



THE HONG KONG  
POLYTECHNIC UNIVERSITY

香港理工大學

Pao Yue-kong Library

包玉剛圖書館

---

## Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

**By reading and using the thesis, the reader understands and agrees to the following terms:**

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

### IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact [lbsys@polyu.edu.hk](mailto:lbsys@polyu.edu.hk) providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

**MULTICLASS MULTILANE TRAFFIC MODELS  
AND  
INTEGRATED OPTIMAL CONTROL STRATEGIES  
FOR FREEWAYS**

TIANLU PAN

PhD

The Hong Kong  
Polytechnic University  
2018



The Hong Kong Polytechnic University  
Department of Civil and Environmental Engineering

**MULTICLASS MULTILANE TRAFFIC MODELS**  
**AND**  
**INTEGRATED OPTIMAL CONTROL STRATEGIES**  
**FOR FREEWAYS**

Tianlu PAN

A thesis submitted in partial fulfillment of  
the requirements for the degree of Doctor of  
Philosophy

June 2017

## **Certificate of originality**

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

Signature:

Name of the student: PAN Tianlu

## **Abstract**

Active Traffic Management (ATM) relies on a fast and reliable traffic simulator, for the rapid quantitative assessment of various control strategies under different traffic scenarios, such as traffic control for congested weaving areas and traffic incidents. The single-lane macroscopic traffic models which have commonly been applied for ATM purposes may fail to capture the complex traffic features on multiple lane roadways. These complex traffic features may consist of heterogeneous traffic flow distribution, capacity drop, and moving bottlenecks at different sections of the highway.

Recently, research has revealed that vehicle lane-changing (LC) has significant impacts on road traffic safety since accidents often happen at lane-changing areas such as weaving sections and interchanges of the expressway. In view of this, there is a need to develop a comprehensive traffic model that can capture the effects of different lane-changing maneuvers explicitly on the surrounding traffic. One of the main objectives of this dissertation is to develop multilane traffic flow models to facilitate real-time simulation for various active traffic management applications.

In the literature, the kinematic wave (KW) based approaches have commonly been used for simulating lane-changing maneuvers typically differentiate between mandatory lane-changing (MLC) and discretionary lane-changing (DLC). However, these two lane-changing behaviors were separately investigated by the existing traffic models. In view of this, a novel macroscopic multilane traffic model is firstly proposed in this dissertation to enable simultaneous simulation of MLC and DLC behaviors to capture the multilane traffic dynamics on freeway. A salient feature is that the proposed model does not require extensive traffic data collected by expensive infrastructure but only relies on the traffic data available to most of the traffic management centers. Such a parsimonious data requirement is a significant improvement over the existing traffic models. Other key features of the proposed model are: 1) incorporating the lane-based

fundamental diagrams to encapsulate the relationship between traffic speed-density and lane usage; and 2) modeling the drivers how to perceive the traffic condition spatially ahead to make their lane-changing decisions. To the knowledge of the author, these important features investigated in this dissertation for modeling vehicle lane-changing behaviors have not yet been reported in the literature.

Apart from the above features, freeways are always subject to traffic demand and supply uncertainties, and noisy traffic data. To model the effects of these stochastic elements, a multilane traffic flow model is developed in this dissertation by extending the stochastic cell transmission model (SCTM) to simulate the effects of vehicle lane-changing maneuvers on freeway traffic dynamics. Link (cell)-node junction formulation is developed to propagate the lane-changing traffic. A fundamental speed-density relationship is used to interpolate the cell-lane speed profiles along a freeway corridor with sparse detectors. To the best of the author's knowledge, this is the first macroscopic stochastic multilane traffic model in the literature.

The proposed models in this dissertation can be deployed as useful simulation tools for assessing the dynamic multilane traffic state based on the data available to the traffic management centers in practice. Furthermore, the model has the potential applications for predicting the impacts of various traffic incidents or lane control strategy on the expressway.

In view of the advantages and the promising market prospect of the emerging connected automated vehicles (CAVs), the number of CAVs will be increased rapidly in the coming decade. Meanwhile, the regular human-piloted vehicles (RHVs) may still play a significant role in the roadway traffic. Therefore, it will be very likely that the roadway is to be shared by CAVs and RHVs in the near future.

In the second part of the dissertation, an integrated optimal freeway traffic flow control framework that aims to minimize the total travel cost is devised for freeway traffic

mixed with a given penetration rate of CAVs equipped with the Vehicle Automation and Communication Systems (VACS) and RHVs via en-route Variable Message Signs (VMS). It is assumed that the CAVs would follow full compliance with the control commands through the VACS. In contrast, the drivers of RHVs would make decisions in response to the information disseminated by the en-route Variable Message Signs (VMS). At the upper level of the integrated control framework, the objective is to devise an integrated action of several control strategies such as variable speed limit control (VSLC) and recommendation (VSLR), lane changing control (LCC) and recommendation (VSLR) under various traffic conditions.

At the lower level of the integrated control framework, a multiclass multilane cell transmission model is developed to simulate the traffic flow dynamics mixed with CAVs and RHVs. The impacts of penetrated CAVs on the freeway traffic characteristics and the lane-changing behaviors are captured to design the optimal traffic control strategies. Firstly, the variations in the fundamental diagrams with respect to different penetration rates of CAVs are quantified. Then, the minimum headway acceptance criteria are determined for the lane changing (LC) maneuvers proposed by CAVs and RHVs with different motivations, respectively. An advanced priority incremental transfer (PIT) principle is adopted to evaluate the sending flows. Finally, the cell-lane-specific multiclass flow conservation law is developed to propagate the traffic flow and density on the freeway section.

The effectiveness and the computational feasibility of the proposed optimal control strategies are illustrated via numerical example for a variety of penetration rates of CAVs under various traffic conditions. It is shown that the integrated control strategies can reduce the number of vehicles queuing at the bottleneck, improve the traffic efficiency and alleviate capacity drop. For road traffic safety, the integration of optimal control strategies can drastically reduce the instances of the stop and go traffic, smoothen the traffic flow and suppress the impact of the shockwaves on the freeway sections concerned.

## List of publications

### Journal papers:

**Pan, T.**, Lam, W.H.K., Sumalee, A., and Zhong, R., 2016. Modelling the impacts of mandatory and discretionary lane-changing maneuvers. *Transportation Research Part C*, 68, 403-424.

**Pan, T.**, Sumalee, A., Zhong, R., and Indra-payoong, N., 2013. Short-term traffic state prediction based on temporal-spatial correlation. *IEEE Transactions on Intelligent Transportation Systems*, 14 (3), 1242 – 1254.

Zhong, R., Yuan, F., **Pan, T.**, Chow, A., Chen, C., and Yang, Z., 2016. Linear complementarity system approach to macroscopic freeway traffic modeling: Uniqueness and convexity. *Transportmetrica A: Transport Science*, 12 (2), 142–174.

Zhong, R., Chen, C., Chow, A., **Pan, T.**, Yuan, F., and He, Z., 2016. Automatic calibration of fundamental diagram for first-order macroscopic freeway traffic models. *Journal of Advanced Transportation*, 50 (3), 363–385.

Zhong, R., Sumalee, A., **Pan, T.**, and Lam, W.H.K., 2014. Optimal and robust strategies for traffic management under demand and supply uncertainties: An overview and general theory. *Transportmetrica*, 10 (10), 849-877.

Sumalee, A., **Pan, T.**, Zhong, R., and Uno, N., 2013. Dynamic stochastic journey time estimation and reliability analysis using stochastic cell transmission model: Algorithm and case studies. *Transportation Research Part C*, 35, 263–285.

Zhong, R., Sumalee, A., **Pan, T.**, and Lam, W.H.K., 2013. Stochastic cell transmission model for traffic network with demand and supply uncertainties. *Transportmetrica*, 9 (7), 567-602.

Sumalee, A., Zhong, R., **Pan, T.**, and Szeto, W., 2011. Stochastic cell transmission model (SCTM): A stochastic dynamic traffic model for traffic state surveillance and assignment. *Transportation Research Part B*, 45 (3), 507-533.

# Acknowledgments

First of all, I would like to express my sincere gratitude to my supervisors: Prof. William H.K. LAM and Prof. Agachai SUMALEE for their kind guidance and supervision. I have gained a lot of knowledge from them that are applicable for my future career. Their valuable suggestions and criticisms have greatly enhanced the quality of this research and dissertation.

Many past and present members in transportation group of the Hong Kong Polytechnic University deserve my thanks for their friendship and technical expertise, including Dr. Renxin ZHONG, Dr. Karen M.L. TAM, Dr. Xiao FU, Miss. Piyanit WEPULANON, Mr. Ekalux UA-AREEMITR, Mr. Kaiyan FU and Mr. Junbiao SU. I would also like to thank them for always being there to support me during my study as well as to share the great moments in the last five years with so many unforgettable memories.

Finally, my love and thanks go to my husband Dr. Renxin ZHONG, my lovely daughters, my parents and my mother-in-law. Mere words, in any of the languages that I know, are inadequate to express my gratitude for the unconditional love and support they have given to me.

Lucia Tianlu PAN

# Contents

Abstract.....	5
Chapter 1 Introduction.....	17
1.1 Background.....	17
1.1.1 Freeway lane changing maneuvers.....	18
1.1.2 Vehicle automation and communication systems (VACS).....	21
1.1.3 Traffic flow modeling and control in the presence of mixed RHVs and CAVs.....	24
1.2 Research objectives.....	26
1.3 Dissertation organization and contributions.....	29
Chapter 2 Literature review.....	33
2.1 Modeling freeway lane-changing (LC) maneuvers.....	33
2.2 The cell transmission model (CTM).....	40
2.3 Randomness of the fundamental diagram.....	43
2.4 The stochastic cell transmission model (SCTM).....	44
2.5 Traffic flow modeling of mixed RHVs and CAVs.....	51
2.6 Freeway traffic control for traffic flow mixed with RHVs and CAVs.....	54
2.7. Summary.....	58
Chapter 3 Modeling the impacts of mandatory and discretionary lane-changing maneuvers.....	<b>60</b>
3.1 Introduction.....	60
3.2 Development of the macroscopic multilane cell transmission model.....	61
3.2.1 Lane-specific traffic flow characteristics.....	61
3.2.2 Determination of minimum gap acceptance criterion.....	64
3.2.3 Sending function and lane-changing fraction.....	68
3.2.4 Allocation of spaces for merging flows and dynamics propagation.....	70
3.3 The lane-changing demand estimation algorithm.....	74
3.3.1 Estimation of the initial value of the total MLC demand.....	75
3.3.2 Longitudinal distribution of MLC and DLC demands.....	76
3.3.3 The dynamical adjustment process of lane-changing demand.....	79
3.4 Empirical study.....	81
3.4.1 Description of the test site.....	82
3.4.2 Model calibration and estimation of MLC demand.....	86
3.4.3 Simulation results against the measurement.....	89
3.5 Conclusions.....	98
Chapter 4 Stochastic multilane cell transmission model by assimilating lane speed observation.....	<b>101</b>
4.1 Introduction.....	101
4.2 Hybrid cell transmission model and its extension.....	102
4.3 Speed-density relationship for traffic speed data assimilation.....	106

4.4 Stochastic elements and lane-changing maneuvers .....	108
4.4.1. Lane-changing and junction models .....	108
4.4.2. Modeling lane-changing maneuvers under stochastic environment using junction models .....	109
4.5 Incremental-Transfer (IT) principle and the issue of priority .....	116
4.6 On-/off- ramps and capacity drop .....	117
4.7 An overview of the proposed SCTM with lane-changing maneuvers .....	121
4.8 Numerical tests.....	125
4.8.1 Scenario 1: A capacity drop on lane 2 .....	126
4.8.2 Scenario 2: A slow-moving vehicle on lane 2 .....	130
4.8.3 Scenario 3: A multilane freeway with an on-ramp .....	133
4.8.4 The effects of stochastic elements .....	138
4.9 Conclusions.....	144
Chapter 5 Optimal control strategies for freeway traffic mixed with regular human- piloted vehicles and connected automated vehicles.....	<b>146</b>
5.1. Introduction.....	146
5.2 System architecture .....	148
5.3 Multiclass traffic flow characteristics mixed with CAVs and RHVs.....	149
5.3.1 Headway distribution and fundamental diagram .....	150
5.3.2 The minimum gap acceptance criteria .....	159
5.4 Optimal control strategies .....	164
5.4.1 Optimal control problem formulation.....	164
5.4.2 Traffic-state propagation .....	170
5.4.3 Sending and receiving functions for multiclass multilane traffic flow propagation .....	175
5.4.4. Allocation of available space for merging flows .....	178
5.5 Solution algorithm based on cross-entropy method and control parameterization.....	180
5.5.1 Control parameterization .....	181
5.5.2 Recapitulation of the cross-entropy method for optimization .....	184
5.5.3 Optimal parameter search based on the CEM.....	188
5.6 Numerical example .....	194
5.6.1 Description of the “test site” .....	194
5.6.2 Simulation results.....	196
5.6.2.1 Optimal control for heavy traffic conditions.....	197
5.6.3 Sensitivity analysis.....	204
5.7 Conclusions.....	207
Chapter 6 Summary and recommendation.....	<b>209</b>
6.1 Summary of the key findings.....	209
6.2 Recommendation for future works .....	218
References.....	<b>225</b>

## Nomenclature

Key variables of the lane-changing models		
Elements	Units	Notation description
$v_{f,m}$	mile/hour	Free-flow speed of the fundamental diagram calibrated on lane $m$
$w_{c,m}$	mile/hour	Wave-back speed of congestion of the fundamental diagram calibrated on lane $m$
$\rho_{c,m}$	P.C.U./mile/lane	Critical density of the fundamental diagram calibrated on lane $m$
$\rho_{J,m}$	P.C.U./mile/lane	Jam-density of the fundamental diagram calibrated on lane $m$
$Q_m$	P.C.U./hour/lane	Traffic flow capacity of the fundamental diagram calibrated on lane $m$
$\alpha_0, \alpha_1, \alpha_2$	-----	Parameters of the cumulative distribution function of MLC demand on longitudinal dimension
$c_l, c_f,$	-----	The coefficient of speed difference for evaluating the extra leading gap and extra lag gap
$x_c, x_r$	mile	The remaining distance by which the test section is partitioned as remote, median, and close sections, respectively, with each section corresponds to a specific level of MLC urgency and minimum acceptance criterion length
$\bar{g}_{min}$	mile	The minimal safe gap for the subject vehicle to be provided by the target lane.
$\tau$	-----	The coefficient for evaluating DLC ratio
$\gamma$	-----	A prescribed threshold for the iteration algorithm of the MLC and DLC demand adjustment algorithm
$T_s$	hour	Time duration between two successive simulation steps
$T_c$	hour	Time duration between two successive control steps
$C$	mile	Minimum safe constant gap
$\rho_{i,m}(k)$	P.C.U./mile/lane	Traffic density of cell $(i, m)$ at time step $k$
$\Delta T_{CAV}$	hour	Response time of CAVs and RHVs
$\Delta T_{RHV}$	hour	

$v_{i,m}(k)$	mile/hour	Traffic speed of <i>cell</i> ( <i>i, m</i> ) at time step <i>k</i> .
$S_{i,m}^{i+1}(k)$	P.C.U./hour/lane	The sending function of flow that intends to leave <i>cell</i> ( <i>i, m</i> ) and towards downstream <i>cell package</i> ( <i>i+1</i> ) involving all possible lanes (such as neighboring lane <i>m-1, m+1</i> , and the current lane <i>m</i> ), during the time interval $[k \cdot T_s, (k+1) \cdot T_s)$
$R_{i-1}^{i,m}(k)$	P.C.U./hour/lane	The receiving function of the receiving <i>cell</i> ( <i>i, m</i> ) towards which traffic flows sent by the upstream <i>cell package</i> ( <i>i-1</i> ) are intended to merge during the time interval $[k \cdot T_s, (k+1) \cdot T_s)$
$U_{i-1}^{i,m}(k)$	P.C.U./hour/lane	The total sending function that determines the number of spaces required by flows that intend to merge towards <i>cell</i> ( <i>i, m</i> ) from upstream <i>cell package</i> ( <i>i-1</i> ) at time step during the time interval $[k \cdot T_s, (k+1) \cdot T_s)$
$S_{lc,i,m}^{i+1,\beta}(k)$	P.C.U./hour/lane	The sending function of lane-changing flow that intends to leave <i>cell</i> ( <i>i, m</i> ) and towards downstream <i>cell</i> ( <i>i+1, <math>\beta</math></i> ) during the time interval $[k \cdot T_s, (k+1) \cdot T_s)$ with $\beta=m \pm 1$ , and the lane-changing intention is of <i>lc</i> type, where <i>lc</i> =MLC, DLC represent the MLC and DLC, respectively
$S_{st,i,m}^{i+1,m}(k)$	P.C.U./hour/lane	The sending function of flow that intends to leave <i>cell</i> ( <i>i, m</i> ) and move ahead towards downstream <i>cell</i> ( <i>i+1, m</i> ) during the time interval $[k \cdot T_s, (k+1) \cdot T_s)$
$q_{lc,i-1,\alpha}^{i,m}(k)$	P.C.U./hour/lane	The lane-changing flow which left <i>cell</i> ( <i>i-1, <math>\alpha</math></i> ) and received by the downstream target <i>cell</i> ( <i>i, m</i> ) during the time interval $[k \cdot T_s, (k+1) \cdot T_s)$ with $\alpha=m \pm 1$ and the lane-changing intention is <i>lc</i> type
$q_{st,i-1,m}^{i,m}(k)$	P.C.U./hour/lane	The straightforward vehicle flow which left <i>cell</i> ( <i>i-1, m</i> ) and received by the downstream target <i>cell</i> ( <i>i, m</i> ) during the time interval $[k \cdot T_s, (k+1) \cdot T_s)$
$x(k)$	mile	Remaining distance from the current position of the subject vehicle to the target turning point
$\tilde{g}_{lc,i-1,\alpha}^{i,m}(k)$	mile	Minimum gap acceptance criterion required by sending flow sent from <i>cell</i> ( <i>i-1, <math>\alpha</math></i> ) towards the gap between successive vehicles on <i>cell</i> ( <i>i-1, m</i> ) during the time interval $[k \cdot T_s, (k+1) \cdot T_s)$ with $\alpha=m \pm 1$ , and the lane-changing intention is <i>lc</i> type
$\tilde{\phi}_{lc,i-1,\alpha}^{i,m}(k)$	-----	The minimum acceptance criterion factor normalized from $\tilde{g}_{lc,i-1,\alpha}^{i,m}(k)$

$\bar{G}_{i,m}(k)$	mile	The average space gap length between two successive vehicles on the cell $(i, m)$ at time step $k$
$\tilde{S}_{lc,m}^{tm(r)}(t)$	P.C.U./hour	The total lane changing demand from lane $m$ to terminal lane $tm$ at time $t$ with lane changing intention $lc$ type estimated at the $r^{\text{th}}$ round of iteration
$L_{MLC,m}^{\beta(r)}(t, x)$	P.C.U./hour	The longitudinal cumulative MLC demand to be sent from lane $m$ to target lane $\beta$ (with $\beta=m\pm 1$ ) at the location with a remaining distance $x$ to the target turning point at time $t$ estimated at the $r^{\text{th}}$ round of iteration
$\sigma_m^\beta(t)$	-----	The traffic state dependent parameter of the exponential distribution of $L_{M,m}^{\beta(r)}(t, x)$ .
$N_m^{tm}$	lane	The number of lanes to be crossed from current lane $m$ to the terminal target lane $tm$ .
$I_{MLC,m}^{\beta(r)}(k \cdot T_s, x_n - i \cdot l_i)$	P.C.U./hour	The cumulative MLC demand (originally proposed at time $k$ and at position $x_n$ ) executed at the downstream boundary of cell $i$ estimated at the $r^{\text{th}}$ round of iteration.

<b>Key variables of the integrated optimal control problem</b>		
Elements	Units	Notation description
$\rho_{i,x,CAV}(k)$	P.C.U./mile/lane	Density of CAVs on cell $(i, x)$ during time step $k$
$\rho_{i,x,RHV}(k)$	P.C.U./mile/lane	Density of RHVs on cell $(i, x)$ during time step $k$
$P_{i,x,CAV}(k)$	P.C.U./mile	Proportion/Penetration rate of CAVs on cell $(i, x)$ during time step $k$
$v_{i,x}(k)$	mile/hour	Traffic speed on cell $(i, x)$ during time step $k$
$W_{x,CAV}(k)$	mile/hour	Virtual extra-queue state variables at the upstream boundary of cell $(1, x)$ with respect to vehicle class CAV or RHV during time step $k$
$W_{x,RHV}(k)$	mile/hour	
$S_{i,x,CAV}^{i+1,\bar{x}}(k)$	P.C.U./hour	Sending function of CAVs issued by LCC from cell $(i, x)$ to cell $(i+1, \bar{x})$ during time step $k$
$S_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$	P.C.U./hour	Sending function of RHVs with MLC or DLC demand from cell $(i, x)$ to cell $(i+1, \bar{x})$ during time step $k$
$S_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$	P.C.U./hour	

$s_{st,i,x,CAV}(k)$	P.C.U./hour	Sending function of flow with vehicle class CAV or RHV that intend to leave cell $(i, x)$ straightforwardly and be received by downstream cell $(i+1, x)$ during time step $k$
$s_{st,i,x,RHV}(k)$	P.C.U./hour	
$\tilde{H}_{i,x,CAV}^{i+1,\bar{x}}(k)$	mile	Minimum space headway criteria of $s_{i,x,CAV}^{i+1,\bar{x}}(k)$
$\tilde{H}_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$	mile	Minimum space headway criteria of $s_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$ and $s_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$
$\tilde{H}_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$	mile	
$q_{i,x,CAV}^{i+1,\bar{x}}(k)$	P.C.U./hour	The amount of $s_{i,x,CAV}^{i+1,\bar{x}}(k)$ received by the target cell (or executed LCC)
$q_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$	P.C.U./hour	The amount of $s_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$ and $s_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$ received by the target cells, respectively
$q_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$	P.C.U./hour	
$q_{st,i,x,CAV}(k)$	P.C.U./hour	The amount of $s_{st,i,x,CAV}(k)$ and $s_{st,i,x,RHV}(k)$ received by the target cells
$q_{st,i,x,RHV}(k)$	P.C.U./hour	
$R_{i,x}(k)$	P.C.U./hour	Receiving function of cell $(i, x)$ during step $k$
$K_f K_e K_c$	----	Matrices for calculating fuel consumption, emission cost and electricity cost, respectively.
$\hat{v}_{i,x,l}(K)$	mile/hour	The implemented speed limit on cell $(i, x)$ during control cycle $K$
$\hat{p}_{i,x,CAV}^{i+1,\bar{x}}(K)$	-----	The initial ratio of CAVs that is issued to switch from cell $(i, x)$ to cell $(i+1, \bar{x})$ during control cycle $K$
$\hat{B}_{i,x}^{\bar{x}}(K)$	-----	Variable denoting whether the lane changing recommendation from cell $(i, x)$ to lane $\bar{x}$ would be released during control cycle $K$

Abbreviation list:	
DLC	Discretionary lane-changing
ISL	Implemented speed limit
IT	Incremental-Transfer
LFD	Lane flow distribution
MFD	Macroscopic fundamental diagram for freeway traffic
MLC	Mandatory lane-changing
P.C.U.	Passenger Car Unit
SSL	Spontaneous speed limit
VSL	Variable speed limit
CAV	Connected automated vehicle
RHV	Regular human-piloted vehicle
VACS	Vehicle automation and communication system

# **Chapter 1 Introduction**

This chapter will introduce the following in order of presentation: research background and problems, research objectives and scope, and an outline of the scientific and practical contributions, as well as the dissertation structure.

## **1.1 Background**

Traffic congestion, including traffic-induced air and noise pollutions, are ongoing problems in many densely populated cities worldwide, including Hong Kong and Beijing. Given the topography and urban development pattern of Hong Kong, any available spaces are scarcely used for further expansion of the existing road transport network. By contrast, active traffic management (ATM) (medium- and short-term) dynamically manages the recurrent and non-recurrent traffic congestions based on prevailing traffic conditions to maximize the effectiveness and efficiency of existing infrastructure systems.

ATM is preferable due to the relatively low cost and flexibility of implementation. ATM becomes effective by influencing travel and driving behaviors using route guidance, dynamic speed limits, and ramp metering. Together with the advances in computer science, ubiquitous mobile communications, recent developments of the Internet of things, and vehicle automation and communication systems (VACS) have enabled easier and cheaper collection, storage, analysis, usage, and dissemination of transportation big data to devise better traffic management strategies and subsequently alleviate traffic congestion.

### 1.1.1 Freeway lane changing maneuvers

As emphasized by Kurzhanskiy and Varaiya (2010), the ATM relies on a fast and trusted traffic simulator (with traffic models as core elements) for the rapid quantitative assessment of various control strategies under different traffic scenarios (e.g., traffic control for congested weaving areas and traffic incidents). Kurzhanskiy and Varaiya (2010) defined traffic flow model as dependable when the model is founded on the sound theory of traffic flow and tested for reliability; this model is parsimonious if it only includes parameters that can be estimated (from traffic data).

State-of-the-art traffic flow theories can be roughly categorized into macroscopic and microscopic traffic flow characteristics, whereas the multilane macroscopic traffic flow characteristics are not extremely well defined in the literature. Car following (CF) and lane changing (LC) are two primary driving behaviors. In the literature, CF and LC rules describe longitudinal and lateral interactions of vehicles, respectively. Describing the CF behavior is the vital component of microscopic traffic flow models which have been widely studied for many years. CF models contain several parameters and cannot provide intuitive descriptions of system-level effects of traffic flow. Moreover, calibration and validation require high-resolution vehicle trajectory data. Nevertheless, the CF models are extremely computationally expensive for large-scale network simulation to support ATM (Kurzhanskiy and Varaiya, 2010).

Macroscopic dynamical traffic models based on the kinematic wave (KW) theory are frequently incorporated as simulators for ATM applications. For example, the Lighthill-Whitham-Richards (LWR) model (Lighthill and Whitham, 1955; Richards, 1956), and its discretized version, the cell transmission model (CTM; Daganzo, 1994; Szeto, 2008; Sumalee et al., 2011), are the most widely used. However, these macroscopic models assume that traffic flow is uniformly distributed over lanes by integrating traffic streams that travel on different lanes into a single-flow stream with a uniformly lateral

distribution (Munjal and Pipes, 1971).

The uniform assumption may be inappropriate in the sense that heterogeneous traffic flow distribution, such as lane-specific flow, density, speed, and vehicle type, can be observed on multilane expressways (Carter et al., 1999; Cassidy and Rudjanakanoknad, 2005; Gunay, 2007; Duret et al., 2012). Therefore, single-lane traffic models may not capture other complex traffic features on multilane roadways (e.g., heterogeneous traffic flow distribution, capacity drop, and moving bottlenecks). Moreover, recent research has revealed that vehicle LC has a significant influence on traffic safety because accidents tend to happen in LC areas, such as weaving sections and interchanges (Golob et al., 2004; Cassidy and Rudjanakanoknad, 2005; Lee and Cassidy, 2009; Srivastava and Geroliminis, 2013). However, little attention has been given to modeling the DLC and MLC maneuvers simultaneously under the KW framework in conjunction with model calibration and validation from real-world data (Zheng, 2014).

To quantify complex features of multilane traffic, understanding LC is essential (Mauch and Cassidy, 2002; Ahn and Cassidy, 2007). According to different decision-making processes and their effects on surrounding traffic, LC maneuvers are usually classified as mandatory LC (MLC) and discretionary LC (DLC; Yang and Koutsopoulos, 1996). An MLC is executed when drivers must change lanes to follow a particular path to their destination. A DLC occurs when a driver seeks for better driving conditions to gain speed (or travel time) advantage. Existing modeling efforts of LC can be roughly classified into two themes, namely, modeling the decision-making process and modeling the influence of LC on surrounding traffic (Zheng, 2014). Although significant progress has been observed in modeling these two aspects of LC, a traffic modeling tool that fully describes LC is still lacking.

Remarkably, existing macroscopic approaches, especially the KW based ones, to simulate vehicle LC maneuvers typically **differentiate DLC and MLC** while **separately developing different models** for these two LC behaviors. For example,

some models that concentrate on DLC usually lack an MLC component (Laval and Daganzo, 2006) and vice versa (Hou et al., 2015). Hence, relating these models to each other or additionally providing **a comprehensive model to simulate LC maneuvers at one scale is difficult**. As reviewed by Zheng (2014), an emerging need arises in developing a comprehensive model that captures **the influence of MLC and DLC maneuvers on surrounding traffic** while maintaining the **balance between maximizing the predictive and explanatory powers** of the model and **minimizing its complexity** on account of ATM applications.

The fundamental diagram plays an essential role in macroscopic traffic flow modeling. Most of the existing models are deterministic, which essentially describe average system behaviors (Li et al., 2012, Wang et al., 2013, Wang et al., 2011). However, recent empirical and theoretical research has indicated that a solely deterministic model tends to exclude the prevalent randomness effects of traffic dynamics (Li et al., 2012, Wang et al., 2013, Wang et al., 2011, Jabari and Liu, 2012, Sumalee et al., 2011). On the collection side of empirical data, the randomness comes intrinsically from the error of the measurement devices and the following data process, such as inaccurate reading and data roundoff. This type of randomness is well-understood and can be statistically controlled. The second type of randomness is inherently due to traffic dynamics, namely, the demand and supply uncertainties of a transportation system.

The LWR model is refined to consider the stochastic fundamental diagram by Li et al. (2012). A new stochastic model of traffic flow was proposed by incorporating the source of randomness as the uncertainty inherent in driver gap selection, which is represented by random state-dependent vehicle time headway (Jabari and Liu, 2012). The CTM was extended to consider demand and supply uncertainties (Sumalee et al., 2011, Zhong et al., 2013). **Nevertheless, all these abovementioned studies fail to address the issue of vehicle LC despite its importance**. One of the primary aims of the present study is to extend the stochastic CTM (SCTM) to consider the effects of LC maneuvers on stochastic traffic flow dynamics.

### **1.1.2 Vehicle automation and communication systems (VACS)**

The advancement of cyber-physical systems enables vehicles to form a mobile wireless network on the road called vehicular ad hoc network (VANET), which links vehicles with roadway infrastructures and wireless communication. As a promising technology, VANET offers the following two types of communications: vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I). In view of these advantages and a promising market prospect, enormous efforts have been exerted by the automobile industry and numerous institutions toward the development, testing, and deployment of VACS in recent years because they are believed to bring other benefits to traffic safety and efficiency.

Expectedly, the number of vehicles equipped with specific VACS will be increasing with a growing penetration rate in recent years. The rapid development of VACS may lead to improvements in freeway network performance and traffic flow efficiency. As summarized in Diakaki et al. (2015), the VACS are classified into two categories, depending on whether traffic flow implications have been considered, as shown in Figure 1.1.

For the category without direct traffic flow implications, VACS aim to improve the safety and comfort of the driver, and their operation does not modify the common traffic flow patterns. For VACS with direct traffic flow implications, their operations modify the prevailing traffic flow characteristics in addition to the safety and comfort features. For example, through the aid of VACS, an effective approach is to change the driving behavior from individual CF to a platoon-based driving. A platoon-based driving means a cooperative driving pattern for a group of vehicles wherein a vehicle follows the preceding vehicle with a small and nearly constant safety space/time headway to form the platoon (Ngoduy and Jia, 2016). The literature has validated that the platoon-based driving pattern can significantly improve roadway capacity and introduce numerous benefits (Hall and Chin, 2005; van Arem et al., 2006). A recent experiment at the

California Partners for Advanced Transit and Highways (PATH) showed that connected automated vehicles (CAVs) in platoons can maintain a time gap as small as 0.6 s compared with 1.5 s for regular human-piloted vehicles, thereby possibly nearly doubling the roadway capacity ([Chen et al., 2017](#)).

The reduced reaction time of CAVs can improve traffic flow efficiency via smaller inter-vehicle headway (Levin and Boyles, 2016a; Zhou et al., 2017; Zhu and Ukkusuri, 2017). Stern et al. (2018) demonstrated in the mixed-autonomy single-lane ring road experiments (consisting of 22 RHVs on a 230m ring track) that a reduction in fuel consumption of over 40% can be achieved by the insertion of a CAV in the traffic to dampen the ring instability. However, rather limited research effort has been dedicated to addressing the implications of the emerging VACS on the flow characteristics of traffic mixed with CAVs and RHVs, as well as their potential exploitation for improving traffic flow operations (Diakaki et al., 2015; Talebpour and Mahmassani, 2016). Given that vehicle platooning can significantly reduce the air drag of vehicles inside the platoon, energy consumption and exhaust emissions can also be considerably reduced as a result.

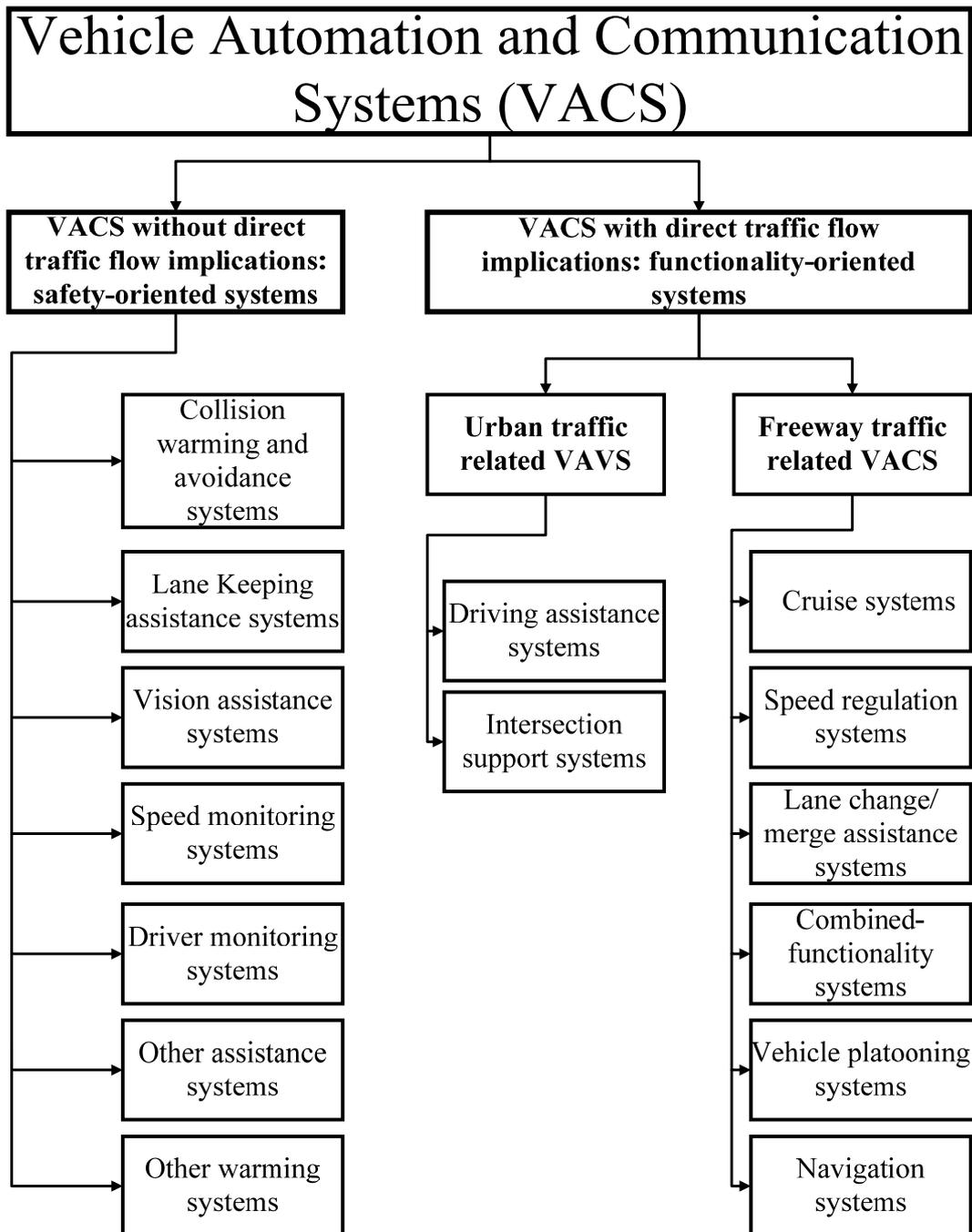


Figure 1.1. VACS taxonomy from a traffic management perspective (Source: revised from Diakaki et al., 2015)

**Most studies reported in the relevant literature belong to the first VACS category. The VACS aim at improving the safety and comfort of the driver.** The California PATH project commenced in 1986. It seeks to improve traffic throughput by deploying platoons on freeways and is probably the first and most well-known in this stream. Other famous projects, including the Grand Cooperative Driving Challenge and EU-

sponsored SARTRE program, aim to utilize state-of-the-art sensor fusion and control methods in conjunction with vehicular communication technologies to promote the development, integration, and deployment of cooperative driving systems. Unlike other projects, the SARTRE program typically uses a truck as the leading vehicle, followed by a series of cars driven autonomously in close formation. Its experiments showed that the platoon could drive at speeds of up to 90 km/hour with a minimum safety gap of no longer than 6 m (Lauer, 2011). In Asia, the representative project is Energy ITS in Japan, which aims to reduce the CO<sub>2</sub> emission from automobiles instead of pursuing efficiency (Tsugawa and Kato, 2010).

All these developments have indicated that the deployment of VACS can improve individual safety, comfort, convenience, and emission of the connected vehicles (Diakaki et al., 2015). The VACS are also currently expected to develop its **potentialities in promoting the global traffic efficiency through traffic control** (Roncoli et al., 2015a; 2015b; 2016). Nevertheless, VACS and CAVs are still in their infant stage.

### **1.1.3 Traffic flow modeling and control in the presence of mixed RHVs and CAVs**

As previously discussed, reliable real-time measurements and trusted estimation of the traffic state are a prerequisite for ATM. State-of-the-art traffic state estimation methods (for regular human-piloted vehicles, RHVs) adopt conventional detector data, such as point-based detectors, smartphones, and vehicles with global positioning system employing various types of traffic flow models via different data-fusion techniques. With the introduction of VACS, vehicles and roadway infrastructures are becoming “connected”, that is, capable of sending and receiving information to/from other vehicles, local/central monitoring unit, and control unit. Vehicles with VACS can act as mobile sensors for traffic state estimation.

