20 veh/mi

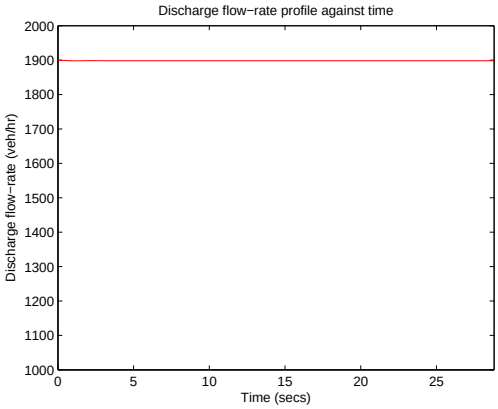
(b) Parameter values



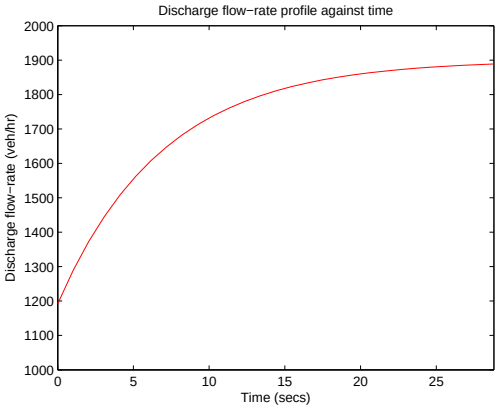
(c) Trad. CTM densities



(d) Mod. CTM densities

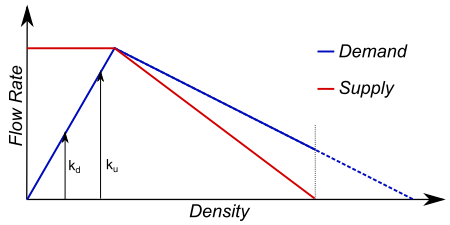


(e) Trad. CTM flow-rate



(f) Mod. CTM flow-rate

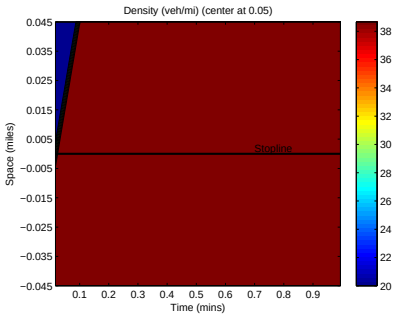
Figure 3.4: Comparison of density contour plots between the traditional and modified CTM in case VII.



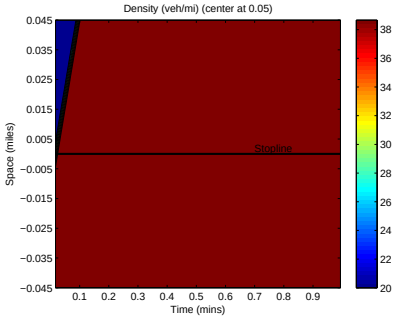
(a) Initial densities

Capacity	1900 veh/hr
Critical Density	54 veh/mi
Jam Density	210 veh/mi
Jam Demand	750 veh/hr
Cell Length	0.01 mi
Initial k_u	40 veh/mi
Initial k_d	20 veh/mi

(b) Parameter values

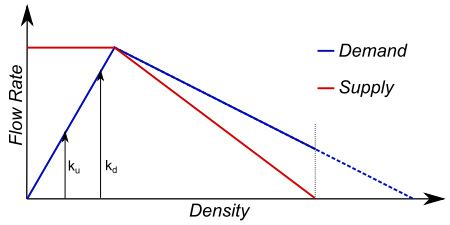


(c) Trad. CTM densities



(d) Mod. CTM densities

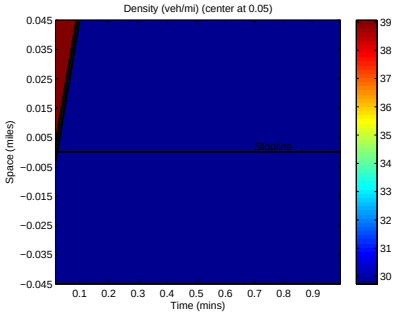
Figure 3.5: Comparison of density contour plots between the traditional and modified CTM in case I.



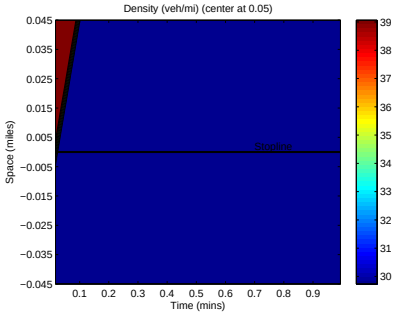
(a) Initial densities

Capacity	1900 veh/hr
Critical Density	54 veh/mi
Jam Density	210 veh/mi
Jam Demand	750 veh/hr
Cell Length	0.01 mi
Initial k_u	30 veh/mi
Initial k_d	40 veh/mi

(b) Parameter values

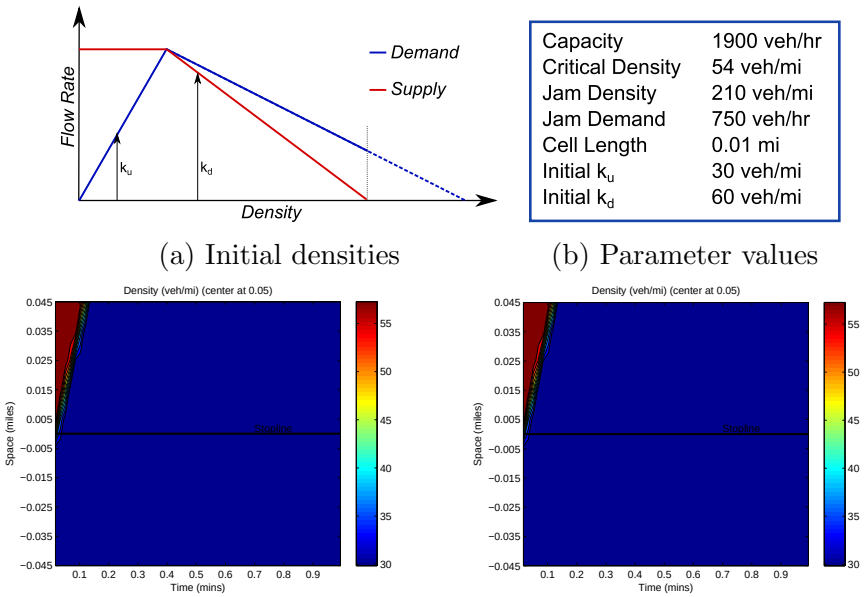


(c) Trad. CTM densities



(d) Mod. CTM densities

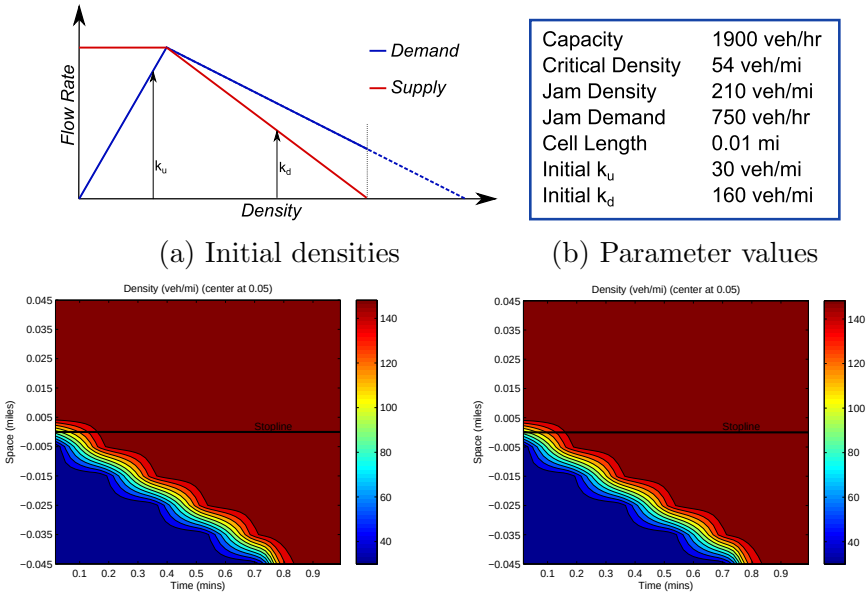
Figure 3.6: Comparison of density contour plots between the traditional and modified CTM in case II.



(a) Initial densities (b) Parameter values

(c) Trad. CTM densities (d) Mod. CTM densities

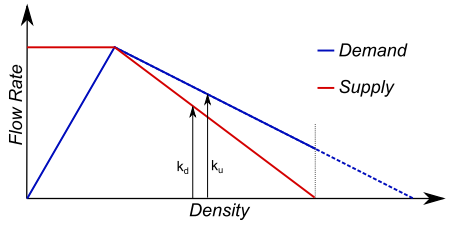
Figure 3.7: Comparison of density contour plots between the traditional and modified CTM in case III.



(a) Initial densities (b) Parameter values

(c) Trad. CTM densities (d) Mod. CTM densities

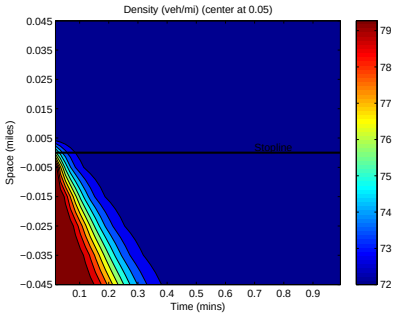
Figure 3.8: Comparison of density contour plots between the traditional and modified CTM in case IV.



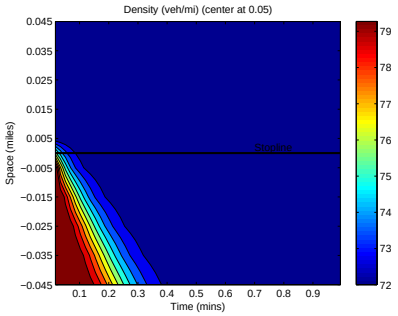
(a) Initial densities

Capacity	1900 veh/hr
Critical Density	54 veh/mi
Jam Density	210 veh/mi
Jam Demand	750 veh/hr
Cell Length	0.01 mi
Initial k_u	80 veh/mi
Initial k_d	72 veh/mi

(b) Parameter values

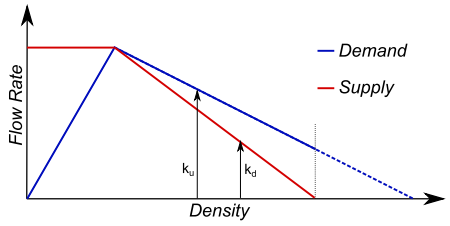


(c) Trad. CTM densities



(d) Mod. CTM densities

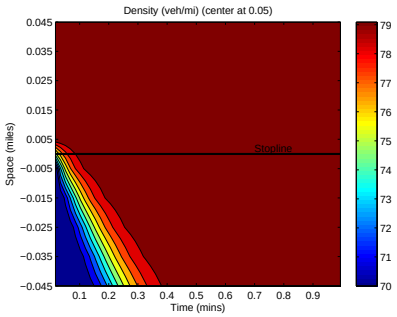
Figure 3.9: Comparison of density contour plots between the traditional and modified CTM in case Va.



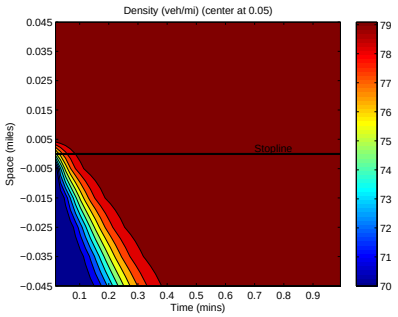
(a) Initial densities

Capacity	1900 veh/hr
Critical Density	54 veh/mi
Jam Density	210 veh/mi
Jam Demand	750 veh/hr
Cell Length	0.01 mi
Initial k_u	70 veh/mi
Initial k_d	80 veh/mi

(b) Parameter values



(c) Trad. CTM densities



(d) Mod. CTM densities

Figure 3.10: Comparison of density contour plots between the traditional and modified CTM in case VI.

3.4 Queue discharge features of the modified CTM

In this section, the modified CTM is analytically solved for the queue discharge scenario under the initial condition given in (Eq. 3.1). From the solutions the queue discharge features of the new model are revealed.

3.4.1 Discharge flow-rate and headway

We can denote the density of the cell upstream to the stop-line by $k(t)$, and the cell length by L. For the queue discharge scenario, cells upstream to the stop-line are always over-critical, and those downstream always under-critical. Thus the in-flux of the studied cell is determined by its supply, $w(k_j - k)$, and its out-flux by its demand, $c^*(k_j^* - k)$. For the purpose of simple analysis, the discrete CTM, (Eq. 2.15), is approximated by the following ordinary differential equation:

$$L\frac{dk}{dt} = w(k_j - k) - c^*(k_j^* - k), \tag{3.8}$$

where the initial condition is $k(0) = k_j$. In addition, the queue discharge flow-rate is given by the out-flux

$$q(0, t) = c^*(k_j^* - k). \tag{3.9}$$

From (Eq. 3.4), we can simplify the right-hand side of (Eq. 3.8) as $(w - c^*)(k_c - k)$. Thus solving (Eq. 3.8), we obtain the time-dependent density in the cell upstream to the stop-line as:

$$k(t) = (k_j - k_c)e^{\frac{c^* - w}{L}t} + k_c, \tag{3.10}$$

which decreases in time since $c^* < w$ and converges to k_c asymptotically. Correspondingly the discharge flow-rate is given by:

$$q(0,t) = q_c - c^*(k_j - k_c)e^{\frac{c^*-w}{L}t}, \tag{3.11}$$

which increases from $q_j = c^*(k_j^* - k_j)$ at $t = 0$ to q_c asymptotically. Thus the jam demand, q_j, is also the initial discharge flow-rate. The resulting discharge flow-rate has the same exponential form as that estimated from observations in (Akçelik et al., 1999; Akçelik and Besley, 2002).

Denoting the cumulative number of vehicles passing the stop-line by $A(t)$, it can be obtained from (Eq. 3.11) as:

$$A(t) = \int_{s=0}^{t} q(0,s)ds = q_c t + \frac{L}{c^* - w} c^*(k_j - k_c)(1 - e^{\frac{c^*-w}{L}t}), \tag{3.12}$$

which is an increasing function in time t. Then the passing time is the inverse function

$$T(a) = A^{-1}(a), \tag{3.13}$$

and the time for vehicle n to pass the stop-line is t_n:

$$t_n = T(n). \tag{3.14}$$

Further, denoting the continuous headway as $h(t)$,

$$h(t) = \frac{1}{q(0,t)} = \frac{1}{q_c - c^*(k_j - k_c)e^{\frac{c^*-w}{L}t}}, \tag{3.15}$$

which decreases from $\frac{1}{q_j}$ to the saturation headway $h_s = \frac{1}{q_c}$. If we denote the discrete headway

of vehicle n by h_n, then after solving t_n from (Eq. 3.14) we have ($n = 1, 2, \cdots$)

$$h_n = t_n - t_{n-1}, \tag{3.16}$$

where we let $t_0 = 0$. From (Eq. 3.14) and (Eq. 3.16) we have

$$q_c h_n + \frac{c^*}{c^* - w}(k_j - k_c) L e^{\frac{c^* - w}{L} t_{n-1}} (e^{h_n} - 1) = 1. \tag{3.17}$$

Since $t_{n-1} \to \infty$, we can see that the discrete headway also converges to the saturation headway; i.e., $h_n \to h_s = \frac{1}{q_c}$.

Figure 3.11 demonstrates the demand and supply functions in Figure 3.11a, the discharge flow-rate in Figure 3.11b, the continuous headway in Figure 3.11c, and the discrete headways in Figure 3.11d.

3.4.2 Lost Times

During the queue discharge process, the start-up lost time, L_1, is defined as the sum of the differences between all headways and the saturation headway (Roess et al., 2010):

$$L_1 = \sum_{n=1}^{\infty} (h_n - h_s). \tag{3.18}$$

From (Eq. 3.16) and (Eq. 3.12), we can see that

$$L_1 = \lim_{a \to \infty} T(a) - h_s a = \lim_{t \to \infty} t - \frac{A(t)}{q_c} = \frac{L}{w - c^*} \frac{c^*}{w}. \tag{3.19}$$

Furthermore, from (Eq. 3.12) we can see that $\lim_{t \to \infty} A(t) = q_c(t - L_1)$. That is, $q_c(t - L_1)$ is the asymptotic line of $A(t)$. Compared with the ideal situation when the discharge flow-rate

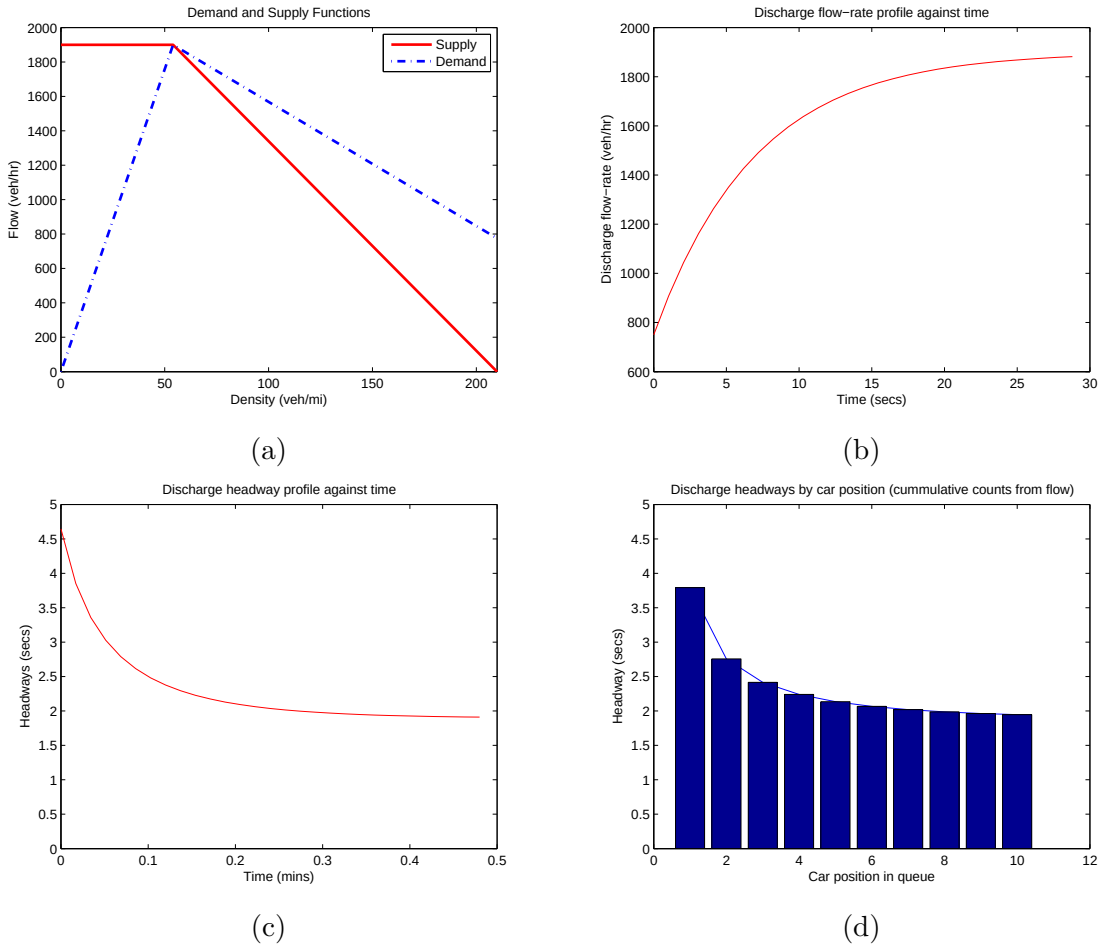


Figure 3.11: Queue discharge flow-rate and headways: (a) demand and supply curves; (b) discharge flow-rate; (c) continuous headways; and (d) discrete headways. The results shown are for $q_c = 1900$, $k_c = 54$, $k_j = 210$, and $q_j = 775$, with upstream fully congested at jam density, and downstream density set to 0 as initial conditions.

is always q_c, there is a lost capacity (as shown in Figure 3.12) of:

$$\Delta C = \lim_{t \to \infty} q_c t - A(t) = q_c L_1. \tag{3.20}$$

This justifies the idea of the effective green time in (Eq. 3.2); that is, if the start-up lost time L_1 is defined as in (Eq. 3.19), then the total number of discharged vehicles will be equivalent if we use the new CTM model or if we use (Eq. 3.2).

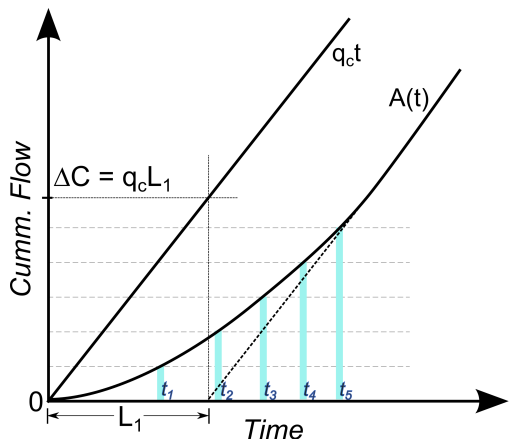


Figure 3.12: Relation between lost capacity and start-up lost time. The figure illustrates how the cumulative flow from the modified model, $A(t)$, is asymptotic to the line $q_c(t - L_1)$. The figure also shows the corresponding passing time of each queued vehicle t_i.

3.5 Calibration and validation

The presented model can be calibrated using minimal additional headway observations if the macroscopic variables associated with traffic behavior at the intersection are known (note that if an appropriate value for L is known, q_j and c^* can be calibrated directly knowing either the headway of the first vehicle, or the total lost time (Eq. 3.19)). However, for robust calibration, detailed headway observations are recommended. Where vehicle headway observations are available, statistical fitting techniques can be used to calibrate the parameters of the model. This is in fact the approach taken in the current study.

3.5.1 Approach

In the traditional CTM, we need to calibrate the three parameters in the triangular fundamental diagram (Eq. 2.11), and the demand and supply functions in (Eq. 2.12)-(Eq. 2.13): the capacity flow-rate (saturation flow-rate), q_c, the critical density, k_c, and the jam density of queued vehicles, k_j; then v_f and w can be calculated as $v_f = q_c/k_c$ and $w = q_c/(k_j - k_c)$.

For the modified CTM, we need to calibrate an additional parameter for the modified demand

function in (Eq. 3.3): the jam demand, q_j. From observations of discharge headways at an intersection, the corresponding passing times for vehicles as they cross the stop-line, t_i, can be obtained. From the modified CTM, we also know the cumulative number of vehicles passing the stop-line as a function of time and the parameters w, c^*, k_c, q_c, k_j and L (Eq. 3.12). However, the relation can be re-written using the unique parameter set: k_c, q_c, k_j, q_j, and L, using (Eq. 3.4)-(Eq. 3.5). The notation $\hat{A}(t)$ is used here to highlight that this is the modeled prediction in the optimization setup:

$$\hat{A}(t) = q_c t - L\left(\frac{q_c - q_j}{q_j}\right)(k_j - k_c)\left(1 - e^{-\frac{q_j}{L(k_j - k_c)}t}\right). \tag{3.21}$$

We can then obtain the predicted cumulative count of vehicles corresponding to actual observed passing times of each vehicle. Thus, an optimization problem can be set up on the cumulative vehicle counts (with passing time as the independent variable) (Figure 3.13). A minimization of the Sum of Squared Errors is used as the optimization criterion here.

$$\begin{aligned} &\underset{X}{\text{minimize}} && \sum_{i=1}^{n}\left(\hat{A}(X, t_i) - i\right)^2, \\ &\text{where,} && X = \{k_c, q_c, k_j, q_j, L\}, \\ &\text{subject to:} && X_L \leq X \leq X_H. \end{aligned} \tag{3.22}$$

Discharge headway data is only applicable for those vehicles that are part of the standing queue before the discharge process starts. This means that the number of data points in the optimization is restricted by the length of the standing queue observed at the location of interest (typically 5-10 vehicles). Since the model requires the calibration of five parameters, thus having a high degree of freedom, there is an inherent risk of over-fitting unless constraints are imposed on the parameters. It is thus imperative to ensure that certain ranges are imposed on the values of the calibration parameters, or that some of the parameters are held constant based on knowledge of traffic behavior at the location (forming the constraints in

the optimization setup).

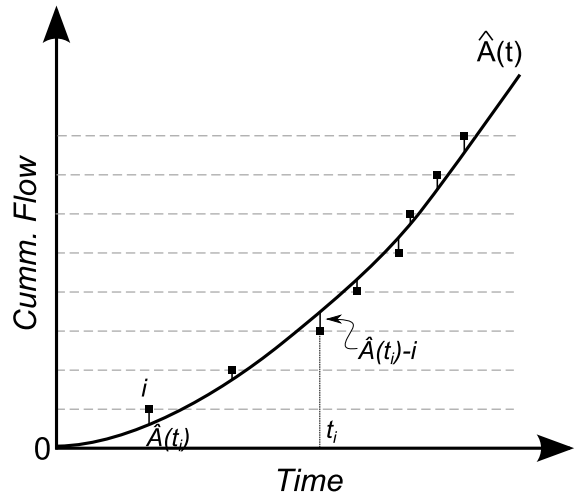


Figure 3.13: The figure illustrates the setup of the calibration process. $\hat{A}(t)$ is the predicted cumulative flow at time t from the model, while the observed passing times for vehicle i are shown as t_i. Since $\hat{A}(t_i)$ represents the modeled cumulative flow corresponding to the observed passing time of the ith vehicle, the objective of the optimization is to minimize $\sum(\hat{A}(t_i) - i)^2$

3.5.2 Calibration

As part of the present study, the linear demand model was calibrated for three data sets: data from various headway studies in past literature, NGSIM (FHWA, 2007) data from Lankershim Blvd., and video recordings from a local study site. Calibration of the model was done for each location individually and employed knowledge of certain features such as free-flow speed and jam density where available. While the fitting technique described herein using headway data is a viable option, headway data points are typically only available for a handful of vehicle positions leading to a possibility of over-fitting. Thus it is important to note, that the recommended calibration process is to first calibrate the fundamental diagram (v_f, q_c, k_c, k_j) from observing steady state traffic flow, and then calibrate the additional q_j and L parameters from observed discharge headways.