A significant concentration of traffic flow research is how much roadway capacity will be improved with respect to the increase in the penetration rate of CAVs. By contrast, the emerging vehicle automation results in an increasing number of CAVs, which can be regarded as “robots” on the road. The driving behavior (e.g., CF and LC) of CAVs and RHVs are different, as mentioned in the previous section. For example, with V2V and V2I communications, CAVs are constantly traveling in platoons. Thus, the CF driving behavior changes from individual CF to cooperative CF. Such type of changes in microscopic driving behavior (e.g., reaction time and acceleration/deceleration) and platoon size affect the traffic flow characteristics comprised of CAVs. The influence of CAVs on roadway capacity have also been studied through simulations and real-world experimental investigations. The cooperative driving pattern of CAVs can increase roadway capacity to double (or higher) and can significantly improve traffic flow stability (Milanes et al., 2014; Milanés and Shladover, 2014).

To the best of the author’s knowledge, most efforts have adopted simulations, whereas limited theoretical research has been conducted. The mechanism on **how the CAVs help improve the throughput is unclear** due to the complicated CF and LC rules specified in the simulations. By contrast, **systematic investigation of the influences of CAVs on collective (macroscopic) traffic flow and sustainability remains insufficient** (Khondaker et al., 2015; Wang et al., 2015, 2016a, 2016b).

Conversely, in computer science, as well as electronics societies, attention is focused on developing VACS without direct traffic flow implications. Research efforts have been dedicated to designing platoon-based cooperative driving by automatic control society (e.g., consensus control and cooperative adaptive cruise control. These studies specifically consider actuator lag and other natural limitations and uncertainties in practical vehicular networking, such as communication topology along with the traffic mobility, transmission range, packet loss, and probabilistic transmission delay. Attention has also been given to motion planning for CAVs, such as autonomous LC decision making and lane change execution based on V2V communication (Luo et al.,

2016; Nie et al., 2017).

However, **limited research efforts** have been made to address the **implications of the emerging VACS on the flow characteristics of traffic mixed with RHVs and CAVs**, as well as **their potential exploitation to improve traffic flow operations** (Diakaki et al., 2015). All these factors impede the development of trusted and fast traffic flow models for ATM to use in traffic flow mixed with CAVs and RHVs. Thus, research and development activities can be further developed.

## **1.2 Research objectives**

This study intends to develop trusted and fast dynamical traffic models to create simultaneous simulations of the MLC and DLC maneuvers while considering the uncertainties from the environment. The prerequisite of the simulation based on the model is discussed through numerical simulations and empirical studies. Further extension of the models is conducted to describe the dynamics of traffic flow mixed with RHVs and CAVs. An integrated freeway traffic flow control framework that aims to improve flow efficiency improve greenness, and ensure safety for freeway traffic mixed with CAVs and RHVs is devised. A reinforcement learning based algorithm is developed to solve the optimal control problem effectively and efficiently. Moreover, the scenarios against various penetration rate of CAVs is evaluated through sensitivity analysis (SA) against the related parameters and decision variables. Achieving these targets sets forth the following research objectives:

To model the DLC and MLC maneuvers simultaneously, a comprehensive macroscopic multilane traffic model is initially developed based on the CTM to propagate macroscopic multilane traffic flow, which enables the following:

- a) Determining the minimum gap acceptance criteria for different LC intentions;
- b) Evaluating whether the traffic flows sent from different directions can be accepted

by a downstream section via an extended incremental transfer (IT) and priority IT (PIT) principle that defines the demand-supply reaction laws;

c) Propagating traffic states on multiple lanes in spatiotemporal dimension.

To provide inputs to the macroscopic multilane traffic flow model, a dynamical LC demand estimation algorithm is developed to enable the following:

- a) An estimation of the longitudinal distribution of MLC and DLC demands;
- b) A mechanism to refine the MLC and DLC dynamically based on the execution of LC flows.

To illustrate the positive and negative influences of LC maneuvers, such as balancing effect and capacity drop, this study would calibrate and validate the proposed macroscopic multilane cell transmission model on some complex freeway weaving sections.

Considering the randomness of data and traffic flow, this study further aims to extend the SCTM to simulate the effects of vehicle LC on traffic dynamics that enables the following:

- a) Random speed-density relationship (regarding mean and variance) for cell-lane speed profile interpolations along a freeway corridor with the installation of sparse detectors;
- b) Link (cell)-node junction formulation to propagate MLC and DLC flows in line with the existing demand-supply reaction laws of the SCTM framework;
- c) Mean and variance estimation of cell-lane traffic states.

The proposed models require no additional data other than that of the CTM (with velocity). Thus, the proposed model can be deployed as a simple simulation tool to access a dynamic macroscopic multilane traffic state from the data available to most management centers, as well as the potential application to predict the influence of traffic incidents or lane control strategies.

To support the integrated freeway traffic control design and subsequently improve the performance efficiency of freeways, this study intends to develop a multilane traffic flow model to simulate traffic flow dynamics mixed with CAVs and RHVs. The proposed multiclass dynamic macroscopic model enables the following:

- a) Determining the variations in the fundamental diagrams with respect to the penetration rate of CAVs;
- b) Identifying minimum headway acceptance criteria for LC maneuvers of RHVs and CAVs;
- c) Evaluating sending flows with an advanced PIT principle;
- d) Propagating the lane-cell specific multiclass traffic flow and density.

With the estimation of the lane cell-specific traffic flow characteristics from the above multiclass multilane traffic flow model, an integrated freeway traffic flow control framework is devised to maximize the throughput, thereby minimizing the total delay. The concept adopted in the proposed control framework employs the integrated action of some control measures, including minimum gap control, variable speed limit (VSL) control, and LC control (LCC) at a macroscopic level. The CAVs are assumed to respond fully to the control, whereas only a certain compliance rate of drivers will follow the control instruction provided by the conventional traffic advisory system. For example, the advisory system would advise drivers in selecting proper lanes (e.g., change lane to left or right or keep lane) from a distance to the incident locations on the freeway or their target off-ramps for RHVs. With such information, drivers can respond to a situation that is still imperceptible to them and therefore allow tactical maneuvers. However, only some of the drivers will follow such lane control instruction.

Finally, a reinforcement learning based algorithm solution algorithm is developed. Evaluation of scenarios against various penetration rates of CAVs is conducted through SA against the related parameter and decision variables to illustrate the following aspects:

- a) The efficiency of LCC policy using CAVs;
- b) The performance assessment of integrated control strategies;
- c) The analysis of the influence of penetration rate and CAV features on the control performance.

### 1.3 Dissertation organization and contributions

Figure 1.2 presents an overview of the dissertation and the relations between the chapters. This dissertation is divided into two parts. After briefly introducing the background materials and literature review, two multilane dynamic traffic flow models are proposed to simulate the influence of MLC and DLC maneuvers simultaneously in Chapter 3 and Chapter 4 under deterministic and stochastic settings, respectively. Then, the second part concentrates on modeling and optimal control of traffic flow mixed with RHVs and CAVs in Chapter 5. The remainder of this dissertation is organized as follows.

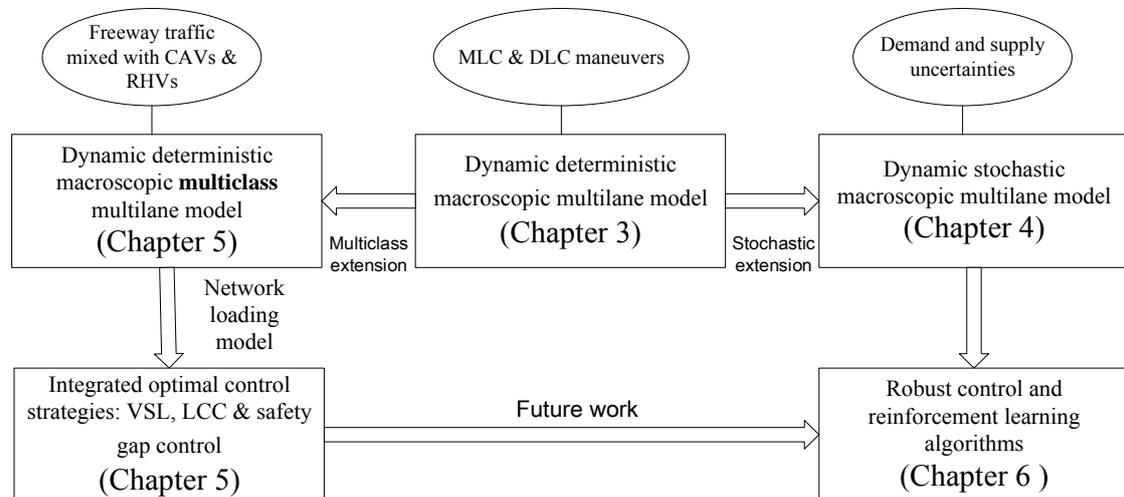


Figure 1.2. Overview of the dissertation

Chapter 2 briefly reviews the literature on several related topics.

Chapter 3 proposes a novel macroscopic multilane model to enable simultaneous simulation of MLC and DLC behaviors to capture multilane traffic dynamics realistically. The model considers lane specific fundamental diagrams to simulate

dynamic heterogeneous lane flow distributions on expressways. Moreover, different priority levels are identified according to different LC motivations and the corresponding levels of urgency. Then, an algorithm is proposed to estimate the dynamic MLC and DLC demands. Finally, the lane flow propagation is defined by the reaction law of the demand-supply functions, which can be regarded as an extension of the IT and PIT principles. The proposed macroscopic multilane CTM is calibrated and validated on a complex weaving section of the State Route 241 freeway in Orange County, California. The empirical result shows the positive and negative influence of LC maneuvers, such as balancing effect and capacity drop, respectively. Moreover, the empirical study verifies that the model requires no additional data other than that of the CTM. Thus, the proposed model can be deployed as a simple simulation tool to access a dynamic macroscopic multilane traffic state from the data available to most management centers, as well as the potential application to predict the influence of traffic incident or lane control strategy.

Chapter 4 extends the SCTM to simulate the effects of vehicle LC on traffic dynamics under various uncertainties by following the multilane traffic flow model developed in Chapter 3. Random speed-density relationship (in terms of mean and variance) is used to interpolate the cell-lane speed profiles along a freeway corridor with sparse detectors. Link (cell)-node junction formulation is proposed to propagate LC vehicles, wherein the LC ratios define virtual node-splitting ratios that are propagated by the IT and PIT principles. Following the operational modes in the SCTM, random events with different probabilities of occurrence are defined to govern the traffic demand-supply reactions. The “actual” flow received by the downstream lane-cell is then a finite mixture of these random events. Accordingly, flow propagations of MLC and DLC maneuvers are proceeded based on the demand-supply reaction laws in line with the existing SCTM framework. Compared with the original SCTM, the key differences in the new model are: i) the traffic states are given in terms of cell and lane; ii) cell-LC ratios are augmented as additional states that define node-splitting ratios; iii) additional process is adopted to resolve cell-lane traffic speed from traffic density estimation to define LC

ratios. However, the link (cell)–node junction formulation integrates the multilane and original SCTMs into a unified framework.

Chapter 5 extends the multilane CTM from a single vehicle class to multivehicle classes to capture the impacts of penetrated CAVs on the freeway traffic characteristics and LC behaviors. First, the variations in the fundamental diagrams with respect to the penetration rate of CAVs are quantified. Then, minimum headway acceptance criteria for LC maneuvers of RHVs and CAVs are identified for the LC maneuvers that are proposed by CAVs and RHVs with different motivations. An advanced PIT principle is then proposed to evaluate the sending flows. Finally, the lane cell-specific multiclass flow conservation law is developed to propagate traffic flow and density.

Based on the multiclass multilane traffic flow model for mixed RHVs and CAVs, Chapter 5 proposes an optimal control framework to improve the efficiency of freeway traffic mixed with RHVs and CAVs. The CAVs would follow full compliance with the control commands through VACS. By contrast, the RHVs that are not connected to the VACS would make decisions in response to the incident alarm and VSL given on the en-route variable message signs (VMS), such as incident message and speed limit. At the upper level of the proposed model, the objective is to devise an integrated action of several control strategies, such as VSL control and LCC to maximize the traffic throughput (or minimize the total traffic delay) on the dual-lane freeway under various traffic conditions for CAVs via VACS and conventional vehicles via VMS. The nominal penetration rate of CAVs is assumed to be provided but subject to uncertainty. The effect of the uncertain ratio of CAVs on traffic flow characteristics is considered in the modified fundamental diagram. The LCC balances the ratio of MLC and DLC demands, whereas the VSL improves the LC efficiency and safety by adjusting the minimum acceptance gap for LC. Advice (e.g., optimal MLC location and VSL) is disseminated upstream such that the drivers can respond to a situation, which is still imperceptible to them, to allow tactical maneuvers. The effectiveness and the computational feasibility of the proposed model are illustrated via numerical example for various penetration

rates of CAVs under different traffic conditions.

Chapter 6 summarizes the study. Some topics for future research are also highlighted in this chapter.

## **Chapter 2 Literature review**

### **2.1 Modeling freeway lane-changing (LC) maneuvers**

The widely used Lighthill-Whitham-Richards (LWR) model (Lighthill and Whitham, 1955; Richards, 1956), and its discretized version, the cell transmission model (CTM; Daganzo, 1994; Szeto, 2008; Sumalee et al., 2011), adopt the triangular fundamental diagram (Newell, 1993). These models are recognized as one of the most straightforward means to explain the evolution of traffic dynamics and features. However, they assume that traffic flow is uniformly distributed over the lanes by integrating traffic streams that travel on different lanes into a single-flow stream with uniform lateral distribution (Munjal and Pipes, 1971). This uniform assumption may be inappropriate, such that heterogeneous traffic flow distribution (e.g., lane-specific flow, density, speed, and vehicle type) can be observed on multilane expressways (Carter et al., 1999; Cassidy and Rudjanakanoknad, 2005; Gunay, 2007; Duret et al., 2012). Therefore, single-lane models may not capture complex traffic features on multilane roadways (e.g., heterogeneous traffic flow distribution, capacity drop, moving bottlenecks, and stop-and-go waves). Moreover, recent research has revealed that vehicle LC has a significant influence on traffic safety because accidents usually happen in LC areas, such as weaving sections and interchanges (Golob et al., 2004; Cassidy and Rudjanakanoknad, 2005; Lee and Cassidy, 2009; Srivastava and Geroliminis, 2013).

Understanding LC is essential to quantifying complex multilane traffic features (Hoogendoorn, 1999; Mauch and Cassidy, 2002; Ahn and Cassidy, 2007). According to different decision-making processes and their influence on surrounding traffic, LC maneuvers are usually classified into mandatory lane-changing (MLC) and discretionary lane-changing (DLC) (Yang and Koutsopoulos, 1996; Zheng, 2014). MLC occurs when a driver must change lane to follow a specific path to a destination, whereas DLC happens when a driver seeks better driving conditions to gain a speed

or time advantage.

Modeling the LC maneuvers is a critical issue in developing microscopic traffic simulation tools (Pipes, 1967; Toledo et al., 2005; Kesting et al., 2007; Sun and Elefteriadou, 2010). Under the simulation framework, LC is usually modeled in the following three steps: 1) LC necessity checking, 2) target lane selection, and 3) gap-acceptance decision. The two most popular microscopic LC algorithms are the rule-based models and discrete choice-based (DCB) models (Toledo et al., 2005; Ben-Akiva et al., 2006; Choudhury et al., 2007a; Choudhury et al., 2007b; Kesting et al., 2007). Rule-based algorithms model LCs from the perspective of heterogeneous drivers with different gap-acceptance conditions<sup>1</sup> and different LC behavior for various situations. Enumerating all possible driving conditions associated with DLC and MLC often results in complex models with numerous parameters (Kesting et al., 2007).

The DCB algorithms simulate driver behavior using the logit or probit models (Ben-Akiva et al., 2006; Toledo et al., 2005; Choudhury et al., 2007a; Choudhury et al., 2007b). These models are usually applied in conjunction with CF models, wherein the attractiveness of a given lane (i.e., its utility) and LC risk regarding acceleration functions of the underlying CF models are measured. A driver's decision becomes either a binary or multichoice selection and utilities for all the alternatives are calculated to obtain the output at each stage during the LC process. Similar to the rule-based models, parameters for gap acceptance in the DCB models should also be extracted from field data and calibrated in the simulation.

The behavior of other drivers is also considered in the rule-based and DCB based

---

<sup>1</sup> Gap-acceptance models are used to model the execution of LC. Before changing lanes, a driver compares the available gaps to the smallest acceptable gap (also known as critical gap). LC is executed if the adjacent gaps for subject vehicles are greater than the critical gaps, that is, the gaps are acceptable. Gaps may be defined either in terms of time or free-space gap.

models. For example, the MOBIL system considers a politeness factor that characterizes the degree of passive cooperativeness and aggressiveness among drivers. Although these models can describe detailed LC behaviors, they usually contain numerous parameters and cannot provide intuitive descriptions of the system-level influence of LC traffic. Model calibration and validation involve high-resolution vehicle trajectory data. Kesting et al. (2007) argued that the empirical investigation of LC behavior for these models is even more difficult than that for car-following (CF) behavior. Also, the gap-acceptance component of these models usually produces too few LCs (Hidas, 2005) and causes an overreaction effect (Laval and Leclercq, 2008). Based on a two-stage test-drive experiment for a variety of drivers on freeways, Keyvan-Ekbatani et al. (2016) studied the decision process of lane changing maneuvers for driving behavior modeling. It is found empirically that the lane change decisions are related to their speed choice.

Meanwhile, several studies have been conducted to understand various LC traffic characteristics at the macroscopic level by extending the kinematic wave (KW) theory and the gas-kinetic theory. By introducing the generalized phase-space density (g-PSD), Hoogendoorn (1999); Hoogendoorn and Bovy (2001a) developed a generic continuum modeling approach to describe the flow operations for a general class of traffic systems ranging from vehicular traffic on motorways to pedestrian flows. The g-PSD generalize traffic density to include both discrete attributes, such as user-class, roadway lane, and destination, and continuous attributes, such as velocity. Considering the traffic as a collection of partially independently moving vehicular platoons traversing along stochastic trajectories, Hoogendoorn (1999); Hoogendoorn and Bovy (2001b) developed a gas-kinetic traffic flow model of multilane traffic. This *multiclass multilane* traffic flow model describes the traffic dynamics, including convection, acceleration, deceleration, and various types of lane-changing, for **free-flowing and platooning** vehicles using the g-PSD.

Laval and Leclercq (2010) proposed a mechanism to describe how stop-and-go

oscillations may result from the reactions of heterogeneous drivers to minor speed variations on a hill. Jin et al. (2006) investigated the degree of first-in-first-out violation among vehicles in multilane traffic. Ahn and Cassidy (2007) and Wang and Coifman (2008) demonstrated that LC maneuvers contribute to traffic oscillations. Laval and Daganzo (2006) proposed a hybrid model of LC traffic theory known as the multilane hybrid (MH) theory. The MH theory models each LC maneuver as a moving bottleneck with bounded acceleration. The bounded acceleration of the LC into a faster-moving lane can explain traffic instabilities at the lane-drop and moving bottlenecks, as well as at merged bottlenecks (Laval et al., 2007). Daamen et al. (2010) analyzed the merging maneuver with empirical data at a freeway on-ramp to reveal the merge location and its relation to gap acceptance and the relaxation phenomenon.

Coifman et al. (2006) found that LC maneuvers cause additional delay in queues. Patire and Cassidy (2011) investigated the LC behavior for an uphill expressway. Carey et al. (2013) extended the CTM to consider LC maneuvers for dynamic traffic assignment purpose. A lane changing model was proposed from a combination of the route, speed, and keep-right incentives by influencing the CF behavior for relaxation and synchronization (Schakel et al., 2012). The model fits well in the lane-flow distributions under a free-flow condition, whereas the fit was unclear under congested conditions.

Cassidy and Rudjanakanoknad (2005) found that LC maneuvers can cause capacity drop at a freeway merge bottleneck. Leclercq et al. (2016a) devised analytical formulas to estimate the effective capacity at freeway merges to evaluate the capacity drop at merges for a single-lane freeway. Under the same modeling framework, Leclercq et al. (2016b) extended the derivation of analytical expressions for estimating the capacity drops at merges of multilane freeways by dividing the mandatory lane-changing and discretionary lane-changing maneuvers into two non-overlapping local merging areas. Noting that the capacity drop is strongly correlated with the congestion state, i.e., the queue discharge rate is a function of vehicular speed in traffic jams, Yuan et al. (2017) developed a Lagrangian kinematic wave model with such varying capacity drop to

support the design of freeway control strategies.

By contrast, LC maneuvers have positive long-term effects on the performance of transportation systems. The heterogeneity among different lanes is usually the reason that motivates LC (Shvestsov and Helbing 1999; Laval and Daganzo 2006; Jin 2010a; Patire and Cassidy, 2011). However, LC maneuvers might have a balancing effect that they could smooth out differences between the lanes under certain situations. This balancing effect can be beneficial in an entire traffic system to achieve high efficiency. For example, Cheu et al. (2009) found that LC can reduce the overall system queuing delay through simulation.

Jin (2010a) claimed that all these studies do not provide a simple approach for analyzing the influence of LC maneuvers and the corresponding traffic dynamics at the aggregate level. When changing its lane, a vehicle affects traffic on its current and target lanes. Based on this observation, Jin (2010a) also proposed an extended KW model to capture such lateral interactions by introducing a new LC intensity variable, that is, LC traffic causes effective additional density, which is determined by the LC choices of drivers and their characteristics in a road section during a time interval. With a modified fundamental diagram, the impacts of LC traffic on the overall traffic flow can be captured by the extended KW model. Jin (2010b) used a set of vehicle trajectory data collected from a freeway section of the Next Generation SIMulation (NGSIM) project to calibrate LC intensity and found that the LC intensity is an exponential function of traffic density. However, this approach was mainly developed for freeway segments without on-/off-ramps, and several restrictive assumptions were enforced (Zheng, 2014).

Tang et al. (2009) proposed a macroscopic model of LC and showed its consistency with CF behavior on a two-lane highway. The effect of LC on the stable region and the propagation speeds of the first- and second-order waves were investigated using the linear stability theory. Both models proposed by Jin (2010a) and Tang et al. (2009) did

not represent lane-specific behavior. Sheu and Ritchie (2001) and Sheu (2004) defined LC fraction from the original lane to adjacent lanes and the return-LC fraction from adjacent lanes to the original lane in the downstream. They used the corresponding upstream and downstream traffic counts that were directly detected from point detectors in state space formulation, wherein the Kalman filtering approach was adopted to estimate all traffic variables, including traffic volumes and LC fractions. Knoop and Buisson (2015) reviewed the methodologies to calibrate and validate both microscopic and macroscopic probabilistic discretionary lane changing models using a likelihood component to capture the stochastic effects. It was found that it is best to use physically interpretable measures that help define the minimum quality of the model during validation.

As previously discussed, vehicle LC maneuvers often result in traffic flow instability. Recent research has found that even more significant effect on surrounding traffic can be caused by the LC maneuvers of heavy vehicles although they only account for a minority of traffic stream. Heavy vehicles produce a disproportionate effect on passenger vehicle drivers, especially under heavy traffic conditions. Despite the increasing number of heavy vehicles on freeways and its pronounced effect on traffic flow, previous studies have focused on the LC behavior of passenger car drivers. The differences between the heavy vehicles and passenger cars are primarily accounted through the differences in vehicle length and acceleration/deceleration capabilities (Gipps, 1986; Ahmed, 1999; Hidas, 2005; Toledo, 2009; Suzuki and Zheng, 2014).

From an empirical aspect, Knoop et al. (2012) investigated the relationship between the number of LCs as the operational characteristics of the origin and target lane. The empirical findings by Duret et al. (2012) highlighted that the flow distribution over different lanes should be included for a multilane traffic model. Such lane flow distribution can simulate the lateral feature of traffic flow due to the heterogeneity in driver behaviors, vehicle types, and the corresponding vehicle-class specific control (e.g., driving ban for trucks). Thus, the speed–density relationship should be in a lane-specific

manner when developing the multilane traffic model. This approach is parallel to conventional macroscopic traffic flow models that adopt a uniform fundamental diagram to represent the longitudinal dynamics of traffic flow.

Remarkably, the existing models at the same scale often follow different modeling approaches (e.g., microscopic, macroscopic, and macroscopic). Thus, relating these models to one another is difficult. For example, a macroscopic model usually lacks a microscopic basis and vice versa. Therefore, an ideal multiscale modeling approach should emphasize not only the model quality at each scale but also the coupling between different scales. Hybrid models are proposed to combine microscopic, macroscopic, and macroscopic models (Leclercq, 2007; Leclercq and Moutari, 2007; Ni, 2011). Hybrid models allow a traffic manager to “zoom in” to examine low-level details and to “zoom out” to have an overview system-wide performance within the same simulation process at different levels of detail.

Notably, Leclercq (2007) and Leclercq and Moutari (2007) found that the way traffic is represented (i.e., microscopic or macroscopic modeling) does not influence the simulation results. Microscopic (car following) and macroscopic (flow conservation) approaches lead to the same solutions under certain assumptions. The model properties are governed by the behavior rule (e.g., the fundamental diagram) rather than by traffic representation. The microscopic-to-macroscopic and macroscopic-to-microscopic interfaces were claimed to be compatible with LC modeling. The expression of this extension in the microscopic scheme can develop a global MH model, including LC. Meanwhile, incorporating the microscopic modeling approach into macroscopic modeling also enables the macroscopic models (e.g., the hybrid CTM proposed by Laval and Daganzo (2006)) to capture LC relaxation by adopting the CF rule (Laval and Leclercq, 2008; Leclercq et al., 2007).

The existing LC models are mostly developed for freeways and are unlikely to be applied to arterial lane changes. Several operation models were recently developed to

fill this gap. A CTM with LC and vehicle tracking framework was proposed for international land ports of entry to simulate lane-specific queuing while allowing vehicles to change their lanes between queues (Cheu et al., 2009). This method can be regarded as a combination of the CTM modeling framework and the random utility theory, wherein the CTM is used to propagate flow dynamics while the LC probabilities are defined using certain random utility theories. Liu and Chang (2011) extended the CTM in a lane-specified manner and proposed the LC potential and maneuvers using the “min” operator in conjunction with the blockage matrix that involves driver behavior to describe arterial LC traffic. Sheu and Ritchie (2001) and Sheu (2004) defined LC fraction from the original lane to adjacent lanes and the return-LC fraction from the adjacent lanes to the original lane in the downstream using the corresponding upstream and downstream traffic counts that are directly detected from point detectors in the state space formulation. They adopted the Kalman filtering approach to estimate all traffic variables, including traffic volumes and LC fractions.

## 2.2 The cell transmission model (CTM)

Daganzo (1994, 1995) proposed CTM as a deterministic dynamic traffic flow model that discretizes the LWR model in time and space. As shown in Figure 2.1, the model divides a freeway segment into several homogeneous, consecutively numbered cells with length  $l_i$ , where  $i$  is a cell index. Ideally, the cell length should be equal to the distance traveled by free-flowing vehicles during one simulation time increment; that is,  $l_i = v_f T_s$ , where  $v_f$  is the free-flow speed of the freeway segment, and  $T_s$  is the simulation time increment. The model assumes that the traffic state is homogenous within one cell. The number of vehicles on cell  $i$  at time  $k\Delta t$ , which is considered as the state variable for the cell, can be evaluated according to the following flow conservation equation:

$$n_i(k+1) = n_i(k) + (y_i(k) - y_{i+1}(k)), \quad (2.1)$$

where  $y_i(k)$  is the inflow volume from cell  $i-1$  to  $i$  during the time interval  $[kT_s, (k+1)T_s)$ . The flow from cell  $i-1$  into cell  $i$  is the minimum of the following three quantities:

$$y_i(k) = \min\left\{n_{i-1}(k), Q_i, (w_c / v_f)[N_i - n_i(k)]\right\}, \quad (2.2)$$

where  $N_i$  is the maximum number of vehicles that can be present in cell  $i$ , with  $N_i = \rho_j l_i$ .  $Q_i$  is the maximum number of vehicles that can flow into cell  $i$  during time interval  $[kT_s, (k+1)T_s)$ , with  $Q_i = Q_M T_s$ .  $w_c$  is the backward wave speed when traffic is congested (mile/hour); and  $v_f$  is the free-flow speed (mile/hour).  $(w_c / v_f)[N_i - n_i(k)]$  is the jam-limited volume, which depends on the amount of available space of cell  $i$ . In this case, the CTM corresponds to a trapezoidal fundamental diagram shape, as shown in Figure 2.2, where  $Q_M$  is the maximum allowable flow rate (vehicle/hour),  $\rho_j$  is the jam density (vehicle/mile), and  $\rho_c$  is the critical density (vehicle/mile).

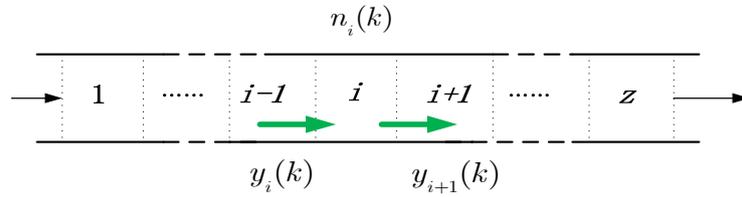


Figure 2.1. Cell partition of a freeway segment

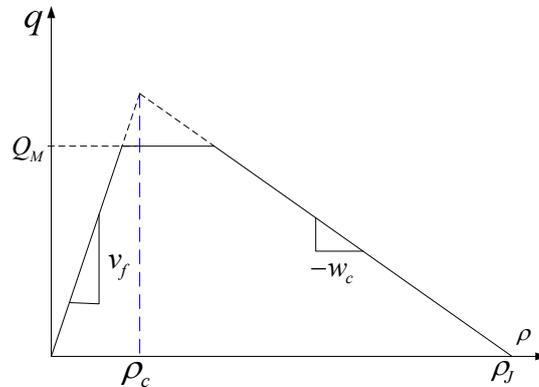


Figure 2.2. Trapezoidal fundamental diagram for CTM and modified CTM (MCTM)

For traffic surveillance and control purposes, traffic density naturally characterizes the congestion level. For example, the mapping from traffic volume to traffic density is multi-valued, as shown in Figure 2.2. A flow rate of 5000 vehicles/hour may correspond to either the free-flowing traffic condition (i.e., the left-hand side of the fundamental diagram) or congested traffic condition (i.e., the right-hand side of the fundamental diagram). By contrast, the mapping from traffic density to traffic condition is unique. Thus, Munoz et al. (2003) developed MCTM for traffic surveillance. This model permits the CTM to adopt non-uniform cell lengths and results in greater flexibility in partitioning the freeways. In the MCTM, the density of cell  $i$  evolves as

$$\rho_i(k+1) = \rho_i(k) + \frac{T_s}{l_i} \left( q_{i,in}(k) - q_{i,out}(k) \right), \quad (2.3)$$

where  $\rho_i(k)$  is the vehicle density in cell  $i$  at time index  $k$ ;  $q_{i,in}(k)$  and  $q_{i,out}(k)$  are the flow rates (in vehicles per unit time) entering and leaving cell  $i$  during the time interval  $\left[ kT_s, (k+1)T_s \right)$ , respectively. The model defines some piecewise affine sending and receiving functions of traffic volumes to describe the interactions between neighboring cells and the shockwaves of a freeway segment.  $q_i(k)$  is determined by considering the minimum of two quantities:

$$q_{i,in}(k) = \min(S_{i-1}(k), R_i(k)), \quad (2.4)$$

where  $S_{i-1}(k)$  is the maximum flow that can be supplied by cell  $i-1$  under free-flow conditions over the interval  $\left[ kT_s, (k+1)T_s \right)$ , and  $R_i(k)$  is the maximum flow rate that can be received by cell  $i$  under congested conditions over the same time interval. :

$$S_{i-1}(k) = \min(v_{f,i-1}\rho_{i-1}(k), Q_{M,i-1}) \quad (2.5)$$

$$R_i(k) = \min(Q_{M,i}, w_{c,i}(\rho_{J,i} - \rho_i(k))) \quad (2.6)$$

Equations (2.4)–(2.6) yield

$$q_{i,in}(k) = \min(v_{f,i-1}\rho_{i-1}(k), Q_{M,i-1}, Q_{M,i}, w_{c,i}(\rho_{J,i} - \rho_i(k))). \quad (2.7)$$

The definitions of sending and receiving functions are useful when the model is extended to handle general network topologies (Daganzo, 1995; Munoz et al., 2003).

Lo et al. (2001) further extended the CTM to track the path flows for dynamic traffic assignment. As a discrete LWR model version, the CTM suffers from most of the drawbacks of the LWR model in describing traffic flow.

### 2.3 Randomness of the fundamental diagram

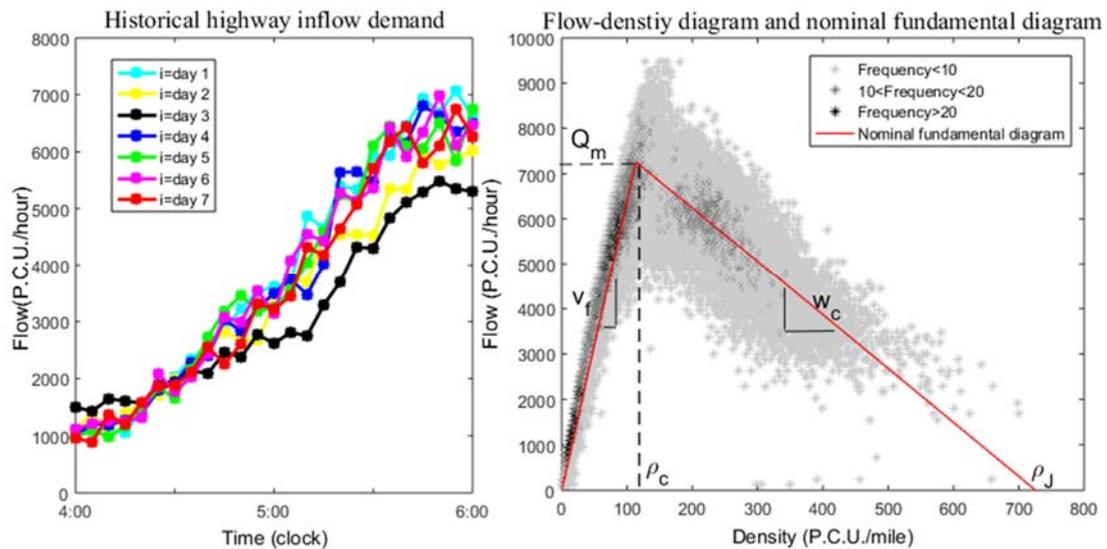


Figure 2.3. Stochastic inflow demand and fundamental diagram (Source: Pan et al. (2013). The data was collected by a detector station on a four-lane freeway segment (with one lane for on/off-ramps) of Interstate 210 located in Monrovia, Los Angeles. The data was downloaded from the Performance Measurement System (PeMS).)

The fundamental diagram plays an essential role in (mesoscopic and macroscopic, or the hybrid approach) traffic flow modeling. Most existing models are deterministic, which essentially describe average system behaviors (Li et al., 2012; Wang et al., 2013; Wang et al., 2011). However, recent research has indicated that a solely deterministic model is unlikely to include the prevalent dynamical randomness effects of traffic flow (Sumalee et al., 2011; Li et al., 2012; Wang et al., 2013; Wang et al., 2011; Jabari et al., 2012). In collecting empirical data, the randomness comes intrinsically from the error of measurement devices and subsequent data process (e.g., inaccurate reading and data roundoff). This type of randomness is well-understood and can be statistically controlled. The second type of randomness is inherently due to traffic dynamics (i.e., the demand and supply uncertainties of a transportation system). The random demand

profiles and stochastic supply functions (e.g., the variations of day-to-day inflow demand at non-rush hour on seven days as shown in the left-hand side of Figure 2.3) and uncertain parameters in the flow–density fundamental diagrams (the flow–density data collected for 54 days as shown in the right-hand side of Figure 2.3) are considered as exogenous inputs of stochastic CTM (SCTM).

The LWR model is refined to consider the stochastic fundamental diagram of Li et al. (2012). A new stochastic model of traffic flow was proposed by incorporating the randomness source as the uncertainty inherent in the gap choice of drivers, which is represented by the random-state dependent vehicle time headway (Jabari et al., 2012). The CTM was extended to consider the demand and supply uncertainties (Sumalee et al., 2011; Zhong et al., 2013). Nevertheless, the previous studies did not address the issue of vehicle lane-changing regardless of its importance. Thus, the present study aims to extend the earlier works on the SCTM, thereby considering the effects of lane-changing maneuvers on stochastic traffic flow dynamics.

## **2.4 The stochastic cell transmission model (SCTM)**

The CTM and MCTM rely on the steady-supply functions, as well as the assumption that the inflow demand and supply functions are deterministic. However, as explained in the previous chapter, this assumption is not realistic due to the inherent demand and supply uncertainties. The Monte Carlo method is a simple way to extend the CTM and handle the demand and supply stochasticities. However, the potentially high computational cost and the techniques required to reduce the result variances may be the apparent disadvantages and difficulties. Sumalee et al. (2011) proposed a stochastic dynamic traffic flow model (i.e., SCTM), which extends the switching-mode model (SMM), to model the effects of the demand and supply uncertainties on traffic flow dynamics and its propagation. The SCTM also considers stochastic parameters of the fundamental flow-density diagram, as well as the stochastic travel demand.

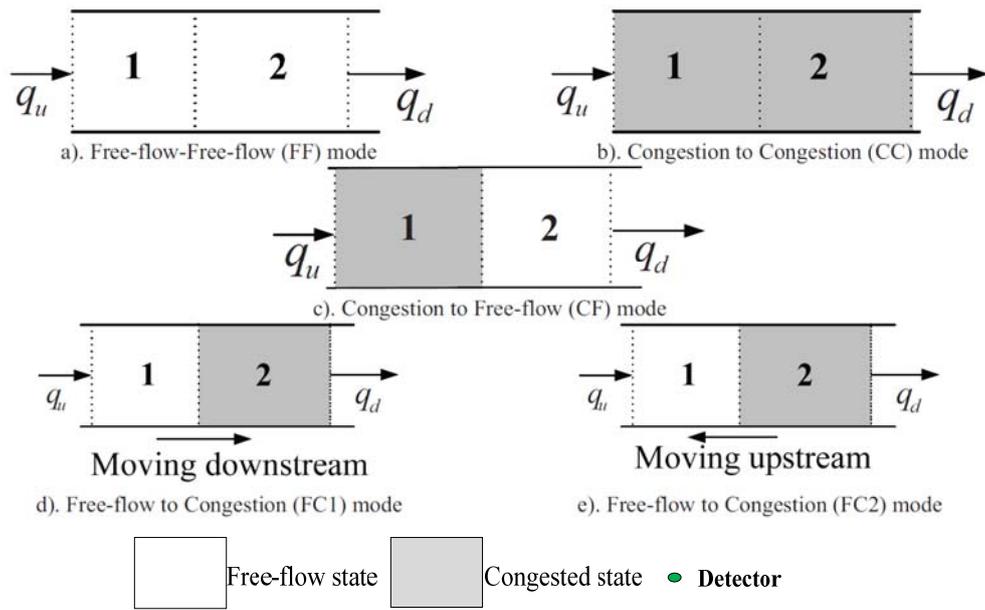


Figure 2.4. Freeway segment with two cells for the SMM and SCTM

The SCTM by Sumalee et al. (2011) is a stochastic dynamic traffic flow model for a freeway segment while Zhong et al. (2011) extended this model to the network case. The SCTM considers the stochastic characteristics of the fundamental diagram (i.e., uncertain flow–density relationship) and admits the stochastic travel demand as the exogenous input, as shown in Figure 2.4. For the supply side, the means and variances of different parameters (i.e., free-flow speed, jam density, critical density, and backward-wave speed) are calibrated based on the statistics of the observed data (e.g., the grey dots on the right side of Figure 2.4) for each section equipped with detectors.

The SCTM defines the random inflows (i.e., uncertain demand) as well as the random parameters of the fundamental flow–density diagram (i.e., uncertain supply functions) as boundary variables. It accepts the means and variances of the boundary variables as exogenous inputs, and the means and variances of the traffic densities, the freeway segment outflow, and probabilities of its operational modes are then calculated as outputs based on the measured boundary conditions. Specifically, the stochastic flow propagation between adjacent cells under each mode can be determined by the flow–density relationship at each simulation time step. Each of these stochastic flow profiles

is associated with a certain probability corresponding to an operational mode. The flow propagated from the upstream to a downstream cell in the next time step is a mixture distribution of the stochastic flows of the operational modes. Sumalee et al. (2011) validated the SCTM performance in estimating stochastic traffic densities and dynamic travel time distribution against empirical freeway traffic data.

Figure 2.4 shows the SCTM application in representing a freeway segment without on- or off-ramp, which is divided into two cells. Similar to the SMM proposed by Muñoz et al. (2006), five probabilistic events (herein referred to as five operational modes) are defined in the SCTM to represent the freeway state under different congestion levels. The five operational modes include two steady-state modes: FF and CC modes; and three transient modes: CF, FC1, and FC2 modes. The division of FC1 and FC2 modes depends on the moving direction of the wave-front; if the wave-front is moving downstream, then the freeway segment is in the FC1 mode; otherwise, it is in the FC2 mode. The SCTM defines the probabilities of occurrences for these five events due to the random demand and supply. Each mode of the SCTM can be represented by the following bilinear system:

$$\begin{aligned} \rho(k+1) &= (A_0 + \sum_{i=1}^2 A_i \omega_i(k)) \rho(k) \\ &+ (B_0 + \sum_{i=1}^2 B_i \omega_i(k)) \lambda(k) + Bu(k), \end{aligned} \quad (2.8)$$

where  $B, A_i, B_i, i = 0, 1, 2$ , are constant matrices to be defined later, and  $\omega_i(k), \forall k \in N$  are the second-order processes consisting of mutually uncorrelated real-valued random variables.

In the FF mode,  $\omega_i(k)$  is set as the free-flow speed  $v_{f,i}(k)$  in Equation (2.8), and the state equation can be expressed as

$$\rho(k+1) = (I + \sum_{i=1}^2 A_i v_{f,i}(k)) \rho(k) + Bu(k), \quad (2.9)$$

where  $A_1 = \begin{bmatrix} -\frac{T_s}{l_1} & 0 \\ \frac{T_s}{l_2} & 0 \end{bmatrix}$ ,  $A_2 = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{T_s}{l_2} \end{bmatrix}$ ,  $B = \begin{bmatrix} \frac{T_s}{l_1} & 0 \\ 0 & 0 \end{bmatrix}$ . Equation (2.9) is a special case of Equation (2.8), with  $B_i, i = 1, 2$  as null matrices and  $\lambda(k)$  as a null vector.

In the CC mode,  $\omega_i(k) = w_{c,i}(k)$ , and vector  $\lambda(k) = (\rho_{J,1}(k), \rho_{J,2}(k))^T$ . The state equation is then

$$\begin{aligned} \rho(k+1) &= (I + \sum_{i=1}^2 A_i w_{c,i}(k))\rho(k) \\ &\quad + \sum_{i=1}^2 B_i w_{c,i}(k)\lambda(k) + Bu(k), \end{aligned} \quad (2.10)$$

where  $A_1 = \begin{bmatrix} -\frac{T_s}{l_1} & 0 \\ 0 & 0 \end{bmatrix}$ ,  $A_2 = \begin{bmatrix} 0 & \frac{T_s}{l_1} \\ 0 & -\frac{T_s}{l_2} \end{bmatrix}$ ,  $B = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{T_s}{l_2} \end{bmatrix}$ ,  $B_i = -A_i, i = 1, 2$ .

In the CF mode,  $\omega_1(k) = w_{c,1}(k)$ ,  $\omega_2(k) = v_{f,2}(k)$ , and the vector  $\lambda(k) = (\rho_{J,1}(k), Q(k))^T$ . The state equation is then

$$\begin{aligned} \rho(k+1) &= (I + \sum_{i=1}^2 A_i \omega_i(k))\rho(k) \\ &\quad + (B_0 + \sum_{i=1}^2 B_i \omega_i(k))\lambda(k) + Bu(k), \end{aligned} \quad (2.11)$$

where  $A_1 = \begin{bmatrix} -\frac{T_s}{l_1} & 0 \\ 0 & 0 \end{bmatrix}$ ,  $A_2 = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{T_s}{l_2} \end{bmatrix}$ ,  $B_0 = \begin{bmatrix} 0 & -\frac{T_s}{l_1} \\ 0 & \frac{T_s}{l_2} \end{bmatrix}$ ,  $B_1 = -A_1$ ,  $B_2 = 0$ ,  $B = 0$ .

In the FC1 mode,  $\omega_1(k) = v_{f,1}(k)$ ,  $\omega_2(k) = 0$ , and  $\lambda(k)$  is a null vector. The state equation is then

$$\rho(k+1) = (I + A_1 \omega_1(k))\rho(k) + Bu(k), \quad (2.12)$$

where  $A_1 = \begin{bmatrix} -\frac{T_s}{l_1} & 0 \\ \frac{T_s}{l_2} & 0 \end{bmatrix}$ ,  $B = \begin{bmatrix} \frac{T_s}{l_1} & 0 \\ 0 & -\frac{T_s}{l_2} \end{bmatrix}$ .

In the FC2 mode,  $\omega_1(k) = 0$ ,  $\omega_2(k) = w_{c,2}(k)$ , and  $\lambda(k) = (0, \rho_{J,2}(k))^T$ . The state equation is

$$\rho(k+1) = (I + A_2 \omega_2(k))\rho(k) + B_2 \omega_2(k)\lambda(k) + Bu(k),$$

where  $A_1 = 0$ ,  $A_2 = \begin{bmatrix} 0 & \frac{T_s}{l_1} \\ 0 & -\frac{T_s}{l_2} \end{bmatrix}$ ,  $B_1 = 0$ ,  $B_2 = \begin{bmatrix} 0 & -\frac{T_s}{l_1} \\ 0 & \frac{T_s}{l_2} \end{bmatrix}$ ,  $B = \begin{bmatrix} \frac{T_s}{l_1} & 0 \\ 0 & -\frac{T_s}{l_2} \end{bmatrix}$ .

The evaluation of mean and variance of the above bilinear system was studied by Sumalee et al. (2011). The corresponding occurrence probabilities of the five modes can be defined as follows.

FF mode:  $P_{FF}(k) \triangleq \Pr(\rho_u(k-1) < \rho_{c,1}(k-1) \cap \rho_d(k-1) < \rho_{c,2}(k-1))$

CC mode:  $P_{CC}(k) \triangleq$

$$\Pr(\rho_u(k-1) \geq \rho_{c,1}(k-1) \cap \rho_d(k-1) \geq \rho_{c,2}(k-1))$$

CF mode:  $P_{CF}(k) \triangleq$

$$\Pr(\rho_u(k-1) \geq \rho_{c,1}(k-1) \cap \rho_d(k-1) < \rho_{c,2}(k-1))$$

FC mode:  $P_{FC}(k) \triangleq 1 - (P_{FF}(k) + P_{CC}(k) + P_{CF}(k))$ , with the wave-front moving downstream as

$$\begin{aligned} P_{D|FC}(k) &\triangleq \Pr(v_{f,1}(k-1)\bar{\rho}_1(k-1) \\ &\leq w_2(k-1)(\rho_{J,2}(k-1) - \bar{\rho}_2(k-1))), \end{aligned}$$

and the wave-front moving upstream as  $P_{U|FC}(k) = 1 - P_{D|FC}(k)$ . Then, the probabilities of the FC1 and FC2 to occur at time step  $k$  are as follows.

FC1 mode:  $P_{FC1}(k) \triangleq P_{D|FC}(k)P_{FC}(k)$

FC2 mode:  $P_{FC2}(k) \triangleq P_{U|FC}(k)P_{FC}(k)$ ,

where  $\rho_{c,i}$  is the critical density,  $w_i$  the backward congestion wave speed, and  $\rho_{J,i}$  is the jam density of cell  $i$ .  $\bar{\rho}_i$  is the joint density of cell  $i$ , which is defined as a finite mixture distribution of the five modes. The mean and covariance matrix of the joint traffic density vector are denoted as  $\bar{\rho}(k)$ ,  $E(\bar{\rho}(k)|\theta(k))$ , and  $Var(\bar{\rho}(k)|\theta(k))$ , where  $\theta(k) = \{\theta_s(k)\}$ ,  $\theta_s(k) = (\rho_s(k), P_s(k))$ , and  $\rho_s(k)$  denotes the vector of cell densities of mode  $s$ . The probability density function (PDF) of the joint traffic density  $f(\bar{\rho}(k)|\theta(k))$  is defined as

$$f(\bar{\rho}(k)|\theta(k)) = \sum_s P_s(k)f(\bar{\rho}(k)|\theta_s(k)). \quad (2.13)$$

Under the mixture model (6), the expectation  $E(\bar{\rho}(k)|\theta(k))$  can be expressed as follows:

$$E(\bar{\rho}(k)|\theta(k)) = \sum_s P_s(k)E(\rho_s(k)). \quad (2.14)$$

Let  $\mu_s(k) = E(\rho_s(k))$  and  $\mu(k) = E(\bar{\rho}(k)|\theta(k))$ . Then,  $\mu(k) = \sum_s P_s(k)\mu_s(k)$ .

To evaluate  $Var(\bar{\rho}(k)|\theta(k))$ , the covariance matrix of  $\rho_s(k)$  is defined as follows:

$$\psi_s(k) = E((\rho_s(k) - \mu_s(k))(\rho_s(k) - \mu_s(k))^T).$$

Then, the covariance matrix  $Var(\bar{\rho}(k)|\theta(k))$  can be evaluated as

$$Var(\bar{\rho}(k)|\theta(k))$$

$$= \sum_s P_s(k)(\psi_s(k) + \mu_s(k)\mu_s^T(k)) - \mu(k)\mu^T(k). \quad (2.15)$$

In the original SCTM, the uncorrelated assumption is enforced to simplify the probability evaluation and traffic flow propagation. This uncorrelated assumption is no longer valid when the spatiotemporal correlations are incorporated into the model, which in turn affects two major SCTM components (i.e., evaluations of occurrence probabilities of different modes and traffic flow propagation). Here, the scenarios introduced in Section 3.3 by Sumalee et al. (2011) are investigated. To begin with, define  $X_{ud}$  as  $X_{ud} = [X_u, X_d]^T = [\tilde{\rho}_u(k) - \rho_{c,1}(k), \tilde{\rho}_d(k) - \rho_{c,2}(k)]^T$ , where time index  $k$  to  $X_{ud}$  are omitted to save notation. Then the PDF of  $X_{ud}$  is a bivariate normal distribution because all its components are assumed to be normally distributed

$$\text{pdf}(X_{ud}; \mu_{ud}, \Sigma_{ud}) = \frac{1}{(2\pi)^{|\Sigma_{ud}|^{1/2}}} e^{-Q(X_{ud}; \mu_{ud}, \Sigma_{ud})/2}, \quad (2.16)$$

where  $\mu_{ud} = (\mu_u, \mu_d)^T = (\bar{\rho}_u - \bar{\rho}_{c,1}, \bar{\rho}_d - \bar{\rho}_{c,2})^T$  is the expectation of  $X_{ud}$ , and  $\Sigma_{ud}$  is the covariance matrix, with

$$\Sigma_{ud} = \begin{bmatrix} \text{Var}(X_u) & \text{Cov}(X_u, X_d) \\ \text{Cov}(X_u, X_d) & \text{Var}(X_d) \end{bmatrix},$$

and  $Q(X_{ud}; \mu_{ud}, \Sigma_{ud}) = (X_{ud} - \mu_{ud})^T \Sigma_{ud}^{-1} (X_{ud} - \mu_{ud})$ . The probability of the occurrence of FF mode is

$$\begin{aligned} P_{FF}(k) &\triangleq \Pr(\tilde{\rho}_u(k) < \rho_{c,1} \cap \tilde{\rho}_d(k) < \rho_{c,2}) \\ &= \Pr(X_u < 0 \cap X_d < 0) \\ &= \int_{-\infty}^0 \int_{-\infty}^0 \text{pdf}(X_{ud}; \mu_{ud}, \Sigma_{ud}) dX_u dX_d. \end{aligned} \quad (2.17)$$

The probabilities of occurrence of other modes can be similarly defined:

$$P_{CC}(k) = \int_0^{+\infty} \int_0^{+\infty} \text{pdf}(X_{ud}; \mu_{ud}, \Sigma_{ud}) dX_u dX_d, \quad (2.18)$$

$$P_{CF}(k) = \int_0^{+\infty} \int_{-\infty}^0 \text{pdf}(X_{ud}; \mu_{ud}, \Sigma_{ud}) dX_u dX_d, \quad (2.19)$$

$$P_{FC}(k) = 1 - (P_{FF}(k) + P_{CC}(k) + P_{CF}(k)). \quad (2.20)$$

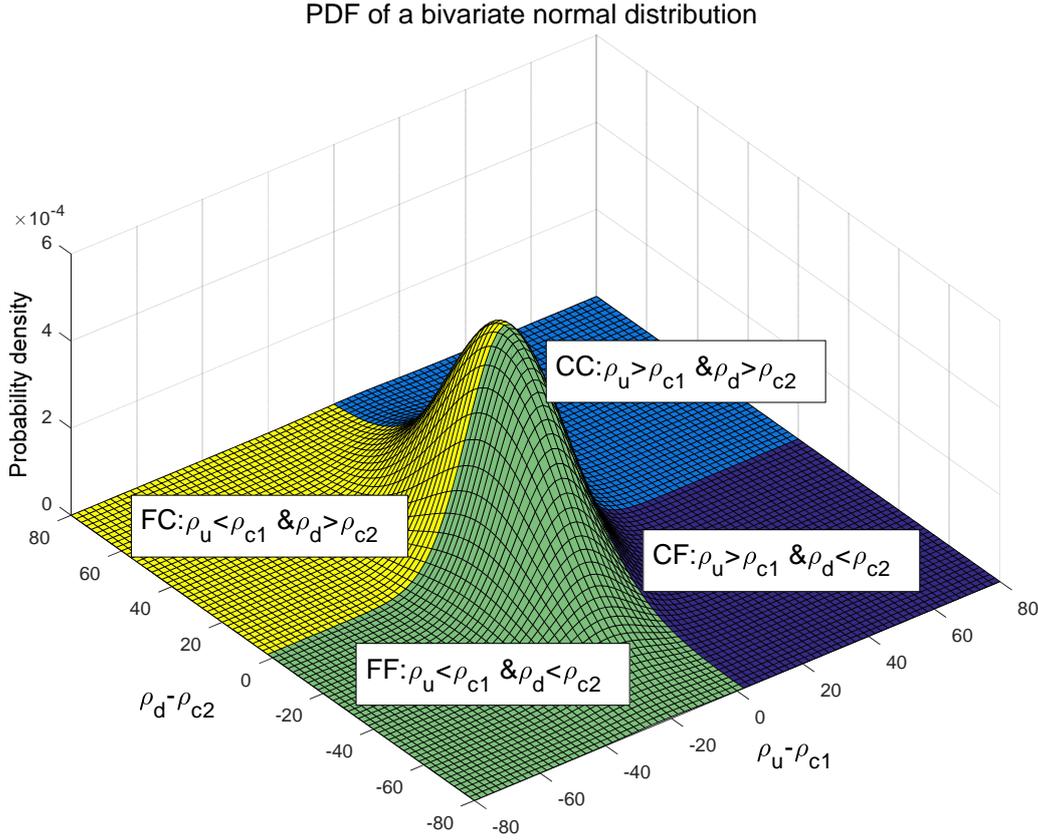


Figure 2.5. PDF and mode occurrence probabilities

A demonstration of the PDF of the bivariate normal distribution discussed above is shown in Figure 2.5. The occurrence probabilities of different modes can be evaluated by the area of the corresponding regimes (e.g., the yellow cover denotes the probability of FF mode). The evaluation of the autocorrelation matrix of the bilinear system should be adjusted accordingly. For example, the covariance of two supply functions is (Equation (29) in Sumalee et al. (2011)):  $Cov(\omega_{s,i}(k), \omega_{s,j}(k)) = E(\omega_{s,i}(k)\omega_{s,j}(k)) - E(\omega_{s,i}(k))E(\omega_{s,j}(k))$ ,  $i \neq j$ , which is now nontrivial due to the spatially correlated assumption of the supply functions. For the purpose of simulation of traffic dynamics, it can be achieved by solving the recursive equations by a computer program. Therefore, all the analytical equations will not be detailed due to the limited space. Interested readers can refer to Zhong et al. (2014) for the detailed derivation of the mean and variance equations of traffic dynamics, which incorporates the spatial correlations of supply functions with application to optimal and robust strategies for

freeway traffic management under demand and supply uncertainties.

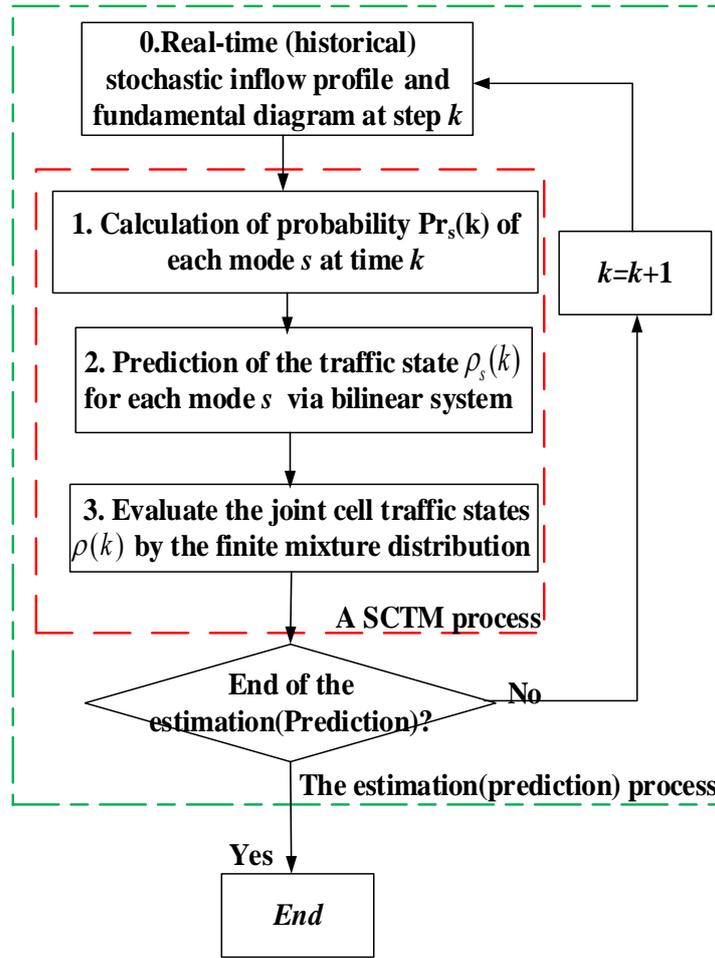


Figure 2.6. The implementation process of the SCTM

## 2.5 Traffic flow modeling of mixed RHVs and CAVs

The number of CAVs equipped with VACS is expected to increase rapidly in the coming decade. Meanwhile, the regular human-piloted vehicles (RHVs) still play a significant role in the market in the coming decade (Levin et al., 2016a). **Therefore, the sharing of the road by CAVs and RHVs will be common in the near future. The penetration of CAVs and VACS may lead to improvements in freeway network performance and traffic flow efficiency.**

Diakaki et al. (2015) stated that VACS with direct traffic flow implications (i.e., adaptive cruising control system (ACC) and cooperative ACC) modifies the

macroscopic traffic flow characteristics. Also, the reduced reaction time of CAVs can improve traffic flow efficiency via the small inter-vehicle headway (Levin et al., 2016a). Diakaki et al. (2015) also claimed that VACS should be able to make control decisions for speed, headway, lane assignment, and lane-change/merge maneuvers of vehicles given certain penetration rate of CAVs and RHVs to improve the safety and efficiency of freeways. CAVs will fully respond to the control while only a certain driver compliance rate will follow the control instruction provided by the conventional traffic advisory system. For example, the advisory system would tell drivers to select proper lanes (e.g., change lanes to left or right, or keep lane; drive with a given speed) from a distance to incident locations on the freeway or their target off-ramps for regular vehicles. Thus, these drivers can respond to an unperceived situation, which allows tactical maneuvers. However, only some drivers will follow such lane control instruction.

Under the car-following framework, Jia and Ngoduy (2016a) developed a microscopic traffic model that simulates the cooperative driving behavior in terms of platooning via inter-vehicle communication. An enhanced cooperative microscopic car-following traffic model was proposed in Jia and Ngoduy (2016b) to consider how V2V and V2I communications affect the vehicle cooperative driving, e.g., the local traffic flow stability and shock wave smoothing. Zhu and Ukkusuri (2017a) proposed a cell-based simulation approach to simulate the proactive driving behavior of CAVs by tracking the trajectory of CAVs and adjusting the exit flow of cells containing connected vehicles. To study the car-following behavior of CAVs following RHVs, Zhu and Ukkusuri (2017b) developed an approach to calibrate CAVs' car following behavior in a mixed traffic environment using the simplified car-following model with disturbances. Zhou et al. (2017) developed a cooperative intelligent driver model to examine the system performance of freeway merging under different proportions of CAVs.

Systematic investigation of the influence of CAVs on collective (macroscopic) traffic flow and sustainability have not been sufficiently understood (Khondaker et al., 2015;

Wang et al., 2015, 2016a, 2016b). The average speed of RHVs is approximately equal to the average speed of connected vehicles in a traffic stream (Bekiaris-Liberis et al., 2016; Fountoulakis et al., 2017). Tampère et al. (2009) proposed a kinetic traffic flow model for analyzing the effect of advanced driver-assistance systems (ADASs) on traffic flow patterns. Simulations reveal the benefit of the ADASs in resulting safer and smoother transition from free-flowing to congested traffic and preventing the emergence of stop-and-go congestion patterns. Calvert et al. (2015) developed a platoon-based first-order macroscopic kinematic wave model to include stochastic vehicle specific behavior and interaction in traffic flow.

Based on scalar conservation law, Bekiaris-Liberis et al. (2016) proposed a Kalman filter approach to estimate the density and flow for traffic comprising regular and connected vehicles using the average speed measurements from connected vehicles and spot-sensor-based total flow measurements. Fountoulakis et al. (2017) further investigated this estimation approach through microscopic simulation while revealing different CF behaviors between these two vehicle classes. Wang et al. (2017) extended the second-order traffic flow model and modeled traffic streams with mixed RHVs and CAVs. Based on the model, traffic state is estimated via a nonlinear particle filtering approach. Levin and Boyles (2016) extended the CTM to simulate the mixed traffic flow with regular and connected vehicles. **This multiclass CTM modified the fundamental diagram to consider the variations in capacity and backward wave speed in response to vehicle-class proportions within each cell while LC maneuvers were not considered.** Chen et al. (2017) proposed traffic operational capacity formulations for mixed traffic under equilibrium. They found that a macroscopic capacity can be properly identified for the mixed traffic by considering the proportion of CAVs, the micro/mesoscopic characteristics of the two vehicle classes, and different lane policies.

As reviewed in this section, a **rather limited research effort** has been dedicated to addressing **the implications of the emerging VACS on the flow characteristics of**

**traffic mixed with RHVs and CAVs, as well as their potential exploitation for improving traffic flow operations** (Diakaki et al., 2015). All these impede the development of trusted and fast traffic flow models for ATM of traffic flow mixed with CAVs and RHVs. Thus, research and development activities can be further developed.

## **2.6 Freeway traffic control for traffic flow mixed with RHVs and CAVs**

Several traffic control schemes exist for freeway traffic. Variable speed limit (VSL) control is a widely studied freeway traffic control scheme for handling dense traffic scenario aside from the ramp-metering control (for not extremely dense traffic). Different VSL control technologies have been devised to improve traffic mobility, safety, and environmental impact. At the macroscopic level, VSL control mainly adjusts the speed limit upstream to a bottleneck that is about to be activated to reduce the mainstream arrival flow and thus retard the bottleneck occurrence (Hegyi et al., 2005a, 2005b; Carlson et al., 2010, 2011).

The fundamental diagram shows that the effect of the speed limit is to change its shape. In Figure 2.7, suppose A is the freeway traffic state. When a speed limit is applied, the free-flow speed decreases (to the dot line with B) and the density increases. Thus, the traffic state will be somewhere between B and C. The excess travel demand will move the traffic state to approach state C (i.e., the capacity of the new fundamental diagram). The critical density under the VSL control is higher than the original, which is uncontrolled. Since this flow is lower than the capacity of the freeway without a speed limit control, some space will be left to accommodate the traffic from the on-ramp and breakdown is prevented. More detailed discussions on-ramp metering and VSL control can be found in Carlson et al. (2010, 2011).

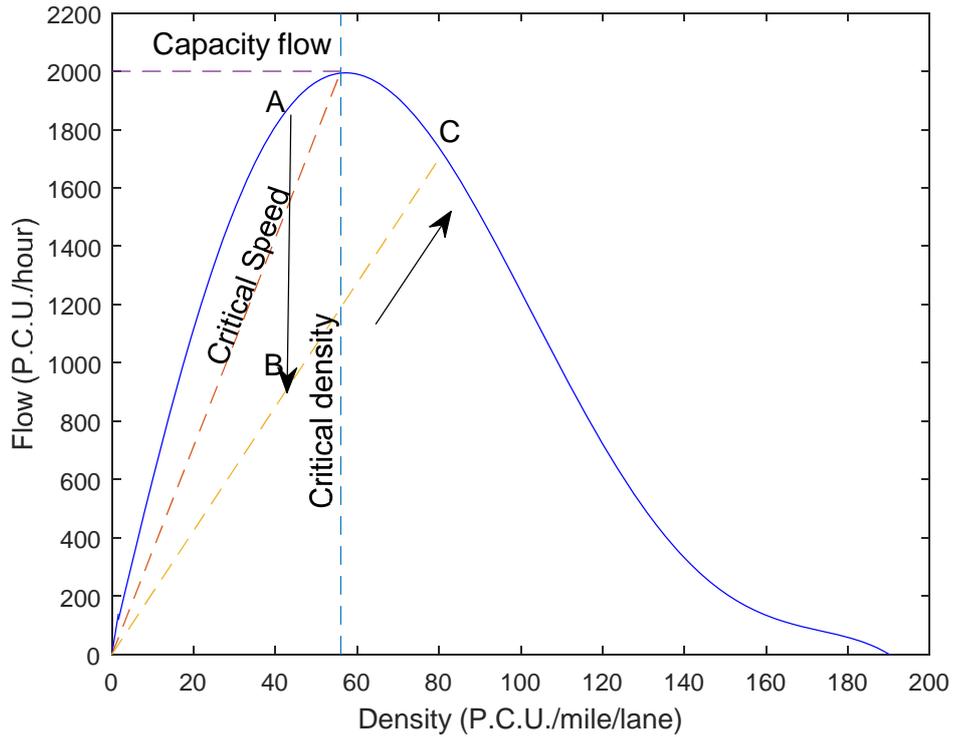


Figure. 2.7. Fundamental diagram of a freeway lane with speed limit control

Some microscopic simulations and field tests suggested that the VSL control techniques cannot generate travel time improvements consistently (see Zhang and Ioannou (2017) and the references therein). The major reasons for this inconsistency were the highly disordered weaving and stochastic traffic conditions at the congested bottlenecks, which might not be captured by conventional macroscopic traffic models as reviewed in Section 2.1.

Recent research has revealed that the weaving sections of a freeway usually admit capacity reductions caused by vehicle LC maneuvers, which have a significant influence on traffic safety because accidents tend to happen in LC areas, such as weaving sections and interchanges. By contrast, a symbiotic relationship between congestion and traffic incidents was reported. Congested traffic condition is one of the main reasons for traffic accidents. Incidents on freeways unexpectedly interrupt traffic flows, and they can be the major cause of unusual bottlenecks and secondary accidents.

These accidents cause even more congestion, which in turn, causes more accidents (Ozbay et al., 1999.).

As discussed in Chapter 1 and the previous sections, the disordered behavior around temporary bottlenecks (caused by incidents) is mainly caused by vehicle LC and the capacity drop phenomenon. This condition renders the traffic dynamics around the bottleneck to be highly unstable. Without considering the LC maneuvers near the incident or bottleneck, the VSL control would have limited effects in improving travel time because the MLC maneuvers would decrease the speed of vehicles in neighboring lanes, whereas the VSL cannot eliminate such capacity drop. To solve this problem, LCC is further required apart from variable speed limit to alleviate congestion at freeway bottlenecks by providing appropriate LC recommendations to upstream vehicles, such that most of the LC maneuvers can be performed away from the incident locations and capacity drop can thus be avoided (Zhang and Ioannou, 2017). Yao et al. (2017) investigated the optimization of traffic flow efficiency through implementing lane changing control to minimize the total travel delay of different vehicle groups on a two-lane freeway stretch. A genetic algorithm was proposed to solve the problems. Through numerical simulation, it is found that vehicles will be suggested to change lanes in bigger gaps to improve collective or group efficiency; while they are supposed to overtake as many vehicles as they can by changing lanes for their benefit.

Several control concepts have been elaborated about traffic control systems with VACS in either a non-cooperative or cooperative manner. Most of these controls pertain to a multilayer structure to address the serious problem complexity, as inspired by Varaiya (1993; see Diakaki et al. (2015) and Roncoli et al. (2016a)). For example, the cruise system automatically adjusts the speed to assist equipped vehicles in following preceding vehicles safely and comfortably. By cooperating the performance of a vehicle-platoon (some vehicles) with V2V communication, the related vehicles can respond smoothly to disruptions in traffic flow and maintain small inter-vehicle gaps. Speed regulation system is another category that assists the regulation of speed

according to legal limits and has recently received considerable attention. For example, the extension of the conventional VSL system that employs I2V communication can disseminate speed limit information, which is determined according to the current speed and position of each vehicle. However, for conventional vehicles, this information was delivered by the variable message sign (VMS) gantry to vehicles upstream to the gantry. Furthermore, LC assistance systems for LC and merging maneuvers (e.g., creating and maintaining an appropriate gap in the target lane) also forms an important category of the VACS.

Diakaki et al. (2015), claimed that VACS should be able to make control decisions regarding speed (e.g., VSL), headway, lane assignment, and LC/merge maneuvers of vehicles to improve the safety and efficiency of the freeway. Although the VSL control has been extensively studied in the literature, Li et al. (2014, 2016) indicated that the following critical issues have not yet been addressed: (1) the optimum placement of the VSL signs/gantries (uniform information) or the personalized VSL information dissemination; (2) the transferability of the VSL control strategy; (3) the effects of driver compliance rate to posted permanent compulsory speed limits on the effectiveness of the VSL control; and (4) the effects of VSL control on safety performance. These issues may deteriorate the performance of VSL control strategies in engineering applications. Nevertheless, any disturbance in the flow (e.g., sudden deceleration, merging, or LC) can create shockwaves in congested traffic, which may result in traffic breakdown.

Microscopic models would be better choices for describing these phenomena. At a microscopic level, the optimal control and model predictive control approaches are developed to regulate traffic speeds through connected VSL gantries and resolve stop-and-go waves at the link level. At the same time, intelligent vehicles control accelerations through vehicle propulsion and brake systems to optimize their local situations for safety performance through CF control (CFC). The disadvantage of this approach is the scalability that the computational burden increases with the number of

cooperative vehicles in the platoon. He et al. (2017) proposed a jam absorbing strategy, which advocates a “slow-in, fast-out” driving strategy, based on a simplified Newell’s car following model. The disadvantage of microscopic approaches is the scalability that the computational burden increases with the number of cooperative vehicles in the platoon. Moreover, the lane-changing maneuvers are not considered in these models.

The macroscopic multilane models provide an alternative balance between single lane macroscopic and microscopic models (Pan et al., 2016). At a macroscopic multilane level, the VSL can be used to reduce speed differences among vehicles traveling in the same lane and adjacent lanes. Given that speed difference is the major incentive for discretionary LC, this reduction in speed differences synchronizes with driver behavior and discourages LC maneuvers, which decreases the probability of collisions. Han et al. (2017) developed the VSL strategies to improve bottleneck discharge rates and reduce system delays using CAVs under various penetration rates. In contrast to the VMS-only strategy, the CAV-based strategies can effectively impose dynamic control over continuous time and space even with a small number of CAVs.

## **2.7. Summary**

As reviewed, most existing KW based approaches for simulating vehicle LC maneuvers typically differentiate the DLC and MLC while developing different models for these two LC behaviors. For example, some models that concentrate on DLC usually lack an MLC component (Laval and Daganzo, 2006) and vice versa (Hou et al., 2015). Hence, relating these models to one another or give a comprehensive model for simulating LC maneuvers at one scale is difficult. As reviewed by Zheng (2014), developing a comprehensive model is needed to capture the influence of MLC and DLC maneuvers on the surrounding traffic while maintaining the balance between maximizing the predictive and explanatory power of the model and minimizing its complexity.

Based on the hybrid CTM, Chapter 3 develops a comprehensive macroscopic multilane traffic model that simulates DLC and MLC maneuvers simultaneously. On the other hand, the proposed model can be used to infer the turning ratios (e.g., off-ramp traffic demand) at ramps of the freeway, as well as the dynamic traffic demand by Origin-Destination (OD) for the freeway system. The previous studies on KW based macroscopic stochastic traffic flow models did not address the issue of vehicle lane-changing. To this end, Chapter 4 extends the earlier works on the SCTM to consider the effects of lane-changing maneuvers on stochastic traffic flow dynamics.

The state-of-the-art modeling frameworks for traffic flow mixed CAVs and RHVs are either single-lane model without considering lane changing maneuvers or did not consider the variations in capacity and backward wave speed in response to vehicle-class proportions. To remedy this, in the first part of Chapter 5, a multilane CTM is proposed to simulate traffic flow dynamics mixed with CAVs and RHVs via capturing the impacts of penetrated CAVs on the freeway traffic characteristics and the lane-changing behaviors considering drivers' anticipation. Built on this multiclass multilane model, an integrated freeway traffic flow control framework that aims to minimize the total travel cost, improve greenness, and ensure safety for freeway traffic mixed with a given penetration rate of CAVs equipped with the VACS and RHVs via en-route VMS is developed.

# Chapter 3 Modeling the impacts of mandatory and discretionary lane-changing maneuvers

## 3.1 Introduction

This chapter develops a comprehensive macroscopic multilane traffic model that simulates discretionary lane-changing (DLC) and mandatory lane-changing (MLC) maneuvers simultaneously. The lane-changing model is comprised of

- 1) A model for propagating macroscopic multilane dynamics which enables:
  - 1.1) Determining the minimum gap acceptance criteria for different lane-changing intentions;
  - 1.2) Evaluating whether or not the traffic flows sent from different directions can be accepted by a downstream section via an extended Incremental-Transfer (IT) and Priority IT (PIT) principle that defines demand-supply reaction laws;
  - 1.3) Propagating traffic states on multi lanes in both temporal and spatial dimensions.
  
- 2) A dynamical lane-changing demand estimation algorithm that includes:
  - 2.1) An estimation of the longitudinal distribution of MLC demand and DLC demand;
  - 2.2) A process to dynamically refine the MLC and DLC demands based on the execution of lane-changing flows.
  
- 3) An empirical study on a complex weaving section of SR-241 freeway in Orange County, California is conducted by using the traffic data provided by the Caltrans Performance Measurement System (PeMS)<sup>2</sup>. Conclusions are drawn and future works

---

<sup>2</sup> The Freeway Performance Measurement System (PeMS: <http://pems.dot.ca.gov/>) is conducted by the University of California, at Berkeley, with the cooperation of the California Department of Transportation (Caltrans), California Partners for Advanced Transit and Highways, and Berkeley Transportation Systems (Chen, 2003).

are then discussed.

### 3.2 Development of the macroscopic multilane cell transmission model

Modeling lane-changing is a key procedure when developing multilane traffic model since heterogeneous traffic flow on multilane freeways is closely related to lane-changing maneuvers. On the other hand, both lane-changing decision making progress and lane-changing impact are continuously affected by the heterogeneous traffic flow on multilane freeway, as well as exogenous elements such as freeway geometry and traffic control schemes. This section presents a model for simulating macroscopic traffic state in temporal, longitudinal and lateral dimensions. As shown in Figure 3.1, the model consists of a traffic flow component governing the lane specific traffic flow propagation and a lane-changing acceptance mechanism based on gap assessment rules. The traffic flow component extends the sending and receiving functions of the CTM based on lane specific fundamental diagrams and lane-changing demand. In addition, levels of urgency determining the lane-changing priority is proposed.