The first data set used for the calibration process, was obtained from headway studies from

past literature. (Gerlough and Wagner, 1967), (Hung et al., 2002), and Tong and Hung's simulation results (Tong and Hung, 2002) have presented various study results on headway data. The model was calibrated for each of the three sets of headway observations.

NGSIM data of through moving traffic from Lankershim Blvd SB, at its intersection with Universal Hollywood Dr (Figure 3.14b) was used as the second data set. The study site offers clear camera visibility and supports a steady demand for through movements. The location has 3 through lanes (1 shared right turn lane). Only vehicles in a standing queue at start of green phase were recorded. Observations suggested a critical density of 54 veh/mi, jam density of 210 veh/mi, and a saturated headway of roughly 1.8 sec at the location.

Finally, a video recording study site was chosen for additional data collection near the University of California, Irvine campus, at a push button pedestrian crossing signal on E. Peltason (between Los Trancos Drive, and Anteater Drive; shown in Figure 3.14a). The location offered observations of single-lane traffic at an intersection that allowed only through movements. Using a video recording device, the headways of vehicles (by queue positions) were observed over multiple cycles and averaged to obtain the headways vs. queue position curve. Still photographs of queued vehicles at the location suggested the jam density to be in the range of 210 veh/mi. The free-flow speed at the location was observed to be 35 mi/hr (30 mi/hr posted speed limit).

The observation of the critical density from the data was used to fix the k_c value for all 5 experiments as a value of 54veh/mi. Similarly, L was held constant at 0.01 mi. The jam density k_j was also set to 210 veh/mi for 4 out of 5 experiments (the exception being data from Gerlough and Wagner's study where k_j was also an optimization variable). The calibration technique used was the minimization problem in (Eq. 3.22). MATLAB's implementation of the Nelder-Mead heuristic search (Nelder and Mead, 1965) was used to find the solution to the non-linear optimization, and the results were verified against Generalized Reduced Gradient technique.

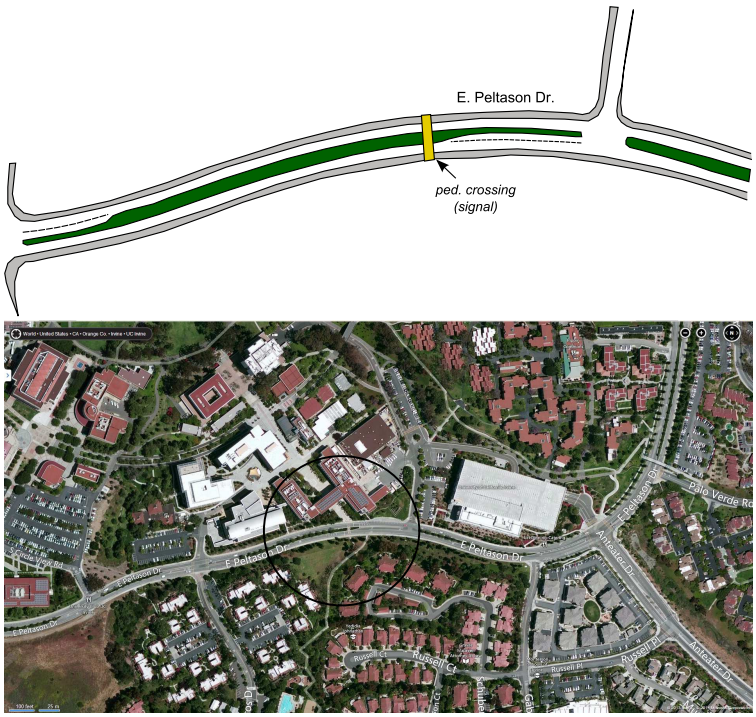
The final form of the optimization problem for the datasets used was:

$$
\begin{aligned}
\underset{\{k_c,q_c,k_j,q_j,L\}}{\text{minimize}} \quad & \sum_{i=1}^{n} \left(q_c t - L \left(\frac{q_c - q_j}{q_j} \right) (k_j - k_c) \left(1 - e^{-\frac{q_j}{L(k_j - k_c)} t} \right) - i \right)^2, \\
\text{subject to:} \quad & k_c = 54, \\
& 1200 \leq q_c \leq 2400, \\
& 80 \leq k_j \leq 280, \\
& k_j = 210, \text{(for all cases excluding Gerlough and Wagner)}, \\
& 0 \leq q_j \leq 1500, \\
& L = 0.01.
\end{aligned}
\tag{3.23}
$$

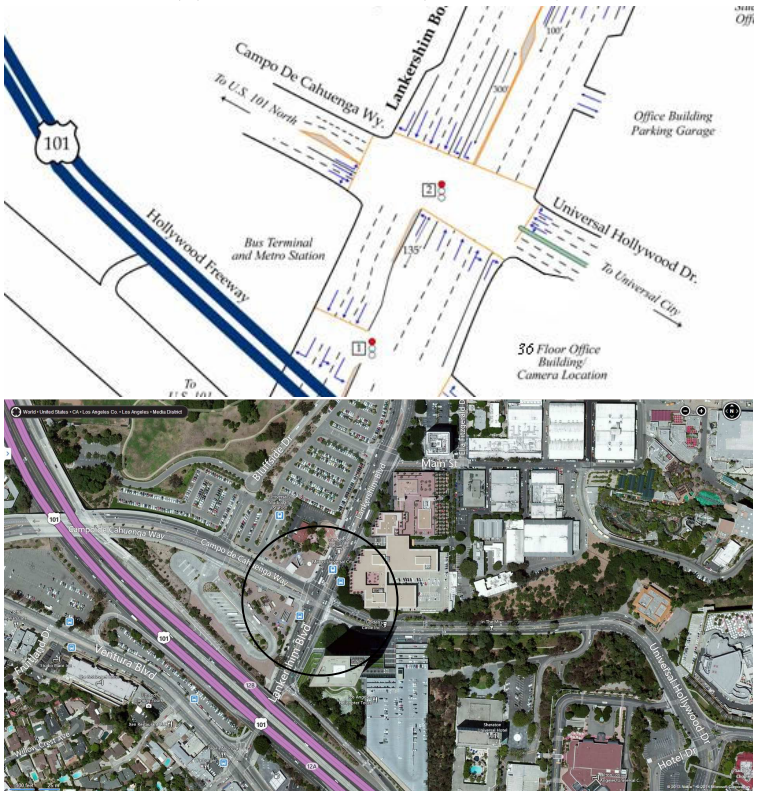
Table 3.1 presents results from the calibration process. The reported sum of squared errors (SSE) and R^2 values in the table were calculated for headways and not the inverse function that the fitting was performed on. The SSE, thus, is the sum of squared differences between the model-predicted headway and the actual observed headway for each queue position. Similarly, R^2 was calculated as $1-(SSE/SS_{tot})$ where SS_{tot} is the sum of squared differences of observed headway with mean headway (proportional to variance of observed data). Figure 3.15 shows the calibrated demand supply functions and the derived discharge headways as compared against observations for each of the cases. The calibration results suggest that the modified CTM indeed replicates realistic features of queue discharge to high accuracy.

		Gerlough (1967)	Hung et. al. (2002)	Tong, Hung (2002)	NGSIM	Recordings
Capacity	(v/h/l)	1530	1760	1775	1950	1900
Critical Density	(v/m/l)	54	54	54	54	54
Jam Density	(v/m/l)	147	210	210	210	210
Jam Demand	(v/h/l)	730	733	630	775	750
L	(mi)	0.01	0.01	0.01	0.01	0.01
R^2		0.99	0.99	0.98	0.96	0.91
SSE	(sec^2)	0.04	0.02	0.07	0.14	0.45

Table 3.1: Headway fitting - Calibrated parameters for the various data sources along with the SSE obtained for each case.

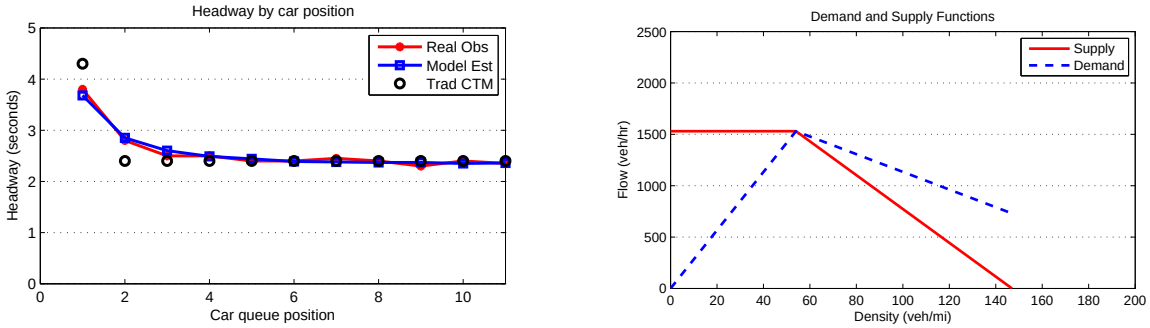


(a) E. Peltason Dr (Recording)

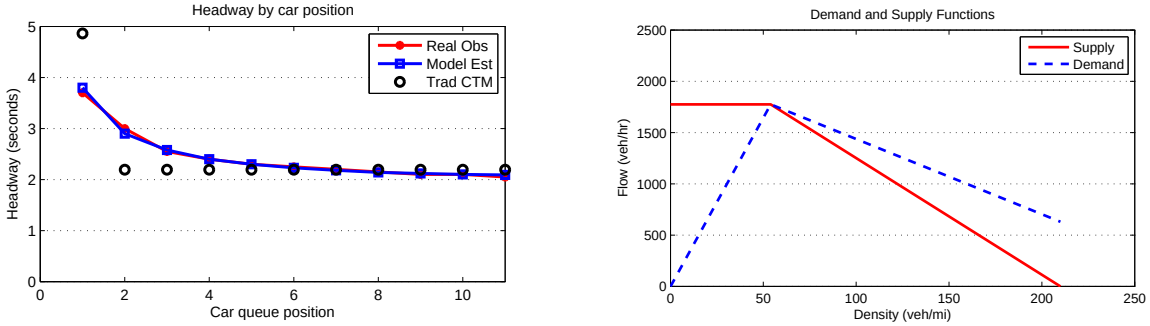


(b) Lankershim Blvd (NGSIM)

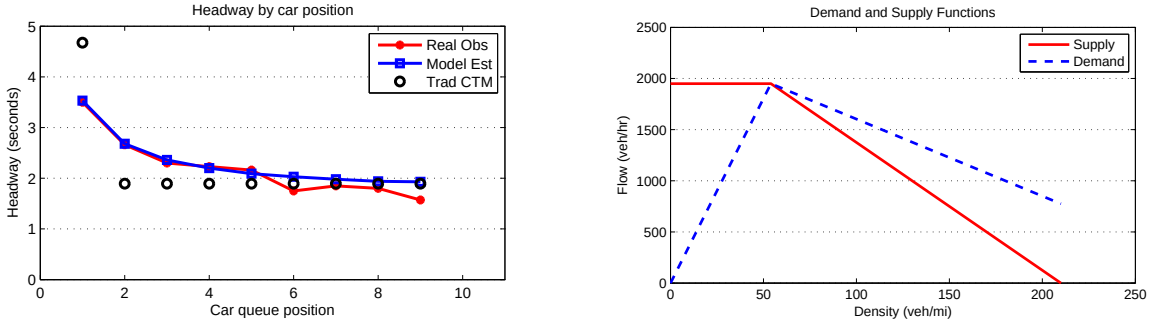
Figure 3.14: Schematic diagrams and aerial maps of the two study sites: 3.14a. E. Peltason Dr. in Irvine, CA; and 3.14b. Lankershim Blvd. and Universal Hollywood Dr. in Los Angeles, CA (Maps coutesy Bing)



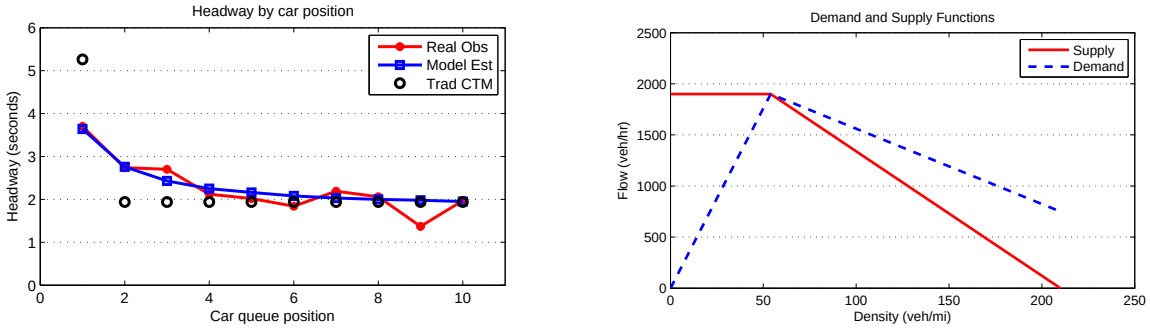
(a) Gerlough, Wagner (1967)



(b) Hung et. al. (2002)



(c) NGSIM



(d) E. Peltason (Video Recording)

Figure 3.15: Headway calibration for the various data sources showing the fitted headway curves and the corresponding demand-supply functions. The graphs depicting headways provides a comparison between observed headways, headways predicted by traditional CTM with lost time, and headways predicted by a calibrated modified CTM.

3.6 Conclusions

Signalized intersection capacity is typically estimated through calibration of lost times of vehicles discharging from a queue at the intersection. This lost time in turn is calculated from the differences between the headways observed by the leading vehicles and the saturation headway. The increased headway for the first few vehicles in the queue is attributed to the reaction times of drivers, as well as the bounded acceleration process involved. In order to adjust for the lost discharge capacity during the initial seconds of the green phase due to higher headways, typically an assumption is made that puts the discharge capacity at zero for an initial period, followed by saturated capacity discharge.

In this chapter, the use of a modified Cell Transmission Model with a new demand function has been proposed, in order to capture realistic queue discharge features at an intersection when vehicles are accelerating away from an initial standing queue. The new model suggests that the demand immediately upstream of the discharge location is indeed not equal to the capacity, due to the vehicles' acceleration being bounded. Through the modeling of the reduced demand, the model is able to also capture the dynamics of traffic behavior at intersections. Further, the modification introduces a single additional parameter thus keeping the calibration process simple. The possibility of obtaining analytical results, as well as the computational tractability offers the model distinct advantages over more complex models.

It is first shown analytically that the resulting headway and discharge flow-rate profiles over time obtained through this queue discharge model are good representations of the expected shapes of those curves qualitatively. The model is then calibrated for data sets obtained from past headway studies as well as newly obtained data, and is shown to model the headways with good accuracy. A discussion on the similarities and comparisons of the existing 'Lost Time' based model and the currently proposed model reveals strengths of both the models. This discussion further leads to a more rigorous definition of the macroscopic 'lost time' and

a means of estimating it from the model itself.

The discharge features and the lost capacity as captured by the model is a combination of the effects of vehicle reaction times, and the bounded acceleration process. While the acceleration is the most critical component to be modeled, it is important to note that the effects of reaction time are also caught as they get internalized within the discharge patterns. Indeed, since reaction time is a truly microscopic feature that might depend on the stimuli to which the driver is reacting, such as the perceived behavior of a leading vehicle or the change in signal phase, a macroscopic model would not be able to distinguish the various properties of reaction time and model them correctly. Since the discharge property encompasses all the details of the reaction process, the presented model can be considered to capture all elements that are important from a macroscopic framework.

An interesting property of the modified CTM lies in the importance of choosing a reasonable time-step size Δt (or cell length L which is also constrained by Δt). The time-step size in the model is loosely related to the reaction time and the acceleration rates of the vehicles. A very small Δt would imply near-instantaneous reactions from drivers and vehicles, and thus result in a flattening of the headway curve with all discharging vehicles having identical headways. Conversely, a large Δt would imply that the initial vehicle would have a very large headway. Studying the model under various scenarios suggests that Δt should be chosen to be in the order of 1-2 seconds which is slightly higher than the time-steps used for microscopic models of traffic.

Queue discharge dynamics and the effects of bounded acceleration can also be captured through higher-order models and hybrid models. Bounded acceleration extensions to the CTM framework (Lebacque, 1997, 2003; Leclercq, 2002) can also be used to catch the behavior of vehicles accelerating from a queue at intersections to produce similar results. However, such models do not lend themselves easily to mathematical analysis, and typically require higher degrees of calibration efforts. To the best of the author's knowledge, there have

not been any calibration and validation studies that compare performances of such models against observed data at arterial intersections, which highlights their complex structure. It would be an interesting extension to compare the performance of such higher order models against the currently proposed model calibrated from observed data. Similarly, though the piecewise linear demand function is a reasonable approximation of the effect of acceleration of vehicles, exploration of other shapes of the demand function could be an interesting study.

This study opens up a debate on macroscopic modeling of intersections that can capture the lost capacity phenomenon at signals due to bounded accelerations and reaction times. Among the many possible extensions of this model, addition of an intersection model could offer a framework for analyzing the effects of the lost capacity over more complex networks of arterial roadways and intersections. Another possible extension could be exploring the portability of the calibrated parameters between locations, as well as studying how factors such as driver aggression and control measures affect the various parameters of the model. The proposed model is able to offer insights into such aspects as the capacity of the network, and prospectively through further extensions, observations such as capacity reductions in network fundamental diagrams and impacts of autonomous vehicles.

Chapter 4

LCBA-CTM for Bounded Acceleration and Lane Changing at Active Freeway Bottlenecks

4.1 Introduction

Researchers have long observed that capacity (maximum possible flow-rate, or throughput) is not a static feature at bottlenecks, and that there is a reduction in the achievable capacity due to activation of bottlenecks. This phenomenon is commonly referred to as 'Capacity Drop', while the capacity before and after congestion are typically called the pre-congestion or uncongested capacity and post-congestion or congested capacity (Bank, 1991). The difference between the uncongested and congested capacities is usually referred to as the capacity drop magnitude.

The variable nature of capacity had long been suspected (Gazis et al., 1961; Drake et al., 1967; Ceder and May, 1976). However, the existence of capacity drop downstream of bottlenecks

was first confirmed in (Hall and Agyemang-Duah, 1991; Bank, 1991), and has since been observed and confirmed multiple times. The extent of the drop, typically measured as a percentage of the uncongested capacity, seems to vary from one location to another and has also been actively debated. While (Hall and Agyemang-Duah, 1991; Bank, 1991) had observed drops of roughly 5% (6% in the former, 3% in latter), later studies observed higher drops. (Cassidy and Bertini, 1999) reported capacity drops close to 12%, (Srivastava and Geroliminis, 2013) observed capacity drops close to 15%, and (Chung et al., 2007) reported up to 18% drop in capacity at active bottlenecks.

In recent years, there has been much interest in understanding the underlying features of traffic that cause this capacity drop. An understanding of the mechanism offers an opportunity to predict and model capacity drop, as well as to mitigate it through control strategies such as ramp metering and variable speed limits. As early as (Hall and Agyemang-Duah, 1991), capacity drop was observed to be associated with upstream queueing and with the acceleration process of slowed vehicles as they depart from the bottleneck. (Cassidy and Rudjanakanoknad, 2005) were able to observe associations between lane changing and capacity drop. These and other subsequent studies have attributed the capacity drop phenomenon to both lane changing (Laval and Daganzo, 2006; Jin, 2010a, 2013), and bounded acceleration (Leclercq, 2007a,b) of vehicles discharging from an active bottleneck. It has been suggested that capacity drop can not be truly modeled in its entirety without accounting for both these factors.

While there have been many attempts at modeling capacity drop through either lane changing or bounded acceleration models, there exists no scalable macroscopic model for capacity drop that incorporates the effects of both lane changing vehicles and bounded acceleration together. This present study attempts to fill this gap. The study uses a lane-drop bottleneck to show how the Cell Transmission Model (Daganzo, 1995a) can be extended through components incorporating lane-changing and bounded acceleration, to present a complete

model of capacity drop at such a location. The chapter starts by first reviewing the Cell Transmission Model in the following subsection.

4.1.1 Cell Transmission Model

The LWR model (Lighthill and Whitham, 1955; Richards, 1956) laid the foundation for the majority of work on macroscopic modeling of traffic that followed since. In the LWR model, traffic, denoted through density k, speed u, and flow-rate q, is treated as a continuous fluid, and is modeled as:

$$\frac{\partial k}{\partial t} + \frac{\partial Q(k)}{\partial x} = 0. \tag{4.1}$$

The above equation is a combination of the traffic conservation law, the fundamental law of traffic flow, and the fundamental flow-density relationship of traffic. The LWR model can either be solved analytically or numerically using discretization schemes.

(Daganzo, 1995a) proposed a discretized framework, the Cell Transmission Model (CTM), for solving the LWR model, which has since become one of the most accepted frameworks for macroscopic modeling of traffic. Space is discretized into cells (length Δx) as shown in Figure 4.1a, while time is discretized into time steps (length Δt). CTM introduces the concept of demand (or sending flow), and supply (or receiving flow) as functions of the densities in the cells. The boundary flux (flow-rate across the shared boundary of any two cells) is determined by the demand in the upstream cell and supply in the downstream cells. For the present study, the triangular flow-density relationship (Munjal et al., 1971) would be used as the fundamental relationship for traffic. The equilibrium flow as a function of density can then be defined as:

$$Q_e(k) = min\{v_f k, w(k_j - k)\}. \tag{4.2}$$

And the corresponding supply and demand relations (Figure 4.3a) are:

$$\text{Demand} = \quad D(k_i) \quad = min\{v_f k_i, q_c\}, \tag{4.3}$$

$$\text{Supply} = \quad S(k_i) \quad = min\{q_c, w(k_j - k_i)\}, \tag{4.4}$$

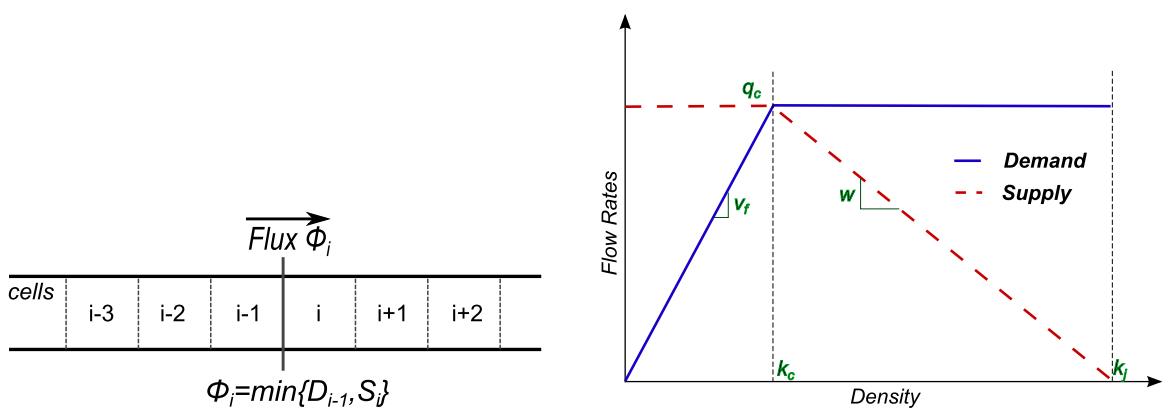
$$\text{Boundary Flux} = \quad \Phi_i \quad = min\{D(k_{i-1}), S(k_i)\}, \tag{4.5}$$

where, v_f is the free flow speed, q_c is the capacity flow-rate, k_j is the jam density, $-w$ is the speed of the congested shock wave, and Φ_i is the flux entering cell i.

The density in each cell is then updated for the next time step according to:

$$k_i(t + \Delta t) = k_i(t) + \frac{\Delta t}{\Delta x} \left(\Phi_i(t) - \Phi_{i+1}(t) \right), \tag{4.6}$$

Δx and Δt follow the Courant-Friedrich-Lewy condition (Courant et al., 1967 [orig.: 1928]).



(a) Cells and Boundary Flux

(b) Trad. Demand, Supply

Figure 4.1: An illustration of the Cell Transmission Model. The figure on the left shows the discretization scheme, while the figure on the right shows the demand and the supply functions used with CTM corresponding to the triangular fundamental diagram.

4.1.2 Lane changing models

The behavior of weaving and lane-changing vehicles at a freeway section has an important effect on freeway performance. Weaving maneuvers can be disruptive to the traffic flow depending on the prevailing conditions. This has led to a lot of interest by researchers towards modeling the lane changing behavior at various levels of detail. At the microscopic level, lane changing behavior typically deals with each individual vehicle's lateral movement process during lane-changing (Gipps, 1986; Kesting et al., 2007). The same can also be extended and modeled from a driver behavior perspective where lane-changes are typically first categorized as discretionary or mandatory, and then estimated based on numerous traffic, vehicle and driver characteristics (Ahmed et al., 1996; Choudhury et al., 2007; Toledo and Zohar, 2007). At a macroscopic level, the lane-changing can be modeled at an aggregate level as exchange of flow across lane boundaries as a derivative of either density perturbations between lanes, or increased utility due to speed differences (Michalopoulos et al., 1984; Holland and Woods, 1997; Daganzo, 2002; Munjal and Pipes, 1971a,b; Gazis et al., 1962). Lane changing can also be implemented through hybrid models, treating the lane changing vehicles as moving bottlenecks with respect to their impact on the target lane (Laval and Daganzo, 2006). Such hybrid models can show how lane changing can lead to capacity drop (Laval et al., 2007; Laval and Leclercq, 2008).