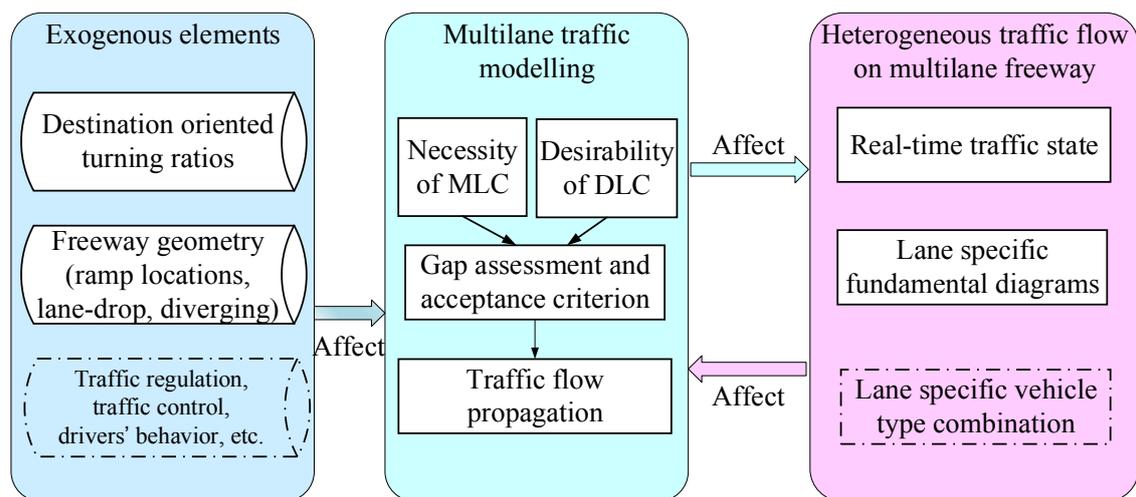


Figure 3.1. A modeling framework of multilane traffic

#### 3.2.1 Lane-specific traffic flow characteristics

Parallel to the conventional macroscopic traffic flow models which adopt a uniform **macroscopic fundamental diagram** (MFD) to propagate the longitudinal dynamics of traffic flow, a properly defined lane specific fundamental diagram would be a better choice for simulation of multilane traffic. In this chapter, the triangular lane specific flow-density equation is adopted:

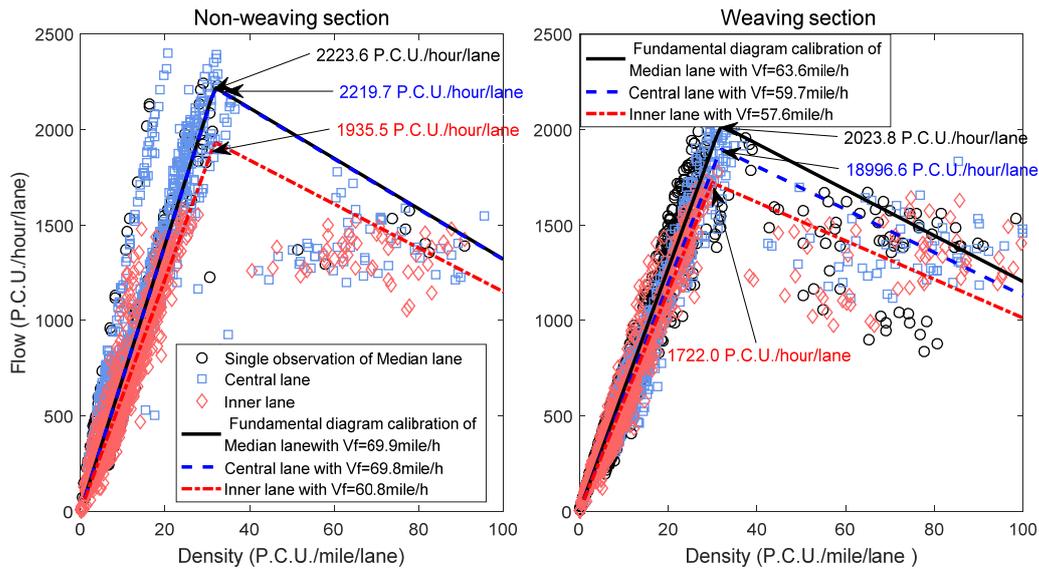
$$q_m(t) = \begin{cases} v_{f,m} \cdot \rho_m(t), & \text{if } \rho_m(t) \leq \rho_{c,m} \\ w_{c,m} \cdot (\rho_{J,m} - \rho_m(t)), & \text{if } \rho_m(t) > \rho_{c,m} \end{cases}, \quad (3.1)$$

in conjunction with a flow-speed-density relationship:

$$v_m(t) = q_m(t)/\rho_m(t) \quad (3.2)$$

where  $\rho_m(t)$  (P.C.U./mile/lane),  $v_m(t)$  (mile/hour) and  $q_m(t)$  (P.C.U./hour/lane) denote the traffic density, speed and flow of lane  $m$  at time  $t$ , respectively. While  $v_{f,m}$  (mile/hour),  $w_{c,m}$  (mile/hour),  $\rho_{c,m}$  (P.C.U./mile/lane) and  $\rho_{J,m}$  (P.C.U./mile/lane) denote free flow speed, wave-back speed of congestion, critical density and jam density of the fundamental diagram of lane  $m$ , respectively. The MFD is built to map the average flow (or speed) to the total effective density, which is a weighted summation over all vehicle type-specific densities with respect to the passenger car unit (P.C.U.) values.

To elaborate the necessity of the lane specific MFDs, historical traffic data is extracted from ten detector stations embedded in a 2.1-mile section on the State Route 241 (SR-241) in Orange County, California from the PeMS database, see Figure 3.8b. The data collected by six detector stations which are beyond 1 mile away from the off-ramp intersection is used to calibrate the MFDs of the non-weaving section (ID 1216230-ID1216220 in Figure 3.8b), while the data by the detectors within 1 mile from the intersection is adopted to calibrate the MFDs of the weaving section (ID 1211425-ID1211586 in Figure 3.8b). Details of the test site will be outlined in the empirical study. Based on the lane specific vehicle count and occupancy observed on March 20, 2014 and June 2, 2015, a set of MFDs were calibrated by the method of least squares in line with the literature.



a). Observed and calibrated on non-weaving section

b). Observed and calibrated on weaving section

Figure 3.2. Lane specific MFDs

First of all, it is shown in Figure 3.2 that the MFDs is very different from one lane to another lane in terms of free-flow speed, capacity, etc. on non-weaving and weaving sections respectively. This is because of the heterogeneity occurs in diver behaviors (due to merging/diverging surrounding on/off ramps, lane specific speed limit), vehicle type composition and the corresponding vehicle-class specific control, e.g., driving ban for trucks<sup>3</sup> on the median lane. This heterogeneity indicates that a properly defined lane specific MFD is essential to consider the effect of multiclass traffic and vehicle-class specific control, while the lane specific MFD would be an effective calibration and validation tool for a multilane traffic model. Since vehicle classes and the corresponding vehicle-class specific control on different lanes significantly affect the

---

<sup>3</sup> For example, General Laws of Massachusetts-Chapter89-Section 4C: On any highway with more than one passing lane in the same direction, heavy commercial vehicles, except buses, shall be restricted in ordinary operation to the right-hand travel lane, and in overtaking and passing shall be restricted to the next adjacent passing or travel lane, and shall not use any other lanes except in an emergency.

MFD, it is not surprising that the lane specific MFD will be also beneficial to capture the impact of lane-based traffic control strategies on traffic patterns (Duret et al., 2012). The regression lines of the MFD calibrated for traffic at the central lane, and the median lane almost coincide with each other for the non-weaving area. This indicates that the central lane and the median lane on the non-weaving area (between ID 1216230-ID1216220 as shown in Figure 3.8b) of the studied section on SR-241 share similar traffic flow characteristics as they are far from the weaving areas (between ID 1211425-ID1211586 as shown in Figure 3.8b). Some data samples are not close to the regression line due to data scattering. Data scattering of loop detectors is a typical problem in calibrating fundamental diagrams for freeways, see, e.g. Sumalee et al. (2011), Pan et al. (2013), and Zhong et al. (2015) and the references therein. Due to the interruptions of MLC maneuvers, free-flow speed reduction and the capacity reduction could be observed from the MFDs calibrated on weaving sections (Figure 3.2b) compared with the MFDs of the corresponding lanes on non-weaving section (Figure 3.2a) (e.g., (Cassidy and Bertini, 1999)). For example, a free-flow speed reduction around 10-20 km/hour can be observed from the weaving section. The capacity of central lane reduces more than 300 P.C.U./hour in Figure 3.2b), wherein the FDs are zoomed to highlight the lane heterogeneity and capacity reduction.

### **3.2.2 Determination of minimum gap acceptance criterion**

The execution of each individual lane-changing maneuver is a trade-off by taking into account necessity, desirability and possibility of all potential lane-changing maneuvers (Zheng et al., 2013, Gipps, 1986), see e.g., Figure 3.3. Necessity reflects the level of lane-changing urgency, which is related to the distance between the subject vehicle which is proposing a MLC maneuver and its target turning point. Desirability is induced by the speed advantage possessed by adjacent vehicles compared with the current speed of the subject vehicle which is proposing a DLC maneuver. Finally, both the necessity

of MLC and desirability of DLC confront the same issue, i.e., whether the prospective space gaps provided by target lane are long enough to ensure a safe lane-changing maneuver.

In Yang and Koutsopoulos (1996), a minimum gap acceptance criterion is first introduced as a benchmark to assess whether the adjacent space gaps on target lane are large enough to accommodate a subject vehicle that intends to change lane with MLC demand. The benchmark involves both the lead gap and lag gap as shown in Figure 3.3. The lead gap denotes the space between the rear of leading vehicle and the front of the subject vehicle while the lag gap denotes the space between the front of following vehicle and the rear of the subject vehicle. In the literature, the minimum lead (or lag) gap acceptance criterion is determined by several factors, such as the lane-changing type, the travelling speeds of the lane-changing vehicles, the level of lane-changing urgency, the number of lanes to be crossed, the length and mechanical characteristics of subject vehicles, drivers' attitudes such as patience or aggression, *etc.* (Gipps, 1986, Ahmed, 1999, Hidas, 2005). In this chapter, in line with the macroscopic nature of the proposed traffic flow model and the within cell homogenous assumption of the CTM, the lead gap, lag gap, and subject vehicle length are aggregated as a minimum acceptance criterion for lane-changing acceptance assessment, as depicted in Figure 3.3.

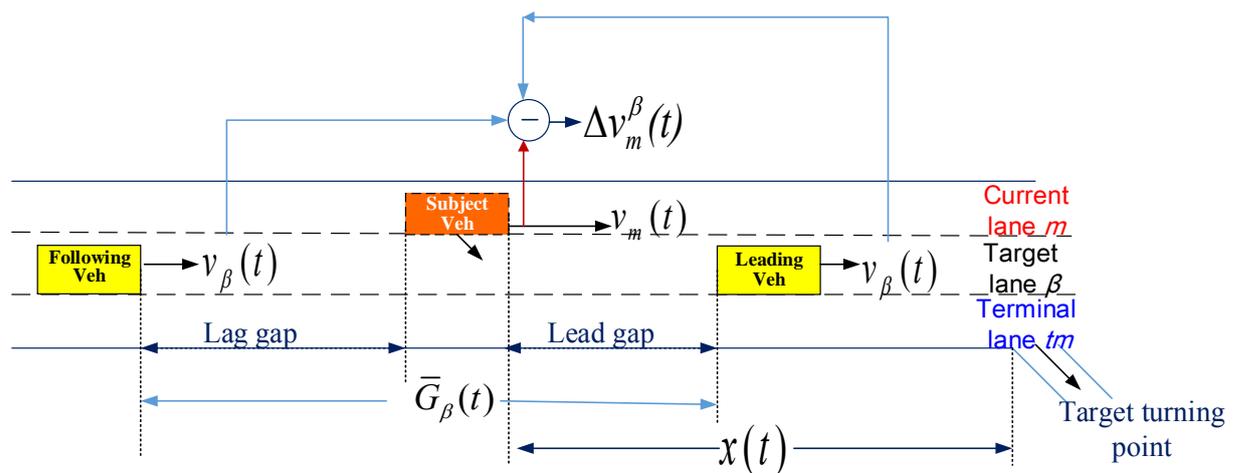


Figure 3.3. Lead gap, lag gap, and adjacent space gap.

The subject vehicle traveling on current lane  $m$  makes its final lane-changing decision towards the target lane  $\beta$  by comparing the minimum gap acceptance criterion

$\tilde{g}_{LC,m}^\beta(t)$  (feet) and the actual space gap  $\bar{G}_\beta(t)$  (feet) provided by the target lane. For the subject vehicle with DLC intention, the minimum gap acceptance criterion  $\tilde{g}_m^\beta(t)$  is determined by the minimal safe gap and the speed difference  $\Delta v_m^\beta(t)$  (mile/hour) between the current lane speed  $v_m(t)$  and the target lane speed  $v_\beta(t)$ <sup>4</sup>.

Apart from gap acceptance criterion and lane speed difference, the subject vehicle with MLC intention is also affected by the remaining distance. The remaining distance  $x(t)$  (mile) refers to the distance from the current position of the subject vehicle to its target off-ramp intersection or downstream traffic incident spot. This distance is directly related to the driver's assessment of the level of urgency associated with her/his MLC intention (Gipps, 1986). Based on Yang and Koutsopoulos (1996), the level of urgency is considered to follow three sequential stages as the vehicle approaches its target turning point. These three stages are separated by two critical positions  $x_r$  and  $x_c$ . Take a vehicle intending to execute a MLC as an example, the target turning point is considered to be **remote** as long as the remaining distance  $x(t) > x_r$ , and **close** if  $x(t) < x_c$ , where  $x_r$  and  $x_c$  are distances that define the range within which the minimum gap acceptance criterion  $\tilde{g}_{MLC,m}^\beta(t)$  linearly varies from the upper bound to the lower one based on current speed difference  $\Delta v_m^\beta(t)$ :

$$\tilde{g}_{MLC,m}^\beta(t) = \begin{cases} c_l \cdot [(v_m(t) - v_\beta(t))] + c_f \cdot [(v_\beta(t) - v_m(t))] + \bar{g}_{\min}, & \text{if } x(t) > x_r \\ \{c_l \cdot [(v_m(t) - v_\beta(t))] + c_f \cdot [(v_\beta(t) - v_m(t))]\} \cdot \frac{x(t) - x_c}{x_r - x_c} + \bar{g}_{\min}, & \text{if } x_c \leq x(t) \leq x_r \\ \bar{g}_{\min}, & \text{if } x(t) < x_c \end{cases} \quad (3.3)$$

---

<sup>4</sup> In practice, the speeds of leading vehicle and following vehicle are different, and such difference might be either slight or significant, depending on the space gap length between them and the traffic condition. This simplification is due to the within cell homogenous assumption of the CTM.

where the operational symbol  $[z]$  is defined as below:

$$[z] = \begin{cases} z, & \text{if } z > 0 \\ 0, & \text{if } z \leq 0 \end{cases}$$

The constant  $\bar{g}_{\min}$  (feet) is the minimal safe gap to be provided by the target lane for the subject vehicle.  $c_l$  and  $c_f$  (feet·hour/mile) are constant presenting the relationship between speed difference with extra lead gap and extra lag gap, when the subject vehicle is still not too aggressive to execute a lane change.

As shown in Equation (3.3), in the remote stage, unless the speed difference between related lanes under consideration is negligible, i.e.,  $\Delta v_m^\beta(t) = v_\beta(t) - v_m(t) \approx 0$ , the driver usually prefers a relatively larger gap at the commencement of lane-changing maneuver. Such a large gap is due to a risk-adverse attitude under the non-urgent situation. In the second stage, the minimum gap acceptance criterion decreases linearly with respect to the reduction of remaining distance from  $x_r$  to  $x_c$ . Finally, after passing the critical distance point  $x_c$ , i.e., the vehicle gets so close to the target turning point that the MLC has to be executed, the minimum gap acceptance criterion  $\tilde{g}_{MLC,m}^\beta(t)$  achieves the lowest value which can just assure the subject vehicle can be safely accommodated by its target lane.

The decrease in the minimum gap acceptance criterion in the last two stages is based on the driver's expectation that the following vehicle in target lane will passively (or voluntarily) slow down to increase the gap before or after the subject vehicle (with MLC intention) enters (Hidas, 2005, Yang and Koutsopoulos, 1996). However, for vehicles with DLC intention, the minimum gap acceptance criterion  $\tilde{g}_{DLC,m}^\beta(t)$  does not decrease as MLC does, because DLC is not related with the level of urgency. In conclusion, a trade-off between minimum gap acceptance criterion and adjacent gap length is the prerequisite in determining the acceptance/rejection result for each lane-changing intention, no matter how the minimum gap acceptance criterion varies with the level of lane-changing urgency or speed difference. On average it would take 2-4

seconds (which will be adopted in the empirical study) for a driver to execute a lane change when the origin lane is stopped and the target lane is freely flowing, as indicated by empirical studies in Laval and Daganzo (2006), Laval and Leclercq (2008), Jin (2010b).

### **3.2.3 Sending function and lane-changing fraction**

The continuum kinematic wave theory (KW) was first applied to model the multilane dynamics by Munjal and Pipes (1971), and was extended to a discrete version known as hybrid CTM by Laval and Daganzo (2006). These KW multilane models are capable of modeling the macroscopic dynamics by balancing the flow intended to merge and the available space provided by the target cell using incremental-transfer (IT) theory. However, there are some unmodeled elements would potentially hinder the previous models representing real traffic:

1. The heterogeneity in the lane specific fundamental diagrams is not considered.
2. The IT theory does not consider the minimum gap acceptance criterion.
3. Unacceptable small gaps scattered among vehicles might be aggregated and considered as an acceptable space by using conventional macroscopic multilane traffic flow models based on IT principle.
4. No account is taken to model different levels of priority associated with (MLC and DLC ) lane-changing intentions and straightforward flows when all of these streams are expecting to merge into the same lane within a prescribed short section.

The foregoing two subsections intend to address the first three issues while this and the forthcoming sections try to handle the last one.

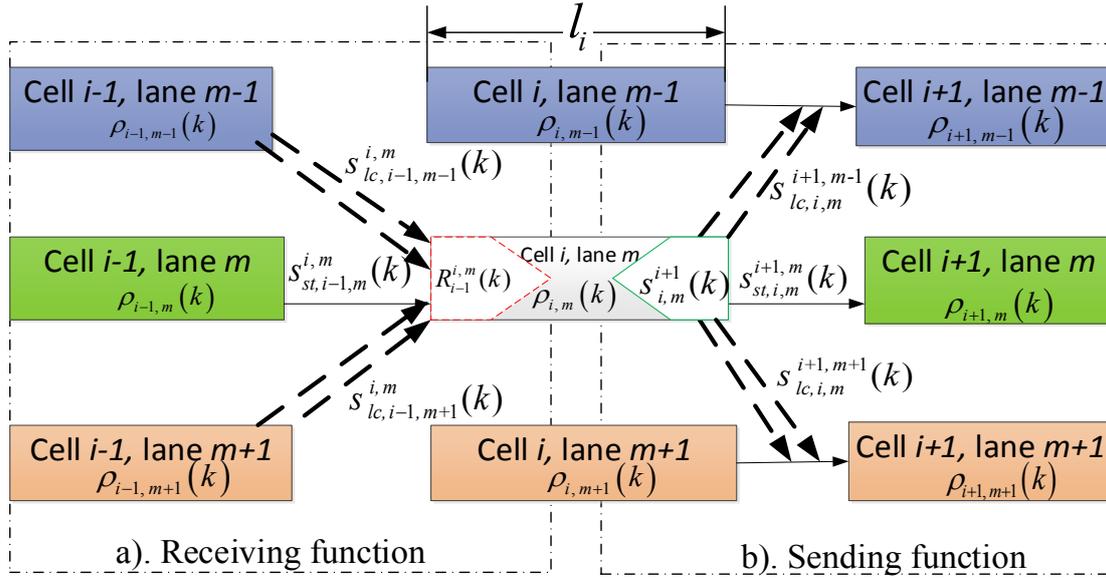


Figure 3.4. Merging and diverging induced by lane-changing maneuvers.

To begin with, consider a multilane freeway section shown in Figure 3.4 which is divided into several cell packages, such as  $i-1$ ,  $i$  and  $i+1$ , along the longitudinal dimension. Each cell package includes multiple cells associated with lanes ranging from the shoulder to the median lane, such as  $m-1$ ,  $m$ , and  $m+1$ . In line with the hybrid CTM, it is assumed that lane-changing events may occur only at the upstream and/or downstream boundaries of a cell. Based on the lane specific fundamental diagrams, the sending function  $s_{i,m}^{i+1}(k)$  (P.C.U./hour/lane) intending to leave cell  $i$  on lane  $m$ , which will be denoted as *cell* ( $i, m$ ) hereafter, during simulation time interval  $[kT_s, (k+1)T_s)$  can be calculated as:

$$s_{i,m}^{i+1}(k) = \begin{cases} v_{f,i,m} \cdot \rho_{i,m}(k), & \text{if } \rho_{i,m}(k) < \rho_{c,i,m} \\ Q_{i,m}, & \text{if } \rho_{i,m}(k) \geq \rho_{c,i,m} \end{cases} \quad (3.4)$$

where  $\rho_{i,m}(k)$  (P.C.U./mile/lane) denotes the dynamic traffic density of *cell* ( $i, m$ ) estimated on time step  $k$ , and  $T_s$  denotes the duration of a simulation time step.  $s_{i,m}^{i+1}(k)$  would move towards all possible directions associated with *cell package* ( $i+1$ ), i.e., *cell* ( $i+1, m-1$ ), *cell* ( $i+1, m$ ), and *cell* ( $i+1, m+1$ ). This algorithm of multilane sending function extends the sending function algorithm developed for single lane section in Daganzo (1994). In this chapter, the subscripts of variables denote the lane change type, source cell and lane, while the superscripts denote the corresponding target

cell and lane, respectively. According to the types of lane-changing intentions and the target cells, the total sending function  $s_{i,m}^{i+1}(k)$  is composed of the following five items:

$$s_{i,m}^{i+1}(k) = \sum_{lc=MLC,DLC} \sum_{\beta=m-1,m+1} s_{lc,i,m}^{i+1,\beta}(k) + s_{st,i,m}^{i+1,m}(k), \quad (3.5)$$

where  $lc$  denotes MLC and DLC, respectively, whilst  $\beta = m \pm 1$  refers to the two adjacent lanes as illustrated in Figure 3.4b.  $s_{lc,i,m}^{i+1,\beta}(k)$  denotes the lane-changing demand be of type  $lc$  that intends to leave *cell* ( $i, m$ ) and be sent towards *cell* ( $i+1, \beta$ ) during the time interval  $[kT_s, (k+1)T_s)$ , while  $s_{st,i,m}^{i+1,m}(k)$  denotes the straightforward demand that intends to leave *cell* ( $i, m$ ) to *cell* ( $i+1, m$ ). An algorithm for estimation of these lane-changing demands will be devised in Section 3.3. On the other hand, these sending functions might not be totally or even partially received when considering the gap and the available space provided by target cells. The corresponding receiving algorithm will be devised in the following section.

### 3.2.4 Allocation of spaces for merging flows and dynamics propagation

In view of the disadvantages of the IT principle as previously discussed, an acceptance/rejection mechanism of lane-changing demand is introduced in this subsection. Considering *cell* ( $i, m$ ) as the target cell as illustrated in Figure 3.4a, the average space gap (*feet*) between successive vehicles in *cell* ( $i, m$ ) at time step  $k$  is estimated based on the *within cell homogenous assumption* of the CTM:

$$\bar{G}_{i,m}(k) = \frac{5280 \cdot l_i - \rho_{i,m}(k) \cdot l_i \cdot L_p}{\rho_{i,m}(k) \cdot l_i}, \quad (3.6)$$

where  $L_p$  (*feet*) represents the vehicle length of Passenger Car Unit (*P.C.U.*), and  $l_i$  (*mile*) denotes the longitudinal cell length of *cell package* ( $i$ ). This equation quantifies the average gap length between successive vehicles traveling on *cell* ( $i, m$ ) by excluding the space occupied by vehicles in current cell ( $1 \text{ mile} = 5280 \text{ feet}$ ). Based on this gap and the gap acceptance rule, the sending function  $s_{lc,i-1,\alpha}^{i,m}(k)$  might be

received totally or partially by *cell* ( $i, m$ ) if  $\bar{G}_{i,m}(k) \geq \tilde{g}_{lc,i-1,\alpha}^{i,m}(k)$ , where  $\tilde{g}_{lc,i-1,\alpha}^{i,m}(k)$  is the minimum gap acceptance criterion required by this stream of sending function  $s_{lc,i-1,\alpha}^{i,m}(k)$ , and  $\tilde{g}_{lc,i-1,\alpha}^{i,m}(k)$  can be estimated by Equation (3.3). Otherwise, the sending flow will keep the current lane.  $\alpha = m \pm 1$  refers to the two adjacent lanes from which the subject vehicles intend to travel towards *cell* ( $i, m$ ) as illustrated in Figure 3.4a. Since a lane-changing vehicle generally requires a larger space than a straightforward vehicle which is not intended to change lane, to quantify the space (in terms of *P.C.U.*) required by each stream of the lane-changing demand, the minimum gap acceptance criterion  $\tilde{g}_{lc,i-1,\alpha}^{i,m}(k)$  is further normalized to be the relative minimum gap acceptance criterion factor using Equation (3.7) as follows:

$$\tilde{\Phi}_{lc,i-1,\alpha}^{i,m}(k) = \frac{\tilde{g}_{lc,i-1,\alpha}^{i,m}(k)}{L_p} \quad (3.7)$$

As it can be inferred from Equation (3.3),  $\tilde{\Phi}_{lc,i-1,\alpha}^{i,m}(k) \geq \frac{\bar{g}_{\min}}{L_p} > \frac{L_p}{L_p} = 1$ . In the conventional CTM, the sending demands might not be totally received by the downstream section due to the limited available space (Daganzo, 1994, Munoz et al., 2003). Extending this to the multilane model, whether the sending functions to be sent from different directions, such as *cell* ( $i-1, m+1$ ), *cell* ( $i-1, m$ ), and *cell* ( $i-1, m-1$ ), can be received by the target *cell* ( $i, m$ ) also depends on its available space defined by the receiving function, on the condition that the minimum gap acceptance criterion is fulfilled. Considering the lane-specific fundamental diagram depicted in Figure 3.1, the receiving function of the target *cell* ( $i, m$ ),  $R_{i-1}^{i,m}(k)$  (*P.C.U./hour*) during the time interval  $[kT_s, (k+1)T_s)$  is given as follow:

$$R_{i-1}^{i,m}(k) = \begin{cases} Q_{i,m}, & \text{if } \rho_{i,m}(k) < \rho_{c,i,m} \\ w_{c,i,m} \cdot (\rho_{j,i,m} - \rho_{i,m}(k)), & \text{if } \rho_{i,m}(k) \geq \rho_{c,i,m} \end{cases} \quad (3.8)$$

The receiving function  $R_{i-1}^{i,m}(k)$  quantify the available space provided by *cell* ( $i, m$ ), because this cell is proposed to allocate both the straightforward flow  $s_{st,i-1,m}^{i,m}(k)$  and the lane-changing flows  $s_{lc,i-1,\alpha}^{i,m}(k)$  (with  $\alpha=m \pm 1$ , and  $lc=MLC, DLC$ ) to be sent

from upstream cells on adjacent lanes. The flow propagation rule can be defined by extending the IT principle considering the gap acceptance criterion. To be specific:

$$\begin{aligned}
q_{lc,i-1,\alpha}^{i,m}(k) &= \begin{cases} s_{lc,i-1,\alpha}^{i,m}(k), & \text{if } U_{i-1}^{i,m}(k) \leq R_{i-1}^{i,m}(k) \ \& \ \tilde{g}_{lc,i-1,\alpha}^{i,m}(k) \leq \bar{G}_{i,m}(k) \\ \frac{s_{lc,i-1,\alpha}^{i,m}(k)}{U_{i-1}^{i,m}(k)} R_{i-1}^{i,m}(k), & \text{if } U_{i-1}^{i,m}(k) > R_{i-1}^{i,m}(k) \ \& \ \tilde{g}_{lc,i-1,\alpha}^{i,m}(k) \leq \bar{G}_{i,m}(k) \\ 0, & \text{if } \tilde{g}_{lc,i-1,\alpha}^{i,m}(k) > \bar{G}_{i,m}(k) \end{cases} \\
q_{st,i-1,m}^{i,m}(k) &= \begin{cases} s_{st,i-1,m}^{i,m}(k), & \text{if } U_{i-1}^{i,m}(k) \leq R_{i-1}^{i,m}(k) \\ \frac{s_{st,i-1,m}^{i,m}(k)}{U_{i-1}^{i,m}(k)} R_{i-1}^{i,m}(k), & \text{if } U_{i-1}^{i,m}(k) > R_{i-1}^{i,m}(k) \end{cases}
\end{aligned} \tag{3.9}$$

where  $q_{lc,i-1,\alpha}^{i,m}(k)$  and  $q_{st,i-1,m}^{i,m}(k)$  denote the flows actually received by *cell*  $(i, m)$  from the lane-changing demands and straightforward demand, respectively. Note that in Equation (3.9), the sending flows need to compete for the downstream supply of *cell*  $(i, m)$  when the available space of the target cell is insufficient to accommodate all the qualified the sending functions.

Regarding the minimum gap acceptance criterion required by different movements as the ‘‘priority levels’’ in the conventional IT principle, the competitive strength of  $s_{lc,i-1,\alpha}^{i,m}(k)$  for the supply of the target *cell*  $(i, m)$  can be evaluated by the first term of Equation (3.9), wherein  $U_{i-1}^{i,m}(k)$  is thus defined:

$$U_{i-1}^{i,m}(k) = s_{st,i-1,m}^{i,m}(k) + \sum_{lc=MLC,DLC} \sum_{\alpha=m\pm 1} s_{lc,i-1,\alpha}^{i,m}(k) \cdot \tilde{\phi}_{lc,i-1,\alpha}^{i,m}(k), \tag{3.10}$$

$\forall s_{lc,i-1,\alpha}^{i,m}(k)$

with  $\tilde{g}_{lc,i-1,\alpha}^{i,m}(k) \leq \bar{G}_{i,m}(k)$ . As that the space to be occupied by each lane changing vehicle which composing  $s_{lc,i-1,\alpha}^{i,m}(k)$  is  $\tilde{\phi}_{lc,i-1,\alpha}^{i,m}(k)$  times than the space to be occupied by straightforward vehicle which composing  $s_{st,i-1,m}^{i,m}(k)$ , in Equation (3.10) the straightforward flow is not scaled by any factor. Finally, the density  $\rho_{i,m}(k+1)$  (*P.C.U./mile*) of *cell*  $(i, m)$  is evaluated according to the flow conservation equation of

the hybrid CTM as follow:

$$\begin{aligned} \rho_{i,m}(k+1) = \rho_{i,m}(k) + \frac{T_s}{l_i} & \left( q_{st,i-1,m}^{i,m}(k) + \sum_{lc=MLC,DLC} \sum_{\alpha=m\pm 1} q_{lc,i-1,\alpha}^{i,m}(k) \right) \\ & - \frac{T_s}{l_i} \left( q_{st,i,m}^{i+1,m}(k) + \sum_{lc=MLC,DLC} \sum_{\beta=m\pm 1} q_{lc,i,m}^{i+1,\beta}(k) \right) \end{aligned} \quad (3.11)$$

and the evaluation is operated overall cell package on all lanes on the test section.

To sum up, a flowchart of the proposed dynamic macroscopic multilane CTM is presented in Figure 3.5. The proposed model enables the traffic state propagation in temporal (the iteration of  $k$ ), longitude (the iteration of  $i$ ), and lateral (the iteration of  $m$ ) dimensions. Firstly, the sending function of a cell is determined according to the lane-specific fundamental diagrams. For lane-changing flows, the minimum gap acceptance criterion associated with levels of lane-changing urgency are also defined. Next, considering a specific lane-changing intention, a gap acceptance/rejection mechanism is developed for assessing the relationship between the minimum gap acceptance criterion and the average space gap provided by the target cell to determine whether this lane-changing intention could be executed, as depicted in the purple box in Figure 3.5. To quantify the space required by a lane-changing vehicle in terms of P.C.U., the relative minimum gap acceptance criterion factor is introduced<sup>5</sup>. In line with the conventional IT principle, this factor can be regarded as “adverse-priority” allocated to the related lane-changing demand when competing for the downstream supply. This is consistent with a general traffic rule that lane-changing vehicle should not harm the benefit of the following vehicle on the target lane. Finally, a flow conservation equation is used to calculate traffic density.

---

<sup>5</sup> To be pointed, the evaluation of sending function, minimum gap acceptance criteria, average space gap and receiving function are all evaluated independently based on the estimated real-time traffic states.

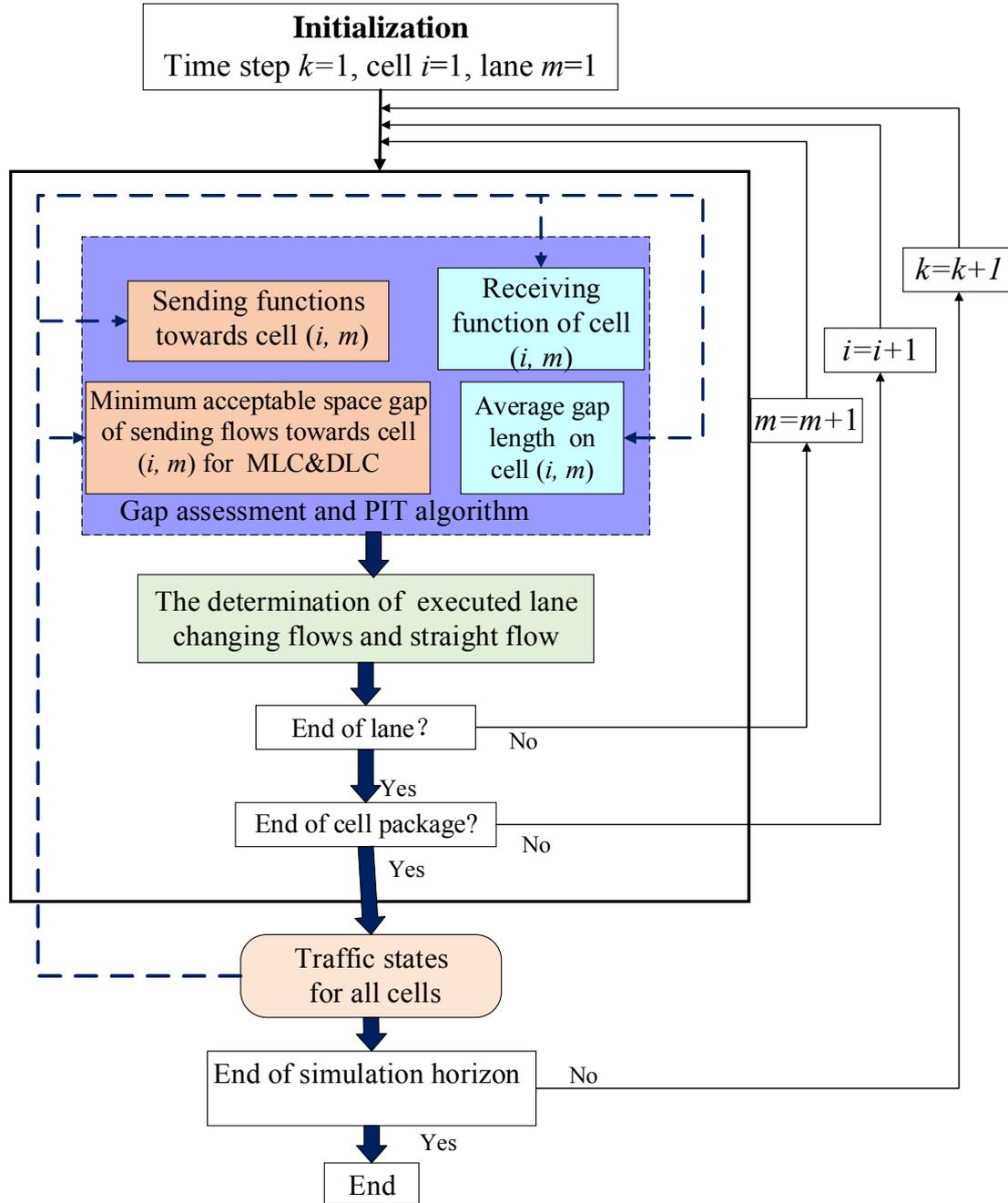


Figure 3.5. Flowchart of dynamic multilane cell transmission model.

### 3.3 The lane-changing demand estimation algorithm

Note from the model development in Section 3.2 that the proposed model presumes each lane-changing demand, say  $s_{lc,i-1,\alpha}^{i,m}(k)$  or  $s_{lc,i-1,m}^{i,\beta}(k)$  as mentioned in Section 3.2, is given regardless of its moving direction and lane-changing category. This is

possible if the trajectory of each individual vehicle can be measured. However, measuring the trajectory of each individual vehicle is barely accessible for large scale applications in reality. On the other hand, even the benchmark vehicle trajectory database provided by the NGSIM project is subject to kinds of errors as previously mentioned. Acknowledging the lane-changing demand cannot be directly measured in general, it is necessary to establish an algorithm to simultaneously estimate the time-dependent MLC and DLC lane-changing demands based on the aggregated dynamic traffic state and historical data. Furthermore, dynamic lane-changing demand estimation is also important for the adjustment of lane-changing flows considering the space gap acceptance/rejection mechanism.

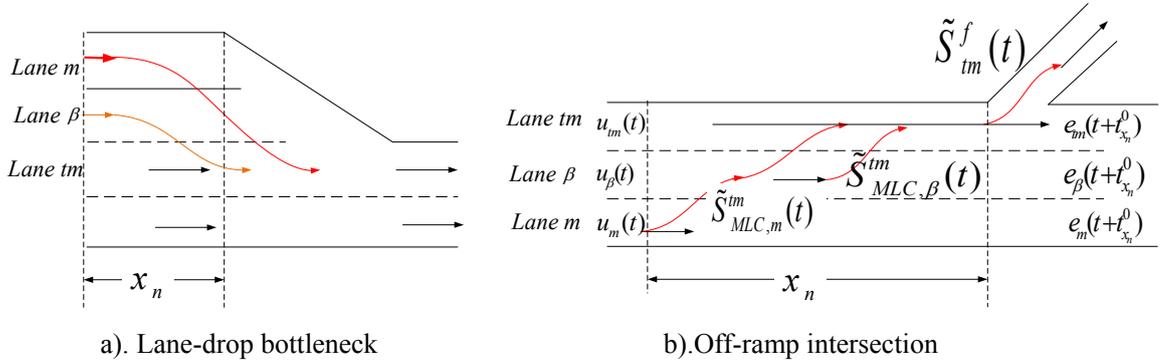


Figure 3.6. The total MLC demand under different scenarios.

### 3.3.1 Estimation of the initial value of the total MLC demand

The MLC demand depends on the topology of the weaving section as shown in Figure 3.6. Bottlenecks due to lane drop is a typical scenario investigated in the literature as shown in Figure 3.6a. Under this scenario, the traffic flow traveling on lane  $m$  must cross over lane  $\beta$  before arriving at the bottleneck. In this section, lane  $tm$  denotes the terminal lane of a specific MLC maneuvers, while lane  $\beta$  denotes the lane between lane  $m$  and lane  $tm$ , and adjacent to lane  $m$ . The total MLC demand  $\tilde{S}_{MLC,m}^{tm}(t)$  to be sent from lane  $m$  to lane  $tm$ , and also  $\tilde{S}_{MLC,\beta}^{tm}(t)$  to be sent from lane  $\beta$  to lane  $tm$ , can be directly deduced from the dynamic traffic flow on lane  $m$  and lane  $\beta$ .

For the off-ramp intersection scenario shown in Figure 3.6b, the initial value of total MLC demand associated with related lanes can be obtained by solving a system of linear equations given by (3.12) based on the flow conservation. The solution of the variables  $\tilde{S}_{MLC,m}^{tm(0)}(t)$  and  $\tilde{S}_{MLC,\beta}^{tm(0)}(t)$  can be identified as the initial values of total MLC demand sent by lane  $m$  and lane  $\beta$ , while  $\tilde{S}_{tm}^f(0)(t)$  denotes the estimated sending flow towards the off-ramp. Let iteration number  $r=0$ , the lane specific flow conservation is presented as below:

$$\begin{aligned}\tilde{S}_{MLC,m}^{tm(r)}(t) + e_m(t + t_{x_n}^0) &= u_m(t) \\ \tilde{S}_{MLC,\beta}^{tm(r)}(t) + \tilde{S}_{MLC,m}^{tm(r)}(t) + e_\beta(t + t_{x_n}^0) &= u_\beta(t) + \tilde{S}_{MLC,m}^{tm(r)}(t) \\ \tilde{S}_{tm}^f(r)(t) + e_{tm}(t + t_{x_n}^0) &= u_{tm}(t) + \tilde{S}_{MLC,m}^{tm(r)}(t) + \tilde{S}_{MLC,\beta}^{tm(r)}(t)\end{aligned}\quad (3.12)$$

Note that the values of  $u_i(t)$  (with  $i=m, \beta$ , or  $tm$ ),  $e_j(t + t_{x_n}^0)$  (with  $j=m, \beta$ , or  $tm$ ) can be regarded as the measured traffic flow rate if there happen to have detectors or the simulated flows by other simulation blocks at these boundaries, as it will be shown in the empirical study.  $x_n$  is the farthest remaining distance away from the target turning point (which is regarded as the reference point whose remaining distance is 0) where the MLC has just been activated, and  $t_{x_n}^0$  is the average time spent by the vehicles to travel from  $x_n$  to 0 (location of target turning point) via all the three lanes. Similar to the consistent check of the boundary conditions in the CTM (Munoz et al., 2003), the values of the total MLC demand still need to be dynamically adjusted by considering the DLC maneuvers to adapt to the current traffic condition. Meanwhile, there is a need to allocate this total demand along a lane to each cell till it arrives the desired turning point. The forthcoming sections will address these issues.

### 3.3.2 Longitudinal distribution of MLC and DLC demands

As depicted in Figure 3.6b, assume the traffic platoon  $u_m(t)$  passes a location denoted  $x_n$  on lane  $m$  at time  $t$ , and a part of  $u_m(t)$  say  $\tilde{S}_{MLC,m}^{tm(r)}(t)$  would like to switch to its

terminal lane  $tm$ , which is connected to the desired off-ramp. The cumulative MLC demand from lane  $m$  to the first target lane  $\beta$  ( $\beta = m \pm 1$ ) follows an exponential distribution with respect to the remaining distance from  $x_n$  to another downstream location  $x$ , which is defined by  $L_{MLC,m}^{\beta(r)}(t, x)$  (P.C.U./hour) as follows with  $x \in [0, x_n]$ :

$$L_{MLC,m}^{\beta(r)}(t, x) = \begin{cases} e^{\left(-[x-x_c]^2/(\sigma_m^\beta(t))^2\right)} \cdot \tilde{S}_{MLC,m}^{tm(r)}(t), & \text{if } x > x_c \\ \tilde{S}_{1,m}^{tm(r)}(t), & \text{if } x \leq x_c \end{cases} \quad (3.13)$$

$$\sigma_m^\beta(t) = \alpha_0 + \alpha_1 \cdot N_m^{tm} + \alpha_2 \cdot \bar{\rho}_\beta(t)$$

where  $x_c$  is the critical remaining distance, at which location all the related drivers have to execute their MLC maneuvers (Yang and Koutsopoulos, 1996), was introduced in Section 3.2.2. The function  $\sigma_m^\beta(t)$  reflects the level of urgency related to the lane change decision making mechanism of  $\tilde{S}_{MLC,m}^{tm(r)}(t)$  by switching from the initial lane  $m$  to lane  $\beta$ . This function is affected by two variables, i.e.,  $N_m^{tm}$  and  $\bar{\rho}_\beta(t)$ , where  $N_m^{tm}$  denotes the number of lanes to be crossed from current lane  $m$  to terminal lane  $tm$  on the lateral dimension,  $\bar{\rho}_\beta(t)$  denotes the average density of lane  $\beta$  along the longitudinal dimension from  $x_n$  to the target turning point 0 at time  $t$ , and  $\alpha_0$ ,  $\alpha_1$  and  $\alpha_2$  are the related parameters.

Equation (3.13) is a simplification of the exponential probabilistic distribution of cumulative MLC demand proposed in Yang and Koutsopoulos (1996). In their original definition, it is assumed that the driver can predict the short-term future traffic condition from her/his current location at the very moment till the time she/he reaches the destination ( $x=0$ ). However, this is unlikely to be true especially under congested traffic condition. This renders the original model difficult to be calibrated. By Equation (3.13), drivers only use the current traffic condition spatially ahead that they could perceive to make their lane-changing decisions. The effect of the perceived traffic condition decreases exponentially with respect to the distance away from her/his current position. While the calibration method for parameters  $\alpha_0$ ,  $\alpha_1$  and  $\alpha_2$  is not introduced in

Yang and Koutsopoulos (1996), there is a need to introduce a curve fitting based algorithm to calibrate these parameters in the empirical study, see section 3.4.2.

Similarly, the longitudinal cumulative probability distribution of MLC demand from lane  $\beta$  to lane  $tm$ , with respect to the remaining distance from  $x_n$  to another downstream location  $x$  is estimated as:

$$L_{\text{MLC},\beta}^{tm(r)}(t, x) = \begin{cases} e^{\left(-[x-x_c]^2/(\sigma_{\beta}^{tm}(t))^2\right)} \cdot \left(\tilde{S}_{\text{MLC},\beta}^{tm(r)}(t) + L_m^{\beta(r)}(t, x)\right), & \text{if } x > x_c \\ \tilde{S}_{\text{MLC},\beta}^{tm(r)}(t) + L_m^{\beta(r)}(t, x), & \text{if } x \leq x_c \end{cases}$$

$$\sigma_{\beta}^{tm}(t) = \alpha_0 + \alpha_1 \cdot N_{\beta}^{tm} + \alpha_2 \cdot \bar{\rho}_{tm}(t)$$
(3.14)

In theory, the proportion of MLC demand assigned to *cell* ( $i, m$ ) (heading for cell ( $i+1, \beta$ )) is evaluated as:

$$S_{\text{MLC},i,m}^{i+1,\beta(r)}(k) = L_{\text{MLC},m}^{\beta(r)}(k \cdot T_s, x_n - (i+1) \cdot l_i) - L_{\text{MLC},m}^{\beta(r)}(k \cdot T_s, x_n - i \cdot l_i). \quad (3.15)$$

However, lane-changing demand might not be received by the target cell according to current traffic condition. Under such circumstance, the remaining part has to postpone the lane-changing maneuvers. Thus  $S_{\text{MLC},i,m}^{i+1,\beta(r)}(k)$  is refined as:

$$S_{\text{MLC},i,m}^{i+1,\beta(r)}(k) = L_{\text{MLC},m}^{\beta(r)}(k \cdot T_s, x_n - (i+1) \cdot l_i) - I_{\text{MLC},m}^{\beta(r)}(k \cdot T_s, x_n - i \cdot l_i) \quad (3.16)$$

where  $I_{\text{MLC},m}^{\beta(r)}(k \cdot T_s, x_n - i \cdot l_i)$  denotes the cumulative MLC demand which was originally proposed at time  $k$  and position  $x_n$  but actually be executed at the downstream boundary of cell  $i$ , which can be evaluated as:

$$I_{\text{MLC},m}^{\beta(r)}(k \cdot T_s, x_n - i \cdot l_i) = \sum_{w=1}^{i-1} q_{\text{MLC},w,m}^{w+1,\beta(r)}(k) \quad (3.17)$$

Apart from the MLC demand which has not been executed, the remaining sending flow of *cell* ( $i, m$ ), i.e.,  $S_{i,m}^{i+1(r)}(k) - \left(\tilde{S}_{\text{MLC},m}^{tm(r)}(k \cdot T_s) - I_{\text{MLC},m}^{\beta(r)}(k \cdot T_s, x_n - i \cdot l_i)\right)$ , should either go straightly towards *cell* ( $i+1, m$ ) or seek for a better driving condition. Speed difference between adjacent lanes, and drivers' desire for traveling faster might trigger DLC maneuvers. The DLC demand for the 1<sup>st</sup> round of estimation can be thus estimated:

$$\tilde{S}_{\text{DLC},i,m}^{i+1,\beta (r)}(k) = \frac{\max(0, v_{i,\beta}(k) - v_{i,m}(k))}{v_{f,i,m} \cdot \tau} \times \left\{ s_{i,m}^{i+1 (r)}(k) - \left( \tilde{S}_{\text{MLC},m}^{tm (r)}(k \cdot T_s) - I_{\text{MLC},m}^{\beta (r)}(k \cdot T_s, x_n - i \cdot l_i) \right) \right\} \quad (3.18)$$

where  $\tau$  can be interpreted as the average time a driver takes to decide and execute a lane change when the origin lane is stopped and the target lane is freely flowing Laval and Daganzo (2006), Laval and Leclercq (2008). Along the longitudinal dimension, the total DLC flow sent by lane  $m$  to lane  $\beta$  from remaining distance  $x_n$  to 0 is estimated by:

$$\begin{aligned} \tilde{S}_{\text{DLC},m}^{\beta (r)}(t) &= \tilde{S}_{\text{DLC},m}^{\beta (r)}(k \cdot T_s) \\ &= \sum_{w=1}^{i_n-1} \tilde{S}_{\text{DLC},w,m}^{w+1,\beta (r)}(k) \end{aligned} \quad (3.19)$$

where the segment from  $x_n$  to 0 is partitioned into  $i_n$  cells. This estimation of initial values of total DLC demand enables a dynamical adjustment of the MLC demand, which will be depicted in Section 3.3.3.

### 3.3.3 The dynamical adjustment process of lane-changing demand

Considering both MLC and DLC maneuvers, the  $r^{th}$  iteration of flow conservation for the off-ramp scenario in Figure 3.6b, can be rewritten as

$$\begin{aligned} \tilde{S}_{\text{MLC},m}^{tm (r+1)}(t) + \tilde{S}_{\text{DLC},m}^{\beta (r)}(t) + e_m(t + t_{x_n}^0) &= u_m(t) + \tilde{S}_{\text{DLC},\beta}^m(r)(t) \\ \tilde{S}_{\text{MCL},\beta}^{tm (r+1)}(t) + \tilde{S}_{\text{MLC},m}^{tm (r+1)}(t) + \tilde{S}_{\text{DLC},\beta}^m(r)(t) + \tilde{S}_{\text{DLC},\beta}^{tm (r)}(t) + e_{\beta}(t + t_{x_n}^0) \\ &= u_{\beta}(t) + \tilde{S}_{\text{MLC},m}^{tm (r+1)}(t) + \tilde{S}_{\text{DLC},m}^{\beta (r)}(t) + \tilde{S}_{\text{DLC},tm}^{\beta (r)}(t) \\ \tilde{S}_{tm}^{f (r+1)}(t) + \tilde{S}_{\text{DLC},tm}^{\beta (r)}(t) + e_{tm}(t + t_{x_n}^0) &= u_{tm}(t) + \tilde{S}_{\text{MLC},m}^{tm (r+1)}(t) + \tilde{S}_{\text{MLC},\beta}^{tm (r+1)}(t) \end{aligned} \quad (3.20)$$

Following the iteration as depicted by Equations (3.13)-(3.19), the estimation of vector

$$\left[ \tilde{S}_{\text{MLC},m}^{tm (r)}(t), \tilde{S}_{\text{MLC},\beta}^{tm (r)}(t), \tilde{S}_{tm}^{f (r)}(t) \right]^T \text{ becomes}$$

$$\left[ \tilde{S}_{MLC,m}^{tm(r+1)}(t), \quad \tilde{S}_{MLC,\beta}^{tm(r+1)}(t), \quad \tilde{S}_{tm}^{f(r+1)}(t) \right]^T$$

by the end of round  $r$ , where  $r$  is the iteration counter. If the gap between these two vectors as mentioned above is less than or equals to the prescribed threshold  $\gamma$ , i.e.,

$$\frac{\left\| \left[ \tilde{S}_{MLC,m}^{tm(r+1)}(t), \tilde{S}_{MLC,\beta}^{tm(r+1)}(t), \tilde{S}_{tm}^{f(r+1)}(t) \right]^T - \left[ \tilde{S}_{MLC,m}^{tm(r)}(t), \tilde{S}_{MLC,\beta}^{tm(r)}(t), \tilde{S}_{tm}^{f(r)}(t) \right]^T \right\|}{\left\| \left[ \tilde{S}_{MLC,m}^{tm(r)}(t), \tilde{S}_{MLC,\beta}^{tm(r)}(t), \tilde{S}_{tm}^{f(r)}(t) \right]^T \right\|} \leq \gamma, (3.21)$$

then the iteration terminates. Otherwise, the algorithm repeats. This dynamical adjustment process of lane-changing demand is illustrated in Figure 3.7.

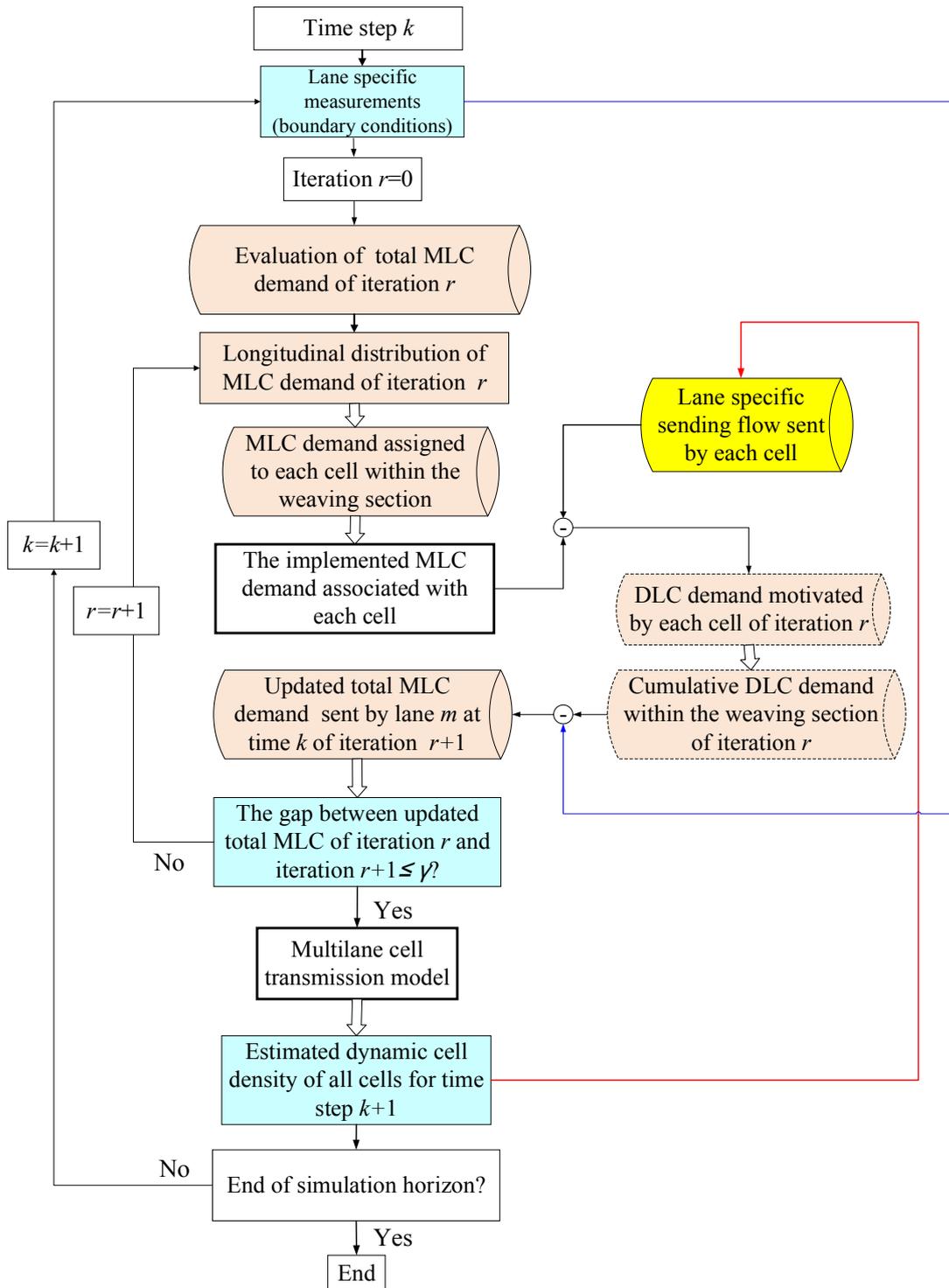


Figure 3.7. The flow chart for MLC and DLC demand estimation

### 3.4 Empirical study

In this section, traffic dynamics of a complex weaving section be of three lanes with significant MLC is investigated. Based on the traffic data provided by lane specific loop

detectors, the proposed macroscopic multilane cell transmission model is calibrated. After that, the proposed model is validated on this section to reveal negative impacts and positive effects of the MLC and DLC maneuvers under various traffic conditions such as congestion onset and dissolve processes. Moreover, this empirical study will also verify that the model requires no additional data other than the CTM. Thus, it can be deployed as a simple simulation tool for accessing traffic state from data available to most management centers.

### **3.4.1 Description of the test site**

The region of interest for simulation is a section of freeway SR-241 southbound. This section, located in the east of Orange County, California, stretches from 33 °47'41.63"N, 117 °43'56.92"W (as indicated by **OR** in Figure 3.7 to 33 °47'6.69"N, 117 °44'47.81"W (position **EN** in Figure 3.8. The main road is composed of three lanes at the first 0.9 miles, and reduces to two lanes when SR-241 forks to SR-241 and SR-261 as depicted in Figure 3.8. After this diverging node, the main section stretches another 0.1 miles until the position EN on SR-241. The first 0.3 miles of SR-261 with two lanes connected to the inner lane and the central lane of the main road is also modeled in this empirical study. An off-ramp to Chapman Avenue is connected to the main road of SR-241 through an intersection that is around 0.2 miles upstream to the diverging node. Overall, a vehicle may exit this weaving section to one of the three directions, i.e., Chapman Avenue towards Santiago Canyon community, State Route-261 towards Irvine city, or continues traveling on State Route-241 towards the southern part of Orange County. The gantry information board, which is located 0.75 miles upstream to the diverging node, reminds the drivers about the remaining distance away from the downstream off-ramp locations and diverging point. Therefore, this 1-mile section between **OR** and **EN** is considered to be the weaving section.

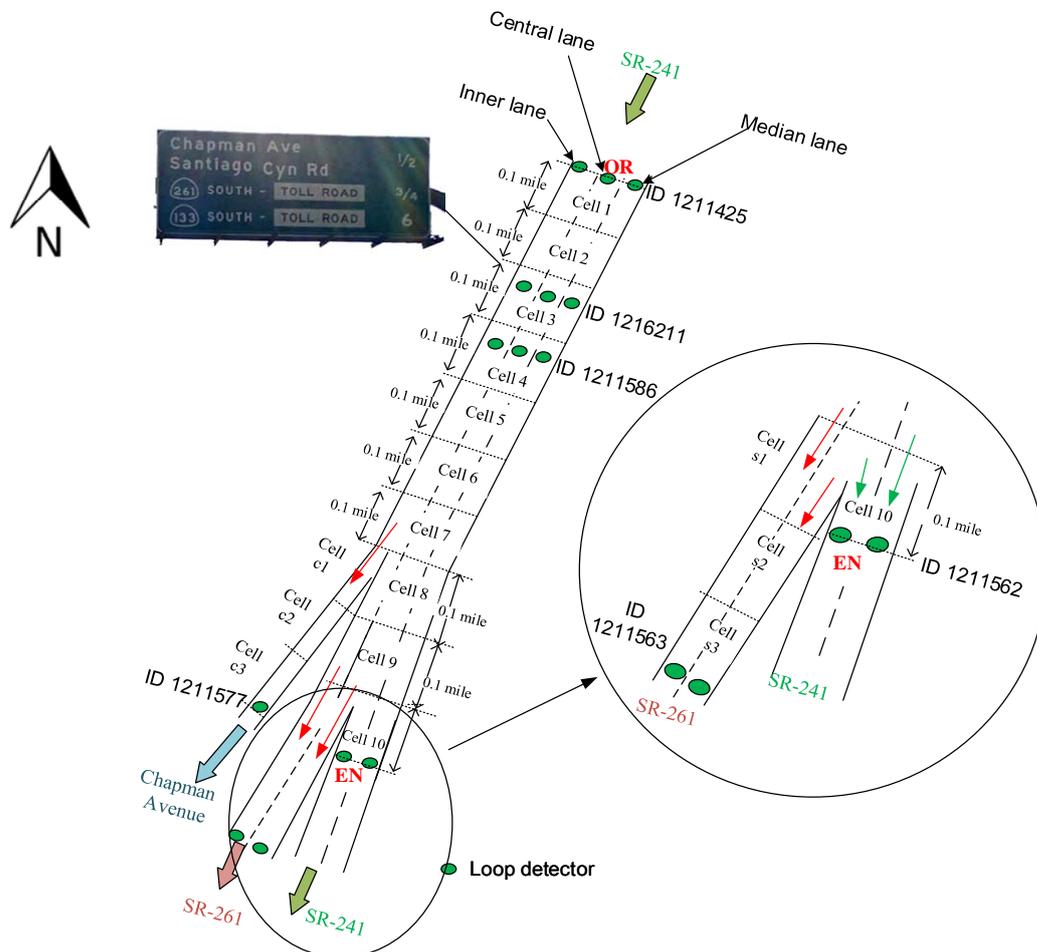
For modeling purpose, the main road of the test section is partitioned into 10 cells with equal length of 0.1 miles as demonstrated in Figure 3.8 to fulfill the Courant–Friedrichs–Lewy (CFL) condition that is a vehicle cannot travel across more than one cell during one simulation time interval (3 sec in this empirical study) with the maximal free-flow speed (74.3 miles/hour for the median lane). The off-ramp intersection between the diverging node and detector station 1211563 on SR-261 is partitioned into three cells, and so does the off-ramp section on Chapman Avenue by detector station 1211577. To show the capacity difference between the weaving and non-weaving sections, the lane-specific MFDs of the 1-mile segment upstream to **OR**, which is densely instrumented with loop detectors, are also calibrated. Based on the lane-based detector stations, the PeMS is capable of updating the aggregated traffic data, including traffic volume, occupancy, speed, and truck proportion every 5 minutes<sup>6</sup>. As it has been demonstrated in Section 3.2.1, the lane specific MFDs of both weaving section and non-weaving section are calibrated based on the data collected on SR-241 on June 2, 2015 and March 20, 2014, where ID 1211425 is regarded as the boundary of the two sections (Figure 3.8b).

The selected data of the two days covers different traffic conditions, e.g., all lanes were free-flowing on June 2, 2015, while they suffered from serious congestion on the morning of March 20, 2014 due to the congestion spilled back from the SR-261. And in particular, the data recorded on the two days has much less data loss and error due to fewer ill-functioning of detectors compared with most of the other days. This weaving section is chosen since, as it will be shown later, the heavy traffic during the peak hours would trigger an obvious queuing, complex DLC and MLC maneuvers due to queuing and weaving, as well as the subsequent congestion dissolve processes can be observed on March 20, 2014. For example, high MLC demand and DLC demand would induce

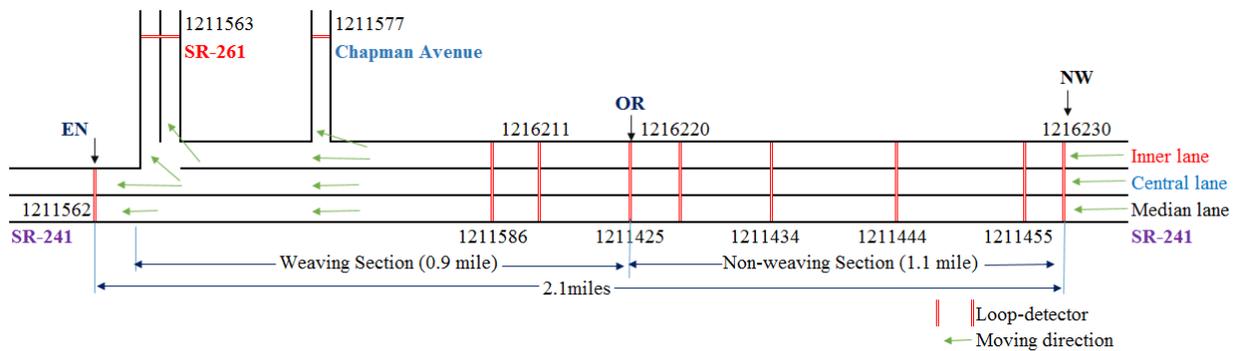
---

<sup>6</sup> The raw data of PeMS, such as traffic volume and occupancy, are updated every 30 seconds. However, on the one hand, the raw data does not provide the values of truck proportion; and on the other hand, the data loss of raw data greatly affect the analysis on lane flow distribution.

a drop of discharging rates on the weaving section during the congestion period due to the complex topology and heavy traffic as it will be shown later in Section 4.3.



a. A weaving section of the SR-241 and its detector configuration



b. The schematic graph of the non-weaving section and weaving section  
Figure 3.8. Topology of the test site

Detector station ID	Related cells	Inner lane			Central lane			Median lane		
		$v_f$ (mile/hour)	$Q_m$ (P.C.U./hour/lane)	$w_c$ (mile/hour)	$v_f$ (mile/hour)	$Q_m$ (P.C.U./hour/lane)	$w_c$ (mile/hour)	$v_f$ (mile/hour)	$Q_m$ (P.C.U./Hour/lane)	$w_c$ (mile/hour)
1211425	cell 1	58.3	1903.1	11.4	65.8	2002.9	11.8	74.3	1857.8	10.6
1216211	cell 2-3	58.5	1804.0	10.7	55.4	1929.7	11.4	66.5	2083.7	11.6
1211586	cell 4-9	59.1	1958.4	11.8	55.4	2204.9	13.8	66.5	1941.4	11.4
1211562	cell 10	—	—	—	45.6	N/A	N/A	59.1	1956.0	11.3
1211563	cell s1-s3	58.2	1476.0	12.8	—	—	—	55.0	1632.0	14.2
1211577	cell c1-c3	77.8	N/A	N/A	84.2	N/A	N/A	79.2	N/A	N/A

Table 3.1. Calibration of fundamental diagrams

### 3.4.2 Model calibration and estimation of MLC demand

Assume that each pair of lane-specific traffic flow and density measured by the same detector satisfies a specific triangular relationship as introduced in Section 3.2.1. In this empirical study, cell parameters of the main road are calibrated using conventional least squares method in line with the first-order macroscopic traffic flow literature as shown in Table 3.1. Their graphical counterparts have been demonstrated in Section 3.2.1. For those cells without detectors, the calibration adopts the data observed by the most adjacent detectors on the main road. Note that some parameters cannot be calibrated due to the limited amount of congestion data at certain locations. Alternatively, one can apply the automatic calibration algorithm by Zhong et al. (2015) to interpolate the parameters for those with no measurement devices. The MFDs of SR-261 and off-ramp sections (after the weaving section) are similarly calibrated.

To estimate the initial time-dependent longitudinal distribution of MLC demand by Equation (3.13), the median lane traffic flow data observed during the non-rush hour on June 2, 2015 (excluding March 20, 2014) is firstly applied to calibrate the parameters concerned, i.e.,  $\alpha_0$ ,  $\alpha_1$  and  $\alpha_2$  and  $x_c$ . The main reasons for using this day are:

- a) The median lane would most possibly send out MLC demand (towards the fork and off-ramp intersection) to central lane without receiving flows with MLC intention in return under light traffic conditions of the median lane in conjunction with normal conditions (e.g., incident free conditions) of the central lane.
- b) According to the lane specific traffic speed measurements by the adjacent detecting stations, the speed difference between the median lane and the central lane during non-rush hour is always less than 5 miles/hour on June 2, 2015, therefore the DLC demand should be negligible based on the assumption that a DLC is mainly motivated by sufficiently large speed advantage of the adjacent lanes.
- c) The traffic state of the central lane maintains free-flowing for the whole day, whose density is lower than 25 *P.C.U./mile/lane* during non-rush hour in particular. Therefore

the space between the vehicles traveling on central lane would be no less than 200 feet on average during the non-rush hours. While noting that the space of 200 feet (61 meters) is large enough to receive almost all the MLC flows.

d) The vehicles traveling on the median lane must move across two lanes if they would like to exit the freeway via Chapman Avenue off-ramp while they only need to cross one lane to the SR 261. This topology feature enables the calibration of the weighting associated with the number of lanes to be crossed by MLC flows, i.e.,  $\alpha_1$ .

To sum up, the gradual reduction of traffic flow on the median lane along the longitudinal dimension is mainly caused by MLC maneuvers towards the fork and off-ramp intersection. In other words, the longitudinal lane flow distribution on the median lane would reflect the cumulative effect of MLC maneuvers for a short time period such as five minutes within which the traffic state would not admit a sharp change.

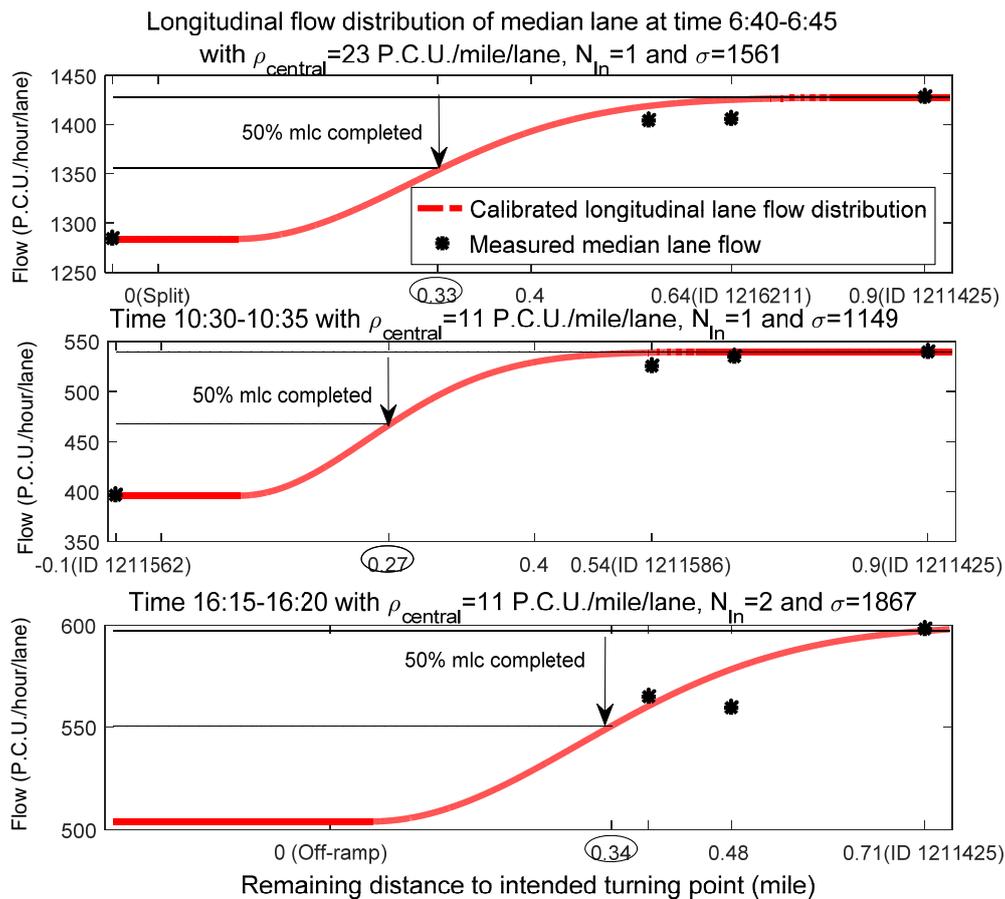


Figure 3.9. The longitudinal variation of median lane flow due to MLC maneuvers: a) During 6:40-6:45, with  $\rho_{central} = 23$  P.C.U./mile,  $N_{ln} = 1$  and  $\sigma = 1561$ ; b) During 10:30-10:35, with  $\rho_{central} = 11$  P.C.U./mile,  $N_{ln} = 1$  and  $\sigma = 1149$ ; c) During 16:15-16:20, with  $\rho_{central} = 11$  P.C.U./mile,  $N_{ln} = 2$  and  $\sigma = 1867$ .

The calibration procedures for the parameters  $\alpha_0$ ,  $\alpha_1$  and  $\alpha_2$  can be summarized as follow: The first step is to identify  $\sigma(t)$  through a nonlinear curve fitting algorithm based on the longitudinal variation of median lane flows measured at  $t$  simultaneously. The curve fitting algorithm tries to identify  $\sigma(t)$  in the first equation of (3.13) such that the distance between the fitted curve (red line) and the measured data (black star) is minimized. Then a linear regression algorithm is applied to identify the parameters  $\alpha_0$ ,  $\alpha_1$  and  $\alpha_2$  by acknowledging the average dynamic traffic density of central lane  $\rho_{central}(t)$  (central lane as target lane) on weaving section and number of lanes  $N_{ln}$  to be crossed to reach the terminal lane (central or inner lane).

Figure 3.9 demonstrates the cumulative impact of MLC maneuvers on the longitudinal lane flow variation of the median lane on SR-241 on June 2, 2015, with respect to the remaining distance to the target turning point, e.g., the diverging point to SR 261 or the off-ramp intersection connected to Chapman Avenue. For example, the utmost and the middle figures assume that the MLC is initialized at the first detection station (ID 1211425), while the diverging point before detection station ID 1211562 is regarded as the destination in the morning<sup>7</sup> of June 2, 2015. Comparing the longitudinal lane flow variation of the median lane during 6:40-6:45 and 10:30-10:35, it can be observed that drivers with MLC intentions, might decide to change lane earlier during 6:40-6:45 than 10:30-10:35 due to a higher average density on the central lane, which is reflected by a larger  $\sigma$  during 6:40-6:45. To be specific, during 6:40-6:45, 50% MLC demand can be completed with a remaining distance of 0.33 miles to the target turning point while this becomes closer (0.27 miles) to the intended turning point during 10:30-10:35 due to the lower level of urgency (lower traffic density after peak hour). On the other hand,

---

<sup>7</sup> According to data observation, the mandatory lane-changing ratios assigned to Chapman Avenue and SR-261 does not maintain stationary. In the morning, the traffic leaving SR-241 via SR-261 dominates the mandatory lane-changing behaviors, while in the afternoon Chapman Avenue becomes busier than SR-261.

although the average density of central lane observed during 16:15-16:20 is close to that of 10:30-10:35, a farther lateral distance to the target turning point (two lanes away from the off-ramp intersection connected to Chapman Avenue on lateral dimension) might strengthen the drivers' sense of urgency to change lanes. Therefore, 50% MLC demand is completed earlier with a remaining distance up to 0.34 miles to the off-ramp intersection during 16:15-16:20 which is indicated by a larger  $\sigma$ . The average reaction time for a DLC is calibrated by Laval and Daganzo (2006). The minimum safe gap is calibrated by Hidas (2005). Here their typical values have been adopted as reported in the literature as indicated in Table 3.2.

Longitudinal cumulative distribution function of MLC demand (Calibrated based on the PeMS data collected on June 2, 2015)	$\alpha_0$	-55.9
	$\alpha_1$ the weighting associated with the number of lanes to be crossed	726.9/lane
	$\alpha_2$ the density of target lane	33.7 mile/P.C.U.
	$x_c$ Critical remaining distance	0.05 mile (265 feet)
	$x_r$ Remote remaining distance	1.0 mile (5280 feet)
DLC demand determination function (Laval and Daganzo, 2006)	$\tau$ the average lane-changing reaction time	3 seconds
Minimum gap acceptance criterion function (Hidas, 2005)	$c_l$ Parameter of the extra leading gap	1.32 feet· hour/mile
	$c_f$ Parameter of the extra lag gap	1.32 feet· hour/mile
	$\bar{g}_{\min}$	37.7 feet

Table 3.2. The parameters in the macroscopic multilane model

### 3.4.3 Simulation results against the measurement

Based on the parameters calibrated in Section 3.4.2, model validation is conducted using the inflow profile and boundary conditions observed on the morning of March 20, 2014, as input profiles to simulate the proposed macroscopic multilane cell transmission model that considers both MLC and DLC maneuvers. It is worthwhile to point out that, in what follows, the measurements in the middle of weaving section, i.e., those given by detector stations **ID 1216211** and **ID 1211586** on March 20, 2014, are

not regarded as boundary conditions in the simulation during this test day. Instead, it was assumed that both the detectors were “missing”, i.e., their measurements were not used for simulation, but would be used for cross-validation.