4.1.3 Bounded acceleration models

First order macroscopic models of traffic such as the LWR model usually imply that vehicles are able to accelerate instantaneously to their desired speeds, thus ignoring the transitional acceleration states. Higher order models such as (Payne, 1971; Whitham, 2011; Helbing, 2001) can model the transitional acceleration states of traffic, and bounded acceleration well. However such models are not mathematically tractable.

(Lebacque, 2003) introduced a two phase model of traffic to capture bounded acceleration of vehicles. In this model, traffic is allowed to exist in one of two phases: a regular LWR phase which models traffic in equilibrium state; and a bounded acceleration phase with traffic in transitional second-order state with constant acceleration rate equal to the bounded acceleration value. Vehicles either operate under equilibrium conditions with constant speed and no acceleration, or, accelerate at a constant rate up to the desired equilibrium speeds.

(Leclercq, 2007a) proposed a 'field of constraints' model where the maximal allowed speed is calculated and externally imposed as a constraint for each segment (divided in time and space) of the freeway. The allowed speed is determined by restricting acceleration to a constant bounded acceleration value 'A', which is dependent on the length of the acceleration zone. The field of constraints is applied to the LWR model by modifying the fundamental relationship.

In the remainder of the chapter, the lane changing and bounded acceleration models that inspire the present work and form the basis for the Lane Changing Bounded Acceleration Cell Transmission Model (LCBA-CTM) framework are first reviewed in Section 4.2. The properties of a lane drop bottleneck are also introduced in the same section. The LCBA-CTM itself is presented in Section 4.3. The section also shows the application of the model to a lane-drop bottleneck and illustrates how the model can predict capacity drop at such a bottleneck. The model is examined under various setups of a constant loading problem in Section 4.4, and is numerically studied through simulations for a variety of starting and boundary conditions in Section 4.5. Finally Section 5.4 concludes with some final discussions.

4.2 Lane changing and bounded acceleration components

4.2.1 Lane changing model

(Jin, 2010a,b) presented a framework for modeling the effect of lane changing vehicles on traffic stream. Lane changing is treated as an aggregate multi-lane-group process where all lanes are considered to be balanced in terms of traffic conditions and traffic behavior outside of the lane-changing region. The underlying idea is that a vehicle changing lanes, contributes to conditions on both its current and target lanes. This can be visualized as the current vehicle, and its 'phantom' double occupying space on both the lanes of interest, such that the lane-changing vehicle influences the behavior of its followers both in the original and the target lanes during the lane-changing process. Thus, the vehicle contributes 'twice' to density during the length of the entire lane-changing process. Further, a new variable, the lane changing intensity is introduced. The idea was also extended through a multi-commodity LWR model for lane-changing incorporating the lane-changing intensity as an endogenous parameter (Jin, 2013).

For the lane changing component, some of the ideas that were presented in (Jin, 2010a,b) are borrowed. The disruptive effect of lane-changing on the flow of traffic is captured by scaling the 'actual density' of traffic in the region by a perceived density factor 'α' ($\alpha \geq 1$) to get the 'perceived density' within the region. This 'percieved density' influences the speeds of the vehicles in the segment. A 'lane-changing intensity' parameter ϵ (Jin, 2010a,b) can also be defined as the ratio of aggregate time spent by vehicles changing lanes, to the time taken for longitudinal traversal of the section. α and ϵ are then related as: $\alpha = \epsilon + 1$.

The equilibrium flow-rate, for the triangular fundamental diagram here, can be computed

from the equilibrium traffic speed as follows:

$$\text{Equilibrium speed} = V^e_{LC}(k) = V^e(\alpha k) \tag{4.7}$$

$$\begin{aligned}\text{LC Equilibrium flow} &= k * V^e(\alpha k)\\ &= k * min\left\{v_f, w\left(\frac{k_j}{\alpha k} - 1\right)\right\}\\ &= min\left\{v_f k, w\left(\frac{k_j}{\alpha} - k\right)\right\}\end{aligned} \tag{4.8}$$

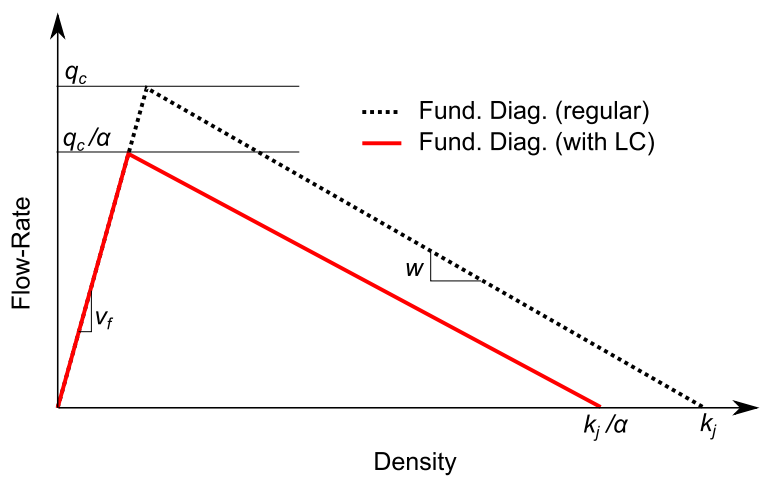


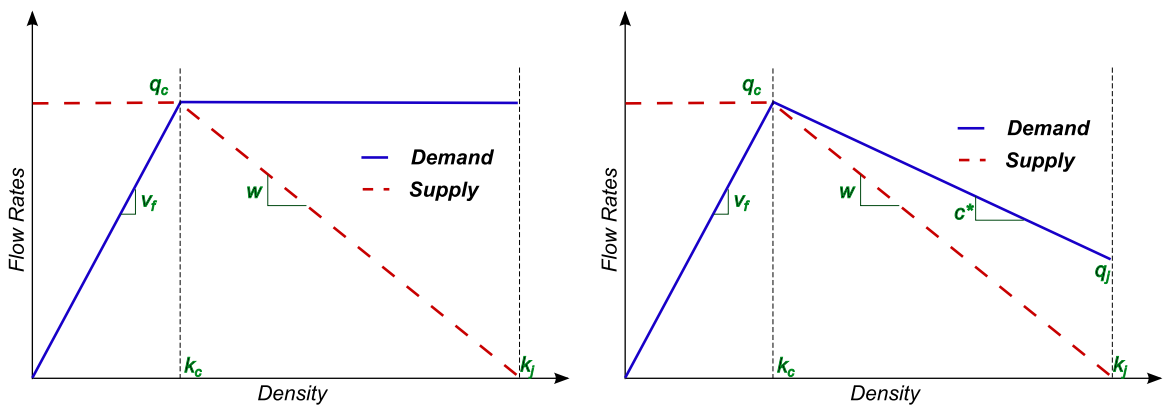
Figure 4.2: The effect of lane changing on the fundamental diagram. The figure shows the flow-density relationship at a location before and after the effects of the lane-changing are applied. For any $\alpha > 0$, the capacity and the jam density both reduce by a factor α.

The resultant effect of lane-changing on the flow-rate (Figure 4.2) is a lateral shift in the congested portion of the curve downwards such that the new density intercept is k_j/α instead of k_j. Consequently, the capacity flow-rate, and correspondingly, the critical density are also reduced by a factor α.

For the present study, the 'perceived density' factor α is assumed to be a constant value for a given region and does not change with time or traffic states. α can also be extended by defining it for all regions, but setting $\alpha = 1$ anywhere outside of the lane-changing zone, in order to simplify the application of the model.

4.2.2 Bounded acceleration and queue discharge model

Based on (Lebacque, 2003), in Chapter 3, a modified demand function was proposed to be used with CTM [also see: Lebacque (1984); Monamy et al. (2012)]. This function updates the demand under congested conditions and makes it linearly decreasing in density. The modified model was shown to capture the discharge behavior of vehicles starting from a standing queue. By reducing the congested demand, the model intrinsically attempts to capture the bounded acceleration behavior of vehicles. As vehicles have to accelerate from a queue, the corresponding demand is lower than capacity due to the additional time and space needed for acceleration to desired speeds.



(a) Trad. Demand, Supply (b) Modified Demand

Figure 4.3: A comparison of the demand and supply functions corresponding to the triangular fundamental diagram for the traditional CTM and the modified CTM for acceleration effects.

Figure 4.3b shows the modified demand function corresponding to the traditional supply and demand shown in Figure 4.3a. The modified demand function can be represented by introducing one additional parameter, either the slope of this congested demand c^*, the demand corresponding to the jam density q_j, or the projected density intercept of the demand line k_j^*. As can be seen here, the congested demand is lower than capacity, and is a function of the density k. More specifically, even when the downstream supply is unrestrictive (downstream is under-critical, and thus supply is equal to capacity flow-rate), the discharge flux at the head of the queue is determined by the new demand function and is lower than

capacity. This means, that attributing to the boundedness of acceleration, a location can not immediately operate at capacity when its upstream is congested, even when downstream is uncongested.

$$\begin{aligned}\text{Modified Demand} = D'(k) &= min\left\{v_f k, q_j + c^*(k_j - k)\right\} \\ &= min\left\{v_f k, q_j + \frac{(q_c - q_j)(k_j - k)}{(k_j - k_c)}\right\}. \\ &= min\left\{v_f k, c^*\left(k_j^* - k\right)\right\}. \end{aligned} \tag{4.9}$$

$$\begin{aligned}\text{Effective speed corresp. to demand} = V'(k) &= Q'(k)/k \\ &= min\left\{v_f, c^*\left(\frac{k_j^*}{k} - 1\right)\right\}. \end{aligned} \tag{4.10}$$

4.2.3 Capacity at lane drop bottlenecks

In (Chung et al., 2007) the capacity drop phenomenon was systematically observed at three bottleneck locations: a merge location; a lane drop location; and a horizontal curve location. Of the three bottleneck types, the lane drop bottleneck and the merge bottleneck are especially of great interest to researchers. These are typically the most commonly occurring active bottlenecks on freeway networks, and have similar physical traits. Both these bottlenecks involve lane changing maneuvers, predominantly left lane changes, that are disruptive to the flow, followed by a region of acceleration downstream to the bottleneck. The study focuses on the lane drop bottleneck, which perhaps has the simpler geometry. The ideas discussed in this chapter can however be extended to the merge bottleneck as well, which shares the lane-changing and bounded acceleration regions with the lane-drop bottleneck, through the addition of a merge model. The horizontal curve bottleneck, and other similar bottlenecks such as a Variable Speed Limit (VSL) induced bottleneck, or a tunnel bottleneck, however have an inherent dissimilarity in the fact that they do not necessitate

any lane-changing maneuvers. Instead, these bottlenecks are associated with an externally imposed speed restriction within a confined length, either through geometrical restrictions or through imposed control strategies. The discussion of such bottlenecks is left to future studies.

The lane drop bottleneck is characterized by the upstream segment having 'm' lanes and the downstream segment having 'n' lanes such that $m > n$ resulting in $(m - n)$ dropped lanes (Figure 4.4a). Vehicles on these $(m - n)$ dropped lanes have to merge into the 'n' continuing lanes somewhere upstream of the lane drop. The resulting left lane changes and the dis-balanced lane densities result in disruption of traffic leading to the activation of the bottleneck.

While the bottleneck is active, there exists a lane changing region immediately upstream of the lane drop location where vehicles switch away from the dropped lanes. Further, there is also the presence of an acceleration zone downstream of the bottleneck location where vehicles starting at congested speeds accelerate up to the free flow speed. Thus, in order to truly capture the dynamics at such bottlenecks, both the lane-changing and the acceleration process need to be accurately modeled.

The simplest treatment of a lane-drop location within the scope of macroscopic traffic modeling, would ignore the effects of lane-changing and of bounded acceleration, modeling the lane-drop as a simple discontinuity in the number of lanes at the lane-drop location. In such a model, the discharge flow-rate across the bottleneck would be governed by the downstream capacity whenever the bottleneck is active. Such a model would not be able to capture the traffic dynamics at the bottleneck correctly and would fail to represent capacity drop. Hybrid and higher order models such as (Laval and Daganzo, 2006) can model both the lane-changing and acceleration components at such bottlenecks, but need extensive calibration efforts and are not scalable like first order macroscopic models.

4.3 A lane changing bounded acceleration model (LCBA-CTM) to capture capacity drop at lane-drop bottlenecks

The present model is built upon the modified CTM proposed in Chapter 3, and adds to it the effect of lane changing behavior inspired by (Jin, 2010a) (refer Sections 4.2.1 and 4.2.2). The demand at the bottleneck location is estimated using a 'perceived density' within the lane-changing region instead of the actual density. In the following subsections, the details of the model is first explained, followed by a demonstration of the application of the model to show capacity drop at a lane-drop bottleneck.

4.3.1 The complete LCBA-CTM model

A lane drop bottleneck, with 'm' upstream lanes, and 'n' downstream lanes (m-n dropped lanes where $m > n$) is used to demonstrate the application of the LCBA-CTM model (Figure 4.4). The roadway is divided into three regions: R1 is upstream of the lane drop and devoid of any lane changing induced by the lane drop; R2 is immediately upstream of the bottleneck where vehicles change lanes in anticipation of the lane drop; and R3 is the downstream region where vehicles accelerate out of the lane drop bottleneck.

The model can be summarized through the following:

- *Approximate lane-drop to be sudden with no lateral dimension.*

  The first assumption is that the lane-drop is sudden and is represented without any dimensions, i.e., there is a discontinuity in the number of lanes between regions R2 and R3 instead of a gradual change over a length of roadway. This is an assumption used for simplifying the geometry of the location, and it's reasonable when the lane-drop

region is very short.

- *Assume a constant lane-changing intensity factor in region R2; i.e., $\alpha > 1$ is constant inside R2. Set this factor α to 1 everywhere else.*

- *Approximate R2 to be of length exactly equal to one CTM cell.*

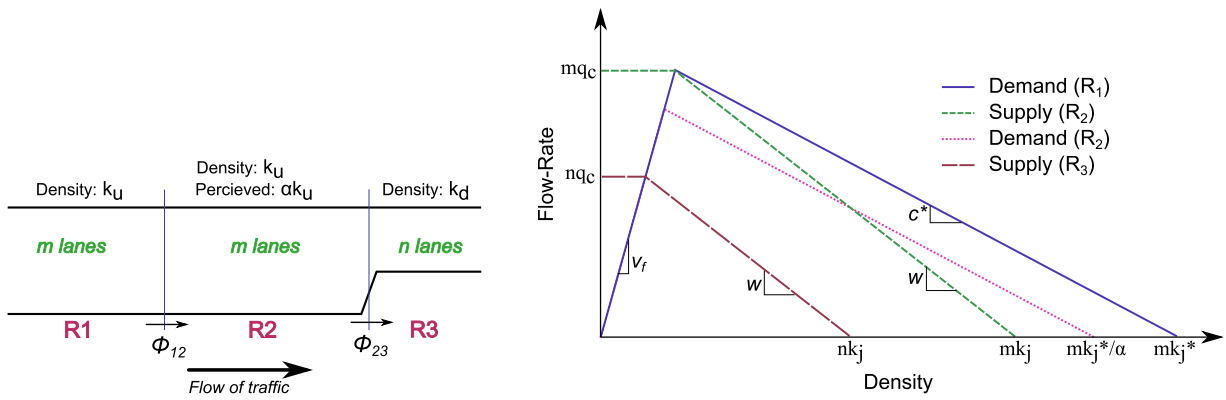
  The effect of the lane-changing vehicles in the lane-changing region, is approximated to exist only in the cell immediately upstream of the lane-drop location. Thus, all other upstream cells are not impacted by the detrimental effect of lane changing on the flow of traffic, and region R2 is exactly one CTM cell length long. Under congestion, this becomes a reasonable assumption as vehicles tend to merge late in the lane-drop region.

- *Compute the demand in cell R2 based on 'perceived density', and supply in R2 cell based on actual density.*

  Finally, it is proposed that the lane-changing process has a direct impact only on the demand within this cell, and not on the supply available in the cell. This reflects the idea that only speeds of the vehicles within this cell are impacted by the lane-changing process as they move to the downstream section, while vehicles entering from the upstream cell are not impacted immediately. This approximation once again reflects the behavior of merging vehicles, the primary contributors for the lane-changing, merging late during congestion and much closer to the lane-drop location. The impact of the lane changing reduces as one moves upstream, away from the lane-drop.

$$\begin{aligned}
\text{LC Demand} &= k * V^d_{LC} = k * V^d(\alpha k) \\
&= k * min\left\{v_f, c^*\left(\frac{k^*_j}{\alpha k} - 1\right)\right\} \\
&= min\left\{v_f k, c^*\left(\frac{k^*_j}{\alpha} - k\right)\right\}
\end{aligned} \tag{4.11}$$

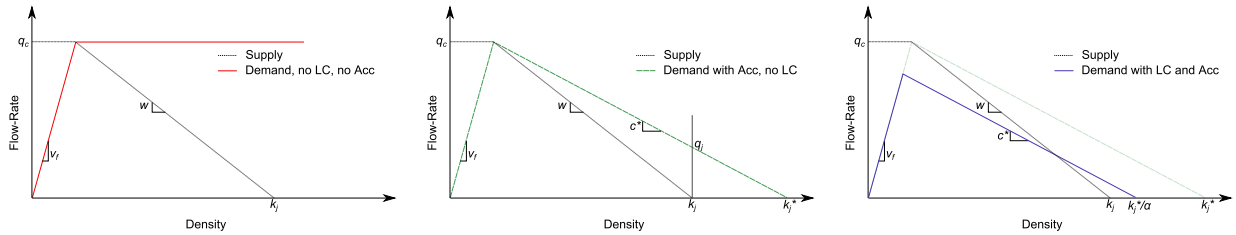
Following the setup, the demand at the bottleneck is determined by the 'perceived density' in the cell upstream of the bottleneck. The demand function used to compute this demand is



(a) Lane Drop Schematic (b) Demand and Supply post LC

Figure 4.4: Illustration of the lane drop bottleneck and the corresponding demand and supply curves. There are m lanes upstream and n lanes downstream. The region is divided into 3 components, R1, R2 and R3. Figure 4.4b shows the demand and supply curves after the lane-changing adjustment has been made.

a linearly decreasing function in density in order to incorporate the effects of the acceleration process. Figure 4.5c shows the lane based supply and demand functions that would be used for each of the cells, with $\alpha > 1$ for the cell immediately upstream of the bottleneck where the lane-changing exists, and $\alpha = 1$ for all other cells (which is equivalent to supply and demand shown in Figure 4.5b without the effect of lane-changing). The relevant lane-aggregate demand and supply functions for regions R1 through R3 of the lane drop bottleneck are shown together in Figure 4.4b.



(a) Traditional Demand (b) With Acc. (c) With Acc. and LC

Figure 4.5: The effects of including the acceleration and lane-changing modifications in to the demand function for a lane-drop setup.

The flux across the boundary between R1-R2 is denoted as Φ_{12}, and the flux across the boundary between R2-R3 is denoted as Φ_{23}. Φ_{12} is determined by the demand in R1 and the supply available in R2, and similarly, Φ_{23} is determined by the demand in R2 and the supply

available in R3. The demand in R2 is calculated using the 'perceived density' while other values are all calculated using actual density. If the densities in regions R1, R2, and R3 are denoted as k_1, k_2 and k_3 respectively (note that typically we would have $k_1 = k_2 = k_U$ and $k_3 = k_D$), the demand and supply in the regions are calculated as:

$$D_1(k) = min\left\{v_f k, c^*\left(mk_j^* - k\right)\right\} \tag{4.12}$$

$$S_2(k) = min\left\{mq_c, w\left(mk_j - k\right)\right\} \tag{4.13}$$

$$D_2(k) = min\left\{v_f k, c^*\left(\frac{mk_j^*}{\alpha} - k\right)\right\} \tag{4.14}$$

$$S_3(k) = min\left\{nq_c, w\left(nk_j - k\right)\right\} \tag{4.15}$$

And the flux across the boundaries, Φ_{12} and Φ_{23} are:

$$\Phi_{12} = min\left\{D_1(k_1), S_2(k_2)\right\} \tag{4.16}$$

$$\Phi_{23} = min\left\{D_2(k_2), S_3(k_3)\right\} \tag{4.17}$$

The combined model can thus be represented through two modifications to the demand function: shifting the congested demand function by a factor α in order to incorporate the effects of lane-changing; and changing the slope of the congested portion of the demand function from 0 to $-c^*$ in order to incorporate the effects of vehicle acceleration.

In the following subsection, the presented model is shown to capture sustained capacity drop at the location, by proving that there exists a stationary state condition for the location where the bottleneck throughput can be lower than the uncongested downstream capacity whenever the bottleneck is active.

4.3.2 Stationary state with capacity drop

It is next shown how the presented model can capture sustained capacity drop at the location, by proving that there exists a stationary state condition for the location where the bottleneck throughput can be lower than the uncongested downstream capacity whenever the bottleneck is active.

Figure 4.6 shows a possible stationary state under the new model such that the upstream is congested, the downstream is uncongested, and the discharge flow-rate is lower than uncongested bottleneck capacity. The downstream operates at a density k_D which is under-critical. The region (R1) upstream of the lane changing region operates at density k_U while Region R2 (lane changing region immediately upstream of the lane drop) has an actual density also equal to k_U like region R1, but has a higher perceived density.

For the bottleneck to be active, k_U must be over-critical, and k_D must be under-critical. When k_U is over-critical, since the densities in R1 and R2 are equal, the flux Φ_{12} would be determined by the supply in R2 (the demand in R1 would always be higher than the supply in R2 as can be seen Figure 4.4b). The flux at the bottleneck location, Φ_{23} however, can depend either on the demand in R2 or the supply in R3 (which is typically equal to the downstream capacity).

In order for a capacity drop to be stable over time, there needs to exist some stationary state conditions between the two boundaries R1-R2 and R2-R3, such that $\Phi_{12} = \Phi_{23} < nq_c$, where nq_c is the uncongested downstream capacity. Such a stationary state condition is shown in Figure 4.6, where the downstream operates at condition 'A' and the upstream operates at condition 'B'. As noted above, Φ_{12} is determined by the supply curve for R2, and Φ_{23} is determined by demand curve R2 (as long as it is lower than the downstream supply). The point of intersection of the curves for supply and demand of R2 gives a condition where the flux across the boundaries are equal, thus resulting in stationary state conditions. Whenever

this intersection point ('B' in the figure) is below the uncongested downstream capacity, there would exist a corresponding capacity drop.

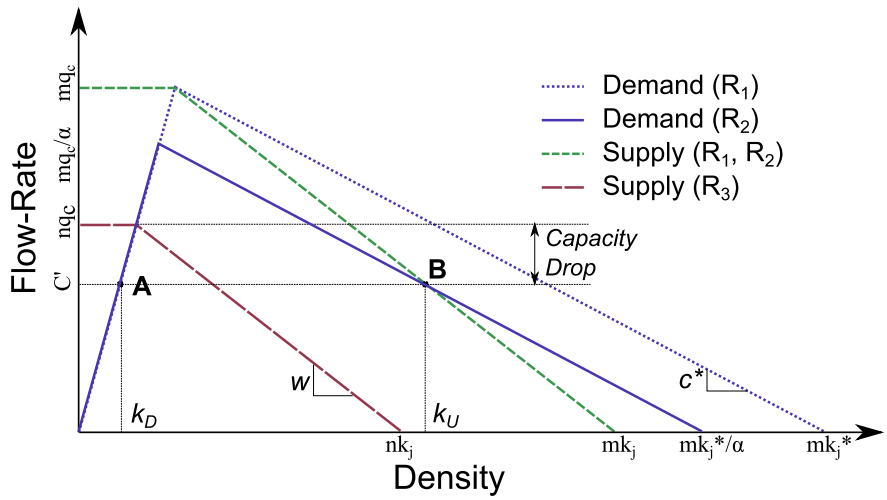


Figure 4.6: Stationary state condition with capacity drop in lane drop setup. The states in regions R1 and R2 are given by B, and the state in the downstream R3 is given by A. The boundary flux between R1 and R2, and the flux between R2 and R3 are both given by the same flow rate. Flux Φ_{12} is determined by the demand from R1 and the supply in R2, while the flux Φ_{23} is determined by the downstream supply in R3 and the demand in R2. The difference between this flow-rate from capacity gives the capacity drop.

The stationary state density in the upstream region k_U^S can be obtained geometrically from the demand supply curves. Since flux Φ_{12} depends on the upstream supply in R2 and should be equal to flux Φ_{23} which depends on the demand in R2, we have:

$$\Phi_{12} = \Phi_{23}$$

$$S_2(k_U^S) = D_2(k_U^S) \tag{4.18}$$

For the stationary state of interest, corresponding to a capacity drop, the upstream should be congested. Thus, we need to only use the congested portion of supply function S_2 and

demand function D_2.

$$w\left(mk_j - k_U^S\right) = c^*\left(\frac{mk_j^*}{\alpha} - k_U^S\right),$$
$$\alpha w k_U^S - \alpha c^* k_U^S = \alpha w m k_j - m c^* k_j^*. \tag{4.19}$$

The stationary state upstream density can be computed from the solution of the above equation.

$$k_U^S = \frac{m\left(\alpha w k_j - c^* k_j^*\right)}{\alpha\left(w - c^*\right)}. \tag{4.20}$$

The operating bottleneck capacity, or the congested capacity, would then be determined from Φ_{12} or Φ_{23} as:

$$\begin{aligned}\text{Reduced capacity} &= w\left(mk_j - k_U^S\right)\\ &= w\left[mk_j - \frac{m\left(\alpha w k_j - c^* k_j^*\right)}{\alpha\left(w - c^*\right)}\right]\\ &= \frac{mwc^*\left(k_j^* - \alpha k_j\right)}{\alpha\left(w - c^*\right)}.\end{aligned} \tag{4.21}$$

In contrast, the original uncongested capacity can be written as [1]:

$$\begin{aligned}\text{Uncongested capacity} &= nq_c\\ &= \frac{nwc^*\left(k_j^* - k_j\right)}{w - c^*}.\end{aligned} \tag{4.22}$$

The difference between the uncongested and the congested capacities gives the capacity drop expected at the location.