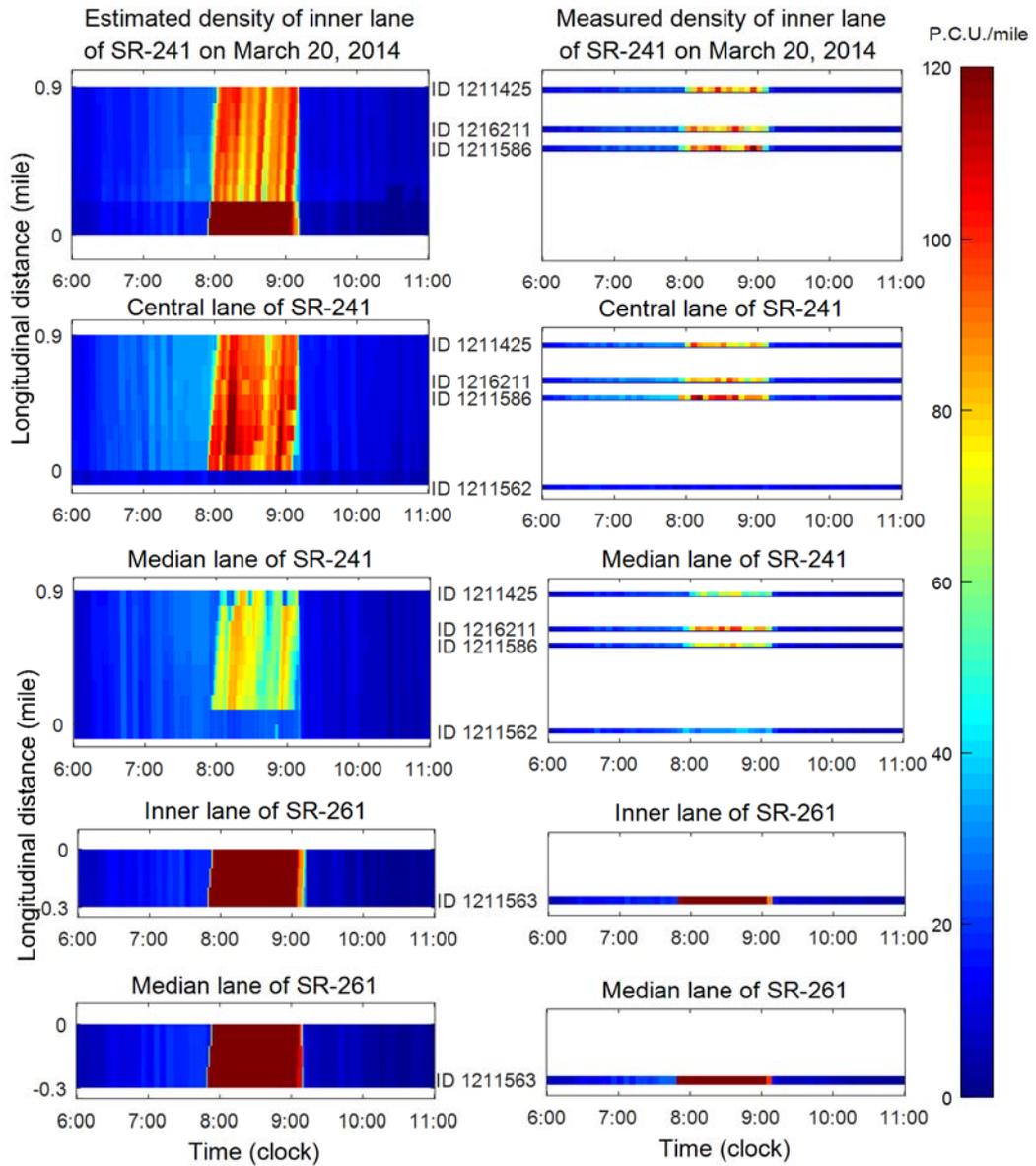


Figure 3.10a. Heat map comparison of the estimated multilane density against measurement counterpart on March 20, 2014

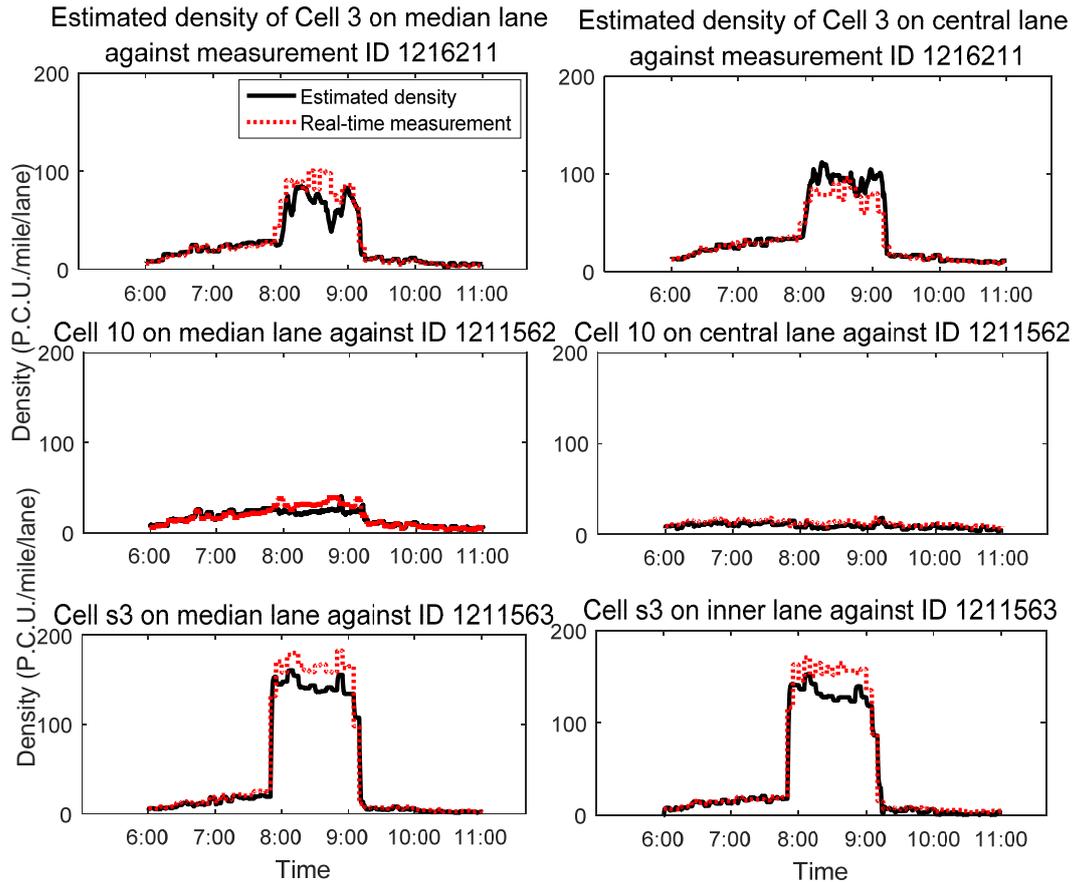


Figure 3.10b. Time series comparison of the estimated multilane density against measurement counterpart on March 20, 2014

Figure 3.10a presents the heat map of the simulated traffic state (left graphs) of the test segment against its observation counterpart (right graphs), wherein the fork of SR-241 and SR-261 is chosen as the reference point (i.e., 0 mile in longitudinal distance). Figure 3.10b depicts the time series comparison of the estimated multilane density against measurement counterpart for detector ID1216211, ID 1211562 and ID 1211563. On the whole, the simulated lane-specific traffic density captures the trend of measurements. Generally speaking, each estimated lane based cell density (with a resolution, i.e., the simulation time step of 3 seconds) is close to its measured counterpart (with a resolution of 5 minutes of PeMS data) under both heavy and light traffic conditions from 6:00-11:00 AM on the test day.

In particular, the transient states, i.e., the longitudinal congestion onset, the lateral

congestion onset and congestion offset, are also well captured by the proposed model. In the beginning, there was no congestion queue along the whole segment. Under such scenario, the lane-flow distribution was almost stable and evenly distributed until 7:50 AM as indicated in Figure 3.10. A congestion was first simultaneously observed around 7:50 AM on both the two lanes of SR-261, i.e., the downstream section after the weaving area, to be specific the fork of SR-241 and SR-261. Then the weaving section suffered from a severe congestion due to the queue spilled back from SR-261. Unlike the conventional CTM, traffic flow characteristics of different lanes along the main road are with some subtle and important differences as demonstrated by the simulation results presented in the left-hand side of Figure 3.10. Since SR-261 is directly connected to the main road SR-241 via its inner lane and central lane as shown in Figure 3.8, the drop of discharging rate due to the congestion on SR-261 directly affected the traffic on these two lanes. In contrast, the traffic on the median lane of SR-241 was not influenced by this congestion during its early stage (7:50-8:00 AM).

The congestion on central lane slows down or partially rejects the MLC flows that intended to merge into the central lane from the median lane, due to the limited available space of the lane and small space gaps between successive vehicles on the congested central lane. Therefore, **the congestion is gradually spread to the median lane**. However, compared with the central lane, the median lane still provides a better driving condition. This imbalance between adjacent lanes (median lane and central lane) triggers some vehicles, which were originally traveling on the central lane of SR 241 and did not intend to leave SR-241 within the current weaving section, switched to the median lane at the upstream cells to seek for a better driving condition so as to avoid congestion. Such DLC maneuvers further increased the traffic density on the median lane while slightly alleviated the congestion on the central lane. Therefore, the central lane was less congested than the inner lane, in particular on *cell 8* and *cell 9*. This phenomenon is known as **the balancing effect of DLC**, i.e., lane changes could smooth out the differences between adjacent lanes.

Also since the drivers tend to switch to the median lane long before they could arrive at the congestion site for better driving condition and to avoid congestion as previously discussed, the congestion on the median lane started from upstream cells, see e.g., detector station ID 1216211, rather than the downstream bottleneck location. This congestion did not cause a severe impact as shown by the simulation and measurements of the detector stations ID 1211586 and ID 1211425 on the median lane. This result is unlike some existing models which assume drivers do not **perceive the downstream traffic condition** to make their lane-changing decisions, i.e., each driver decides to change lane or not based on the perceived traffic condition of her/his current location. Shiomi et al. (2015) realized such an unrealistic assumption on LC maneuver and encouraged future effort to review and to overcome this limitation. The proposed model can be regarded as a possible solution to this stream. Finally, the congestion spreads to all lanes. The proposed model can capture all these phenomena.

During the congestion period, the off-ramp connected to Chapman Avenue helps in alleviating the congestion level of cells upstream to cell 7 on the inner lane as indicated by the simulation. This is because the vehicles still get a chance to exit the freeway from the off-ramp to Chapman Avenue, while those enter the weaving section of SR-261 do not have such a luxury to escape from the congestion. Note that the section of SR-241 downstream to the diverging point never gets congested (both the central and median lanes). This is a typical phenomenon that the downstream section of a bottleneck is usually free-flowing. After the peak hour period, all the lanes return back to free-flowing condition, and the lane-flow distributions go back to almost stable again.

As mentioned previously the measured data of detector station ID 1216211 and ID 1211586 is used for cross-validation, the corresponding lane specific MAPE (Mean Absolute Percentage Error) of the density estimation is reported in Table 3.3, where the MAPE is evaluated as

$$MAPE_{i,m} = \frac{1}{N} \sum_{K=1}^N \frac{|\bar{\rho}_{i,m}(K) - \rho_{i,m}^M(K)|}{\rho_{i,m}^M(K)} \quad (3.22)$$

Here  $\bar{\rho}_{i,m}(K)$  denotes the average density of *cell* ( $i, m$ ) estimated during the time interval  $[kT_s, (k + 1)T_s)$ , in accordance with its measurement counterpart  $\rho_{i,m}^M(K)$  whose resolution is 5 minutes.  $N$  denotes the amount of observations during the whole simulation horizon.  $\rho_{i,m}^M(K)$  is provided by the detector station most close to *cell package*  $i$  on the same lane. The MAPE for lane specific traffic density estimation ranges from around 13.0% to 18.5% which indicate that the proposed model can capture the lane-changing traffic dynamics well. To compare with the CTM without considering lane heterogeneity nor lane-changing maneuvers, both the simulated and measured lane specific traffic densities are aggregated, i.e., summing the densities of the three lanes. Defining the aggregated MAPE as the aggregated absolute value of the discrepancy between the simulated data and measured data over the aggregated measured data, it is found that the aggregated MAPE is around **9%-11%**. This result is more accurate than that of traffic density by the **CTM** which is around **13%-15%**.

	<b>Inner lane</b>	<b>Central lane</b>	<b>Median lane</b>	<b>Aggregated</b>	<b>CTM</b>
Simulated traffic density of <i>cell 3</i> against the measurement of ID 1216211	16.7%	13.2%	18.3%	10.7%	13.4%
Simulated traffic density of <i>cell 4</i> against the measurement of ID 1211586	16.0%	16.5%	18.3%	9.4%	14.6%

Table. 3.3. Cell density estimation against measurement on March 20, 2014

To test the effect of the average reaction time, a One-at-A-Time (OAT) sensitivity analysis is conducted. As shown in Figure 3.11 that the proposed lane-changing model is not sensitive to  $\tau$  for a small change of  $\tau$  ranging from 2.0-4.0 sec in the sense that the MAPE of the estimation does not admit an abrupt change.

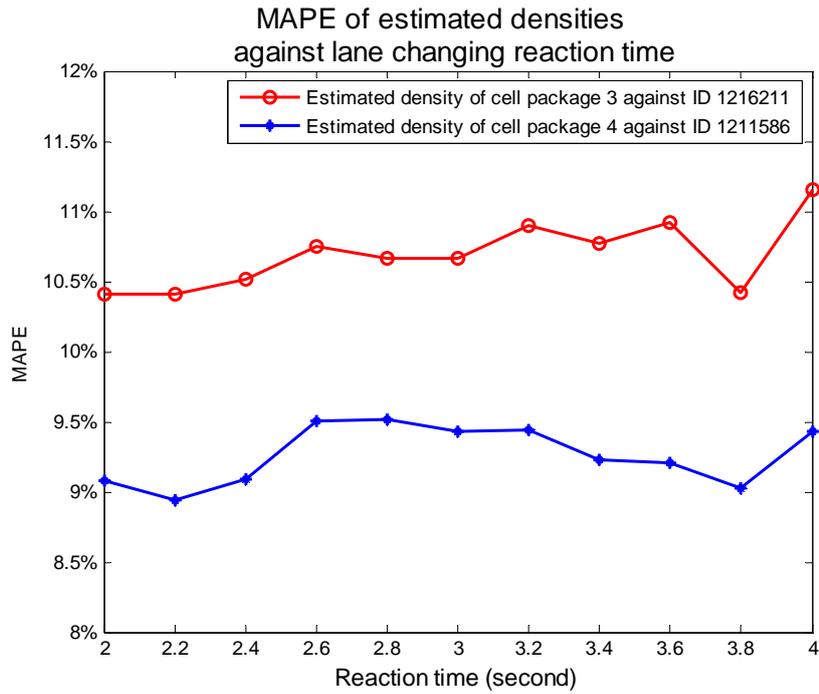


Figure 3.11. One-at-A-Time sensitivity analysis on the lane-changing reaction time

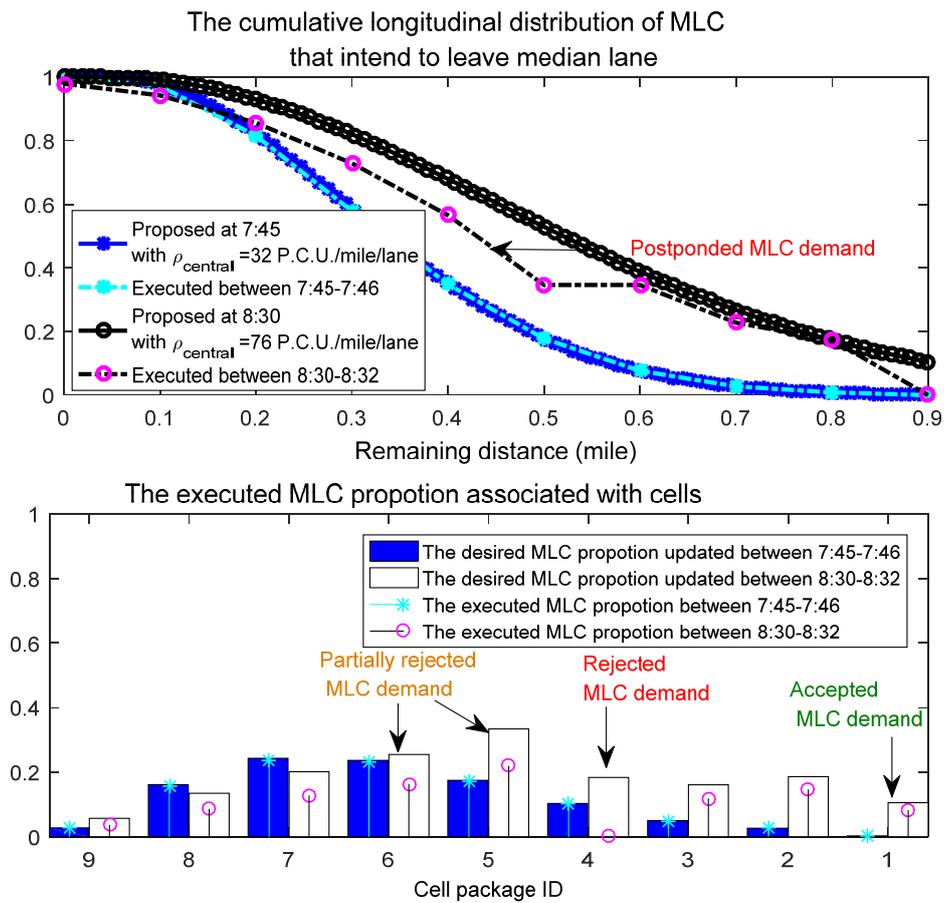


Figure 3.12. Proposed MLC demand and execution on longitudinal dimension

	Inner lane	Central lane	Median lane
The estimated lane flow variation	7.66%	7.60%	2.13%
The measured lane flow variation	6.12%	6.24%	2.55%

Table 3.4. The absolute variation percentage

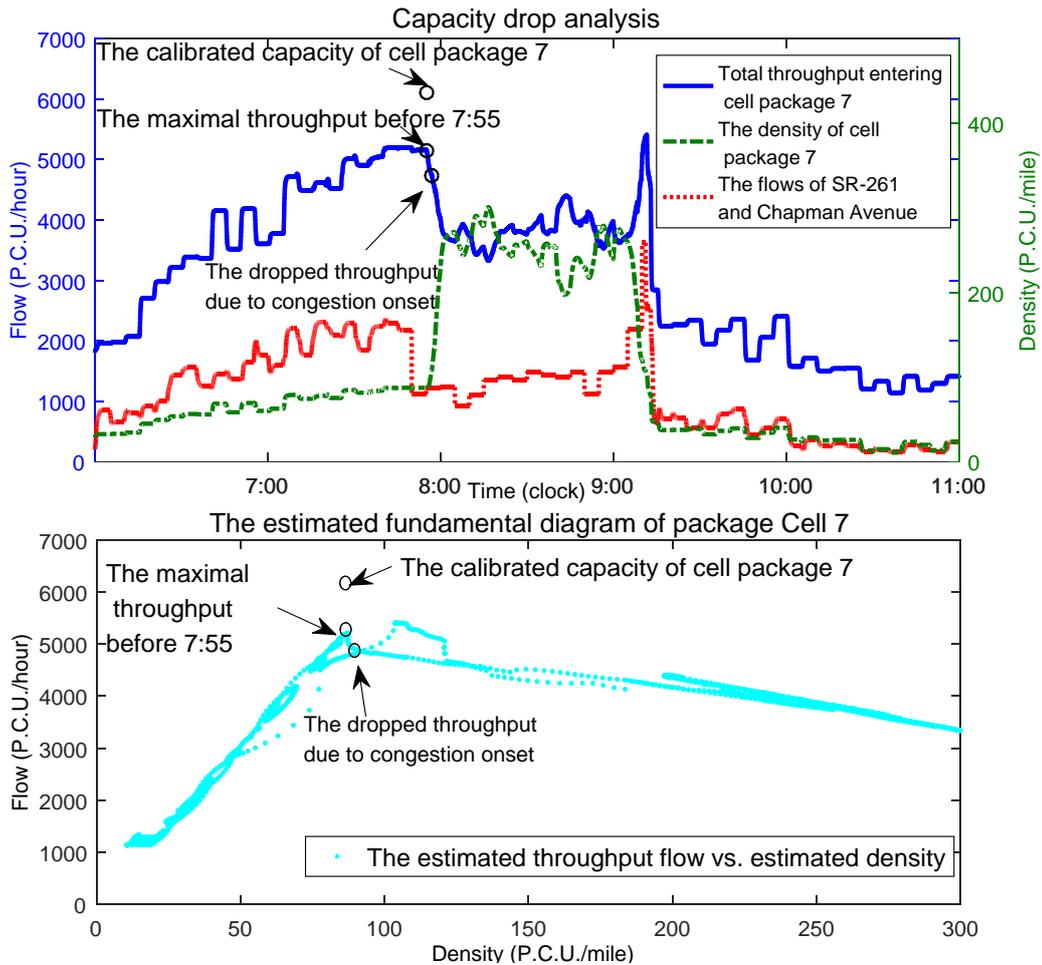


Figure 3.13. Capacity drop identification at *cell package 7* within the weaving section.

As it has been discussed in Section 3.4.1, the longitudinal distribution of MLC demand, which is assumed to be initialized at the beginning of this weaving section (with a remaining distance of 0.9 miles to the diverging point this empirical study), is affected by the perceived level of congestion of the target lane (central lane) as well as the number of lanes to be crossed on the lateral dimension.

To demonstrate this phenomenon, the upper graph of Figure 3.12 depicts the longitudinal cumulative distribution functions of MLC demand initially proposed at different times on the test day, e.g., 7:45 AM and 8:30 AM, against the executed ones. This graph indicates that drivers intend to execute an MLC maneuver earlier when facing the congested condition (i.e., 8:30 AM with  $\rho_{central} = 76P.C.U/mile/lane$ ) than the free flowing condition (i.e., 7:45 AM with  $\rho_{central} = 32P.C.U/mile/lane$ ). This earlier plan of MLC is proposed to ensure the execution of lane-changing maneuvers before arriving the turning point in case that the congested target lane might reject the lane-changing demand due to its limited available space as well as the small gaps between vehicles. This phenomenon is demonstrated in the lower graph of Figure 3.12 that the MLC demand proposed at 8:30 AM was firstly partially rejected by *cell 2*. Therefore, this rejected MLC demand has to postpone to the downstream cells. However, the updated proportion of MLC demand assigned to *cell 3* and the original assignment were still partially rejected by central lane. Similar observations can be drawn from other cells, especially the completed rejection on cell 4.

Finally, cell 9 accepted all the MLC assignment, because people with MLC intention have to squash into the target lane regardless whether there is a minimum acceptable gap or not as discussed in Section 3.2.2. In conclusion, the MLC demand which was sent from the median lane and proposed at 7:45 AM was well executed during 7:45-7:46 AM by arriving the diverging node due to the free-flowing traffic condition on the central lane. However, due to the saturated traffic condition of the central lane at 8:30, which is nearly three times of that at 7:45 AM, the proposed demand was not executed as planned but postponed to downstream cells, and cost longer time compared with the one proposed at 7:45 AM. Such postponement proves the necessity of proposing an early MLC at upstream during congestion.

To compare the simulated and measured lane-changing effects, the cumulative lane flow variation analysis of this weaving section, i.e., from OR to EN, is conducted

wherein the Chapman Avenue off-ramp and SR-261 is regarded as the stretch of the inner lane for comparison. The results are depicted in Table 3.4. It is assumed in this test that the cumulative lane flow difference between the upstream OR and the downstream EN during the simulation horizon is caused by executed MLC and DLC maneuvers (If there is no lane change, the cumulative upstream flow should equal to the downstream by flow conservation, excluding the flow on the freeway at the initial time and the remaining flow at the end time). As shown in Table 3.4, the estimated lane flow variations (in terms of percentage) are very close to their measurement counterparts which prove the accuracy of the proposed lane-changing model.

A conventional way to observe capacity drops is to compare the flow just upstream of the bottleneck and the flow just downstream of the bottleneck as mentioned by Cassidy and Bertini (1999), e.g., Section 3.2. Another way is to depict the flow and density relationship at the bottleneck site. In line with Srivastava and Geroliminis (2013), Figure 3.13 presents a representative throughput time series plot along with the corresponding off-ramp flows and the mainline density at the bottleneck concerned. The total throughput of *cell package 7* can be seen here to decrease from a maximum of about 5150 P.C.U./hour (by aggregating for 3 lanes) between 7:45–7:55 AM, to about 4700 P.C.U./hour after 7:56 AM, when the density of *cell package 7* maintains at critical density. The off-ramp flow was increasing until the weaving section entered the congested period, then both the throughput and the off-ramp flow drop. Note that both the ramp flow and the throughput significantly increases before the breakdown. The congestion blocks both the MLC for the off-ramp and the throughput until both the off-ramp and the throughput admits a sharp increase around 9:15 AM after the congestion was alleviated.

### **3.5 Conclusions**

Modeling lane-changing maneuvers is essential to capture several important

characteristics of multilane traffic flow, e.g., heterogeneous traffic flow distribution, capacity drop, and flow balancing effect. To this end, a novel macroscopic multilane model was proposed to simulate the effect of MLC and DLC maneuvers in this chapter. The proposed model extends the multilane hybrid (MH) theory by incorporating the lane-based fundamental diagrams to capture the relation between speed and lanes which is believed to be missing in most existing models (Keyvan-Ekbatani et al., 2016). Meanwhile, different priority levels were identified according to the lane-changing motivations and the corresponding levels of urgency. In particular, a recursive lane-changing demand estimation algorithm that considers the impact of level of urgency on longitudinal probability distribution function of lane-changing maneuvers was devised. Flow propagations of both MLC and DLC maneuvers were calculated by demand-supply reaction laws based on the extended IT and PIT principles.

Using traffic data from the PeMS, the proposed macroscopic multilane cell transmission model was calibrated and validated on a complex weaving section of the SR241 freeway in Orange County, California. The results indicated that the proposed model can capture the impacts of lane-changing maneuvers on the temporal and spatial traffic state, especially the lateral lane flow distribution and the queuing effect on the longitudinal dimension in conjunction with congestion spreading to adjacent lanes. The MLC demand estimation algorithm and the lane-changing probability distribution function on longitudinal dimension fitted their measurement counterparts in a satisfactory manner. This implies the model can also be used to infer turning ratios (e.g., off-ramp demand) for ramps.

Unlike some existing models which assume drivers do not perceive the downstream traffic condition to make their lane-changing decisions, which is regarded as unrealistic by Shiomi et al. (2015), the proposed model assumes that the drivers use the traffic condition spatially ahead that they could perceive to make their lane-changing decisions. The effect of the perceived traffic condition decreases exponentially with respect to the

distance away from her/his current position. This exponential probabilistic distribution function of the cumulative MLC was calibrated and validated by the empirical results. With the same data source, the proposed model outperforms the CTM in terms of accuracy. Meanwhile, some important effects of lane-changing such as capacity drop and flow balancing effect of DLC were presented in the empirical study which are essentially in agreement with previous findings in the literature. As proven by the empirical study, the proposed model does not require high-resolution traffic data but traffic data available to most of the traffic management centers. This can be regarded as a significant improvement over the existing models.

## **Chapter 4 Stochastic multilane cell transmission model by assimilating lane speed observation**

### **4.1 Introduction**

Freeways are always subject to traffic demand and supply uncertainties particularly in congested segments such as the weaving areas. The measured traffic data is noisy due to various reasons, e.g., detector error and data drop due to transmission. Moreover, the driver behavior is somehow random during the congestion periods. Parallel to the macroscopic counterpart, i.e., from the cell transmission model to the stochastic cell transmission model (Sumalee et al., 2011; Zhong et al., 2013), developing a multilane traffic flow model to consider these random elements is of importance.

On the other hand, from the definition of discretionary lane-changing (DLC) demand, one needs to have cell traffic speed to proceed the calculation of DLC demand for each cell. However, this may not be possible due to limited measurement devices. There is a need to interpolate the cell speeds that are not available from direct measurement by some speed-density relationship diagrams. It may resolve the “missing” cell-lane traffic speeds from a well calibrated fundamental diagram similar to the cell transmission model with velocity (CTM-v).

In this chapter, the stochastic cell transmission model (SCTM) will be extended to simulate the effects of vehicle lane-changing on traffic dynamics. Lane-changing ratios are defined according to lane speed heterogeneity in line with the hybrid CTM. A certain speed-density relationship is used to interpolate the cell-lane speed profiles along a freeway corridor with sparse detectors from both detected and estimated traffic density to overcome the disadvantages of the original hybrid CTM under free-flowing traffic conditions caused by the triangular fundamental diagram. Link (cell)-node

junction formulation is proposed to propagate vehicle lane-changing vehicles, wherein the lane-changing ratios define virtual node splitting ratios which are propagated by the Incremental-Transfer (IT) and Priority IT (PIT) principles.

Following the operational modes in the SCTM, random events with different probabilities of occurrence are defined to govern the traffic demand-supply reactions. The “actual” flow received by the downstream lane-cell is then a finite mixture of these random events. By this manner, flow propagations of both mandatory and discretionary lane-changing maneuvers are proceeded by demand-supply reaction laws in line with the existing SCTM framework. The lane-changing fractions can be well defined by using the measurements of boundary variables and the average execution time of lane changing. Compared with the original SCTM, the new model is of the following key features. i) The traffic states are given regarding both cell and lane. ii) The cell-lane-changing ratios are augmented as additional states that define node splitting ratios. iii) An additional process is adopted to resolve cell-lane traffic speed from traffic density estimation to define lane-changing ratios. However, the link (cell)-node junction formulation integrates the multilane SCTM and the original SCTM into a unified framework. On the practical side, the proposed models do not require high-resolution traffic data which is a significant improvement over the existing models.

## **4.2 Hybrid cell transmission model and its extension**

Laval and Daganzo (2006) proposed the multilane hybrid model, which can be discretized into a lane-specific cell transmission model, to simulate traffic lane-changing behavior. The model termed as hybrid cell transmission model divides each lane of a multilane highway into a series of cells with cell positions denoted by  $(i, l)$ , where  $i$  indicates the position along the travel direction and  $l$  represents the lane number. Vehicle units from cell  $(i, l)$  at time  $t$  can move into cell  $i+1$  in the same lane or the

adjacent lanes. It is assumed that lane-changing events may occur only at the upstream and downstream boundaries of a cell. Between time steps  $t$  and  $t + \Delta t$ , the total number of vehicle units that can be sent from cell  $(i, l)$  to the immediate downstream cells is given by

$$S_i^l(t) = \min \{n_i^l(t), Q_i^l(t)\Delta t\}, \quad (4.1)$$

where  $n_i^l(t)$  is the number of vehicle units in cell  $(i, l)$  at time  $t$  and  $Q_i^l(t)$  is the capacity of cell  $(i, l)$  at time  $t$ . Assume that the probability ratio of these vehicles desiring to move from cell  $(i, l)$  to cell  $(i+1, k)$  ( $k$  may be the same or different lane) is  $p_i^{lk}(t)$ , the number of vehicles desiring to change lane from cell  $(i, l)$  to cell  $(i+1, k)$  is given by

$$\tilde{S}_i^{lk}(t) = S_i^l(t)p_i^{lk}(t). \quad (4.2)$$

The lane-changing probability ratio  $p_i^{lk}(t)$  (which will be formally defined later on) depends on several factors such as the difference in speed and density between lanes  $l$  and  $k$ , the free-flow speed of the cell and the average execution time of vehicles' lane-changing movements. On the other hand, the number of vehicle units that can be received by cell  $(i+1, l)$  is decided by

$$R_{i+1}^l(t) = \min \{Q_{i+1}^l(t)\Delta t, w_{i+1}^l(N_{i+1}^l(t) - n_{i+1}^l(t))\} \quad (4.3)$$

where  $N_{i+1}^l(t)$  is the maximum number of vehicle units that can be accommodated in cell  $(i+1, l)$  at time  $t$ ,  $n_{i+1}^l(t)$  is the number of vehicle units in cell  $(i+1, l)$  at time  $t$ , and  $w_{i+1}^l(N_{i+1}^l(t) - n_{i+1}^l(t))$  is the available space of the cell at time  $t$ . A simplified Incremental-Transfer (IT) principle is applied to allocate the available space in the receiving cell according to the upstream demand in the same and adjacent lanes. A service fraction is defined for the receiving cell  $(i+1, l)$  to accommodate the lane-changing demand as follows:

$$f_{i+1}^l(t) = \min \left\{ 1, \frac{R_{i+1}^l(t)}{\sum_{\forall k} \tilde{S}_i^{kl}(t)} \right\} \quad (4.4)$$

The actual number of vehicle units sent from cell  $(i, k)$  to  $(i+1, l)$  is then

$$s_i^{kl}(t) = f_{i+1}^l(t)\tilde{S}_i^{kl}(t). \quad (4.5)$$

Flow conservation equation is expressed in terms of lane-specific manner:

$$n_i^l(t + \Delta t) = n_i^l(t) - \sum_{\forall k} s_i^{lk}(t) + \sum_{\forall k} s_{i-1}^{kl}(t), \quad (4.6)$$

A density version of the above flow conservation equation is

$$\rho_i^l(t + \Delta t) = \rho_i^l(t) - \frac{\Delta t}{L_i} (\sum_{\forall k} s_i^{lk}(t) - \sum_{\forall k} s_{i-1}^{kl}(t)). \quad (4.7)$$

where  $L_i$  denotes the cell length. To simulate lane-changing events, the probability ratio for discretionary lane-changing (DLC in probability per unit time) was presented by Laval and Daganzo (2006) wherein lane changes are assumed to be triggered by speed differences between adjacent lanes, and drivers' desire for traveling faster:

$$p_{DLC,i}^{lk}(t) = \max \left\{ 0, \frac{v_i^k(t) - v_i^l(t)}{v_{f,i}^l \tau} \Delta t \right\}, l \neq k \quad (4.8)$$

where  $\Delta t$  is the simulation time-step,  $v_i^l(t)$  is the perceived average speed in lane  $l$  at time  $t$ , taken as  $v_{i+1}^l(t - \Delta t)$ . This probability ratio defines the lane-changing behavior of a driver experiencing a speed difference between the adjacent lanes, i.e.,  $\Delta v(t) = v_{,,i}^k(t) - v_i^l(t) \geq 0, \forall k \neq l$ .  $\tau$  can be interpreted as the (average) time that a driver takes to decide and execute a lane change when the origin lane is stopped and the target lane is freely flowing (Laval and Daganzo, 2006; Laval and Leclercq, 2008). In Laval and Daganzo (2006), the value  $\tau=3$  sec is used. Simulation experiments in Cheu et al. (2009) indicated that the queue length increases with the value of  $\tau$ . The lane-changing maneuvers tends to a negligible number with  $\tau \rightarrow \infty$ , i.e., no lane-changing behavior would occur. The the average queue length is likely to be overestimated if the lane-changing behavior is not taken into account. As proposed by Laval et al. (2007), MLCs can be included by using

$$p_{MLC,i}^{lk}(t) = 1, \quad l \neq k.$$

As deemed by Laval and Leclercq (2008), the lane-changing vehicles act as moving bottlenecks with realistic accelerations in the target lane as per the moving bottlenecks model. Empirical studies have revealed the satisfactory performance of the hybrid CTM model (Laval et al., 2007). As commented by Laval and Leclercq (2008) that the complex lane-changing process can be simulated with a parsimonious model that

requires only one extra parameter, i.e.,  $\tau$  the average time to complete a lane-changing maneuver, except for free-flowing traffic conditions due to the triangular fundamental diagram.

The lane-changing probability ratio defined by (4.8) can exceed one if the simulation time increment and the speed heterogeneity are large enough, e.g., suppose that  $\Delta t > 10$  sec while the speed difference is greater than 1/3. This can be regarded as a drawback of the model by Laval and Daganzo (2006) and Cheu et al. (2009). To overcome this, the definition of the above probability ratio may be changed as

$$p_{DLC,i}^{lk}(t) = \min \left\{ 1, \max \left\{ 0, \frac{v_i^k(t) - v_i^l(t)}{v_{f,i}^l \tau} \Delta t \right\} \right\}, l \neq k \quad (4.9)$$

to avoid the ill-defined probability ratio for dual-lane freeways. Still, the summation of lane-changing probability ratios could be greater than one for the case when a driver has two or more target lanes. In this case, the lane-changing probability ratios can be normalized as

$$\bar{p}_{DLC,i}^{lk}(t) = \frac{\max \left\{ 0, \frac{v_i^k(t) - v_i^l(t)}{v_{f,i}^l \tau} \Delta t \right\}}{\sum_{\forall k \neq l} \max \left\{ 0, \frac{v_i^k(t) - v_i^l(t)}{v_{f,i}^l \tau} \Delta t \right\}}, l \neq k,$$

$$\text{if } \sum_{\forall k \neq l} \max \left\{ 0, \frac{v_i^k(t) - v_i^l(t)}{v_{f,i}^l \tau} \Delta t \right\} > 1. \quad (4.10)$$

Another drawback revealed by (4.8) is the discretization effect on the transient response of traffic dynamics, which is also suffered by other macroscopic traffic flow models with or without lane changing, see, e.g., Zhong et al. (2013) and the references therein. Choosing a proper simulation time increment is important in this sense. As for the discrete-time lane-changing models,  $\Delta t$  affects the lane-changing probability ratios too, e.g., (4.8)-(4.10). Generally,  $\Delta t$  should be chosen such that  $\Delta t \leq \tau$ . This implies that the discretization step should not shadow the (average) time needed to execute a lane-changing maneuver, i.e.,  $\tau$ . Otherwise, the model cannot reflect the drivers'

concern on the execution time  $\tau$  so as to make a lane-changing decision. But when a smaller  $\Delta t$  is chosen, the cell length should be also reduced accordingly. The ideal case (as to discretize the fluid dynamics model) is that the mesh size should be chosen as small as possible (Laval and Daganzo, 2006). However, this is not realistic in simulation practice.

A reasonable choice of time increment for simulation of lane-changing models in terms of CTM and SCTM settings is the average time needed to execute lane-changing, i.e.,  $\tau$ . This is because the introduction of cells and the inherent within cell (by lane) homogeneous assumption of the CTM, lane changes can only happen at the upstream and downstream boundaries of cells. A driver facing a lane changing opportunity at a cell boundary has to make his/her decision, that is either to change to another lane or retain in the current lane for the next simulation time step rather than keep executing a lane-changing maneuver at the (same) cell boundary for several simulation time steps.