[1] Other definitions of the downstream capacity, such as reduced or externally defined capacities, can also be easily substituted.

$$\text{Capacity drop magnitude} = \frac{nwc^*\left(k_j^* - k_j\right)}{w - c^*} - \frac{mwc^*\left(k_j^* - \alpha k_j\right)}{\alpha\left(w - c^*\right)},$$
$$= \frac{wc^*}{\alpha\left(w - c^*\right)}\left[\alpha k_j\left(m - n\right) - k_j^*\left(m - \alpha n\right)\right]$$
$$\text{Capacity drop ratio} = 1 - \frac{m}{\alpha n}\left[\frac{k_j^* - \alpha k_j}{k_j^* - k_j}\right]. \tag{4.23}$$

This capacity drop ratio is dependent on the various characteristics of the demand and supply curves (m, n, w, k_j, c^*, and k_j^*), but also depends on the lane changing intensity ($\epsilon = \alpha - 1$).

Note that α is physically bounded in the range $\left[1, k_j^*/k_j\right]$ as can be seen from the demand-supply curves. Any value higher than k_j^*/k_j would mean that there would be no value of k for which the supply and demand in R2 are equivalent. It is also easy to show that α should be greater than $\frac{mk_j^*}{nk_j^* - nk_j + mk_j}$ for capacity drop to exist. For any lower value of α, the location would continue to operate at the regular downstream capacity.

$$\text{If } \alpha > \frac{mk_j^*}{nk_j^* - nk_j + mk_j} \rightarrow \text{Capacity Drop,}$$
$$\text{else} \rightarrow \text{Operate at full downstream capacity.} \tag{4.24}$$

4.3.3 Need for both lane-changing and acceleration components

The model suggests that lane-changing and acceleration both need to be correctly modeled in order to capture capacity drop accurately. Further, lane-changing should be modeled in a way that its effect on the disruption of flow is accounted for. This would mean that conventional lane-based lateral flow models of lane-changing, such as (Munjal and Pipes, 1971a,b; Michalopoulos et al., 1984), would not be adequate for reproducing sustained capacity drop.

In the absence of the acceleration model, the demand generated due to over-saturated traffic

upstream of the bottleneck would be equal to capacity of the upstream segment after being adjusted for the lane changing factor, independent of the actual density. This adjusted capacity would be equal to mq_c/α. For a large enough α, it is possible that this upstream demand becomes lower than the downstream uncongested capacity. In such a situation, it is possible to observe a sustained capacity drop. However, the computed magnitude of the drop would not be accurate due to ignoring the effect of vehicle acceleration. For values of α not large enough, the downstream supply would determine the bottleneck discharge, and there would be no capacity drop in such a scenario.

$$D_2(k_u) = min\left\{v_f k_u, \frac{m v_f k_c}{\alpha}\right\} \tag{4.25}$$

With the existence of capacity drop being determined by:

$$\begin{aligned} \text{if } \alpha > m/n, &\implies D_2(k_u) < S_3(k_d) \implies \text{Capacity drop} \\ \text{else}, &\implies D_2(k_u) \geq S_3(k_d) \implies \text{No capacity drop} \end{aligned} \tag{4.26}$$

In the absence of lane-changing, the discharge flow would still be decided by the demand immediately upstream of the bottleneck as long as it is lower than the downstream capacity. Thus, it is indeed possible to have a discharge flow-rate lower than the capacity of the downstream segment. However, due to the absence of lane-changing, the flow-rate into the upstream segment would be determined by the supply available in the segment, and would be lower than the flow across the lane-drop boundary. This state can not be sustained over time, and the density immediately upstream of the bottleneck would gradually decrease till the demand becomes greater than downstream capacity and the bottleneck starts operating at full downstream capacity. In such a case, the capacity drop can, therefore not be sustained.

Using a value of $\alpha = 1$ in (Eq. 4.21),

$$\begin{aligned}\text{New Capacity} &= \frac{mwc^* \left(k_j^* - \alpha k_j\right)}{\alpha \left(w - c^*\right)} \\ &= \frac{mwc^* \left(k_j^* - k_j\right)}{\left(w - c^*\right)} = mq_c \end{aligned} \tag{4.27}$$

Since this value is higher than the uncongested downstream capacity, this state is not feasible. The only possible stationary state corresponds to the bottleneck operating at full downstream uncongested capacity, resulting in no capacity drop.

4.3.4 Downstream lane-changing region

For the majority of the discussions above, it is assumed that lane-changing maneuvers occur exclusively upstream of the lane-drop location. This is a reasonable assumption in most situations as long as the vehicles redistribute approximately equally on all continuing lanes upstream of the lane-drop. It has however been observed in past studies that a fraction of the lane-changes might persist beyond the lane-drop location. (Jin, 2010a) considers a case where downstream lane-changes affects the capacity of a location. The downstream lane-changing is assumed to be disruptive to the flow as it impacts the velocity of the traffic stream and can eventually lead to a reduction in the capacity.

Downstream lane-changing can be shown to cause a capacity reduction such that the bottleneck location can not operate at full capacity as long as lane-changing persists. Upstream lane-changing on the other hand, as shown in the present model, results in a capacity drop where the original uncongested capacity can still be attained under uncongested conditions, but a fraction is lost under congestion. The two models can easily be combined where the downstream lane-changing is effective primarily in reducing the operable capacity under conditions when the downstream might be near congestion, while the upstream lane-changing

causes the capacity drop in conditions were the bottleneck is active.

When including the effects of downstream lane-changing, the supply function for the downstream region R3 would be updated to account for lane-changing. If the downstream lane-changing impact is α_d, the new supply function and the corresponding uncongested capacity would be:

$$S_{R3}(k) = min\left\{\frac{nq_c}{\alpha_d}, w\left(\frac{nk_j}{\alpha_d} - k\right)\right\} \tag{4.28}$$

$$\text{Uncongested Capacity} = \frac{nq_c}{\alpha_d} = \frac{nwc^*\left(k_j^* - k_j\right)}{\alpha_d\left(w - c^*\right)}. \tag{4.29}$$

The congested capacity (if capacity drop does exist) remains unaffected, and the capacity drop magnitude can be computed using the new uncongested capacity instead.

Note that the effect on capacity due to traffic dynamics such as onset of congestion is referred to as capacity drop. This capacity drop is determined strictly by the impacts of lane-changing and acceleration on the demand function upstream. In contrast, the lane-changing downstream impacts the downstream supply. Such a behavior is referred to as capacity reduction, which is independent of the traffic conditions. Other external players, such as geometric curves may also impact the supply downstream and thus cause a capacity reduction.

4.3.5 A numerical example

The following shows a simple numerical example for a lane drop location and the corresponding capacity drop that would be observed at such a location. A 3 lanes to 2 lanes lane drop bottleneck (Figure 4.7a) is considered for the example. The location is assumed to be an active bottleneck not affected by any conditions existing downstream. Representative values describing the lane fundamental diagram for the location are set to be given by: $q_c =$

2000vph, $k_c = 35$vpm, $k_j = 200$vpm. k_j^* is set to 250vpm for the modified demand function. The demand and supply function parameters are based on calibration observations reported in Chapter 3. The resulting lane fundamental diagram is shown in Figure 4.7b.

Further, $\epsilon = 0.09$, $\alpha = 1.09$ are set based on (Jin, 2010a). The uncongested capacity, congested capacity, and capacity drop at the bottleneck can be computed (Eqns. 4.21-4.23) as:

$$\text{Uncongested capacity} = nq_c$$

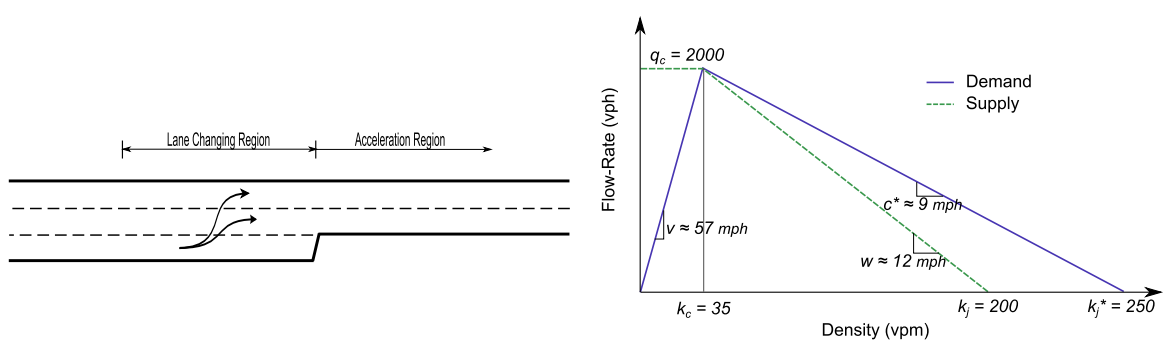
$$= 2 * 2000 = 4000vph. \tag{4.30}$$

$$\text{Reduced capacity} = \frac{mwc^* \left(k_j^* - \alpha k_j\right)}{\alpha \left(w - c^*\right)}$$

$$= 3523vph. \tag{4.31}$$

$$\text{Capacity drop} = 4000 - 3523 = 477vph \approx 12\% \tag{4.32}$$

The model predicts a 12% capacity drop which is consistent with observations in past studies as the ones covered in Section 5.1.



(a) A lane drop schematic for 3 to 2 lanes. (b) Sample lane demand, supply functions.

Figure 4.7: Illustration of the lane drop bottleneck and corresponding lane based base (before adjusting for lane-changing impact) demand and supply curves for the numerical example in Section 4.3.5.

4.4 Constant Loading Problem

In order to understand the behavior of the model under various conditions, the solutions to various scenarios of the constant loading problem applied to the model are examined. The constant loading problem is defined on a lane-drop set up, where the links are loaded with an initial density, and a constant demand is imposed on the upstream boundary.

The merge cell (region R2) is analyzed here, under constant demand conditions, and with the downstream segment supply defined. For the analysis, the uncongested and congested downstream capacities are denoted as C_U and C_C respectively, the stationary state density (previously k_U^S) as simply k_S as we are only interested in the upstream merge cell, and the upstream demand and downstream supply boundary conditions as D_0 and S_0 respectively.

The following are the scenarios that need to be explored:

1. Unrestricted Downstream supply, $S_0 = nq_c$.
    a. High Demand, $D_0 > C_U$.
    b. Medium Demand, $C_U > D_0 > C_C$ (Stationary state also depends on initial density).
    c. Low Demand, $C_C > D_0$.
2. Downstream supply between congested and uncongested capacity, $C_C < S_0 < C_U$ (Stationary state also depends on initial density).
3. Downstream supply below congested capacity, $S_0 < C_C$.

For the merge cell, the inflow and the outflow are given by Φ_{12} and Φ_{23} as defined in Figure 4.4a. Imposing the constant demand D_0 and constant supply S_0 on the cell boundaries, if k

is the density in the merge cell, we have:

$$\begin{aligned}\Phi_{12} &= min\left\{D_0, S_2(k)\right\} \\ &= min\left\{D_0, mq_c, w\left(mk_j - k\right)\right\} \qquad (4.33)\\ \Phi_{23} &= min\left\{D_2(k), S_0\right\} \\ &= min\left\{v_f k, c^*\left(\frac{mk_j^*}{\alpha} - k\right), S_0\right\} \qquad (4.34)\end{aligned}$$

And the density evolution can be approximated (as an ODE) through:

$$L\frac{dk}{dt} = \Phi_{12} - \Phi_{23}. \qquad (4.35)$$

4.4.1 Case 1: $S_0 \geq C_U$, $D_0 > C_U$: Onset of congestion, Bottleneck activation

We start with what is perhaps the most interesting scenario, with unrestrictive downstream supply and upstream demand higher than downstream uncongested capacity. This is the setup that typically leads to the activation of a bottleneck. The scenario is shown in Figure 4.8. In this scenario, the only stationary state possible is when density is equal to k_S. For any value of density higher than k_S, the outflow would be greater than the inflow, resulting in a net reduction of density, while for any value of density lower than k_S the outflow is lower than inflow resulting in an increase in density. Thus, in both cases, traffic state moves towards the stationary state conditions with $k = k_S$.

Proof of stability of stationary state

The stability of the stationary state corresponding to the dropped capacity can be analyzed

by perturbing the density in the cell by a small amount in either direction.

In the neighborhood of the stationary state condition, we have the following relation for density in the merge cell:

$$L\frac{dk}{dt} = w(mk_j - k) - \frac{c^*}{\alpha}\left(mk_j^* - \alpha k\right). \tag{4.36}$$

Solving the above, the density evolution with time can be computed as:

$$\begin{aligned} k &= C_0 e^{(c^*-w)t/L} + \frac{m\left(\alpha w k_j - c^* k_j^*\right)}{\alpha\left(w - c^*\right)} \\ &= C_0 e^{(c^*-w)t/L} + k_S \end{aligned} \tag{4.37}$$

where L is the CTM cell length, and C_0 is the constant of integration which depends on the initial conditions.

Let the new perturbed density be $k' = k_S + \delta$ where $|\delta| << k_S$. Therefore, in the above equation, $k = k'$ at $t = 0$.

$$k = \delta e^{(c^*-w)t/L} + k_S \tag{4.38}$$

Since $w > c^*$, the exponential term approaches 0 as t goes to ∞, and the density monotonically converges to k_S, the system is asymptotically stable. Thus, given enough time, the state would always stabilize at k_S and this state is a stable stationary state.

Traffic dynamics during bottleneck activation and onset of congestion

During onset of congestion, the bottleneck discharge flow can be computed from knowing the density evolution. The discharge flow rate (given by the outflow from the cell considered here) increases to the uncongested capacity as the upstream density increases up to nk_c.

The variables k_{C1}, k_{C2} and k_D, where applicable (Figures 4.8 - 4.10), refer to the densities where the inflow and the outflow curves match the capacity flow rate or the demand values. Similarly, variables k_{S2} and k_{S3}, where applicable (Figures 4.9 - 4.12), represent possible stationary state densities in addition to k_S.

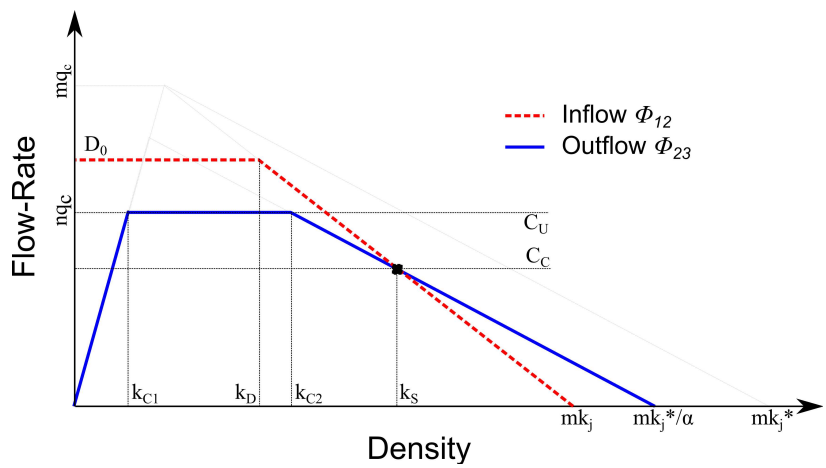


Figure 4.8: Case 1: Constant loading with demand greater than uncongested downstream capacity. This scenario typically corresponds to the onset of congestion. There is a single stationary state possible and it is stable.

$$k_S = \frac{m}{\alpha}\left(\frac{\alpha w k_j - c^* k_j^*}{w - c^*}\right) \qquad k_{C1} = \frac{C_U}{v_f}$$
$$k_D = mk_j - \frac{D_0}{w} \qquad k_{C2} = \frac{mk_j^*}{\alpha} - \frac{C_U}{c^*}$$

If $k_D \leq k_{C2}$,

$$L\frac{dk}{dt} = \begin{cases} D_0 - v_f k, & \text{for } k \in [0, k_{C1}] \\ D_0 - nq_c, & \text{for } k \in (k_{C1}, k_D] \\ w(mk_j - k) - nq_c, & \text{for } k \in (k_D, k_{C2}] \\ w(mk_j - k) - \frac{c^*}{\alpha}(mk_j^* - \alpha k), & \text{for } k \in (k_{C2}, k_S] \end{cases} \tag{4.39}$$

If $k_D > k_{C2}$,

$$L\frac{dk}{dt} = \begin{cases} D_0 - v_f k, & \text{for } k \in [0, k_{C1}] \\ D_0 - nq_c, & \text{for } k \in (k_{C1}, k_{C2}] \\ D_0 - \frac{c^*}{\alpha}(mk_j^* - \alpha k), & \text{for } k \in (k_{C2}, k_D] \\ w(mk_j - k) - \frac{c^*}{\alpha}(mk_j^* - \alpha k), & \text{for } k \in (k_D, k_S] \end{cases} \tag{4.40}$$

The solution of the above equations, depending on the initial conditions, would give the evolution of the density within the merge cell, as well as the bottleneck discharge flow rates (from the outflow corresponding to the density).

4.4.2 Case 2: $S_0 \geq C_U$, $C_U \geq D_0 \geq C_C$: Dependence on initial state

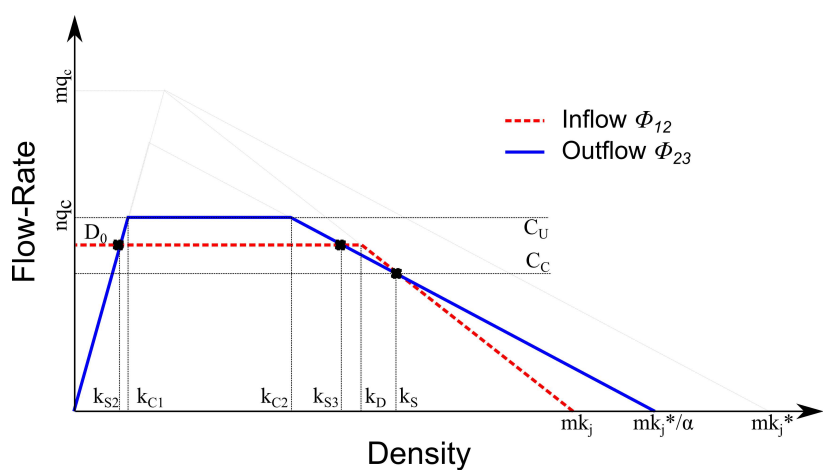


Figure 4.9: Case 2: Constant loading with demand smaller than the uncongested downstream capacity, but greater than the congested capacity. There are two stable stationary states, one with upstream uncongested, and one with upstream congested. There is also an unstable stationary state possible in this scenario.

In the second scenario, we consider a situation where the upstream demand into the merge cell (R2) is higher than the congested downstream capacity, but lower than the uncongested downstream capacity. This scenario results in a bi-stable system, such that the final sta-

tionary state depends on the initial density in the region. There are a total of 3 possible stationary states as shown in Figure 4.9. In addition to state corresponding to density k_S, there would be a second stable stationary state corresponding to density k_{S2}, an uncongested traffic state where the corresponding steady state flow-rate matches the upstream demand flow. Both these states can be shown to be stable as a slight increase in the density corresponds to a net outflow from the cell (outflow > inflow), counteracting the density increase, and a slight decrease in the density corresponds to a net inflow into the cell, once again counteracting the density change. A third state k_{S3}, however, is a possible unstable stationary state where any perturbation to the density would move traffic towards one of the two stable stationary states.

The final stable stationary states that the system would resolve to, depends on the initial density within the merge cell. If the initial density is greater than k_{S3}, then the final resting state would be k_S which corresponds to the dropped capacity and congested upstream conditions. If the initial density is lower than k_{S3}, the final state would be k_{S2} corresponding to uncongested traffic with the bottleneck flow being determined by the upstream demand D_0. For the unique case where density is exactly equal to k_{S3}, we have an unstable stationary state corresponding to congested traffic.

In the case that the upstream demand is exactly equal to the uncongested downstream capacity, stationary states can exist for a range of densities $[k_{C1}, k_{C2}]$, corresponding to a bottleneck flow equivalent to downstream uncongested capacity. Similarly, the special case with the upstream demand exactly equal to the congested downstream capacity would lead to two possible stable stationary states, one corresponding to uncongested traffic and the other to congested traffic conditions.

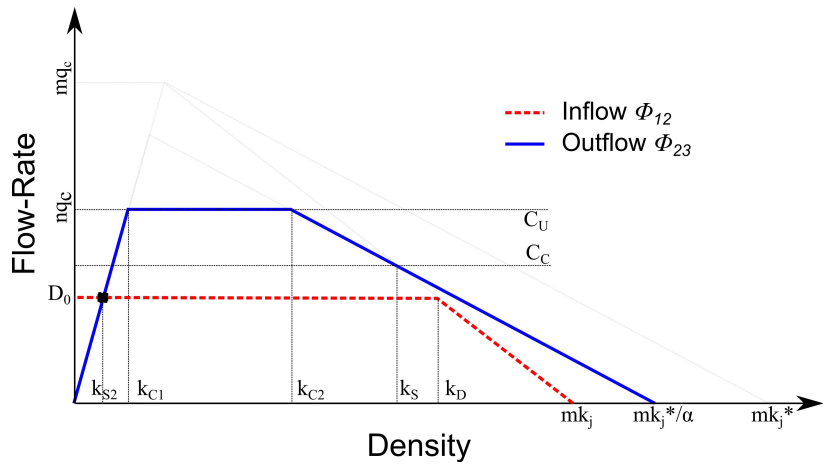


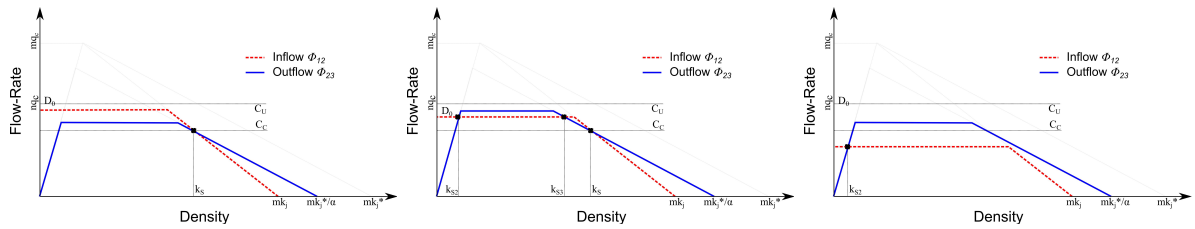
Figure 4.10: Case 3: Constant loading with demand lower than uncongested downstream capacity. This scenario covers the typical recession of congestion period where the bottleneck gets relieved due to low demand. There is only a single stationary state possible under this scenario that corresponds to uncongested states with densities corresponding to an equilibrium flow rate matching the upstream demand.

4.4.3 Case 3: $S_0 \geq C_U$, $C_C > D_0$: Recession of congestion

Next we consider the situation where the demand is lower than the congested capacity of the bottleneck (Figure 4.10). Under these conditions, the bottleneck would eventually get uncongested regardless of initial state converging to the stable stationary state corresponding to density k_{S2}. For all values of the density higher than the uncongested density corresponding to a flow rate equal to the demand, k_{S2}, the outflow is higher than the inflow, thus reducing the density within in the merge region towards k_{S2}. Similarly, for all densities lower than the stationary state density, the inflow is greater than the outflow, resulting in an increase in the density towards k_{S2}.

4.4.4 Case 4: $C_C < S_0 < C_U$

Reducing the supply to a value between the uncongested and the congested downstream capacities makes the present bottleneck inactive, as capacity flow is no longer possible regardless of the demand conditions. This setup results in three new scenarios, depending on



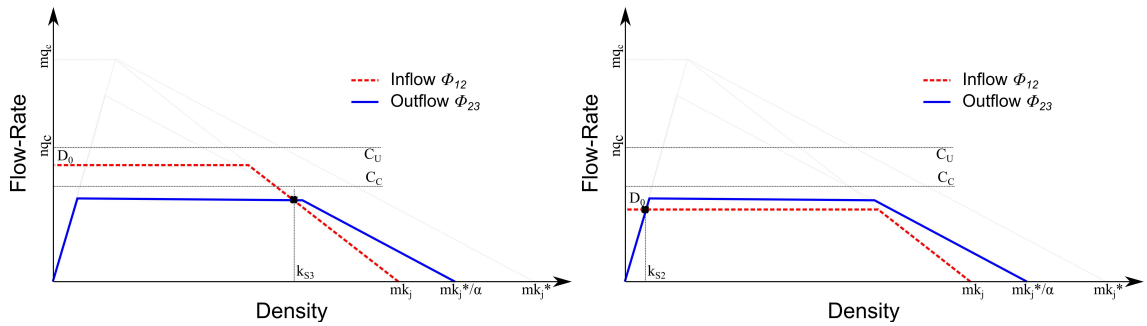
(a) High Demand (b) Med Demand (c) Low Demand

Figure 4.11: Case 4: Constant loading where the downstream supply is lower than the uncongested downstream capacity, but larger than the congested downstream capacity. There can be three sub-scenarios within this case depending on the upstream demand level as shown in the sub-figures above. The stationary states possible are similar to those for the corresponding cases with supply higher than uncongested downstream capacity.

the demand, that are very similar to the case where the downstream supply is unrestricted. The stationary states are determined in a similar manner as earlier. A high demand results in a single stable congested stationary state while a low demand results in a single stable uncongested stationary state. A third situation with demand only slightly lower than supply results in three possible stationary states, a stable uncongested state, a stable congested state and an unstable congested state. While the behavior of the bottleneck is very similar to earlier cases examined, the actual discharge flow rate under this set up would of course always be lower than the uncongested downstream capacity.

4.4.5 Case 5: $S_0 \leq C_C$

If the downstream supply is restricted below the congested capacity, two possible scenarios can be attained based on whether the demand upstream is higher, or lower than the downstream supply (Figure 4.12). If the demand is lower, the only possible stationary state occurs corresponding to uncongested upstream traffic, while higher demand results in a single stationary state with congested traffic conditions. In the special case where the demand is exactly equal to the downstream supply, instead of distinct stationary state densities, we would have a range of stationary state densities corresponding to both uncongested and



(a) High Demand (b) Med Demand

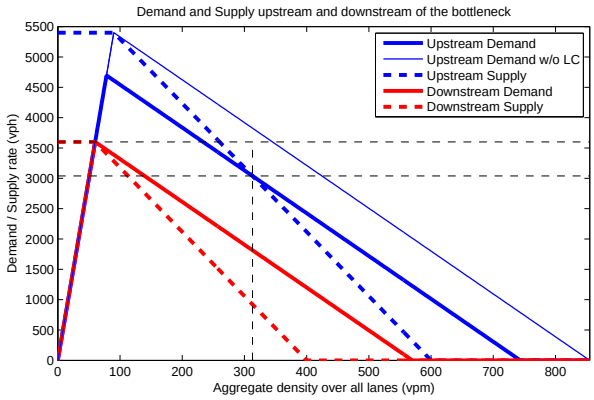
Figure 4.12: Case 5: Constant loading where the downstream supply is lower than the congested downstream capacity. Under high upstream demand conditions, the resultant stationary state is stable and corresponds to the available downstream supply regardless of the value of the demand. If the demand is lower than the supply available, then traffic converges to a stable uncongested stationary state determined by the value of the demand.

congested upstream traffic conditions.

4.5 Numerical solutions for generalized bottleneck conditions

In order to further test the robustness of the proposed model, the traffic behavior is assessed under more generalized conditions through numerical simulation, studying a longer stretch of roadway. Using a simple lane-drop setup with the upstream link consisting of 3 lanes, and the downstream link consisting of 2 lanes, multiple scenarios are enumerated with varying combinations of initial and boundary conditions. The initial conditions considered always reflect a fixed uniform density upstream of the lane-drop location, and a (potentially) different uniform density downstream. Keeping an active bottleneck in mind as the motivating scenario, only set ups where the downstream supply is unrestrictive are considered. The upstream boundary condition is imposed through constant upstream demand levels.

Each studied scenario is simulated through a CTM model setup, the calibration parameters



(a) Demand and Supply curves

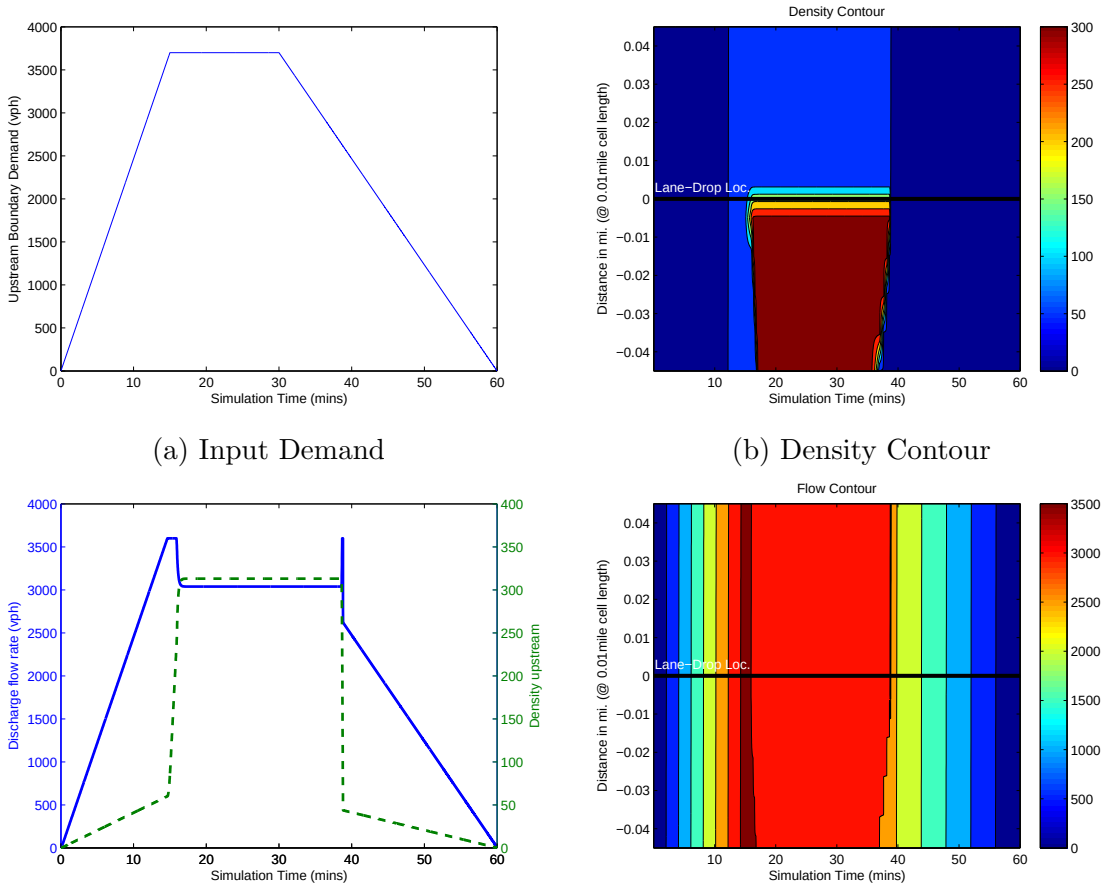
Param	Value	Param	Value
v_{ff}	60 *mph*	q_c	1800 *vph*
k_c	30 *vpm*	w	10.6 *mph*
k_j	200 *vpm*	c*	7.1 *mph*
q_j	600 *vph*	k_j*	285 *vpm*
L	0.01 *mi*	*Uncon. Cap.*	3600 *vph*
m	3	*Cong. Cap.*	3038 *vph*
n	2	*Cap. Drop*	562 (15.6%)
α	1.15	k^S_U	313 *vpm*

(b) Model Parameters

Figure 4.13: The demand and supply functions for the entire lane-drop location based on the calibration parameters, shown as a table on the right, used for the simulation under various scenarios.

used for which are shown in Figure 4.13. The first setup is a simplistic approximation of a typical rise and fall in demand that is often seen around peak periods of traffic. The demand is assumed to steadily rise, linearly in this approximation, to a value that slightly exceeds the downstream capacity, enough to cause a breakdown in flow leading to the activation of the bottleneck. This high demand is maintained for a while, and is eventually followed by the congestion recession period where it reduces again linearly. The demand pattern, along with the resulting bottleneck features and the flow and density contours for the region are shown in Figure 4.14. From the time series plot of the bottleneck throughput in Figure 4.14c, it can be seen that the flow-rate at the bottleneck increases as the demand upstream increases, leading the bottleneck to operate at the uncongested downstream capacity initially. However, since the demand is set to be slightly higher than the downstream capacity, the upstream density keeps increasing, eventually leading to breakdown. This breakdown point corresponds to the density for which the upstream demand is equal to the uncongested downstream capacity in Figure 4.13a. Once the breakdown happens, the bottleneck capacity reduces to the congested capacity and the bottleneck continues to operate at this flow-rate till the demand eventually decreases in the congestion recession period. Once the demand is low enough, the merge cell queue is dissipated, temporarily at a rate equal to the uncongested capacity, thus clearing

the bottleneck completely.



(a) Input Demand

(b) Density Contour

(c) Bottleneck Throughput

(d) Flow Contour

Figure 4.14: Numerical solution showing bottleneck traffic behavior corresponding to onset and recession of congestion. The upstream demand used, shown in 4.14a is a rough approximation for the rise and fall of demand conditions during peak periods. 4.14c shows the corresponding bottleneck throughput, showing the increase of flow rate up to the uncongested capacity, capacity drop to the congested capacity, as well as the recession of congestion. 4.14b and 4.14d show the density and flow contours for the bottleneck region.

As mentioned above, the boundary condition relevant to the setup is the upstream demand to the section. Typically, only two ranges of values of demand would be of interest: demand lower than, and demand equal to or higher than the downstream capacity. In our case however, the downstream section capacity itself is not a fixed value and has two states: uncongested capacity, and congested capacity. Thus, the upstream demand ranges of interest are now: a. Lower than congested bottleneck capacity; b. Higher than congested capacity

but lower than uncongested capacity at bottleneck; and c. Higher than the uncongested bottleneck capacity.

The resulting bottleneck discharge flow-rates, flow contours, and density contours for the bottleneck region are shown in Tables 4.1-4.3. Of particular interest to the study, are the scenarios where the upstream gets congested, thus leading to a drop in capacity from the uncongested capacity, 3600 vph, to the congested capacity value of 3038 vph (eg. scenarios with upstream demand - 3650 vph and initial density 40, 55, or 80 vpm). It can also be seen that, whenever the link is initially congested, the discharge flow rate is always equal to the congested capacity as long as the upstream demand is not lower than this capacity. This can be seen in scenarios with upstream demand set to 3200 vph. While this demand is lower than the uncongested capacity of the bottleneck, the operational capacity stays restricted at the congested capacity value as long as the upstream is congested (scenarios iv-b, v-b, vi-b).

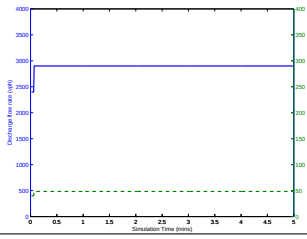
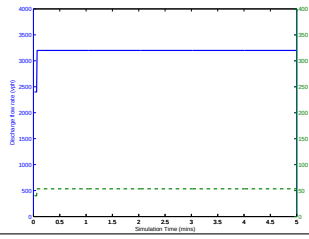
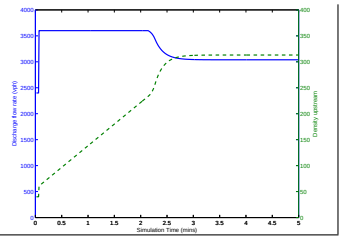
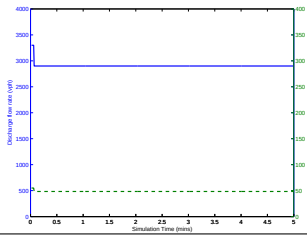
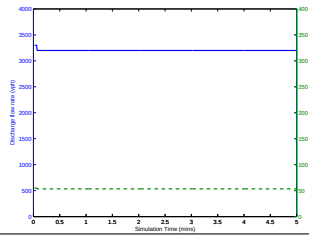
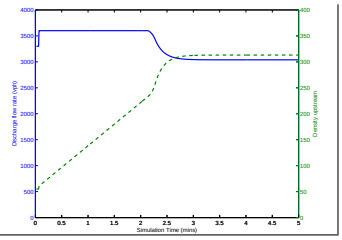
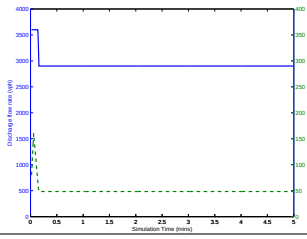
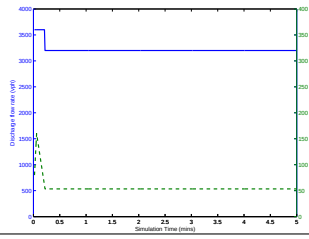
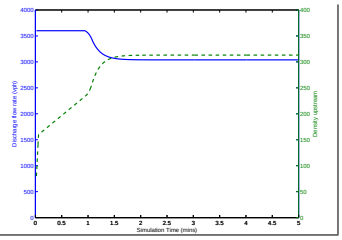
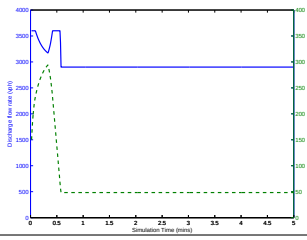
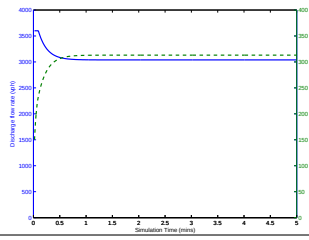
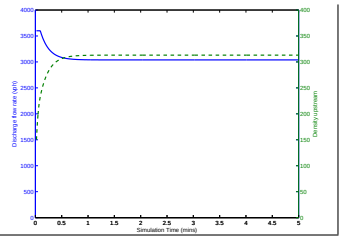
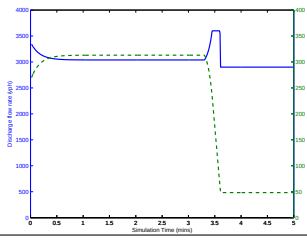
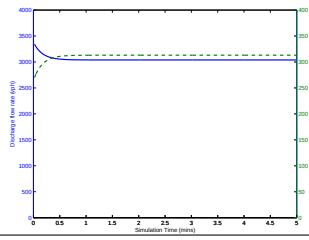
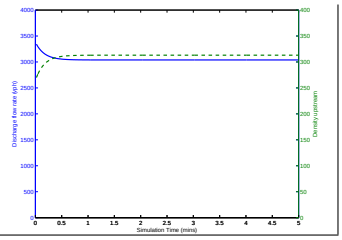
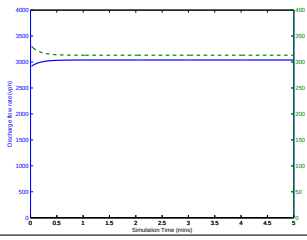
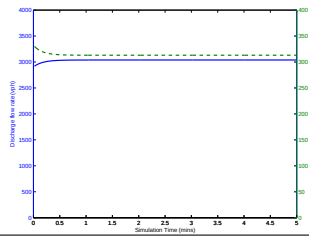
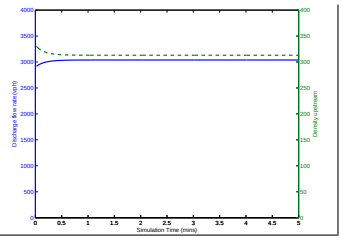
Upstr.	Upstream Demand		
Dens.	a. 2900 vph	b. 3200 vph	c. 3650 vph
(i) 40 vpm			
(ii) 55 vpm			
(iii) 80 vpm			
(iv) 150 vpm			
(v) 270 vpm			
(vi) 330 vpm			

Table 4.1: Discharge flow-rates at the bottleneck, along with the density immediately upstream of the lane-drop location shown for various combinations of initial and boundary conditions.

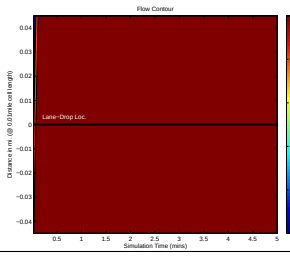
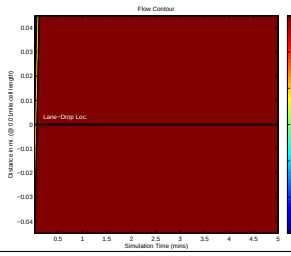
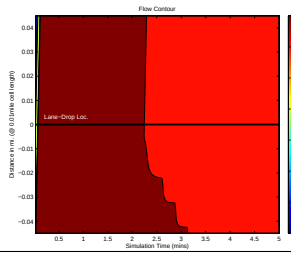
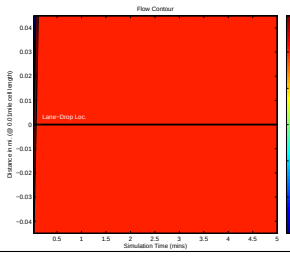
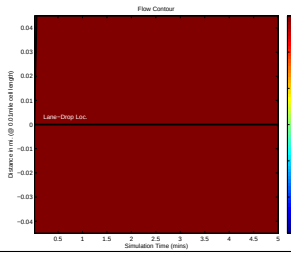
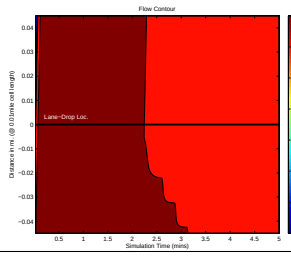
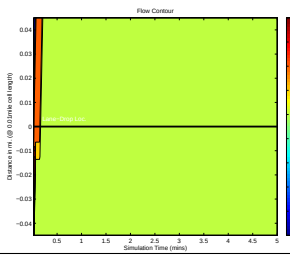
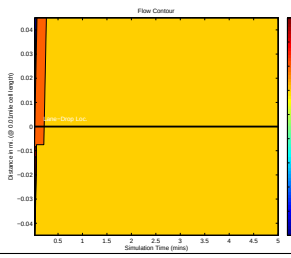
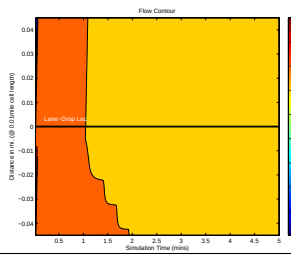
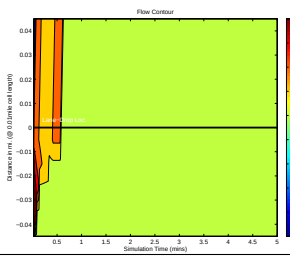
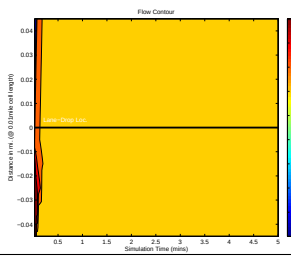
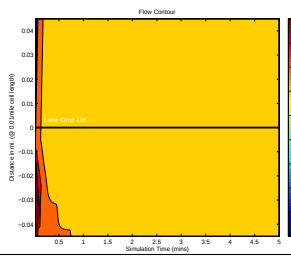
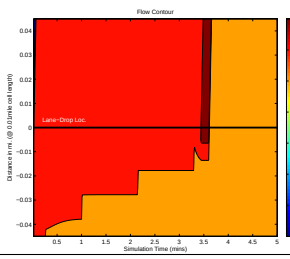
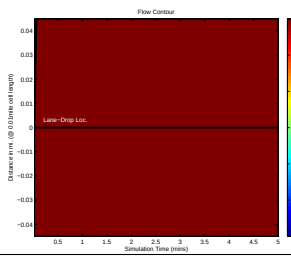
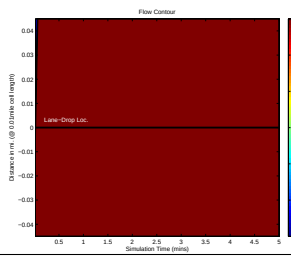
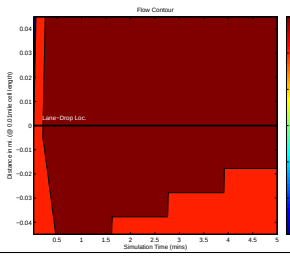
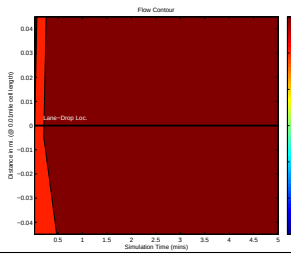
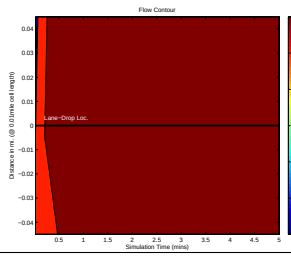
Upstr.	Upstream Demand		
Dens.	a. 2900 vph	b. 3200 vph	c. 3650 vph
(i) 40 vpm			
(ii) 55 vpm			
(iii) 80 vpm			
(iv) 150 vpm			
(v) 270 vpm			
(vi) 330 vpm			

Table 4.2: Flow contours for various combinations of initial and boundary conditions.

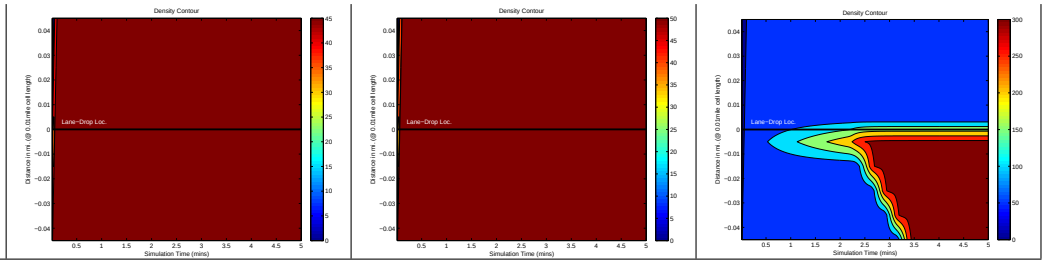
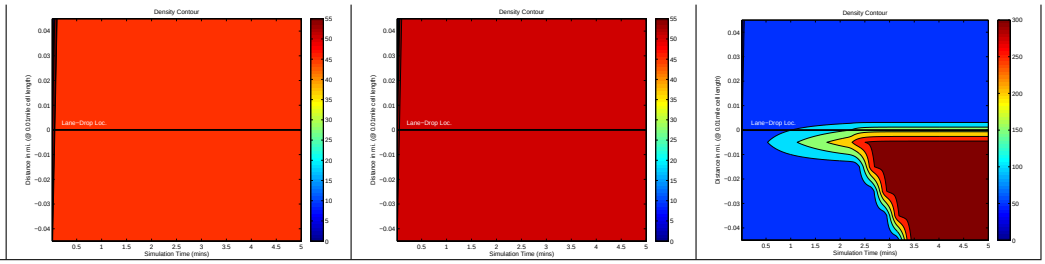
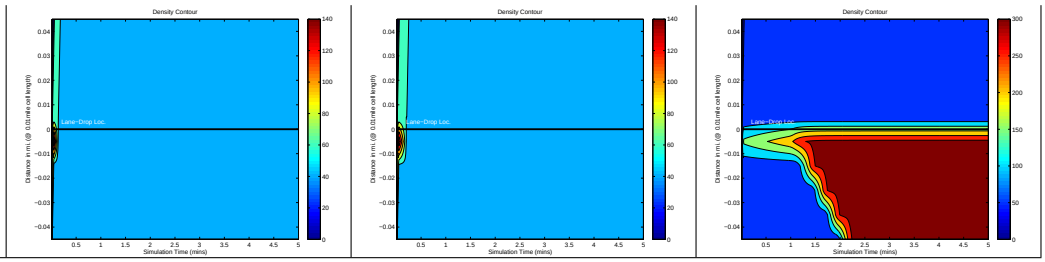
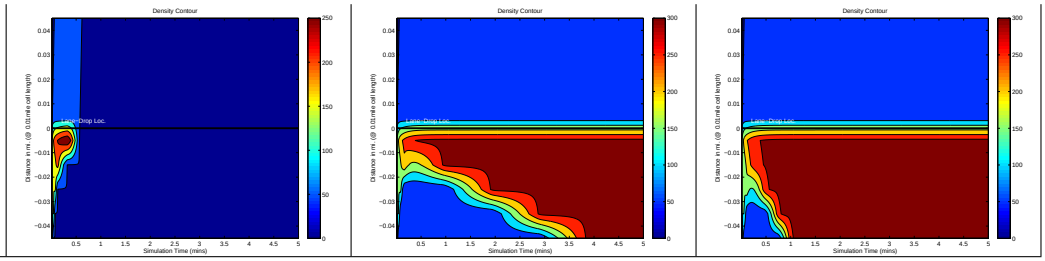
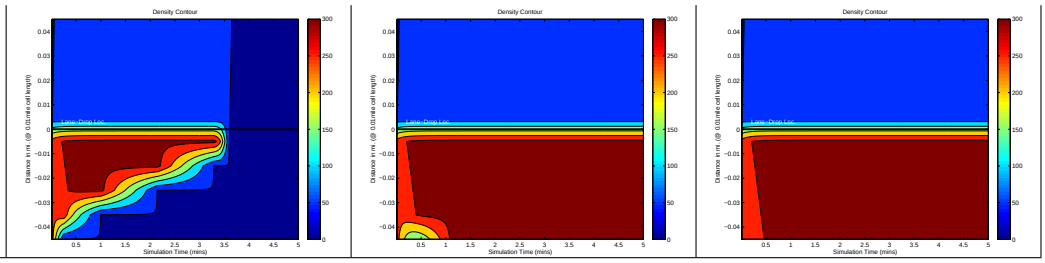
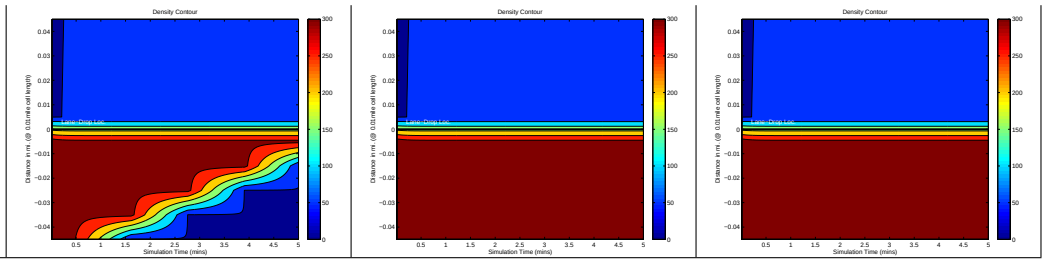
Upstr.	Upstream Demand		
Dens.	a. 2900 vph	b. 3200 vph	c. 3650 vph
(i) 40 vpm			
(ii) 55 vpm			
(iii) 80 vpm			
(iv) 150 vpm			
(v) 270 vpm			
(vi) 330 vpm			

Table 4.3: Density contours for various combinations of initial and boundary conditions.