### **4.3 Speed-density relationship for traffic speed data assimilation**

From the definition of DLC probability ratio (or DLC demand), one needs to have cell traffic speed to proceed the calculation of DLC demand for each cell. However, this may not be possible since it is not realistic to obtain all cell speeds from the conventional point detectors. On the other hand, for those cell speeds that are not directly available from measurement, one can interpolate the cell speeds by the traffic speed-density relationship. However, the triangular fundamental diagram may be not a good choice for such kind of velocity data assimilation especially under free-flowing conditions. Herrera and Bayen (2010) claimed that it is not possible to observe the local density through speed measurements under free-flowing conditions using a triangular fundamental diagram. As found in Laval and Leclercq (2008); Jin (2010a); Work et al. (2010); Del Castillo (2012) that lane-changing traffic model may not well capture the lane-changing flows under free-flowing conditions if a triangular fundamental diagram

was used. Therefore, for velocity data assimilation, one may use the following speed density relation proposed by Ben-Akiva (1996), which has been adopted in Liu and Chang (2011) for modeling lane-changing maneuvers on urban arterials:

$$v_i^l(t) = \begin{cases} v_{f,i} & \text{if } \rho_i^l(t) < \rho_{\min}^l \\ v_{\min} + (v_{f,i} - v_{\min}) \left[ 1 - \left( \frac{\rho_i^l(t) - \rho_{\min}^l}{\rho_j^l - \rho_{\min}^l} \right)^\alpha \right]^\beta, & \text{if } \rho_i^l(t) \in [\rho_{\min}^l, \rho_j^l] \\ v_{\min} & \text{if } \rho_i^l(t) > \rho_j^l \end{cases} \quad (4.11)$$

Another possible choice is the fundamental diagram presented in Jin (2010a) and Del Castillo (2012) which extend the original fundamental diagram adopted in the LWR model to consider the lane-changing effect, see e.g., the Castillo-Benitez models (Del Castillo and Benitez, 1995a; Del Castillo and Benitez, 1995b; Del Castillo, 2012). An important property of the fundamental diagram<sup>8</sup>, i.e., the concavity property, was recognized. The speed-density model adopted in this study was developed for simulating the lane-changing traffic by a macroscopic dynamic model proposed by Jin (2010a) and Del Castillo (2012) to interpolate cell-lane traffic speed that is not directly available from GPS trajectory data.

---

<sup>8</sup> More recently, the role of the concavity condition within the framework of a variational theory of traffic flow has been more precisely described in Daganzo (2005). In this work, it was proven that if the flow-density relationship is strictly concave, the solution of the Lighthill-Whitham-Richard (LWR) model is a set of minimal cost or shortest paths in space-time coordinates and that such a set is the set of kinematic waves. Additionally, the concavity condition has also been required in Gentile et al. (2005) for the dynamic traffic assignment problem and by Jin (2010a) for a macroscopic dynamic model that includes the effect of lane-changing flow. Finally, the existence of a fundamental diagram for networks has also been pointed out by several researchers, e.g. Geroliminis and Daganzo (2008), Daganzo and Geroliminis (2008), wherein the network wide fundamental diagram has been demonstrated to be concave.

$$v_i^l(\rho_i^l) = v_{f,i}^l \left\{ 1 - \exp \left[ 1 - \exp \left( \frac{w_i^l}{v_{f,i}^l} \left( \frac{\rho_{j,i}^l}{\rho_i^l} - 1 \right) \right) \right] \right\} \quad (4.12)$$

Compared with Ben-Akiva's model (11), the Castillo-Benitez model (4.12) is smooth and admits fewer parameters to be calibrated.

**Remark 4.1** The spread of crowd-sourced traffic data, e.g., vehicle trajectory data, floating car data (or probe vehicles), portable smart devices, and social networks, is emerging traffic data in the era of ubiquitous sensing with "big data". The new data sources offer another possibility (compared with the model proposed in Chapter 3 that adopts conventional point detector data for simulation) to develop a real-time lane-changing traffic model by evaluating the cell-lane speed profiles along a freeway corridor and the lane-changing probability.

## 4.4 Stochastic elements and lane-changing maneuvers

### 4.4.1. Lane-changing and junction models

Schnetzler et al. (2012) proposed a method to schematize a link of a multilane freeway using junction models. Note that under the setting of the hybrid multilane CTM, lane changes can only happen at the boundaries of cells. To this end, lane-changing maneuvers between two adjacent lanes can be regarded as a sequence of diverges and merges. Then the lane-changing flow can be propagated by a pinpoint junction flow conservation. This schematization means that there is no conceptual difference between a multilane freeway with lane-changing flows and a junction with merging and diverging flows as illustrated in Figure 4.1. By this schematization, a unified (junction) approach can be developed for modeling freeway traffic dynamics with lane-changing maneuvers. In Schnetzler et al. (2012), the splitting models proposed by Daganzo (1995), Jin and Zhang (2003), Lebacque and Khoshyaran (2002), Lebacque and

Khoshyaran (2005) were tested and extended to adopt the lane-changing maneuvers. Flow propagation is calculated by a demand-supply reaction of junctions. This idea, being consistent with the link-node formulation of the network SCTM by Zhong et al. (2013), is adopted in this section to extend the SCTM framework to capture the dynamics of stochastic traffic flow with lane-changing maneuvers to result in a unified framework<sup>9</sup>.

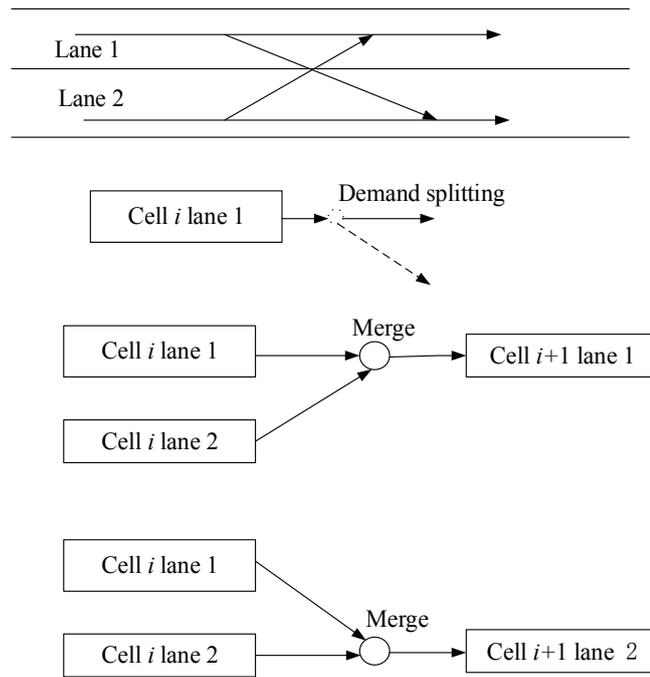


Figure 4.1. An illustration of junction representation of lane-changing

#### 4.4.2. Modeling lane-changing maneuvers under stochastic environment using junction models

---

<sup>9</sup> Indeed this is not realistic due to the fact that lane-changing maneuvers can take place whenever and wherever a driver can find a proper chance to change his/her lanes. The lane changes cannot be modeled by pinpoint junctions. However, due to the introduction of cells and the inherent within cell homogeneous assumption of the CTM, lane changes can only happen at the upstream and downstream boundaries of cells as aforementioned. Under this circumstance, the pinpoint junction approach to lane-changing is feasible.

As the lane-changing behavior is formulated in a unified link-node framework, the lane-changing probability ratio indeed acts as splitting ratio in the CTM/SCTM for network cases as previously explained. In this sense, the lane-changing probability ratio in this “network” lane-changing SCTM can be deterministic or random splitting ratio due to different assumptions and objectives of the simulation.

The deterministic lane-changing ratio can be explained as follows: The drivers cannot perceive the randomness effects caused by demand and supply uncertainties directly. For a driver on the road who would like to change his/her lane, the traffic condition he/she is facing is deterministic, which is regarded as a realization of the random field generated by demand and supply uncertainties. He/she makes a lane-changing decision based on the traffic condition, e.g., traffic speed and density, he/she can perceive, which is, of course, a realization of the random field (i.e., deterministic). Therefore, this lane-changing probability ratio wherein the drivers make their lane-changing decisions based on real traffic condition they perceived should be deterministic.

As it can be seen later, when the link-node SCTM formulation is used, the deterministic lane-changing ratio will be the deterministic splitting ratio (i.e., potential demand) as if node demand splitting in the network case. The supply uncertainty indeed affects the propagation of this demand in time and space as it will be discussed later in this chapter. Therefore, the definition given by Equation (4.8) (or Equation (4.10)) is still applicable to define the potential lane-changing demand. On the other hand, if it is assumed that drivers can perceive all the stochastic elements including the statistics of traffic speed  $v_i^k(t)$ , cell free-flow speed  $v_{f,i}$ , congestion backward wave speed (or shock wave speed for jammed traffic)  $w_{c,i}$ , and jam density  $\rho_{J,i}$ , which may make the definition of lane-changing ratio more general but at the price of loss of reality and increase of complexity. Therefore, the probability for a discretionary lane-changing maneuvers can be defined by following the probabilistic DLC ratio proposed by Laval and Daganzo

(2006) as follows:

$$p_{DLC,i}^{lk}(t) = \Pr\left\{\frac{v_i^k(t)-v_i^l(t)}{v_{f,i}^l\tau}\Delta t > 0\right\}, l \neq k. \quad (4.13)$$

However, in light of the speed-density relation of the form (4.12), it is challenging for us to obtain the distribution of cell-lane traffic speed  $v_i^k(t)$  in terms of an analytical way or by the characteristic function (or Fourier transform (Ng and Waller, 2010)) due to its hard nonlinearity. Nevertheless, the purpose of introducing the speed-density relation (4.12) is to interpolate the values of cell-lane speed for the cells without a detector (wherein a detection can be viewed as a realization of the corresponding random variable). It is also not realistic for the drivers to consider all the stochastic elements to make a lane-changing decision, e.g., it would be meaningless for a driver to consider the effect of possible congestion, i.e.  $w_{c,i}$  and  $\rho_{j,i}$  involved in (4.12) to obtain the statistics of cell-lane speed when the traffic condition he/she is facing is free-flowing. As a result, the deterministic cell-lane speed of (4.12) can be used to proceed the analysis so as to simplify the problem and to reduce the computational effort. Based on this argument, the lane-changing probability defined by (4.13) involves only one random variable  $v_{f,i}$  whose statistics can be easily obtained from real data. However,  $p_{DLC,i}^{lk}$  is not well defined after this simplification. To this end, it is revised as

$$p_{DLC,i}^{lk}(t) = \Pr(v_{f,i}^l\tau < (v_i^k(t) - v_i^l(t))\Delta t | v_i^k(t) - v_i^l(t) > 0), l \neq k, \quad (4.14)$$

that is the probability that the distance traveled after changing the lane (by the simulation time increment  $\Delta t$ ) be larger than the distance traveled by retaining the current lane (by the amount of time needed to execute the lane change  $\tau$ ) conditioned on the speed of the target lane is larger than that of the current lane. Remind that  $\tau$  is the average time takes to decide and execute a lane change, this probability describe the chance that the lane-changing maneuvers can save the travel time. Again, for the case when there are more than one target lane,  $\sum_{k \neq l} p_{DLC,i}^{lk}(t)$  may be larger than 1. Subsequently, the probability may be modified and be properly defined, that is

$$\tilde{P}_{DLC,i}^{lk}(t) = \frac{p_{DLC,i}^{lk}(t)}{\sum_{k \neq l} p_{DLC,i}^{lk}(t)}. \quad (4.15)$$

Then the probability for vehicles retain their lanes will be given by  $1 - \sum_{\forall k \neq l} \tilde{P}_{DLC,i}^{lk}(t)$ .

As depicted in Figure 4.2, the lane-changing maneuvers can be represented in the SCTM with the junction (or node) formulation wherein the demand splitting means potential demand to retain its lane and execute lane changes. Vehicles intend to change their lanes may not be received by the downstream cell-lane will retain their current lanes (or "queuing" in the current lane). Whether the DLC maneuvers will be further pursued depends on traffic conditions of the target lanes in the forthcoming simulation time steps. This is different from the MLCs, or actual traffic diverge. The readers are referred to Schnetzler et al. (2012) for more detailed discussion. This section would focus on extending the link-node hybrid CTM with lane-changing and speed interpolation to consider the stochastic elements aforementioned. The sending function of cell  $(i, l)$  can be defined as

$$S_i^l(t) = \text{mix} (v_{f,i}(t)\bar{\rho}_i^l(t), Q_i^l(t)) \quad (4.16)$$

where *mix* denotes the finite mixture distribution: Let  $\bar{\rho}_i$  be the joint density of cell  $i$ , which is defined as a finite mixture distribution. Denote the joint traffic density vector, its mean and covariance matrix as:  $\bar{\rho}(t)$ ,  $E(\bar{\rho}(t)|\theta(t))$ , and  $Var(\bar{\rho}(t)|\theta(t))$ , where  $\theta(t) = \{\theta_s(t)\}$ ,  $\theta_s(t) = (\rho_s(t), P_s(t))$ , and  $\rho_s(t)$  denotes the vector of cell densities of mode  $s$  comprising the finite mixture distribution. The probability density function (PDF) of the joint traffic density  $f(\bar{\rho}(t)|\theta(t))$  is defined as:

$$f(\bar{\rho}(t)|\theta(t)) = \sum_s P_s(t)f(\bar{\rho}(t)|\theta_s(t)). \quad (4.17)$$

Under the mixture model (17), the expectation  $E(\bar{\rho}(t)|\theta(t))$  is given by

$$E(\bar{\rho}(t)|\theta(t)) = \sum_s P_s(t)E(\rho_s(t)). \quad (4.18)$$

Let  $\mu_s(t) = E(\rho_s(t))$  and  $\mu(t) = E(\bar{\rho}(t)|\theta(t))$ . Then  $\mu(t) = \sum_s P_s(t)\mu_s(t)$ . To evaluate  $Var(\bar{\rho}(t)|\theta(t))$ , the covariance matrix of  $\rho_s(t)$  is defined as

$$\psi_s(t) = E((\rho_s(t) - \mu_s(t))(\rho_s(t) - \mu_s(t))^T).$$

Then the covariance matrix  $Var(\bar{\rho}(t)|\theta(t))$  can be evaluated as:

$$Var(\bar{\rho}(t)|\theta(t)) = \sum_s P_s(t)(\psi_s(t) + \mu_s(t)\mu_s^T(t)) - \mu(t)\mu^T(t). \quad (4.19)$$

The finite mixture distribution definition of the sending function (4.16) means that: if cell  $(i, l)$  is free-flowing at time  $k$ , the amount of traffic to be sent out is  $v_{f,i}(t)\bar{\rho}_i^l(t)$ ; if cell  $(i, l)$  is congested, the amount to be sent out is  $Q_i^l(t)$ . The probabilities for these two events are  $P_1^S(t) = \Pr(\bar{\rho}_i^l(t) \leq \rho_c^l(t))$ ,  $P_2^S(t) = 1 - P_1^S(t)$ , respectively. The demand splitting is then multiplied with this  $S_i^l(t)$  with the lane-changing probabilities.

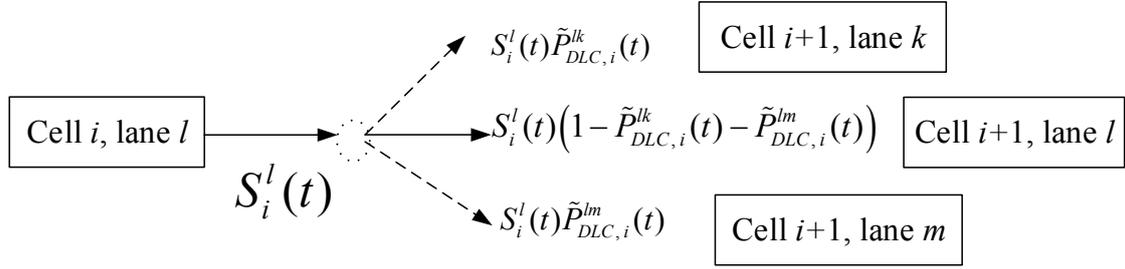


Figure 4.2. Demand splitting for a lane changing

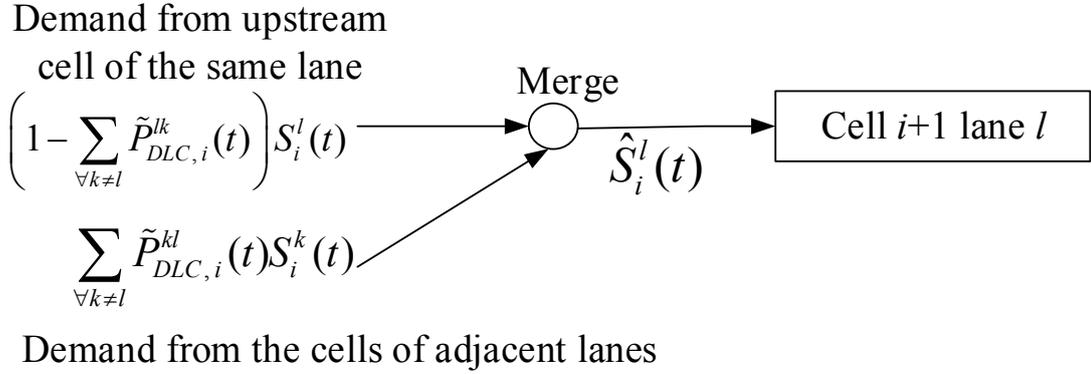


Figure 4.3. Lane-changing as a virtual merge

After the demand splits from source cell-lane, it will virtually “merge” into a particular cell-lane as depicted in Figure 4.3. The word “virtual” merge means that the potential demand which intends to make DLC would cancel the lane change or seek for future chances in the forthcoming simulation time steps (but keep on traveling in the current lane) if they fail to change lane in the current time step. This can be regarded as a key

difference between the actual merge (including MLC and traffic merge where the demand will queue up if they fail to merge) and “virtual” merge. Whether the lane-changing flow pattern can be merged into cell  $(i + 1, l)$  depends on the traffic state of cell  $(i + 1, l)$ . From the link-node formulation, the total travel demand to cell  $i+1$  is defined as  $\hat{S}_i^l(t)$ . Then the “actual” lane-changing flow can be determined by comparing  $\hat{S}_i^l(t)$  with the receiving function of cell  $(i + 1, l)$ . To this end, the four events are define below for calculating the flow propagation:

1. Cell  $(i + 1, l)$  is in free-flowing condition ( $F_{i+1}^l$ ), and  $\hat{S}_i^l(t)$  is less than its capacity. In this case, all the potential lane-changing demand  $\sum_{\forall k \neq l} \tilde{P}_{DLC,i}^{kl}(t) S_i^k(t)$  will be loaded onto cell  $(i + 1, l)$ . The corresponding probability is defined as:

$$P_1^{LC}(t) = \Pr\left(F_{i+1}^l(t) \cap \left(\hat{S}_i^l(t) < Q_{i+1}^l(t)\right)\right).$$

2. Cell  $(i + 1, l)$  is in free-flowing condition ( $F_{i+1}^l$ ), and  $\hat{S}_i^l(t)$  is greater than or equal to the capacity of cell  $(i + 1, l)$ . In this case, an amount of vehicles equals to  $Q_{i+1}^l(t)$  will be loaded. The amount of lane-changing demand  $\frac{\tilde{P}_{DLC,i}^{kl}(t) S_i^k(t)}{\hat{S}_i^l(t)} Q_{i+1}^l(t)$  from lane  $k$  to lane  $l$  can be accepted by the target cell-lane. The probability for this event is defined as:

$$P_2^{LC}(k) = \Pr\left(F_{i+1}^l(t) \cap \left(\hat{S}_i^l(t) \geq Q_{i+1}^l(t)\right)\right).$$

3. Cell  $(i + 1, l)$  is in congested condition ( $C_{i+1}^l$ ) and  $\hat{S}_i^l(t)$  is less than its available space. In this case, all the potential lane-changing demand  $\sum_{\forall k \neq l} P_i^{kl}(t) S_i^k(t)$  will be loaded onto cell  $(i + 1, l)$ . This probability of this event is defined as:

$$P_3^{LC}(t) = \Pr\left(C_{i+1}^l(t) \cap \left(\hat{S}_i^l(t) < w_{c,i+1}(t) \left(\rho_{j,i+1}^l(t) - \rho_{i+1}^l(t)\right)\right)\right).$$

4. Cell  $(i + 1, l)$  is in congested condition ( $C_{i+1}^l$ ) and  $\hat{S}_i^l(t)$  is greater than or equal to the available space of cell  $(i + 1, l)$ . In this case, an amount of vehicles that equals to the available space of cell  $(i + 1, l)$ , i.e.  $w_{c,i+1}(t) \left(\rho_{j,i+1}^l(t) - \rho_{i+1}^l(t)\right)$  will be

loaded onto it. The amount of lane-changing demand  $\frac{P_i^{kl}(t)S_i^k(t)}{\hat{S}_i^l(t)}w_{c,i+1}(t)(\rho_{j,i+1}^l(t) - \rho_{i+1}^l(t))$  from lane  $k$  to lane  $l$  can be accepted by the target cell-lane. The probability for this event is defined as:

$$P_4^{LC}(t) = \Pr\left(C_{i+1}^l(t) \cap \left(\hat{S}_i^l(t) \geq w_{c,i+1}(t)(\rho_{j,i+1}^l(t) - \rho_{i+1}^l(t))\right)\right).$$

By applying the independent argument, their probabilities of occurrence are:

$$P_1^{LC}(t) = \Pr(\bar{\rho}_{i+1}^l(t) < \rho_{c,i+1}^l(t))\Pr(\hat{S}_i^l(t) < Q_{i+1}^l(t)),$$

$$P_2^{LC}(t) = \Pr(\bar{\rho}_{i+1}^l(t) < \rho_{c,i+1}^l(t))\Pr(\hat{S}_i^l(t) \geq Q_{i+1}^l(t)),$$

$$P_3^{LC}(t) = \Pr(\bar{\rho}_{i+1}^l(t) \geq \rho_{c,i+1}^l(t))\Pr\left(\hat{S}_i^l(t) < w_{c,i+1}(t)(\rho_{j,i+1}^l(t) - \rho_{i+1}^l(t))\right),$$

$$P_4^{LC}(t) = \Pr(\bar{\rho}_{i+1}^l(t) \geq \rho_{c,i+1}^l(t))\Pr\left(\hat{S}_i^l(t) \geq w_{c,i+1}(t)(\rho_{j,i+1}^l(t) - \rho_{i+1}^l(t))\right),$$

with  $\sum_y P_y^{LC}(k) = 1$ . In line with Zhong et al. (2013), the probabilistic density function (PDF) for each lane-changing flow is defined as a joint merge flow of the lane-changing, denoted by  $\bar{s}_i^{kl}(t)$ :

$$g_i^{kl}(\bar{s}_i^{kl}(t)|\varphi_i^{kl}(t)) = \sum_y P_y^{LC}(t) \cdot g_i^{kl}(\bar{s}_i^{kl}(t)|\varphi_{i,y}^{kl}(t)), \quad (4.20)$$

where  $\varphi_i^{kl}(t) = \{\varphi_{i,y}^{kl}(t)\}$ ,  $\varphi_{i,y}^{kl}(t) = (P_y^{LC}(t), \varphi_{i,y}^{kl}(t))$ , with  $\varphi_i^{kl}$  contains four events previously. To be more specific, the lane-changing flows for each discretionary lane-changing maneuver corresponding to the four events are:

$$s_{i,1}^{kl}(t) = \tilde{P}_{DLC,i}^{kl} S_i^k(t), \quad s_{i,2}^{kl}(t) = \frac{\tilde{P}_{DLC,i}^{kl} S_i^k(t)}{\hat{S}_i^l(t)} Q_{i+1}^l(t),$$

$$s_{i,3}^{kl}(t) = \tilde{P}_{DLC,i}^{kl} S_i^k(t), \quad \text{and}$$

$$s_{i,4}^{kl}(t) = \frac{\tilde{P}_{DLC,i}^{kl} S_i^k(t)}{\hat{S}_i^l(t)} w_{c,i+1}(t)(\rho_{j,i+1}^l(t) - \rho_{i+1}^l(t)). \quad (4.21)$$

Traffic dynamics can be evaluated by the density version of flow conservation which is stated as:

$$\rho_i^l(t + \Delta t) = \rho_i^l(t) - \frac{\Delta t}{L_i} \left( \sum_{\forall k} \bar{s}_i^{lk}(t) - \sum_{\forall k} \bar{s}_{i-1}^{kl}(t) \right). \quad (4.22)$$

Evaluation of the mean and variance of traffic density defined by (4.22) is

straightforward for it is of a linear form. As a more general case has been discussed by Sumalee et al., (2011) in details, it is therefore omitted here for brevity. The flow propagation proposed in this section is similar to that of Zhong et al. (2013) which calculates the flow propagation by pairing up two neighboring cells.

#### 4.5 Incremental-Transfer (IT) principle and the issue of priority

The above flow propagation for lane-changing maneuvers can be viewed as a stochastic extension of the IT principle proposed by Daganzo et al. (1997) and Daganzo (2005). An interesting issue is the non-uniqueness of the intersection flows. It is reported that the non-uniqueness of the intersection flows can be resolved probabilistically by the stochastic network loading framework, e.g., the link-node merge/diverge formulation in the SCTM (Corthout et al., 2012). Just as a stochastic approach to the dynamic traffic assignment problem yields a unique solution in distributional terms, a stochastic dynamic network loading model could replace non-unique intersection flows by a unique distribution (Floterod et al., 2011).

Finally, the issue of priority is also interesting and of importance in lane-changing (e.g., the HOV lanes) and node models (Tampere et al., 2011, Floterod and Rohde, 2011). This issue has been thoroughly discussed in Corthout et al. (2012) for macroscopic first-order intersection (or junction, node) models. Both the hybrid CTM and the multilane SCTM can be easily extended to consider the priority ratios in line with Corthout et al. (2012), Zhong et al. (2013).

Take the service fraction in the hybrid CTM for example, a new service fraction considering priority ratios can be defined as

$$f_{i+1}^l(t) = \min \left\{ 1, \frac{\alpha_i^{kl}(t)R_{i+1}^l(t)}{\sum_{\forall k} \alpha_i^{kl}(t)\tilde{S}_i^{kl}(t)} \right\}, \quad (4.23)$$

where  $\alpha_i^{kl}$  is a strictly positive finite priority parameter that determines the priority of

each competitor of the supply  $R_{i+1}^l(t)$ . The actual demand that can be sent from cell  $(i, k)$  to  $(i+1, l)$  is then

$$s_i^{kl}(t) = \min \left( \tilde{S}_i^{kl}(t), \frac{\alpha_i^{kl}(t)\tilde{S}_i^{kl}(t)}{\sum_{\forall k} \alpha_i^{kl}(t)\tilde{S}_i^{kl}(t)} R_{i+1}^l(t) \right). \quad (4.24)$$

The priority ratio means the competitive strength of demand from cell  $(i, k)$  for supply  $R_{i+1}^l(t)$  of cell  $(i+1, l)$ . The definition of the minimum operator in (4.24) implies that priority must not be used when the total supply exceeds the total demand. In light of this, the priority ratios can be introduced to the probabilistic events which define the stochastic lane-changing flow by revising the events with demand exceeds supply, i.e., events 2 and 4 in the previous section. The probabilities, i.e.  $P_i^{LC}(t)$ ,  $i = 1,2,3,4$ , would retain the same whereas the lane-changing flows for each lane-changing maneuver corresponding to the four events are:

$$\begin{aligned} s_{i,1}^{kl}(t) &= P_i^{kl}(t)S_i^k(t), \\ s_{i,2}^{kl}(t) &= \frac{\alpha_i^{kl}(t)P_i^{kl}(t)S_i^k(t)}{\tilde{S}_i^l(t)} Q_{i+1}^l(t), \\ s_{i,3}^{kl}(t) &= P_i^{kl}(t)S_i^k(t), \text{ and} \\ s_{i,4}^{kl}(t) &= \frac{\alpha_i^{kl}(t)P_i^{kl}(t)S_i^k(t)}{\tilde{S}_i^l(t)} w_{c,i+1}(t) (\rho_{j,i+1}^l(t) - \rho_{i+1}^l(t)), \end{aligned} \quad (4.25)$$

where  $\tilde{S}_i^l(t)$  is defined as the product of the priority ratios and the lane-changing demands as in (4.24).

## 4.6 On-/off- ramps and capacity drop

The lane-changing rule as mentioned in Section 4.1-4.5 is only applicable for DLCs. However, MLC and DLC would simultaneously exist in a freeway corridor concerning different locations, e.g., on-/off- ramps, lane blockage incidents, and traffic regulations as depicted in Figure 4.4. A potential approach is to extend the MLC model developed in Chapter 3. However, as it can be observed from the development of the MLC model in Chapter 3 that the model is too complicated to be extended to consider the stochastic

elements. Note from the literature and the empirical study in Chapter 3 that capacity drop is one of the major consequences of MLC. Using capacity drop models is a common approach to simulate the impact of MLC on the freeway traffic flow, see Kontorinaki et al. (2016) for an overview. This section will refine the stochastic lane-changing model developed in previous sections to simulate MLC maneuvers by a capacity drop model. On the other hand, the original SCTM consider the random fluctuations in the fundamental diagram with emphasis on the random capacity. Therefore, a capacity drop model describing the effects of MLC maneuvers would better suit the SCTM framework for multilane traffic.

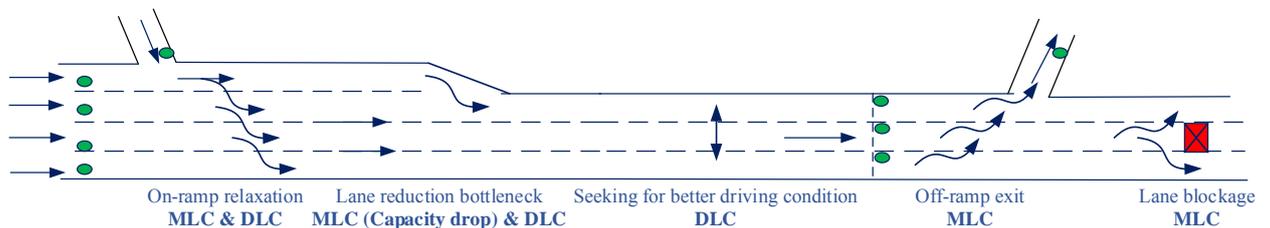


Figure 4.4. Lane-changing maneuvers concerning different locations and traffic scenarios

For the on-ramp and the lane-drop cases, see, e.g., Figure 4.5, a capacity drop model proposed by Leclercq et al. (2011) in conjunction with the aforementioned lane-changing rule will govern the traffic dynamics. The capacity drop model takes into account the capacity reduction due to the lane-changing traffic while the lane-changing rule is used to reproduce lane flow distributions around the weaving areas. It is a consensus that vehicle lane-changing maneuver is one of the major reasons that cause capacity drop especially in the areas such as the areas near a merging junction and lane-drops, upstream to a diverging junction, inside a weaving section or around a cloverleaf interchange and a toll plaza. Leclercq et al. (2011) derived that the change of capacity at merges with respect to the demand on the on-ramp when the main road has only one lane. Under uncongested on-ramp case, the relative capacity drop  $\phi$  can be roughly estimated as:

$$\phi = 1 - (q_1 + \lambda_0)/Q_M, \quad (4.26)$$

where  $\lambda_0$  is the on-ramp demand,  $q_1$  is the total upstream flow of the shoulder lane

before the beginning of the insertion section and  $Q_M$  is the capacity given by the fundamental diagram, saying  $Q_M = v_f \cdot w_c \cdot \rho_f / (v_f + w_c)$ , which is the fixed maximal traffic flow that could be achieved at the specific location of the freeway. For the uncongested on-ramp case, the actual on-ramp flow  $q_0$  should be equal to  $\lambda_0$ , i.e.  $\lambda_0 = q_0$ . For the congested on-ramp case, the relative capacity drop  $\phi$  can be evaluated approximately by:

$$\phi = 1 - \frac{\left(1 + \frac{1}{\alpha}\right)q_0}{Q_M}, \quad (4.27)$$

where  $\alpha$  is the average merging ratio. The experimental results presented in Leclercq et al. (2011) show that the analytical capacity drop model is promising by providing an efficient and elegant way to estimate the upper (when the freeway is uniformly congested) and lower (when congestion only occurs on shoulder lane) bounds of the capacity drop, while this variant value is defined as effective capacity in Leclercq et al. (2011). As claimed in Leclercq et al. (2011) that by adopting this capacity drop model, the ramp flows can be evaluated by the conventional Incremental-Transfer (IT) and/or Priority IT (PIT) principles. These results can be adopted here to adjust the capacities of the weaving areas, e.g., the cells with on-ramps, to increase accuracy. However, lacking a lane flow distribution model is deemed to be the main limitation of the model proposed by Leclercq et al. (2011) when the main road has several lanes. The lane-changing model proposed in this chapter can be used as a supplement to the capacity drop model to reproduce lane flow distributions at merges.

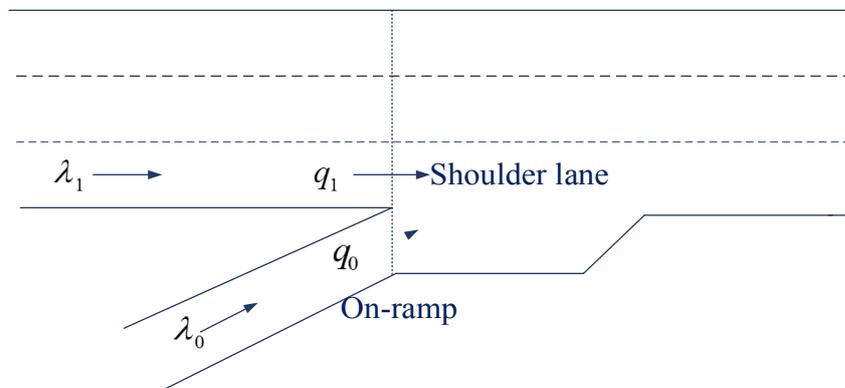


Figure 4.5. Lane-changing near the on-ramp

For the MLC around an off-ramp (or lane blockage incident locations) as depicted in Figure 4.4, Lee and Cassidy (2009) claimed that the driver's decision to attempt an MLC is affected by several factors such as the distance to the desired destination, the number of lanes to be crossed. In Chapter 3, an exponential probabilistic distribution function was proposed to describe the cumulative MLC demand. This function assumes that drivers only use the current traffic condition spatially ahead that they could perceive to make their MLC decisions. The effect of the perceived traffic condition decreases exponentially with respect to the distance away from her/his current position (Yang and Koutsopoulos, 1996). The level of urgency function in the model considers similar factors as in Lee and Cassidy (2009), e.g., the number of lanes to be crossed in reaching the desired destination and the difference in densities between the driver's target lane and her current one. All these models can provide a more accurate description on the MLC maneuvers at the price of increasing the model complexity and having more parameters to be calibrated, which have introduced difficulty in extending these models to stochastic case.

Simulations in Schnetzler et al. (2012) have indicated that as soon as a lane is highly congested, except for the case of insertion at the tail of the queue, drivers postpone their unsatisfied wish to change lanes or to make an MLC at the last minute. Part of the off-ramp demand joins the tail of queue at the shoulder lane to get to the off-ramp while the other part chooses to travel on Lane 1 till they have to make an MLC. If its maneuver fails, then this vehicle has to slow down (or even stop) which in turn blocks the direct upstream traffic on lane 1. These MLC at low speed cause the traffic to slow down on the adjacent lanes. A similar observation was achieved by Zhang and Ioannou (2017) for MLC under incident scenarios. When drivers approach the incident spot along the affected lanes but without being aware that their lanes are blocked, they have to slow down to see MLC chances. These MLC maneuvers at low speed cause the traffic on the adjacent lanes to slow down leading to a low discharging rate of the temporary bottleneck. Zhang and Ioannou (2017) and Kontorinaki et al. (2016) suggest adding a

capacity drop to traffic flow models to take into account such kind of phenomenon. To sum up, the capacity would drop to  $\tilde{Q}_M = (1 - \epsilon)Q_M$  once a queue forms, where  $0 < \epsilon < 1$  can be a constant or a Gaussian random variable in line with the SCTM.

#### **4.7 An overview of the proposed SCTM with lane-changing maneuvers**

An overview of the proposed lane-changing models and their relationship is depicted in the framework as Figure 4.6. Both models are based on the measurements of boundary variables and cell-lane speeds (if available). Similar to the conventional CTM, measured traffic flow variables are used to calibrate the flow-density diagram but at the lane level. The speed measurements (at the boundaries) and cell-lane speeds (if available) are used to interpolate the speeds for all cell-lanes. Traffic speeds from both measurement and interpolation are adopted to define lane-changing (probability) ratios which are regarded as node demand splitting ratios in the flow models. The node demand is then propagated to the downstream cell-lanes based on the IT or PIT principles. The lane-changing CTM updates the traffic state based on deterministic flow conservation in conjunction with the flow propagation laws defined by the flow-density diagram. On the other hand, the lane-changing SCTM defines probabilistic events based on the demand and supply reaction and calculates the overall effect of the probabilistic events by finite mixture distribution to propagate the flow in a stochastic sense. Pseudo code realizations of the proposed lane-changing models are given in Figures 4.7 and 4.8.