4.6 Conclusions

In this chapter, a macroscopic model for capacity drop that incorporates a lane-changing and a bounded acceleration component into the Cell Transmission Model was proposed. It is shown that the proposed Lane Changing, Bounded Acceleration CTM (LCBA-CTM) can explain, and model the capacity drop phenomenon at lane drop bottlenecks. The lane-changing component is based on treating each lane-changing vehicle as contributing to the density of both its original and target lanes. This is modeled by defining a lane-changing intensity ratio, and defining 'perceived density' as the actual density scaled up by the lane-changing intensity. While calculating the bottleneck flux or flow-rate, the upstream demand from the lane-changing region in the model is computed from the 'perceived density' in place of the actual density. The bounded acceleration component uses a modified demand function linearly decreasing in density for over-saturated conditions. The reduced demand reflects the lower flow-rate of vehicles discharging from a queued condition due to the bounded nature of acceleration, and consequently, the lower discharge speeds compared to traditional CTM. With the two components in place, the congested portion of the demand function undergoes two transitions, a rotation to $-c^*$ slope to accommodate effects of acceleration, and a lateral translation to accommodate effects of lane-changing. The final model is shown to support the existence of stationary states that capture the capacity drop phenomenon at lane-drop bottlenecks and through analysis of the model, the stationary states corresponding to the bottleneck acting at the uncongested and the congested capacities are shown to both be stable to small perturbations in density. LCBA-CTM is also shown to be able to model traffic dynamics at the bottleneck.

A numerical example is used as proof of concept for the model where representative values of the model parameters lead to capacity drop magnitude well within the range observed in past studies. Constant loading problem analysis, as well as numerical simulation results are used to show the sturdiness of the model under varying initial and boundary condition setups.

The most interesting set up corresponding to the onset of congestion is explored in more detail using numerical simulation results. A full calibration effort however, would require recording trajectories and extracting lane-changing behavior at active lane-drop bottleneck sites, and remains outside of the scope of this chapter.

In order to keep the proposed model easily tractable, a few simplifying assumptions are made. First key assumption is related to the macroscopic behavior of vehicles near a lane-drop bottleneck. The study assumes the existence of a lane-changing region strictly upstream of the lane-drop, followed by an acceleration region strictly downstream. Further, the lane drop is considered sudden and the lane changing region is assumed to be effectively of a single CTM cell length in terms of how it impacts traffic flow. Together with the assumption that only the demand in the lane-changing region gets affected by the lane-changes, the assumption reflects a suggestion that vehicles change lanes very close to the lane-drop location under congested conditions.

For the present study, a constant value of the lane-changing intensity is further assumed. The lane-changing intensity ϵ $(= \alpha - 1)$, can be thought of as the ratio of total lane-changing time to total travel time for a region. (Jin, 2010a) shows relations that can be used to derive ϵ for a location knowing the average time taken by vehicles to perform lane-changes, and the length of the lane-changing region. This value is in fact also dependent on the traffic state as can be seen in (Jin, 2010a). Further, the average time taken by a vehicle to change lanes, a key parameter in computing the lane-changing intensity, in turn might also depend on traffic state, traffic composition, and on driver's familiarity with the location. Due to the dependency on traffic states, it is possible to have multiple stationary states corresponding to different values of α, and therefore, have varying capacities and capacity drop values. This could explain why capacity appears to be stochastic in observations, and explain the scatter in the flow-density observations near capacity. While the present model would greatly benefit from an empirical study aimed at investigating such relationships, it is left for future

studies to explore.

Another interesting extension to the study would be to investigate the extent of the lane-changing and bounded acceleration regions. The acceleration process at a lane-drop bottleneck might indeed begin slightly upstream of the lane-drop. Similarly, a dis-balance in lane densities immediately downstream of the lane-drop (rightmost lanes having higher density), might result in downstream lane-changes as well. Related studies could also look at the lateral distribution of densities near a bottleneck, and also investigate the longitudinal distribution of lane changes. Finally, extending the model to merge bottlenecks where the merging behavior is of importance in addition to the lane-changing and acceleration features is also a very interesting study to pursue.

Chapter 5

Deriving Modified Demand Functions from Microscopic Acceleration Models

5.1 Introduction

In Chapter 3 it was shown that a well calibrated modified demand function used within the Cell Transmission Model framework is able to reproduce observed traffic discharge properties such as flow-rate and headways at intersections. Though the chapter also dealt with calibration of the demand function, this was done by fitting a generated headway curve against observed headways for a given location. While this provides a useful tool for calibrating the model, such a calibration can only be performed when headway data is readily available. Thus, the model can not be easily used for predictive modeling purposes such as estimating the impact of technological improvements to the vehicle or driver assist systems on traffic dynamics and on intersection performance.

In this chapter, an alternate calibration technique is proposed where the macroscopic demand function is calibrated through a traffic stream representative microscopic acceleration profile.

Various acceleration models have been proposed in past studies that model acceleration rates of vehicles against their speeds. This relationship is referred to here, as the acceleration profile for a vehicle. The acceleration profile typically maps the bounded acceleration rate of vehicles to their speed. If an aggregate acceleration profile can be obtained for a traffic stream, I show how it can be used to calibrate the modified demand function.

I start with proposing the framework to be used to derive the demand function from acceleration models in Section 5.2. Following the explanation of the framework, in Section 5.3 the demand functions are derived for some popular acceleration models as applied to the triangular fundamental diagram. Finally, in Section 5.4 I conclude with some final thoughts and comments.

5.2 Framework for obtaining demand function

The objective, as stated earlier, is to derive a CTM demand function corresponding to a known acceleration profile. In this section, I outline the assumptions made towards this goal, and the framework developed for the process.

The first assumption to start with, is that the bounded acceleration rate of a vehicle depends only on its current speed, and neither on the past acceleration rates, nor on the desired final speed. Note that this acceleration rate is the maximal acceleration possible for a given speed. The actual acceleration rate of a vehicle might also depend on other aspects of traffic such as the car-following mechanics involved. This is indeed a strong assumption, suggesting say that the instantaneous acceleration of a vehicle currently at a speed of 30 mph in equilibrium wanting to accelerate to 35 mph, is the same as a vehicle currently also at 30 mph but in the process of accelerating from 0 to 60 mph. However, this assumption is in line with many acceleration models, specifically from the kinetic family of models of vehicle acceleration.

This assumption is indeed necessary to the effect that the corresponding macroscopic model does not track vehicle speeds as an independent variable. Thus, the history of the past speeds is not known, and thus, can't be relied on. This assumption allows us to derive a unique valued demand-density function.

If the above is true, we have a unique single-valued acceleration vs. speed relationship that does not depend on the desired speed of the vehicle, or the stage of acceleration it currently is in. Note here, that the relationship can still be dependent on parameters such as the free flow speed v_f. As stated earlier, this assumption holds true for most acceleration models belonging to the kinetic family of models, including the constant acceleration model, the dual regime model, the linear decay model, and Gipps' non-linear acceleration model. For the present study, I consider three of the above models, excluding the dual regime model due to it non-continuous nature.

Since the acceleration process is of interest only corresponding to initially congested traffic conditions, and since bounded acceleration comes into play only under the presence of unhindered downstream conditions, a roadway setup with traffic upstream of a break point in equilibrium conditions with density k_u and speed v_u and downstream uncongested with 0 density initially is used. The upstream density k_u can be any density as long as it satisfies $k_c < k_u \leq k_j$. Let us next assume that we wish to apply a CTM-like discretization of space and time into Δx and Δt cells such that $\Delta x/\Delta t \approx v_f$. Using the ideas of traffic demand and supply, the flow q_0 across the location of the discontinuity depends on the demand in the upstream cell and the supply available in the downstream cell. However, since downstream is considered to be empty, the discharging flow-rate across the boundary would be determined by the demand in the upstream cell.

$$\text{Boundary Flow} = D'(k_u) \tag{5.1}$$

Since the upstream cell is initially in equilibrium traffic condition, we can assume all vehicles in the cell have a speed v_u at time $t = 0$. As noted earlier, the downstream is uncongested and hence, the vehicles in the upstream cell would not be restricted by any car-following mechanism, and instead solely by the acceleration model. From the microscopic acceleration process, the flow-rate across the boundary can be determined knowing how many vehicles cross the boundary within a single CTM time step Δt. If this can be computed, the CTM demand corresponding to the upstream density can be derived from matching the boundary flow-rate.

Two approaches to identifying the boundary flow from knowing the acceleration profiles of vehicles are presented. In the first 'simplified' approach, vehicles within a cell are assumed to respond instantaneously to the acceleration process. In the second 'shockwave approach', the acceleration information is propagated within the cell at the shockwave speed. The following subsections explore these approaches in greater detail.

5.2.1 Simplified approach

As stated earlier, in the simplified approach, vehicles within a cell are assumed to respond instantaneously to the desire to accelerate such that the trigger information travels instantaneously to all vehicles within the cell (Figure 5.2a). This is effectively same as assuming that the density within a cell remains constant during a Δt time-step, but the speed of the block of vehicles changes according to the acceleration rate determined by the bounded acceleration function computed for the current speed within the cell. Thus, the block of vehicles accelerates together as an incompressible fluid within each Δx, Δt cell. Further, within the CTM cell, acceleration rate is considered to be constant during a time-step.

From Figure 5.2a, the last vehicle to cross the boundary within the time-step, would travel a distance x equivalent to the distance covered by a vehicle traveling at an initial speed v

and accelerating at a constant rate A in time Δt.

$$x = v\Delta t + \frac{1}{2}A(v)\left(\Delta t\right)^2 \tag{5.2}$$

The total flow across the boundary can be calculated from the initial density within the cell k and the distance to the last vehicle that crosses the boundary x. The corresponding flow-rate for the time-step would be given as:

$$\text{Boundary Flow} = \frac{xk}{\Delta t} = vk + \frac{1}{2}A(v)\Delta tk \tag{5.3}$$

Since this should also match the demand corresponding to density k,

$$\text{Demand}\left(k\right) = vk + \frac{1}{2}A(v)\Delta tk \tag{5.4}$$

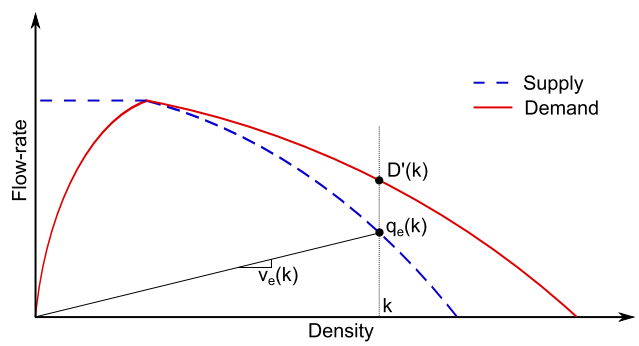
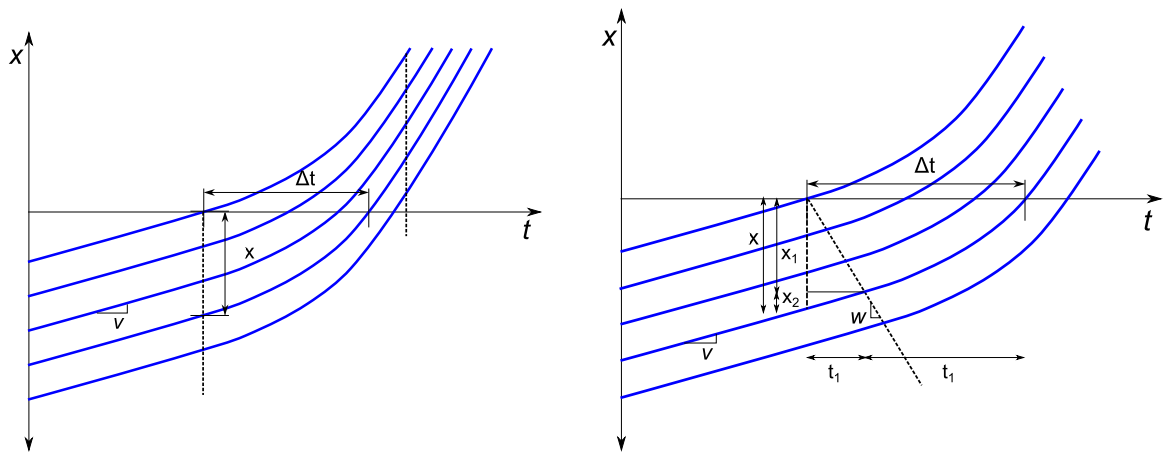


Figure 5.1: Demand Supply Diagram

5.2.2 Shockwave approach

In the shockwave approach, acceleration information travels along a shockwave to vehicles within the cell. Thus, vehicles continue to travel at their initial speeds for a finite time till the shockwave reaches the vehicle,triggering its acceleration process (Figure 5.2b). Once again,



(a) Simple Framework (b) With shockwave

Figure 5.2: The figure shows the simple and the shockwave based frameworks for deriving demand function from acceleration.

the acceleration rate is assumed to be constant for all vehicles in a cell during a time-step when they start accelerating.

The triangular fundamental diagram is used here to show the derivation framework. Using the triangular fundamental diagram reduces the complexity involved as the shockwave speed is constant regardless of the traffic state.

From Figure 5.2b, the last vehicle to cross the boundary within the time-step, would travel a distance $x = x_1 + x_2$ in time Δt, where x_2 is the distance covered by the vehicle at constant speed in time t_1 when the shockwave catches up to the vehicle, and x_1 is the distance covered by the vehicle during its acceleration phase with an acceleration rate of A. The total time-step is broken down into two components:

$$\Delta t = t_1 + t_2 \tag{5.5}$$

The shockwave travels a distance x_1, while the last vehicle crossing the boundary travels a

distance x_2 in time t_1:

$$x_1 = wt_1 \tag{5.6}$$

$$x_2 = vt_1 \tag{5.7}$$

$$x = x_1 + x_2 \tag{5.8}$$

Finally, the distance covered during the acceleration time t_2 can be given as:

$$x_1 = ut_2 + \frac{1}{2}A(v)t_2^2 \tag{5.9}$$

Using the value of x_1 from (Eq. 5.6) and (Eq. 5.8) in (Eq. 5.9) gives us:

$$\frac{1}{2}A(v)t_2^2 + (u+w)t_2 - w\Delta t = 0 \tag{5.10}$$

whose solution yeilds:

$$t_2 = \frac{\sqrt{(wk_j)^2 + 2A(v)\Delta twk^2}}{Ak} - \frac{wk_j}{Ak} \tag{5.11}$$

The demand function once again can be obtained from matching the flow-rate as:

$$\text{Demand}\,(k) = wk_j\left(1 - \frac{t_2}{\Delta t}\right) \tag{5.12}$$

5.3 Demand functions for triangular FD

The triangular fundamental diagram is used to show how the demand function can be obtained corresponding to some popular acceleration functions. Since we are only concerned with the acceleration scenario where the upstream location is congested (density higher than

critical density), only the congested portion of the fundamental relationship is used.

For the triangular fundamental diagram, under congestion we have:

$$q = (k_j - k)w \tag{5.13}$$

$$u = \left(\frac{k_j}{k} - 1\right) w \tag{5.14}$$

$$k = \frac{k_j w}{u + w} \tag{5.15}$$

The constant acceleration, the linear decay acceleration, and Gipps' acceleration models are first examined through the simplified approach, followed by the constant and linear decay acceleration models under the shockwave approach.

5.3.1 Simplified approach

Constant Acceleration Model

In the constant bounded acceleration model, the acceleration bound is considered to be constant and independent of speed. Thus, we have:

$$\text{Acceleration, } A(v) = a_0 \tag{5.16}$$

$$D'(k) = wk_j + \left(\frac{a_0 \Delta t}{2} - w\right) k \tag{5.17}$$

The resulting demand function would be linear in density. The slope of the demand function, and the demand corresponding to the critical density and the jam density can be computed

as:

$$D'(k_c) = q_c + \frac{1}{2}a_0 \Delta t k_c \tag{5.18}$$

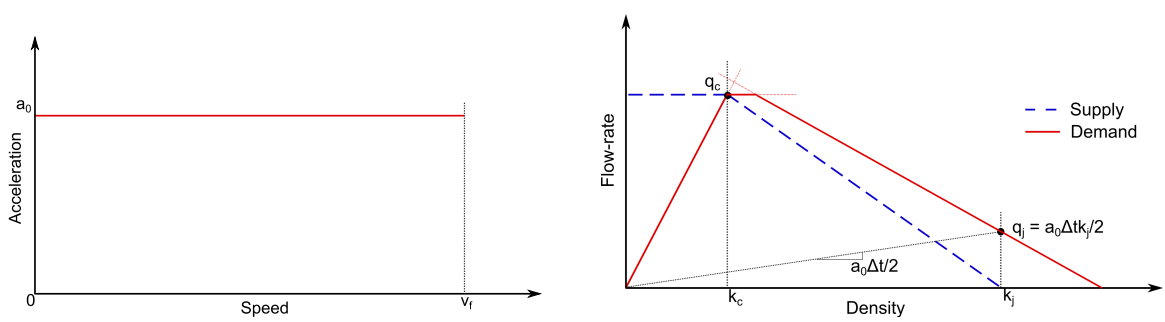
$$D'(k_j) = \frac{1}{2}a_0 \Delta t k_j \tag{5.19}$$

$$\text{Slope of demand curve, } c^* = w - \frac{a_0 \Delta t}{2} \tag{5.20}$$

One complication with the constant acceleration model is noted at values of density close to the critical density. For such densities, the corresponding speed is almost equal to the free-flow speed. The constant acceleration assumption of this model suggests that vehicles have a constant positive acceleration rate even at $k \approx k_c$, $v \approx v_f$.

In order to overcome this issue, a condition is added that the demand can never exceed the capacity flow. The adjusted demand function then becomes:

$$D'(k) = min\left\{k_j w + \left(\frac{a_0 \Delta t}{2} - w\right) k, q_c\right\} \tag{5.21}$$



(a) Constant Acceleration Model

(b) Supply and Demand

Figure 5.3: The figure shows a. The constant acceleration model where acceleration remains at a constant value a_0; and b. The demand and supply functions corresponding to the constant acceleration model applied to a Triangular fundamental diagram

Linear Decay Acceleration Model

The linear decay acceleration model corresponds to a acceleration function that linearly decreases with an increase in speed (Figure 5.4a). The acceleration is maximum when speed is 0, and this maximum acceleration is given by a_0. Additionally, the acceleration is considered to be 0 when speed is equal to the free-flow speed at the location. Formally, the acceleration equation would be denoted as:

$$\text{Acceleration, } A(v) = a_0 \left(1 - \frac{v}{v_f}\right) \tag{5.22}$$

In case of linear decay acceleration model, the cruise control acceleration rate would never take precedence over the linear acceleration function as long as the acceleration function is properly defined. We thus have:

$$D'(k) = \left(\frac{1}{2}a_0\Delta t - \frac{a_0 v \Delta t}{2v_f} + v\right) k \tag{5.23}$$

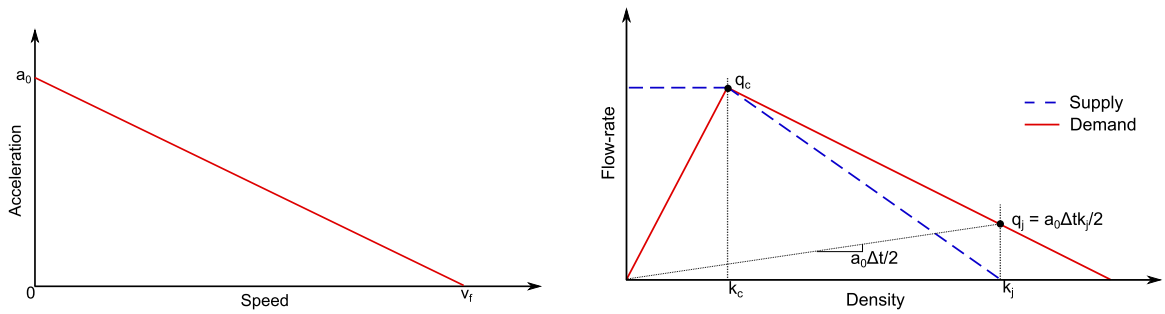
$$= k_j w + \left(\frac{a_0}{2}\Delta t - w\right) k - \frac{w a_0 \Delta t}{2v_f}\left(k_j - k\right) \tag{5.24}$$

$$D'(k_c) = q_c \tag{5.25}$$

$$D'(k_j) = \frac{a_0}{2}\Delta t k_j \tag{5.26}$$

$$\text{Slope of demand curve, } c^* = w - \frac{a_0 \Delta t}{2}\left(\frac{k_j}{k_j - k_c}\right) \tag{5.27}$$

Note that the demand function evaluates to q_c at $k = k_c$, and a positive value $(a_0 \Delta t k_j)/2$ at $k = k_j$. Thus, the demand is linearly decreasing in density k in the congested region (Figure 5.4b).



(a) Linear Decay Acceleration Model (b) Supply and Demand

Figure 5.4: The figure shows a. The linear decay acceleration model where acceleration linearly decreases with increasing speed; and b. The demand and supply functions corresponding to the linear acceleration model applied to a Triangular fundamental diagram

Non-Linear Gipps-like Acceleration Model

A third popular model for acceleration is based on Gipp's model. This acceleration model is non-linear with respect to speed. The acceleration rate is given by the following function:

$$\text{Acceleration, } A(v) = a_0 \left(1 - \frac{v}{v_f}\right) \sqrt{b + \frac{v}{v_f}} \tag{5.28}$$

And the corresponding demand function is given as:

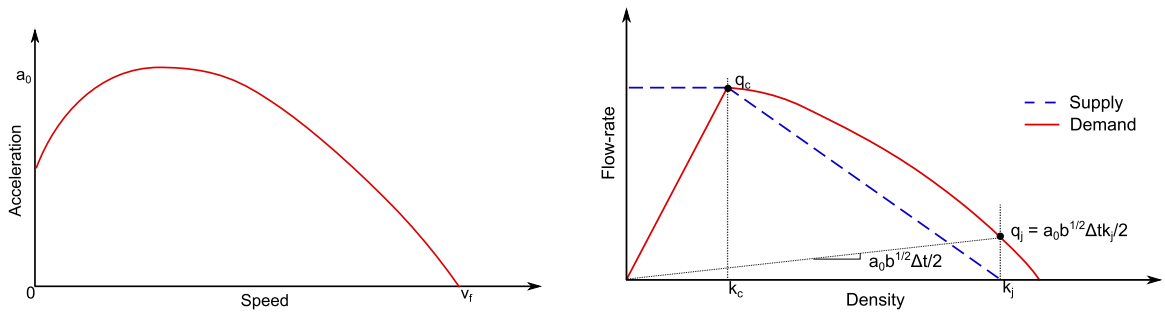
$$D'(k) = \left(\frac{a_0}{2}\left(1 - \frac{v}{v_f}\right)\sqrt{b + \frac{v}{v_f}}\Delta t + v\right) k \tag{5.29}$$

$$= \frac{a_0}{2}\left(1 - \frac{(k_j - k)\, w}{k v_f}\right)\sqrt{b + \frac{(k_j - k)\, w}{k v_f}}\Delta t k + w\,(k_j - k) \tag{5.30}$$

$$D'(k_c) = q_c \tag{5.31}$$

$$D'(k_j) = \frac{a_0}{2}\sqrt{b}\Delta t k_j \tag{5.32}$$

As a result, the demand function for this case is a non-linear demand function.



(a) Gipps' NonLinear Acceleration Model (b) Supply and Demand

Figure 5.5: The figure shows a. Gipps non-linear acceleration model where acceleration is a combination of a linear and a square-root speed component; and b. The demand and supply functions corresponding to the acceleration model applied to a Triangular fundamental diagram

5.3.2 Shockwave approach

The demand functions derived under the shockwave approach assumptions are of the following form:

$$D'(k) = wk_j \left(1 - \frac{t_2}{\Delta t}\right), \tag{5.33}$$

where,

$$\frac{t_2}{\Delta t} = \sqrt{\left(\frac{w}{A\Delta t}\frac{k_j}{k}\right)^2 + \frac{2w}{A\Delta t}} - \frac{w}{A\Delta t}\frac{k_j}{k} \tag{5.34}$$

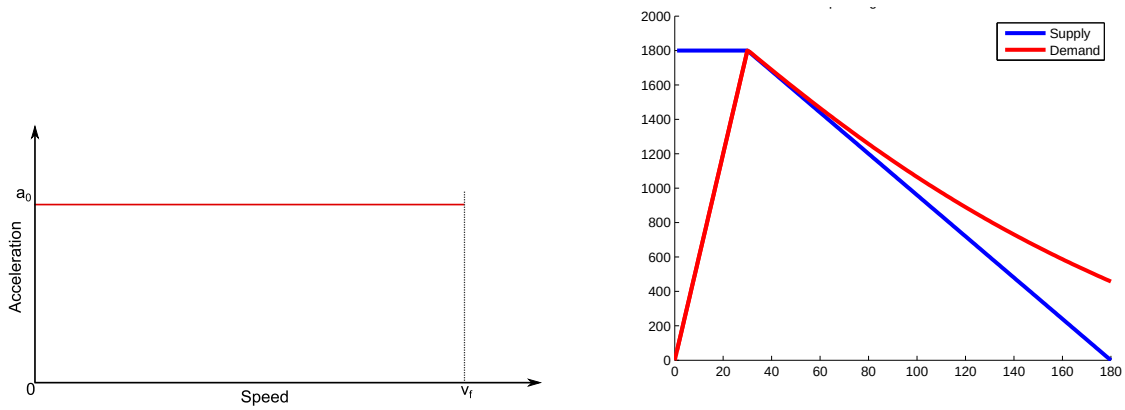
The demand equations corresponding to the constant acceleration and linear acceleration models can be obtained by substituting the value of $A(v)$ above with the following:

For constant acceleration model:

$$A(v) = a_0 \tag{5.35}$$

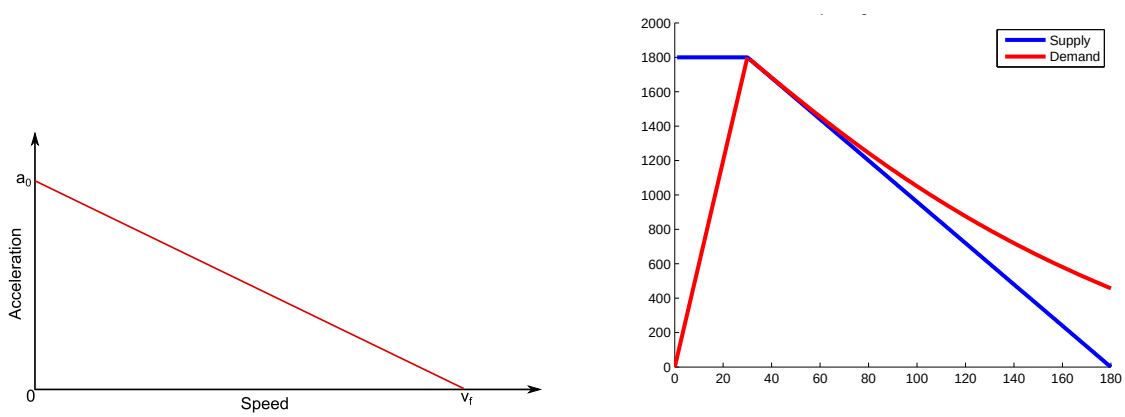
For linear acceleration model:

$$
\begin{aligned}
A(v) &= a_0 \left(1 - \frac{v}{v_f}\right) \\
&= a_0 \left[\frac{(v_f + w)\, k - w k_j}{v_f k}\right]
\end{aligned}
\tag{5.36}
$$



(a) Constant Acceleration Model (b) Supply and Demand

Figure 5.6: The figure shows a. The constant acceleration model where acceleration remains at a constant value a_0; and b. The demand and supply functions corresponding to the constant acceleration model applied to a Triangular fundamental diagram via the shockwave derivation approach.



(a) Linear Decay Acceleration Model (b) Supply and Demand

Figure 5.7: The figure shows a. The linear decay acceleration model where acceleration linearly decreases with increasing speed; and b. The demand and supply functions corresponding to the linear acceleration model applied to a Triangular fundamental diagram via the shockwave derivation approach.

The resulting demand functions with representative values of the various parameters used

are shown in Figure 5.6 for the constant acceleration model, and Figure 5.7 for the linear acceleration model. In each of the models, the demand function has very similar non-linear shapes with respect to density.

5.4 Conclusions and discussions

This chapter introduced a framework for deriving a CTM demand function corresponding to a known acceleration function for a fixed fundamental flow-density relationship. The framework explores two mechanism for drawing the equivalence between the microscopic vehicle trajectories and the corresponding demand function. In the first assumption, density is considered to be constant within the scope of a single CTM cell observed for a time-step, similar to the assumption in general CTM. The speed, however, is assumed to change during each time step based on a fixed acceleration rate computed corresponding to the traffic conditions at the start of the time-step. In the second model, the acceleration trigger information is assumed to travel at the shockwave speed with a cell and time-step unit. Thus, densities within a cell can change within the time-step due to the presence of shockwave. The first technique results in an easier to track demand function for popular acceleration models, while the second technique represents more realistic behavior. However, further investigations are needed to determine the comparative correctness of the techniques.

The framework presented makes certain assumptions that may be relaxed through future endeavors. First, only unique speed-acceleration relationships are explored. While this might not be a perfectly realistic representation of field behavior, this is a common assumption made for various studies and is thus considered accepetable.

A second intrinsic assumption is that vehicles modeled above, are limited only by the bounded acceleration process. While the car-following restrictions may be eliminated as such con-

straints would get captured by the downstream supply restricting the boundary flux, the possibility of a cruise-control constraint and maximal speed constraint are realistic. The framework could be extended by imposing such restrictions within the techniques explored.

One last assumption made is the absence of a reaction time. While the shockwave propagation does account in part the delay caused by reaction times, it should be noted that reaction delays to stimuli such as the change of the signal phase is missed.

Chapter 6

Applications

6.1 Improved Vehicle Accelerations

6.1.1 Introduction

There has been much interest and debate about how technological improvements in vehicles, such as the introduction of connected autonomous vehicles and alternative fuel vehicles, affect traffic flow over networks. As autonomous vehicle penetration rates increase with technological advances, they are expected to have a growing impact on roadway network performance. While it is unanimously believed that such vehicles would have a positive impact on traffic flow and safety, the magnitude of improvements and the trade-offs to be expected are somewhat unclear. Shorter reaction times due to the automation, sensing, and coordination between vehicles, and improved accelerations are among the most important influential aspects that could improve traffic flow.

Due to faster reaction times, it is believed that vehicles would be able to travel safely with shorter following distances, resulting in more tightly packed traffic streams traveling

at a given speed compared to current conditions. The improved obstacle detection and response systems could potentially mean an overall increase in free flow speeds on most roadways, further increasing the capacity flow-rate and the flow of traffic in general. Better reaction times coupled with coordinated driving could also result in more stable flows and fewer occurrences of stop-and-go traffic. Finally, a combination of factors, such as faster reaction times and improved vehicle response and performance, would result in faster possible acceleration for vehicles. Similarly, electric and hybrid vehicles are associated with faster pick up accelerations, and the acceleration rates are expected to improve as battery technologies improve.

In this section, I consider the specific case of assessing how arterial intersection discharge features such as throughput flow-rate and discharge headways as described in Chapter 3 would potentially improve due to the growth of vehicle technologies. More specifically, I use the frameworks described in Chapter 3 and Chapter 5, to assess how improved acceleration rates would influence traffic flow at signalized intersections.

6.1.2 Intersection features with improved vehicular accelerations

For this study, I concentrate on the linear decay acceleration model and the linear demand function that results from applying it in conjunction with the triangular fundamental diagram. The demand function thus obtained is similar to the linearly decreasing demand function used in Chapter 3. While the same exercise can be repeated for other flow-density and speed-acceleration relations, the linear demand function obtained from the above combination is already familiar to us from Chapter 3.

I assume that the acceleration function representing conventional vehicles (human-driver and internal combustion engine) and improved acceleration vehicles (driver-less and electric

vehicles) are given respectively by:

$$\text{Acceleration (conventional)},\ A_h(v) = a_h\left(1 - \frac{v}{v_f}\right) \tag{6.1}$$

$$\text{Acceleration (enhanced)},\ A_a(v) = a_a\left(1 - \frac{v}{v_f}\right) \tag{6.2}$$

where a_h and a_a are the acceleration coefficients representing maximum acceleration rate at 0 speed for conventional and enhanced vehicles. The higher performance of enhanced technology vehicles is represented by a_a being higher than a_h.

From Section 5.3.1 we have the corresponding demand function, $D'(k)$, the slope of the demand function under congested conditions, c^*, and jam demand, $q_j = D'(k_j)$:

conventional vehicles

$$D'_h(k) = k_j w + (a_h \Delta t - w)\, k - \frac{w a_h \Delta t}{v_f}\,(k_j - k) \tag{6.3}$$

$$c^*_h = w - a_h \Delta t \frac{k_j}{k_j - k_c} \tag{6.4}$$

$$q_{j_h} = a_h \Delta t k_j \tag{6.5}$$

enhanced tech. vehicles

$$D'_a(k) = k_j w + (a_a \Delta t - w)\, k - \frac{w a_a \Delta t}{v_f}\,(k_j - k) \tag{6.6}$$

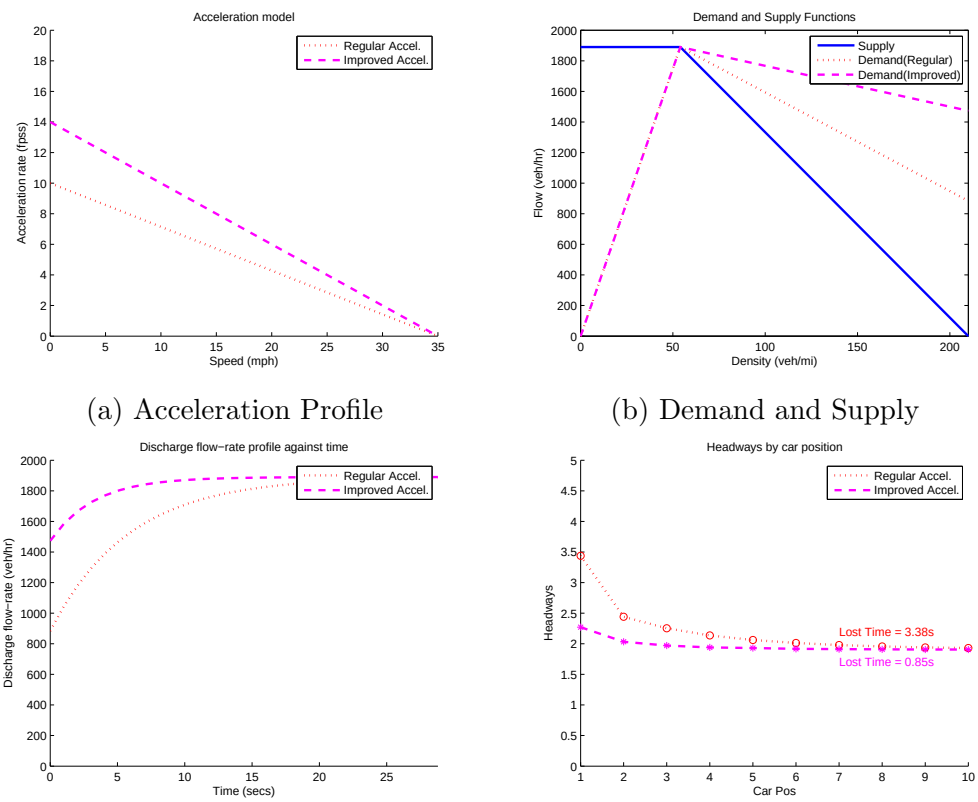
$$c^*_a = w - a_a \Delta t \frac{k_j}{k_j - k_c} \tag{6.7}$$

$$q_{j_a} = a_a \Delta t k_j \tag{6.8}$$

From Chapter 3, we can then compute the corresponding lost times for the two cases (as-

suming CFL number = 1):

$$L_{1_h} = \frac{\Delta x}{w - c_h^*} \frac{c_h^*}{w}$$
$$= \frac{v_f}{a_h} \frac{k_j - k_c}{k_j} - \frac{\Delta x}{w} \tag{6.9}$$
$$L_{1_a} = \frac{v_f}{a_a} \frac{k_j - k_c}{k_j} - \frac{\Delta x}{w} \tag{6.10}$$



(a) Acceleration Profile

(b) Demand and Supply

(c) Discharge Flow-rates

(d) Headways

Figure 6.1: Predicted enhanced technology vehicle discharge properties compared against conventional vehicles. The above figure shows the demand and supply functions, the intersection discharge flow-rates, and headways for two cases with different acceleration rates. A maximum acceleration rate of 10fpss is used to represent conventional vehicles, and a value of 14fpss is used for enhanced acceleration vehicles.

Using values for the demand and supply function parameters calibrated for the NGSIM data as shown in Chapter 3: $v_f = 35mph$, $k_c = 54vpm$, $k_j = 210vpm$, $w = 12.1mph$, $\Delta x = 0.01m$, and setting $a_h = 10fpss$, and $a_a = 14fpss$ gives us lost time values of

6.6s and 2.5s respectively. Effectively, the lost capacity per green phase at an intersection would reduce from $3.4s * 1890vph \approx 1.8veh$ to $0.9s * 1890vph \approx 0.5veh$, with a little over 1 additional vehicle being served by the same green period in the latter case. Figure 6.1 illustrates the impact of the improved acceleration rates on the intersection discharge flow-rate and headways.

6.2 Arterial Network Fundamental Diagram - Signalized Ring Road

6.2.1 Introduction

(Godfrey, 1900) first proposed that the average flow-rate over a roadway network is related to the network density through a Network Fundamental Diagram (NFD) or Macroscopic Fundamental Diagram (MFD). Various studies since have tried to study this property of networks and to calibrate the relationship using field data or derive it analytically (Ardekani and Herman, 1987; Mahmassani et al., 1987; Geroliminis and Daganzo, 2008; Daganzo and Geroliminis, 2008; Cassidy et al., 2011). The MFD has been shown to be a unique relationship in homogeneous networks. The existence of the MFD has been a popular topic for traffic researchers recently, and has inspired various studies towards application of the MFD to understand traffic behavior and to develop network control strategies.

In order to estimate an approximate MFD, I use a signalized ring road setup, simulated through various modifications of the CTM framework, as a representation of a homogeneous road network. The signalized ring road is the simplest approximation of a road network, equivalent to an infinitely long street with uni-directional flow of traffic, where each link is identical in terms of traffic conditions and intersection signal settings. The set up assumes

that all links have the same cycle lengths and green ratios, and all signals are synchronized so that there is no effective offset. While only one directional flow is considered inside the ring, the ring road is an effective approximation of a city grid system where streets have uni-directional traffic flow, with the direction of travel alternating between streets.

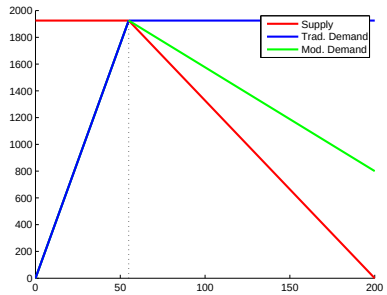
6.2.2 Simulation setup

Network aggregate traffic conditions are generated using CTM based simulations over the signalized ring road corresponding to varying traffic densities. Traffic is assumed to follow a triangular equilibrium flow-density relationship. The fundamental diagram is defined by representative parameters with a free-flow speed of 35 mph, critical density 55 vpm, and jam density equal to 200 vpm. For the modified demand function based CTM, the jam demand is set to 800 vph. The corresponding equivalent lost time is set to 3.8 sec based on Section 3.4.2.

The simulation setup uses a ring road length of 0.1 mi which corresponds to the network with city blocks of similar lengths. The intersection setup assumes a pre-timed signal at the intersection with a constant green ratio, set to 0.5 for the simulations. Since the conflicting movement phases are not designed for, all-red time is set to 0. The green phase therefore is set to exactly one half of the design cycle length.

The simulation is run over 1000 cycle lengths with CTM cell size of set to 0.01 miles and the timestep set to correspond to a CFL number equal to 1.

The fundamental diagram and the parameters of the simulation are shown in Figure 6.2.



(a) Flow-density relationship

Param	Value	Param	Value
v_{ff}	35 *mph*	q_c	1925 *vph*
k_c	55 *vpm*	w	13.3 *mph*
k_j	200 *vpm*	c*	7.8 *mph*
q_j	800 *vph*	k_j*	303 *vpm*
L	0.01 *mi*	*Block Length*	0.1 *mi*
Δt	1.03 sec	*Lost Time*	3.8 *sec*

(b) Simulation Params

Figure 6.2: Simulation parameters showing the demand and supply functions used and the values for each of the parameters used.

6.2.3 Simulation results

Figure 6.3 shows the MFDs obtained for various Cycle Lengths comparing Cell Transmission Model with traditionally defined demand, a Lost Time based CTM, and the modified demand CTM. As is expected, the traditional CTM model overestimates the performance of the network for densities close to critical density. The estimation is especially inaccurate for low cycle lengths, where the start up lost time is a higher proportion of the green time. The lost time based model that uses effective green times in place of actual green times in order to adjust for the lost capacity, is a more accurate representation. However, as can be seen from Figure 6.3, lost time based models might underestimate the aggregate flow-rate for smaller cycle lengths. This is due to the fact that lost time is typically computed independent of the demand volume and cycle length. When either the demand value is low, or the cycle length is small, it is possible that fewer vehicles are served every green phase, than used for computing the lost time. Thus, the lost time computed is an over-estimation of the actual startup lost time for vehicles being served. In such situations, the intersection throughput would be higher than predicted by the lost time model (though still smaller than that predicted by traditional CTM).

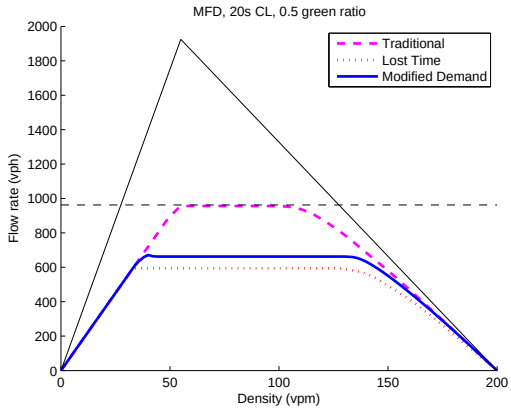
The modified CTM captures the vehicle discharge process more accurately than the other

two models. Catching the exponential form of the discharge flow-rate over time, the flow-rate is correctly calculated to be between values reported by the other two model. While the lost time model is reasonably accurate for medium and large cycle lengths, the modified demand CTM can greatly useful where smaller cycle lengths might be needed. The modified demand CTM also offers a slightly higher accuracy over the lost time model for high traffic density values, where queue spill-back might affect the intersection discharge flow-rate.

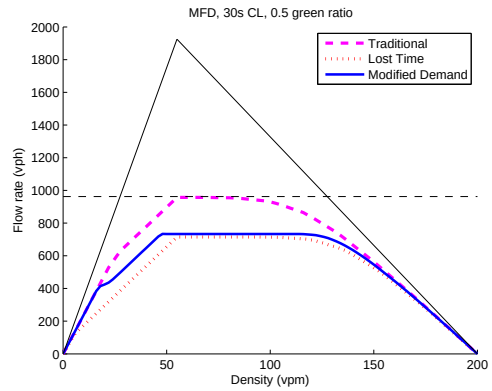
Figure 6.4 shows MFDs for a range of Cycle Lengths (2s to 120s, in 2s increments) for each of the models. The figure further illustrates how each of the three models performs in the range of values for network traffic density and cycle length. The MFD range obtained from the traditional CTM model resembles the expected shape (Jin and Yu, 2015), but overestimates the flow-rates. The accuracy of the lost time CTM model and the modified demand CTM model are similar, and better than the traditional CTM, for most cycle lengths and densities. However, the lost time model predicts lower flow-rates for smaller cycle lengths that expected.

The properties of the models are shown from another perspective in Figure 6.5, where the aggregate flow-rate is reported against cycle length for select densities. The graph can be used to estimate the optimal flow-rate and optimal cycle length for various levels of traffic density. In the case of the traditional CTM where the intersection operates at saturation flow-rate for the entire green period, the optimal flow-rates typically correspond to smaller cycle lengths. For higher cycle lengths, green periods for certain cycles might be wasted based on demand values, thus resulting in reduced aggregate intersection flow-rates. In the case of the lost-time model, for any value of cycle length corresponding to green phase being smaller than the constant lost time, there would be no flow through the intersection. The aggregate flow-rate steadily increases with increasing cycle length for smaller cycle lengths that are high enough to have green times larger than the lost time. The modified demand CTM can be seen as a mix of properties both from the traditional CTM and the lost time model. While the graph can not be taken as conclusive evidence for global optimality as

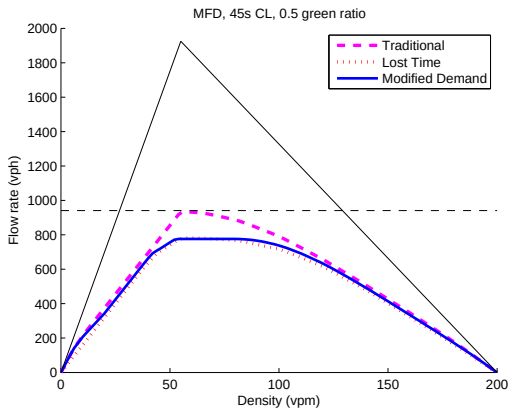
the objective function is highly non-linear, such a simulation study can give a good estimate on what behavior to expect for various cycle lengths, and can be a useful tool for design of intersection signals.



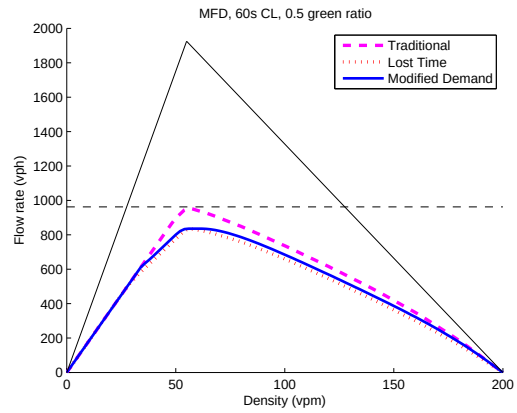
(a) CL = 20s



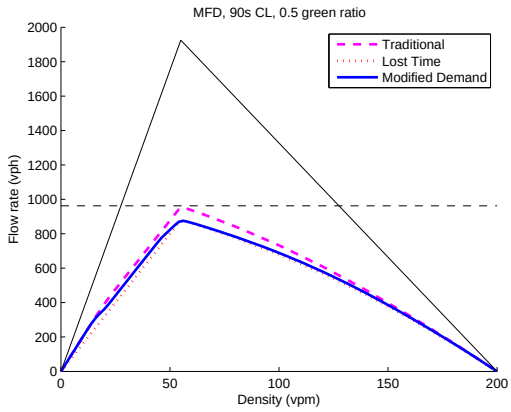
(b) CL = 30s



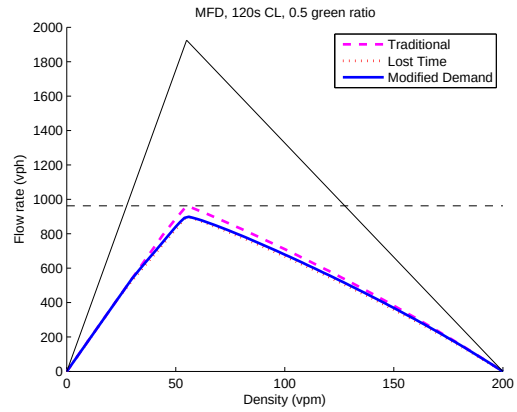
(c) CL = 45s



(d) CL = 60s

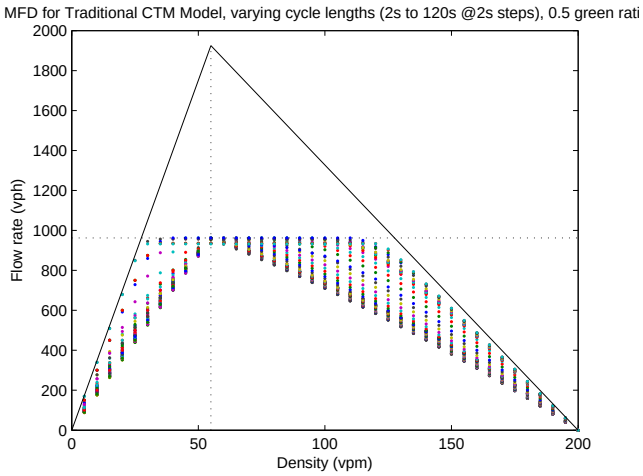


(e) CL = 90s

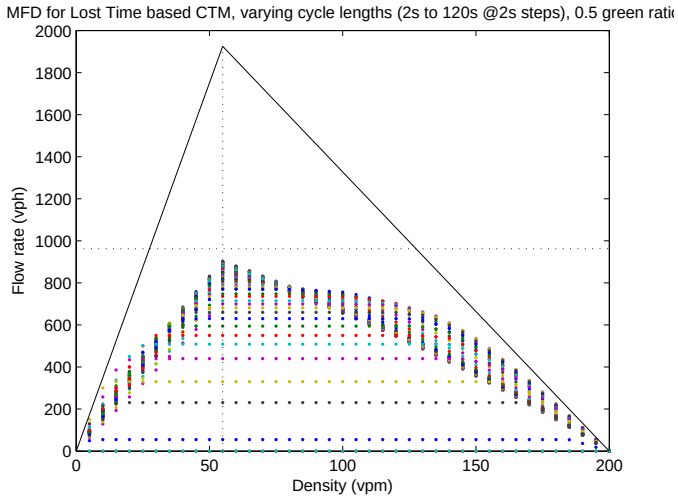


(f) CL = 120s

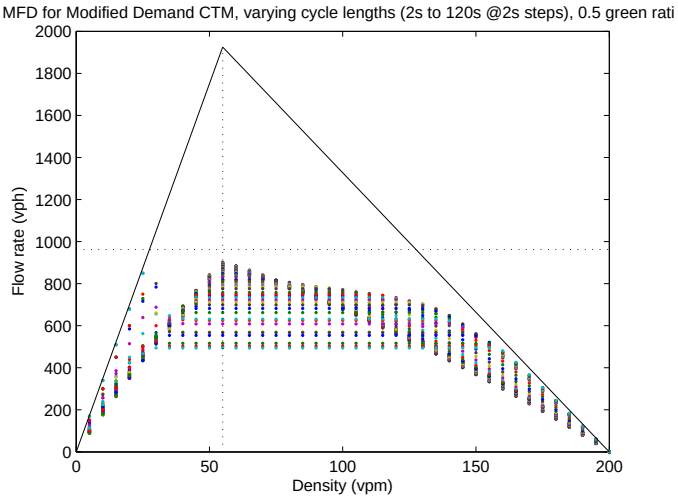
Figure 6.3: MFDs comparing the behavior of the Traditional CTM, Lost Time based CTM, and CTM with Modified Demand function for various representative cycle lengths. The green ratio used in each of the setup is 0.5 with green time rounded down to integer if needed.



(a) Traditional Model

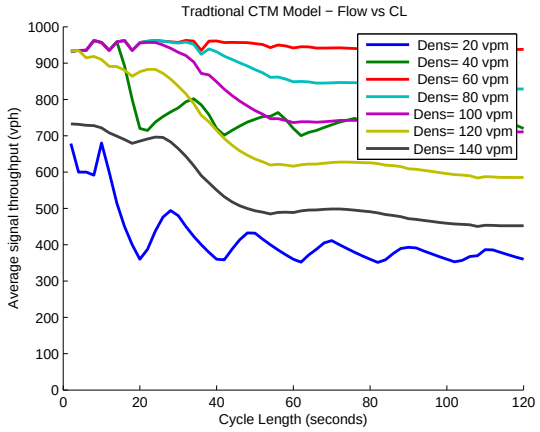


(b) Lost Time Model

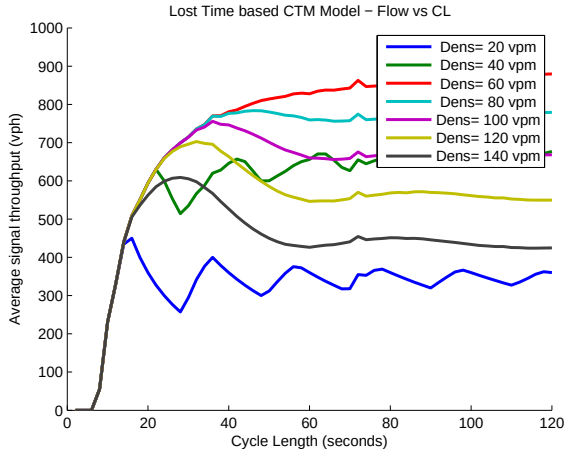


(c) Modified Demand CTM

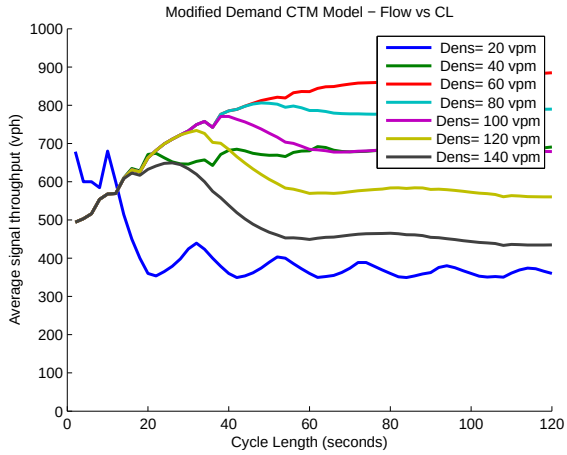
Figure 6.4: MFDs with varying Cycle Lengths for each of the model showing the range of operating conditions that can be obtained by controlling the Cycle Length.



(a) Traditional Model



(b) Lost Time Model



(c) Modified Demand CTM

Figure 6.5: Average flow rate vs. cycle length for various models. The graphs show the impact of cycle length on the aggregate flow rate for varying densities of traffic.

Chapter 7

Concluding Remarks

This chapter presents a summary of the research and its contributions in the first section, followed by a discussion of future research inspired by the study in the second.

7.1 Summary

Through this research, a CTM framework is presented that can model bounded acceleration, as well as its effects on traffic features.

In Chapter 3 it was shown how the impacts of bounded acceleration can be modeled using a modified Cell Transmission Modeling framework. The modified CTM relies on a linearly decreasing congested demand function to capture the acceleration behavior of vehicles. As vehicles accelerate from a low speed to a higher speed, the demand generated by these vehicles is lower than the capacity flow rate, as the vehicles have to spend a finite time accelerating to the free-flow speeds. The model discussed reproduces known observations of discharge vehicle headways and flow-rates accurately. In doing so, the modified CTM is able to provide more a realistic traffic dynamics modeling technique than traditionally used models, while

also offering analytical and computational tractability. Calibration of the model against past headway studies, and against recently collected datasets further validates the model.

In Chapter 4 the modified CTM was further extended to integrate a lane-changing model within the initial model. The integrated LCBA-CTM was shown to model the capacity drop phenomenon at a freeway lane-drop bottleneck and of predicting the capacity drop magnitude. Further, traffic dynamics, such as the onset of congestion, can be modeled realistically through the model. Calibrating the model with representative parameter values showed that the model realistic capacity drop magnitudes in the range of values often reported in studies.

In Chapter 5 a framework was presented allowing demand functions to be derived directly from microscopic acceleration profiles. Chapter 3 and Chapter 4 set up a modeling framework that uses a modified demand function within a CTM framework as a representation of the microscopic vehicle acceleration to model the effects of this acceleration on macroscopic traffic properties. Chapter 5 comes back to connect the acceleration process more directly with the demand function used earlier, thus completing the bridge between acceleration and macroscopic traffic dynamics. While much of the previous work on bounded acceleration has assumed a constant bounded acceleration rate, the present study allows for other acceleration models to be used.

Finally, in Chapter 6, two applications of the framework presented are explored.
First, the impact of a singular aspect, improved acceleration rates, of improved vehicle technologies on roadway network performance is explored. Intersection performance in terms of available capacity usage is compared for two scenarios: traffic stream consisting of only conventional vehicles with representative acceleration rates; and traffic stream composed of enhanced vehicles with increased acceleration rates. If the magnitude of improvement of the vehicle acceleration is known, the presented model can be used as a predictive modeling framework for estimate network performances.

As a second application, Macroscopic Fundamental Diagrams (MFDs) are derived for signalized ring roads comparing the aggregate traffic flow behavior as predicted by traditional CTM, a lost time based CTM, and the modified CTM. The study highlights how traditional CTM's inability to account for the acceleration process can lead to over-estimation of network flows and capacity. While the lost time model predicts accurate behavior for most scenarios, as would be expected, it is inaccurate when the available green time is low. The modified CTM framework improves on the lost time CTM's accuracy under such conditions.

The modeling framework presented in this research provides an innovative near first order macroscopic model that incorporates the effects of bounded acceleration and lane-changing on traffic flow. In doing so, the framework provides perhaps the first complete CTM based model where traffic dynamics can be explained at bottlenecks. The result is an accurate modeling of capacity, including the reduction in capacity observed at signalized intersection, and the capacity drop observed at active freeway bottlenecks.

The framework presented in this study can be very useful towards assessing roadway system performance as well as towards designing control strategies around the improvement of traffic flow at bottlenecks. The analytical tractability lends the model to application in performance optimization setups for design of control strategies, while the lower computational complexity compared to alternative higher order or hybrid models allows for faster modeling of traffic.

7.2 Discussions and Future Research

The application of the model to the signalized intersection setup provides interesting commentary on the impacts of acceleration on capacity. Connecting the microscopic acceleration and driver response behaviors to the macroscopic traffic flow features introduces a new perspective on the startup lost time at signals. While traditionally this lost time is defined

purely at a microscopic level, as a summation of headways from individual vehicles, the application of the lost time in signal design is of a macroscopic nature. The presented study offers a more rigorous definition to the lost time while also offering a means of estimating this lost time from a purely macroscopic perspective.

The modified CTM also re-affirms the benefits of the simple lost time based CTM. At an aggregate level, the lost time model is sufficiently accurate under most circumstances as the average capacity over a green phase is equivalent to that obtained from the modified CTM. The modified CTM however finds its own forte in modeling the evolution of this capacity over time, and can model traffic dynamics more accurately. As stated when deriving the MFD for both the models, the lost time model might also be inadequate where the green time too short, such as when a minor street shares a smaller ratio of the green time at an intersection with a major street. Similarly, the lost time model might also underestimate capacity corresponding to low traffic demands at intersections served by an actuated signal. A low demand at actuated signals results in shorter green times, once again illustrating the weakness of the lost time model.

An interesting aspect of the modified CTM is the importance of picking a reasonable cell length and time-step. While calibration of the model against headway data from various sources show that the values picked should be in the order of roughly 1-2 second time steps, the underlying reason behind this dependence is unclear. The dependence is further highlighted in the fact that as the cell size and time-step size is reduced, the resulting flow rate and headway curves over time become flatter such that the continuous approximation of the modified CTM is equivalent to traditional CTM and allows for unbounded acceleration. Since this time step determines the discretized duration between the flow-rate and the density getting updated, and speeds of the vehicles, it would not be surprising if there was a relation between the time-step and reaction time of drivers. However, further investigation would be needed to make such a connection conclusively.

Some of the more obvious extensions to the modified CTM would be to calibrate the model across a range of intersection and determine the robustness of the model parameters. The acceleration process and reaction times, and thus the demand function, might also depend on interactions between lanes and lane-groups for intersections with multiple lanes. Further, vehicles turning left or right at the intersection might have varying acceleration rates compared to through moving traffic. Future studies could investigate just aspects of the model through more elaborate calibration efforts across a variety of intersections.

The LCBA-CTM model, and its application to the freeway bottleneck suggests that the complete framework can support modeling capacity at freeway bottlenecks, including the capacity drop phenomenon. Indeed, the signalized intersection application of the modified CTM can be considered a special case of the LCBA-CTM where lane-changing factor is set to 1. Thus, the LCBA-CTM framework would be able to model traffic dynamics for both freeway bottlenecks as well as arterial intersections.

There are some simplifying assumptions made in the model. The first main assumption, pertaining to the lane-drop bottleneck, is that the lane-drop is sudden and does not have a lateral dimension. A geometrically reasonable model would instead treat the lane drop as a gradual decrease in lanes steadily over a stretch of the segment. A related assumption that was made was that the lane-changing region is approximated to be within a single CTM cell. Extending the present model to a relaxation of the lane-drop geometry and the lane-changing extent assumption could be interesting future research topics. Such a study might also consider the existence of an explicit upstream acceleration zone extending for the length of the lane-changing region.

Future studies could also explore the properties of the lane-drop bottleneck relevant to the model from an empirical perspective. Among aspects to be explored, would be the true length of the lane-changing region and how it relates to the distance over which the lane-drop lanes gradually are dropped. As was hinted earlier, a similar study might be done on

the extend of the upstream acceleration zone. Further, it would be an interesting study to explore the lateral and longitudinal distribution of the lane-changes within the region, and how to model them.

This study also assumes the existence of a constant lane-changing factor for the lane-changing region. A study aimed at relaxing this assumption and identifying the parameters that affect the lane-changing factor would greatly improve the overall understanding of the mechanics of capacity drop at lane-drop bottlenecks while also potentially shedding more light on the stochastic nature of capacity. A somewhat related extension to this study could also compare the mechanics of lane-changing between lane-drop bottlenecks and merge bottlenecks. Adjust the lane-changing aggression appropriately, the framework could be extended to also cover merge bottlenecks through the addition of a node merge model.

The framework presented in Chapter 5 connects the acceleration profile of vehicles to the demand function. A unique speed-acceleration relationship is assumed for this framework. While this is a reasonable assumption considering most popular acceleration models make such an assumption, it is indeed not realistic. Further, the acceleration models explored within the section are simplistic and meant to demonstrate how the framework can be used. The technique can, and should be extended to incorporate more realistic acceleration profiles. In order to obtain the true demand function, additional constraints should be imposed while deriving the demand, such that the demand never exceeds the capacity, and the vehicle speeds never exceed the free-flow speed during the acceleration process within a cell.

The two setups presented, assume that the information triggering the acceleration process propagates either instantaneously, or at the shockwave speed. While this might be a reasonable approximation, the actual acceleration is a complicated process that involves not just the reaction times for drivers, but also aspects such as anticipation of the acceleration. The reaction times themselves might not be identical for vehicles. As an example, for the case of vehicles discharge from a standing queue at an intersection, the first vehicle reacts to the

stimulus of the change of the signal, while the second vehicle reacts not only to the signal itself, but also to its leading vehicle.

Bibliography

Ahmed, K., Ben-Akiva, M., Koutsopoulos, H., Mishalani, R., 1996. Models of freeway lane changing and gap acceptance behavior. Transportation and traffic theory 13, 501–515.

Akçelik, R., Besley, M., 2002. Queue discharge flow and speed models for signalised intersections. In: Transportation and Traffic Theory in the 21st Century, Proceedings of the 15th International Symposium on Transportation and Traffic Theory. pp. 99–118.

Akçelik, R., Besley, M., Roper, R., 1999. Funaamental relationships for traffic flows at signalised intersections. Research Report ARR 340. ARRB Transport Research Ltd., Vermont South, Australia.

Akçelik, R., Biggs, D., 1987. Acceleration profile models for vehicles in road traffic. Transportation Science 21 (1), 36–54.

Al-Ghamdi, A. S., 1999. Entering headway for through movements at urban signalized intersections. Transportation Research Record: Journal of the Transportation Research Board 1678 (1), 42–47.

Ansorge, R., 1990. What does the entropy condition mean in traffic flow theory? Transportation Research Part B: Methodological 24 (2), 133–143.

Ardekani, S., Herman, R., 1987. Urban network-wide traffic variables and their relations. Transportation Science 21 (1), 1–16.

Aw, A., Rascle, M., 2000. Resurrection of" second order" models of traffic flow. SIAM journal on applied mathematics 60 (3), 916–938.

Bank, J., 1991. Two-capacity phenomenon at freeway bottlenecks: a basis for ramp metering. Transportation Research Record 1320, 64–69.

Bham, G. H., Benekohal, R. F., 2004. A high fidelity traffic simulation model based on cellular automata and car-following concepts. Transportation Research Part C: Emerging Technologies 12 (1), 1–32.

Bloomberg, L., Dale, J., 2000. Comparison of vissim and corsim traffic simulation models on a congested network. Transportation Research Record: Journal of the Transportation Research Board 1727 (1), 52–60.

Cassidy, M., Jang, K., Daganzo, C., 2011. Macroscopic fundamental diagrams for freeway networks: theory and observation. Transportation Research Record: Journal of the Transportation Research Board (2260), 8–15.

Cassidy, M. J., Bertini, R. L., 1999. Some traffic features at freeway bottlenecks. Transportation Research Part B: Methodological 33 (1), 25–42.

Cassidy, M. J., Rudjanakanoknad, J., 2005. Increasing the capacity of an isolated merge by metering its on-ramp. Transportation Research Part B: Methodological 39 (10), 896–913.

Ceder, A., May, A. D., 1976. Further evaluation of single-and two-regime traffic flow models. Transportation Research Record (567).

Chen, L., Jin, W.-L., Hu, J., Zhang, Y., 2008. An urban intersection model based on multi-commodity kinematic wave theories. In: Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on. IEEE, pp. 269–274.

Choudhury, C. F., Ben-Akiva, M., Toledo, T., Lee, G., Rao, A., 2007. Modeling cooperative lane changing and forced merging behavior. In: 86th Annual Meeting of the Transportation Research Board, Washington, DC.

Chung, K., Rudjanakanoknad, J., Cassidy, M. J., 2007. Relation between traffic density and capacity drop at three freeway bottlenecks. Transportation Research Part B: Methodological 41 (1), 82–95.

Courant, R., Friedrichs, K., Lewy, H., 1967 [orig.: 1928]. On the partial difference equations of mathematical physics. IBM journal of Research and Development 11 (2), 215–234.

Daganzo, C. F., 1995a. The cell transmission model, part ii: network traffic. Transportation Research Part B: Methodological 29 (2), 79–93.

Daganzo, C. F., 1995b. Requiem for second-order fluid approximations of traffic flow. Transportation Research Part B: Methodological 29 (4), 277–286.

Daganzo, C. F., 2002. A behavioral theory of multi-lane traffic flow. part i: Long homogeneous freeway sections. Transportation Research Part B: Methodological 36 (2), 131–158.

Daganzo, C. F., Geroliminis, N., 2008. An analytical approximation for the macroscopic fundamental diagram of urban traffic. Transportation Research Part B: Methodological 42 (9), 771–781.

Dion, F., Rakha, H., Kang, Y.-S., 2004. Comparison of delay estimates at under-saturated and over-saturated pre-timed signalized intersections. Transportation Research Part B: Methodological 38 (2), 99–122.

Drake, J. S., Schofer, J. L., May Jr, A. D., 1967. A statistical analysis of speed-density hypotheses. in vehicular traffic science. Highway Research Record (154).

FHWA, 2007. Ngsim - next generation simulation.
URL `http://www.ngsim.fhwa.dot.gov`

Gazis, D. C., Herman, R., Rothery, R. W., 1961. Nonlinear follow-the-leader models of traffic flow. Operations research 9 (4), 545–567.

Gazis, D. C., Herman, R., Weiss, G. H., 1962. Density oscillations between lanes of a multi-lane highway. Operations Research 10 (5), 658–667.

Gerlough, D. L., Wagner, F. A., 1967. Improved criteria for traffic signals at individual intersections. Tech. rep.

Geroliminis, N., Daganzo, C. F., 2008. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. Transportation Research Part B: Methodological 42 (9), 759–770.

Geroliminis, N., Skabardonis, A., 2005. Prediction of arrival profiles and queue lengths along signalized arterials by using a markov decision process. Transportation Research Record: Journal of the Transportation Research Board 1934 (1), 116–124.

Gipps, P. G., 1981. A behavioural car-following model for computer simulation. Transportation Research Part B: Methodological 15 (2), 105–111.

Gipps, P. G., 1986. A model for the structure of lane-changing decisions. Transportation Research Part B: Methodological 20 (5), 403–414.

Godfrey, J., 1900. The mechanism of a road network. Traffic Engineering & Control 8 (8).

Godunov, S. K., 1959. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. Matematicheskii Sbornik 89 (3), 271–306.

Greenberg, H., 1959. An analysis of traffic flow. Operations research 7 (1), 79–85.

Greenshields, B. D., Bibbins, J., Channing, W., Miller, H., 1935. A study of traffic capacity. In: Highway research board proceedings. Vol. 14.

Greenshields, B. D., Schapiro, D., Ericksen, E. L., 1946. Traffic performance at urban street intersections. Tech. rep.

Hall, F. L., Agyemang-Duah, K., 1991. Freeway capacity drop and the definition of capacity. Transportation Research Record (1320), 91–98.

Helbing, D., 2001. Traffic and related self-driven many-particle systems. Reviews of modern physics 73 (4), 1067.

Helbing, D., Johansson, A., 2009. On the controversy around daganzos requiem for and awrascles resurrection of second-order traffic flow models. The European Physical Journal B 69 (4), 549–562.

Holland, E. N., Woods, A. W., 1997. A continuum model for the dispersion of traffic on two-lane roads. Transportation Research Part B: Methodological 31 (6), 473–485.

Hung, W., Tian, F., Tong, H., 2002. Departure headways at one signalized junction in hong kong. In: Proceedings of Better Air Quality: Tales of Pacific Rim Megacities Workshop, Hong Kong.

Jin, W.-L., 2010a. A kinematic wave theory of lane-changing traffic flow. Transportation research part B: methodological 44 (8), 1001–1021.

Jin, W.-L., 2010b. Macroscopic characteristics of lane-changing traffic. Transportation Research Record: Journal of the Transportation Research Board 2188 (1), 55–63.

Jin, W.-L., 2013. A multi-commodity lighthill–whitham–richards model of lane-changing traffic flow. Transportation Research Part B: Methodological 57, 361–377.

Jin, W.-L., Gan, Q.-J., Gayah, V. V., 2013. A kinematic wave approach to traffic statics and dynamics in a double-ring network. Transportation Research Part B: Methodological 57, 114–131.

Jin, W.-L., Yu, Y., 2015. Performance analysis and signal design for a stationary signalized ring road. arXiv preprint arXiv:1510.01216.

Kesting, A., Treiber, M., Helbing, D., 2007. General lane-changing model mobil for car-following models. Transportation Research Record: Journal of the Transportation Research Board 1999 (1), 86–94.

King, G. F., Wilkinson, M., 1977. Relationship of signal design to discharge headway, approach capacity, and delay.

Kunzman, W., 1978. Another look at signalized intersection capacity. ITE journal 48 (HS-024 726).

Laval, J., Cassidy, M., Daganzo, C., 2007. Impacts of lane changes at merge bottlenecks: a theory and strategies to maximize capacity. In: Traffic and Granular Flow05. Springer, pp. 577–586.

Laval, J. A., 2004. Hybrid models of traffic flow: impacts of bounded vehicle accelerations. Ph.D. thesis, University of California, Berkeley.

Laval, J. A., Daganzo, C. F., 2006. Lane-changing in traffic streams. Transportation Research Part B: Methodological 40 (3), 251–264.

Laval, J. A., Leclercq, L., 2008. Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model. Transportation Research Part B: Methodological 42 (6), 511–522.

Lax, P. D., 1954. Weak solutions of nonlinear hyperbolic equations and their numerical computation. Communications on pure and applied mathematics 7 (1), 159–193.

Lebacque, J., 2002. A two phase extension of the lwr model based on the boundedness of traffic acceleration. In: Transportation and traffic theory in the 21st century. Proceedings of the 15th international symposium on transportation and traffic theory. pp. 697–718.

Lebacque, J.-P., 1984. Semi-macroscopic simulation of urban traffic. Proceedings of the International ASME Conference Modeling and Simulation 4, 273–292.

Lebacque, J.-P., 1996. The godunov scheme and what it means for first order traffic flow models. In: Internaional symposium on transportation and traffic theory. pp. 647–677.

Lebacque, J.-P., 1997. A finite acceleration scheme for first order macroscopic traffic flow models. The 8th IFAC Symposium.

Lebacque, J.-P., 2003. Two-phase bounded-acceleration traffic flow model: analytical solutions and applications. Transportation Research Record: Journal of the Transportation Research Board 1852 (1), 220–230.

Leclercq, L., 2002. Modélisation dynamique du trafic et applications à l'estimation du bruit routier. Ph.D. thesis, Villeurbanne, INSA.

Leclercq, L., 2007a. Bounded acceleration close to fixed and moving bottlenecks. Transportation Research Part B: Methodological 41 (3), 309–319.

Leclercq, L., 2007b. A new numerical scheme for bounding acceleration in the lwr model. In: 4th IMA International Conference on Mathematics in Transport.

Lee, J. J., Chen, R., 1986. Entering headway at signalized intersections in a small metropolitan area. Transportation Research Record (1091), 117–126.

Li, L., Wang, F.-Y., 2006. Approximate vehicle waiting time estimation using adaptive video-based vehicle tracking. In: Advances in Machine Vision, Image Processing, and Pattern Analysis. Springer, pp. 105–114.

Lighthill, M. J., Whitham, G. B., 1955. 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 229 (1178), 317–345.

Lo, H., 1999. A dynamic traffic assignment formulation that encapsulates the cell-transmission model. In: 14th International Symposium on Transportation and Traffic Theory.

Mahmassani, H., Williams, J. C., Herman, R., 1987. Performance of urban traffic networks. In: Transportation and Traffic Theory (Proceedings of the Tenth International on Transportation and Traffic Theory Symposium, Cambridge, Massachusetts), NH Gartner, NHM Wilson, editors, Elsevier.

Michalopoulos, P. G., Beskos, D. E., Yamauchi, Y., 1984. Multilane traffic flow dynamics: some macroscopic considerations. Transportation Research Part B: Methodological 18 (4), 377–395.

Michalopoulos, P. G., Stephanopoulos, G., Stephanopoulos, G., 1981. An application of shock wave theory to traffic signal control. Transportation Research Part B: Methodological 15 (1), 35–51.

Monamy, T., Haj-Salem, H., Lebacque, J.-P., 2012. A macroscopic node model related to capacity drop. Procedia-Social and Behavioral Sciences 54, 1388–1396.

Moussavi, M., Tarawneh, M., 1990. Variability of departure headways at signalized intersections. ITE 1990 Compendium of Technical Papers, 313–317.

Munjal, P., Hsu, Y.-S., Lawrence, R., 1971. Analysis and validation of lane-drop effects on multi-lane freeways. Transportation Research 5 (4), 257–266.

Munjal, P., Pipes, L. A., 1971a. Propagation of on-ramp density perturbations on unidirectional two-and three-lane freeways. Transportation Research 5 (4), 241–255.

Munjal, P., Pipes, L. A., 1971b. Propagation of on-ramp density waves on uniform unidirectional multilane freeways. Transportation Science 5 (4), 390–402.

Nelder, J. A., Mead, R., 1965. A simplex method for function minimization. The computer journal 7 (4), 308–313.

Niittymäki, J., Pursula, M., 1997. Saturation flows at signal-group-controlled traffic signals. Transportation Research Record: Journal of the Transportation Research Board 1572 (1), 24–32.

Pacey, G., 1956. The progress of a bunch of vehicles released from a traffic signal. Road Research Laboratory Note RN/2665/GMP.

Papageorgiou, M., 1998. Some remarks on macroscopic traffic flow modelling. Transportation Research Part A: Policy and Practice 32 (5), 323–329.

Payne, H. J., 1971. Models of freeway traffic and control. Mathematical models of public systems.

Pipes, L., 1967. Car following models and the fundamental diagram of road traffic. Transportation research 1 (1), 21–29.

Rakha, H., Snare, M., Dion, F., 2004. Vehicle dynamics model for estimating maximum light-duty vehicle acceleration levels. Transportation Research Record: Journal of the Transportation Research Board 1883 (1), 40–49.

Richards, P. I., 1956. Shock waves on the highway. Operations research 4 (1), 42–51.

Robertson, D. I., 1969. Transyt: a traffic network study tool.

Roess, R., Prassas, E., McShane, W., 2010. Traffic engineering. Prentice Hall.

Searle, J., 1999. Equations for speed, time and distance for vehicles under maximum acceleration. Tech. rep., SAE Technical Paper.

Srivastava, A., Geroliminis, N., 2013. Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. Transportation Research Part C: Emerging Technologies 30, 161–177.

Stephanopoulos, G., Michalopoulos, P. G., Stephanopoulos, G., 1979. Modelling and analysis of traffic queue dynamics at signalized intersections. Transportation Research Part A: General 13 (5), 295–307.

Tian, Z. Z., Urbanik, T., Engelbrecht, R., Balke, K., 2002. Variations in capacity and delay estimates from microscopic traffic simulation models. Transportation Research Record: Journal of the Transportation Research Board 1802 (1), 23–31.

Toledo, T., Zohar, D., 2007. Modeling duration of lane changes. Transportation Research Record: Journal of the Transportation Research Board 1999 (1), 71–78.

Tong, H., Hung, W., 2002. Neural network modeling of vehicle discharge headway at signalized intersection: model descriptions and results. Transportation Research Part A: Policy and Practice 36 (1), 17–40.

Underwood, R., 2008. Speed, volume, and density relationship: Quality and theory of traffic flow, yale bureau of highway traffic (1961) 141–188. New Haven, Connecticut.

Webster, F. V., 1958. Traffic signal settings. Tech. rep.

Whitham, G. B., 2011. Linear and nonlinear waves. Vol. 42. John Wiley & Sons.

Zhang, H., 2003. On the consistency of a class of traffic flow models. Transportation Research Part B: Methodological 37 (1), 101–105.