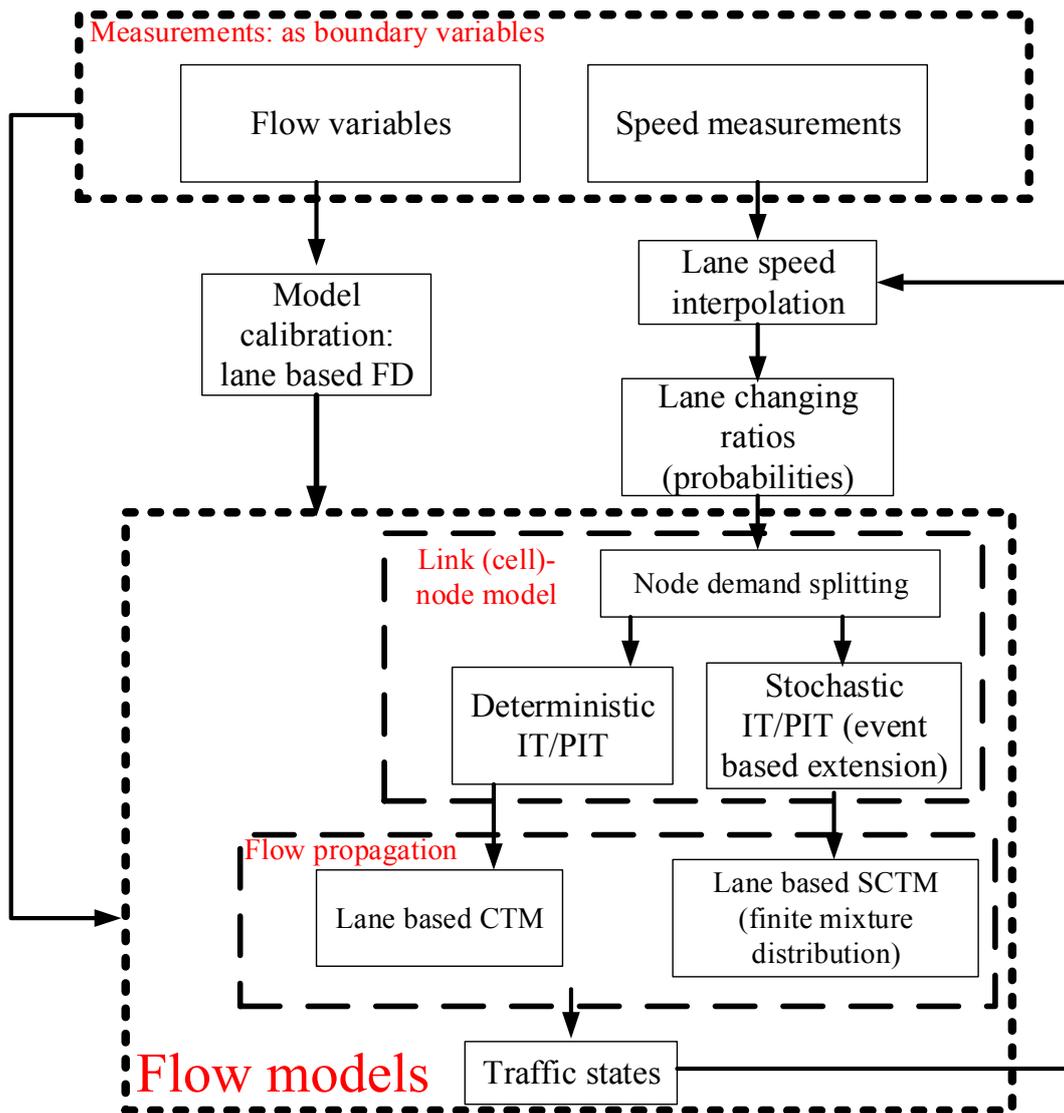


Figure 4.6. An overview of the proposed models and their relationship

**Algorithm** Pseudo code for the proposed CTM with lane-changing

- 1: Calibration of the fundamental diagram for each lane, the average execution time for lane-changing  $\tau$
  - 2: Initialization: Set initial conditions, simulation configurations, dimension of variables, and  $t = 1$
  - 3: Specification and loading the boundary variables and speed-density mappings
  - 4:  $t \leftarrow 2$
  - 5: **while**  $t \leq Iter_{max}$ , **do**
  - 6: CTM with lane changing. Simulate the following recursive equations:
    - 6.1: Define: cell-lane sending function (4.1), receiving function (4.3)
    - 6.2: Resolve speed from density according to (4.12)
    - 6.3: Define: lane-changing probability ratios (4.10) and potential lane-changing demand (4.2)
    - 6.4: Alleviate potential lane-changing demand by IT principle (4.4)-( 4.5) or PIT principle (4.24)-( 4.25)
    - 6.5: Update traffic dynamics according to flow conservation (4.6) or (4.7)
  - 7:  $t \leftarrow t + 1$
  - 8: **end while**
  - 9: **return**
- where  $Iter_m$  is the simulation horizon

Figure 4.7. Pseudo code for the proposed hybrid CTM with lane changing

**Algorithm** Pseudo code for the proposed SCTM with lane-changing

- 1: Calibration of the fundamental diagram for each lane, the average execution time for lane-changing  $\tau$
  - 2: Initialization: Set initial conditions, simulation configurations, dimension of variables, and  $t = 1$
  - 3: Specification and loading the boundary variables and speed-density mappings
  - 4:  $t \leftarrow 2$
  - 5: **while**  $t \leq Iter_{max}$ , **do**
  - 6: SCTM with lane changing. Simulate the following recursive equations:
    - 6.1: Define: cell-lane sending function (4.16)
    - 6.2: Resolve speed from density according to (4.12)
    - 6.3: Define: lane-changing probability ratios (as node demand splitting ratios) (4.10) (in a deterministic manner) or (4.14)-(4.15) (in a stochastic manner) and potential lane-changing demand in line with Figure 3
    - 6.4: Define the probabilistic events with probabilities of occurrence and the potential lane-changing flows by IT principle (4.22) or PIT principle (4.27), accordingly
    - 6.5: Evaluate the finite mixture of the probabilistic events by (4.20)
    - 6.6: Update traffic dynamics according to flow conservation (4.23)
  - 7:  $t \leftarrow t + 1$
  - 8: **end while**
  - 9: **return**
- where  $Iter_m$  is the simulation horizon

Figure 4.8. Pseudo code for the proposed SCTM with lane changing

For multi-scale traffic simulation (with respect to different accuracy requirements and simulation time increments, etc.), e.g., the CTM/SCTM with lane-changing and the original CTM/SCTM, simple interfaces are needed to be defined to translate the boundary conditions when changing the traffic description. For the design of microscopic-macroscopic and macroscopic-microscopic interfaces, Leclercq (2007) proposed a generalized definition of the demand and supply which uses reservoirs that stock vehicles when arriving and spread the right value of flow, which guarantees the consistency of boundary condition in terms of flow. For the multilane-macroscopic interface as depicted in Figure 4.9, if there are measurement devices, the traffic detections can be used as boundary conditions. If no measurement device is available, there is a need to use the aggregate flow and density as the downstream boundary condition for the macroscopic model. For the macroscopic-multilane interface, the same logic can be applied to the case with measurement devices. For the case without measurement device, there is a way to distribute the aggregate flow from the macroscopic model in line with the normalized lane-changing ratios.

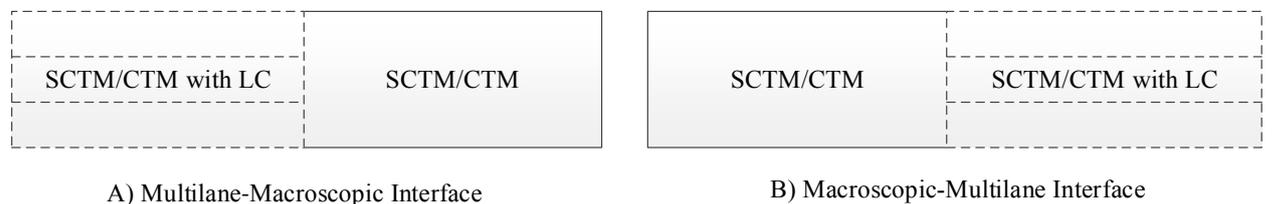


Figure 4.9. Multilane-macroscopic interface and macroscopic-multilane interface

## 4.8 Numerical tests

This section aims to examine the properties of the proposed CTM and SCTM considering lane-changing maneuvers with cell-lane speed interpolation. First, the following simple example will be tested. Consider a freeway segment with two lanes (one narrow shoulder lane for emergency use only) be of 0.8 miles long as depicted in Figure 4.10. The freeway segment is divided into four homogenous cells. It is assumed that the lanes are with the same nominal fundamental diagrams. To be specific, the nominal free-flow speed is  $v_f = 60$  miles/hour, the nominal critical density is

30 P.C.U./mile/lane, the nominal congestion wave speed is 10 miles/hour, and the nominal jam density is 210 P.C.U./mile/lane. Simulation time increment is chosen as 3 sec. It is assumed with uniform arrivals with flow rates be of 1500 P.C.U./hour/lane to both lanes.

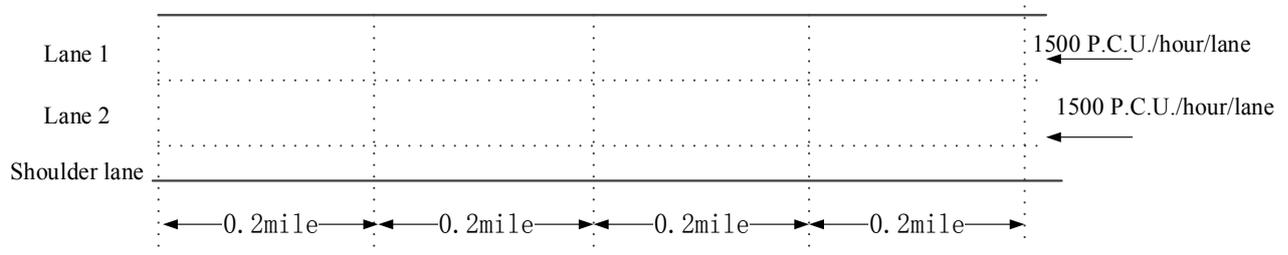


Figure 4.10. A freeway segment with two lanes

#### 4.8.1 Scenario 1: A capacity drop on lane 2

It is assumed that there is a maintenance work on the shoulder lane of cell 3 which yields a new fundamental diagram for cell 3 on lane 2 as  $v_f = 30$  miles/hour, the nominal critical density is 30 P.C.U./mile/lane, the nominal congestion wave speed is 10 miles/hour, and the nominal jam density is 120 P.C.U./mile/lane. Simulation results by Daganzo's CTM without lane changing<sup>10</sup> is depicted in Figures 4.11-4.12. As expected, lane 1 maintains its free-flow condition as lane-changing is not allowed and the inflow to lane 1 is under its capacity. By contrast, congestion builds up on lane 2 and spills back to its first cell. Traffic densities of the first two cells of lane 2 approach to a steady state defined by the jam density of cell 3, while the downstream cell of lane 2 (cell 4, lane 2) is of the lowest traffic density 15 P.C.U./mile/lane (compared with other three cells) as the its inflow is saturated by the bottleneck cell 3, saying 900 P.C.U./hour/lane. To be noticed, cell 3 is of its critical density, 30 P.C.U./mile/lane.

<sup>10</sup> Note that lanes are differentiated in line with Cheu et al. (2009) although lane-changing is not allowed in this case.

Suppose that lane changes are now allowed between the two lanes. There is a need to apply the lane-changing rules and the speed interpolation to run the simulation. Results are illustrated in blue dot lines of Figures 4.11-4.12. Significant changes in traffic densities can be observed by comparing the two simulations. The balancing effect of the lane-changing traffic is obvious. Theoretically, drivers will keep on changing their lanes until no heterogeneous traffic conditions can be observed, i.e., drivers cannot make any benefit by changing their lanes. However, the simulation results reveal that the steady states of traffic densities on lane 1 and lane 2 are not the same. For example, the difference between the steady states of the two lanes of cell 1 is about  $1 P.C.U./mile/lane$ . No further lane change occurs (indeed the value of lane-changing probability is very small due to the fact that speed heterogeneity is small as shown in Figures 4.13-4.14) even though this heterogeneity in traffic condition do exist. This is because that the benefit a driver can obtain from changing his/her lane is too small as compared with the time he/she needed to pay to execute the lane change. However, when the drivers approaches to the bottleneck, they would prefer to change to lane 1 to save time. Again, a difference about  $1 P.C.U./mile/lane$  between the steady states of the two lanes can be observed. However, the cause is different from that of the previous case. In this case, the lane-changing ratios approach to their steady states even though there exists small heterogeneity in traffic densities.

The speed heterogeneity also approaches a steady state. This speed heterogeneity cannot trigger furthermore lane-changing demand as the externality caused by a further new lane change (including the inherent execution time) would be larger than the benefit a driver can obtain from changing his/her lane. The drivers will keep on changing their lanes to maintain the new steady state in terms of traffic density as demonstrated in the figures. Traffic conditions of the cells in the downstream of the bottleneck remain similar to those without lane changing. Note also that traffic densities of cell 4 of the two lanes are quite different. Due to the setting, cell 3 of lane 2 admits a deficient traffic speed while the speed heterogeneity between this cell and cell 4 of

lane 1 (target cell) is so large that all drivers would like to switch to the target cell whenever possible. On the other hand, which is more important in this case, because cell 4 is connected to a sink that no further lane-changing will occur to balance the traffic densities of the two lanes. *Chen et al. (2009)* claimed that the CTM without lane-changing produces larger delay than that produced by the CTM with lane changing. This can be observed from the simulations that the capacity of lane 1 is wasted as vehicles queue up on lane 2 rather than switch to lane 1 to travel faster.

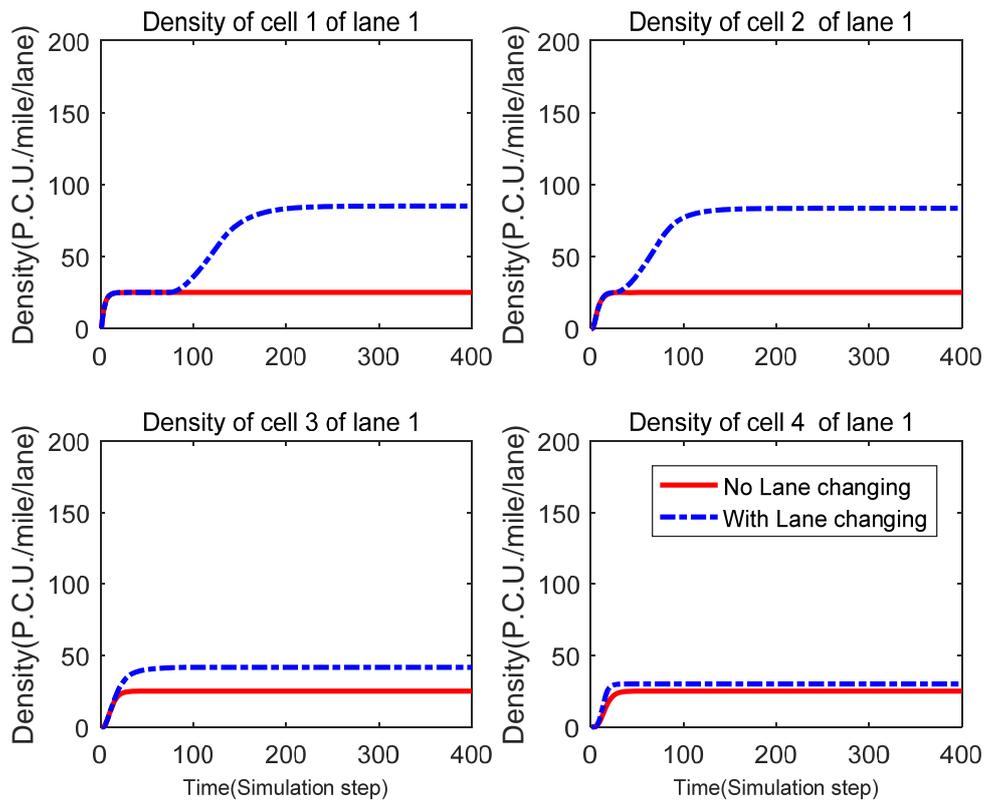


Figure 4.11. Comparison of traffic densities by CTM with and without lane-changing for Scenario 1

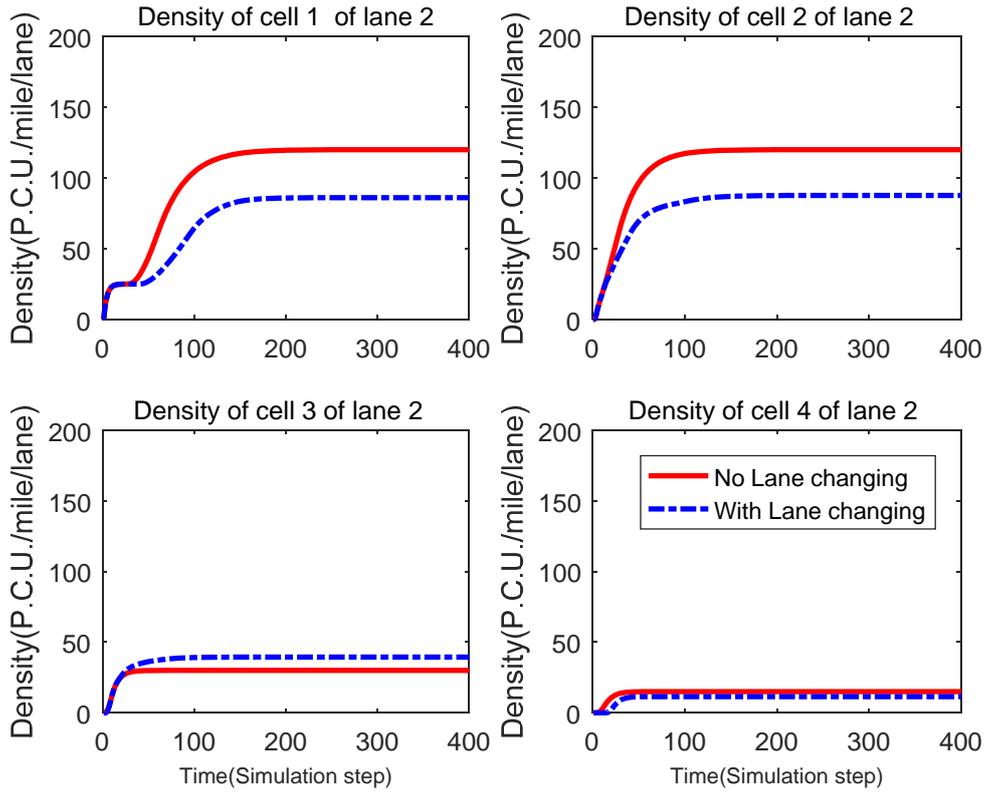


Figure 4.12. Comparison of traffic densities by CTM with and without lane-changing for Scenario 1

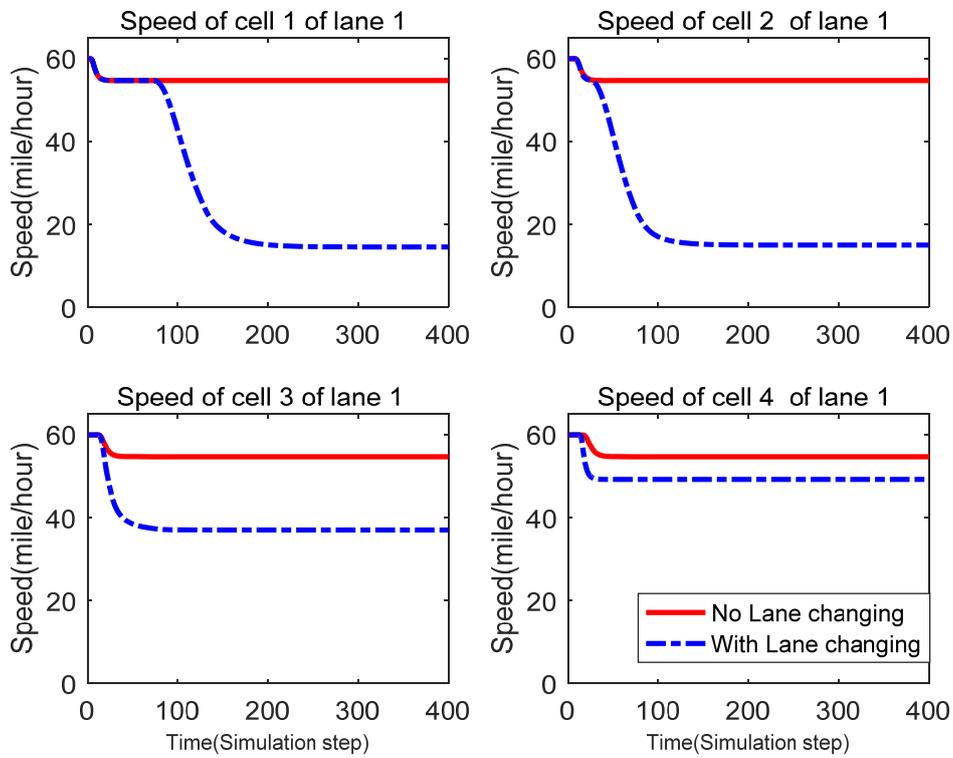


Figure 4.13. Comparison of traffic speeds by CTM with and without lane-changing for Scenario 1

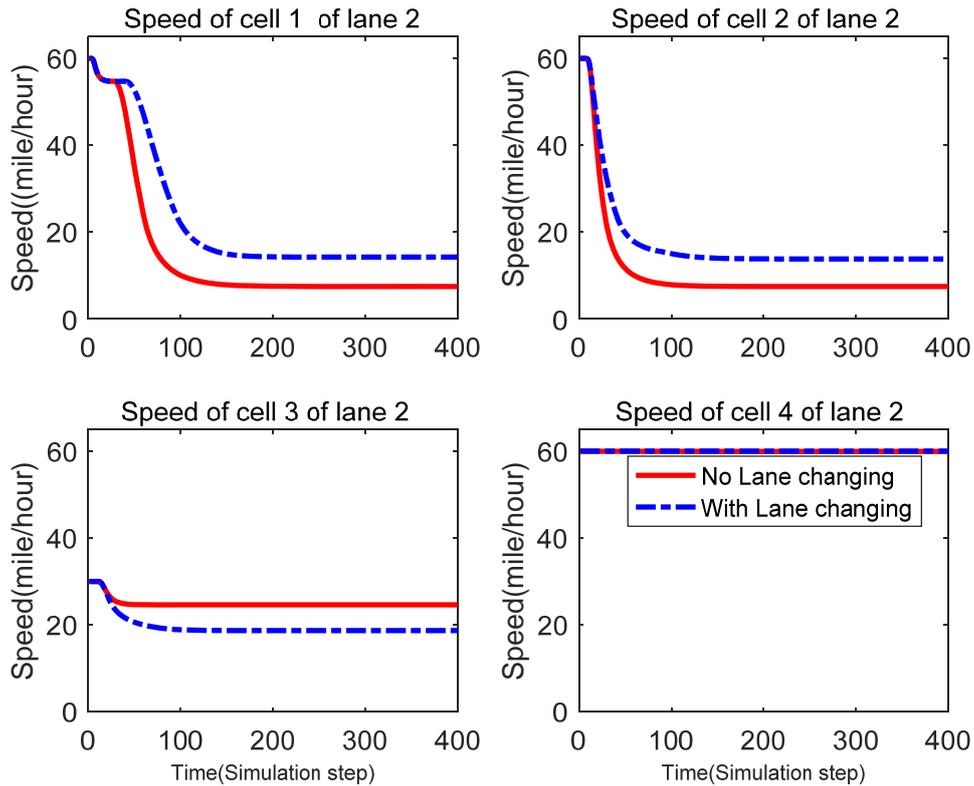


Figure 4.14. Comparison of traffic speeds by CTM with and without lane-changing for Scenario 1

#### 4.8.2 Scenario 2: A slow-moving vehicle on lane 2

The inflow rates of both lanes are increased to 1800 P.C.U./hour/lane. It is assumed that a slow-moving vehicle enters lane 2 at  $t = 101$  simulation time step from the upstream boundary and traverse the freeway segment at an average speed of 40 miles/hour. During the period that this vehicle is traveling on lane 2, it acts as a speed limit for this lane and thus can be regarded as a moving bottleneck. Drivers following the vehicle can retain their lane or switch to lane 1 to travel faster. But have in mind that lane 1 is also operating under its critical condition. It cannot handle too much lane-changing demand or the speed would slow down to that of lane 2.

Simulation results applying CTM without lane-changing are depicted in Figures 4.15-

4.16. Without lane-changing vehicles from lane 2, lane 1 operates under the steady state defined by its supply functions and the demand profile. Lane 2 also operates under the steady state defined by its supply functions and the demand profile until the slow-moving vehicle enters the lane. Then this vehicle acts a moving bottleneck along the lane as reflected by the traffic density evolution. For cells 2-4, traffic densities decrease after the slow vehicle entering the lane for certain time periods. This is because the outflow rate decreases at the bottleneck downstream. Cell traffic densities increase when the moving bottleneck enters their boundaries. Finally, traffic states approach their steady states defined by the supply functions and the demand profile of the freeway segment after the moving bottleneck vanishes.

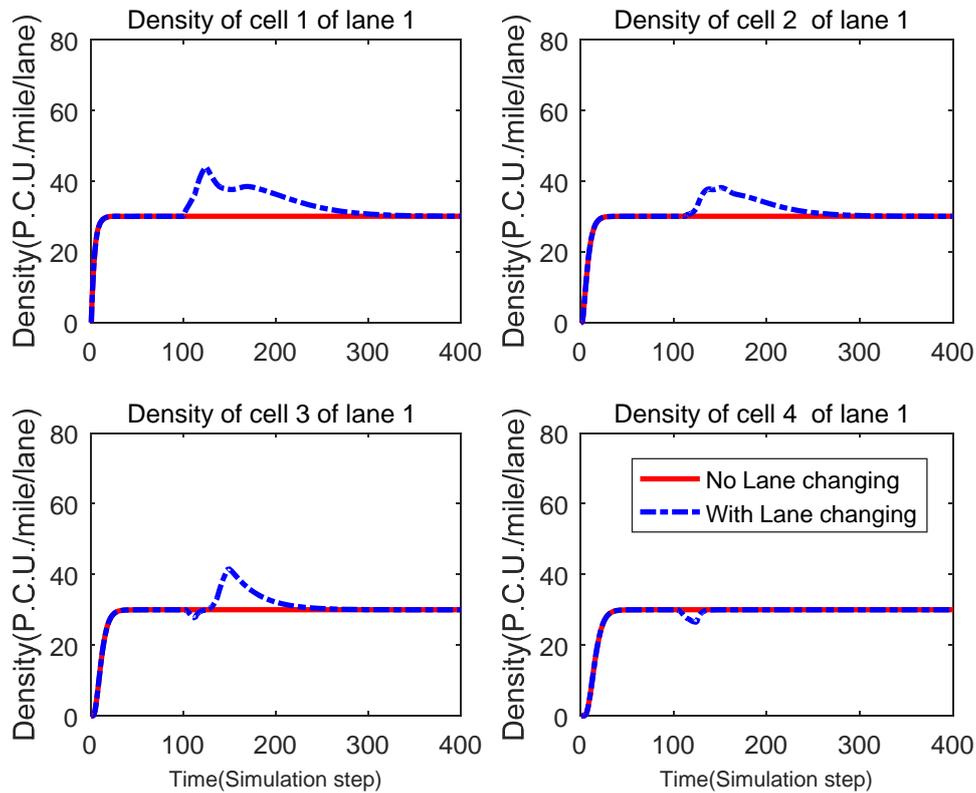


Figure 4.15. Comparison of traffic densities by CTM with and without lane-changing for Scenario 2

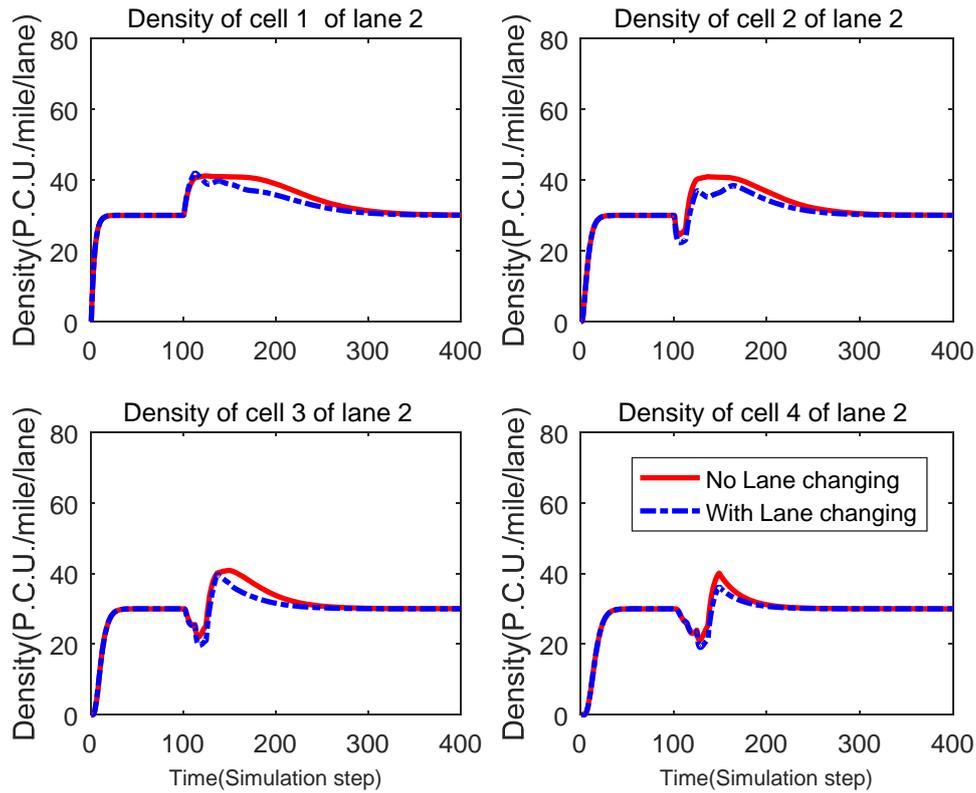


Figure 4.16. Comparison of traffic densities by CTM with and without lane-changing for Scenario 2

For comparison, the effects of lane changes are simulated. Traffic conditions are the same before the slow-moving vehicle appears. However, after the appearance of the slow-moving vehicle, traffic speed of cell 1 of lane 2 reduces to 40 miles/hour. Traffic density increases accordingly. Drivers change their lanes to travel faster. Traffic densities of the downstream cells of lane 2 decreases after the slow vehicle show up which is of the same reason as previously explained, i.e., the outflow rate of the bottleneck decreases. The interesting thing is that cells densities of lane 1 show concave shapes for a certain period, i.e., traffic densities of cells 3 and 4 of lane 1 decreases for a short period. This is because of vehicles following the slow vehicle switch to lane 1. As a result, cell traffic densities of lane 1 increase accordingly. Meanwhile, traffic conditions of the downstream cells of lane 2 become better after the appearance of the slow-moving vehicle, i.e., the downstream cells of the moving bottleneck on lane 2 have better traffic conditions than the corresponding cells on lane 1. The reason is that

the outflow rate is saturated by the moving bottleneck (see Figure 4.16). Better traffic conditions trigger lane-changing maneuvers from lane 1 to lane 2 which in turn yields the concave shapes of cell densities of lane 1. Finally, traffic states approach the steady state defined by its supply functions and the demand profile. These simulation results prove that the proposed CTM with lane-changing and speed interpolation can capture the effects of lane-changing maneuvers.

### 4.8.3 Scenario 3: A multilane freeway with an on-ramp

In this case, a hypothetical multilane freeway is considered with on-ramp as depicted in Figure 4.17. It is assumed that lanes 1, 2, 3 admit the same fundamental diagrams with nominal free-flow speed  $v_f = 60 \text{ miles/hour}$ , nominal critical density  $30 \text{ P.C.U./mile/lane}$ , nominal congestion wave speed  $10 \text{ miles/hour}$ , and nominal jam density is  $210 \text{ P.C.U./mile/lane}$ . The on-ramp is connected to cell 2 of lane 4 which yields the cell admits a fundamental diagram with nominal free-flow speed  $v_f = 30 \text{ miles/hour}$  while others remain the same as the other lanes. It is assumed that uniform arrivals with flow rates be of  $1500 \text{ P.C.U./hour/lane}$  to all the lanes and the ramp admits demand at a flow rate of  $100 \text{ P.C.U./hour/lane}$ . Three scenarios are simulated, i.e., Case 1) the case with no lane changing; Case 2) the case with free lane changing; and Case 3) the case that lane-changing from lane 3 to lane 2 is forbidden. In this scenario, it is chosen that  $\tau = 4 \text{ sec}$  and  $\Delta t = 4 \text{ sec}$ .

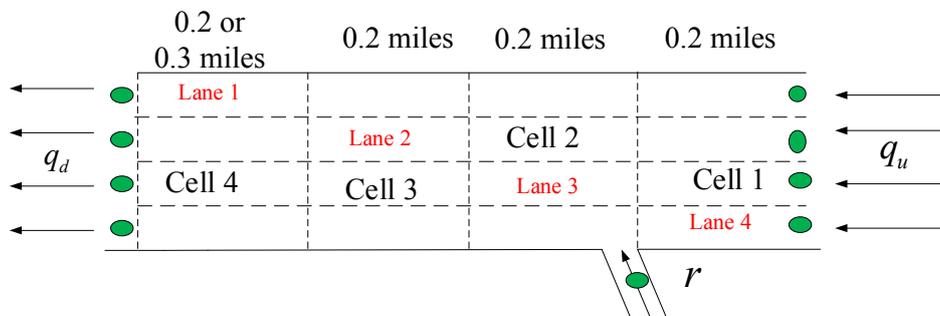


Figure 4.17. A hypothetical multilane freeway with an on-ramp

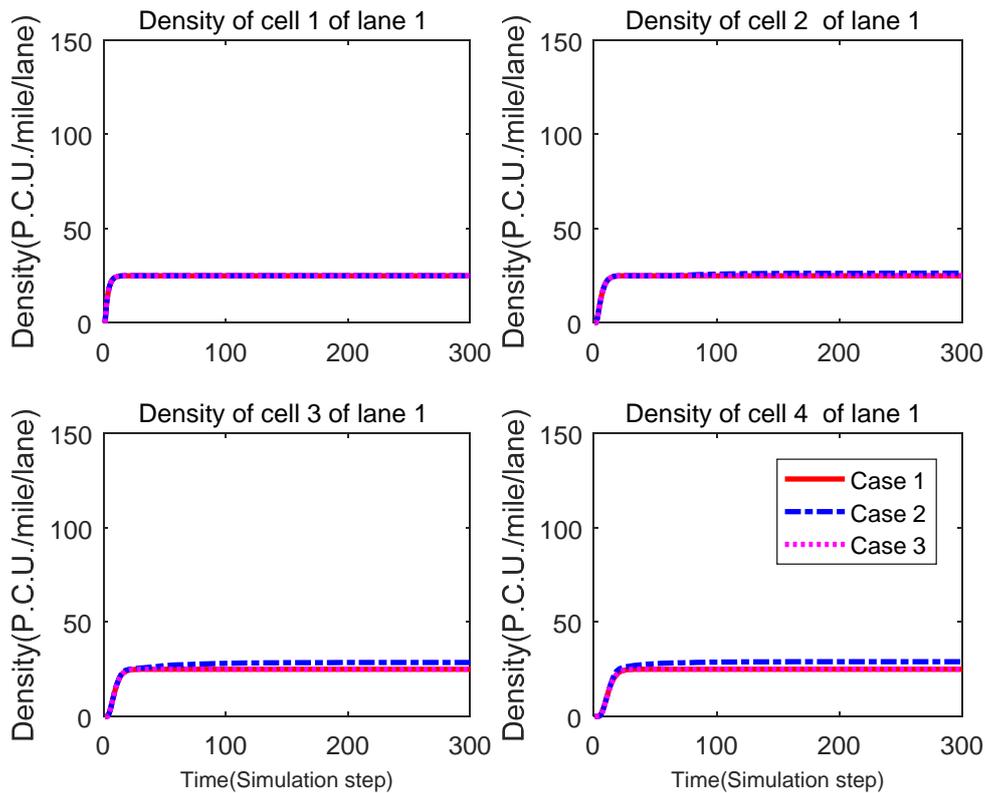


Figure 4.18. Comparison of traffic densities of the three cases for lane 1

Comparisons of traffic densities across the cells and lanes are depicted in Figures 4.18-4.21. For the first case, traffic congestion builds up on lane 4 due to the bottleneck effect of the on-ramp while other lanes are not affected. The analysis is similar to that of Scenario 1, it is thus omitted here for brevity. When the lane-changing policy is changed to Case 2, drivers tend to change to other lanes. While the lane change maneuvers significantly affect the traffic condition of the adjacent lane, i.e., lane 3, they have a minor effect on the traffic condition of lane 1 as illustrated in Figures 4.18-4.21, i.e., more people choose lane 3 rather than lane 2 or lane 1.

The balancing effect of lane-changing makes the heterogeneity of cell-lane traffic densities smaller than those without lane-changing. However, it may still be observed that traffic densities of lane 3 and lane 4 are different for Case 2. These are because: First, some of the vehicles keep on changing lanes till they reach lane 1. Second, the

time required to execute lane change from lane 4 to lane 1 is much larger than that needed to lane 3 while the speed differences among lanes 1,2,3 are not large enough to trigger more lane change maneuvers.

To verify the balancing effect, Case 3 is re-examined for comparison wherein lane changes are restricted from lane 3 to lane 2. The result reflects that traffic conditions on lane 1 and lane 2 reduce to those of Case 1. Traffic density of cell 1 on lane 3 increases while that of cell 1 on lane 4 reduces. These two densities tend to their steady states which are of almost the same value. To see the simulation is not sensitive to the cell length choice, the length of cell 4 is changed to 0.3 miles and denote it as Case 4. Simulation results are compared in Figures 4.22-4.23, which indicate traffic densities remain almost the same as before.

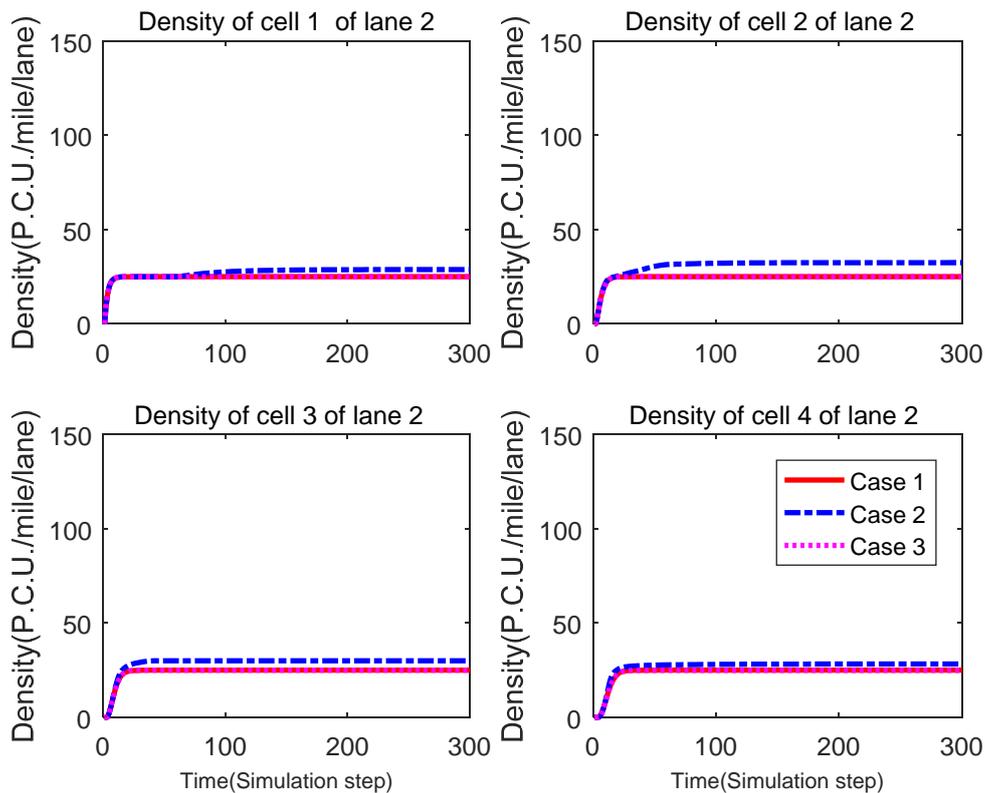


Figure 4.19. Comparison of traffic densities of the three cases for lane 2

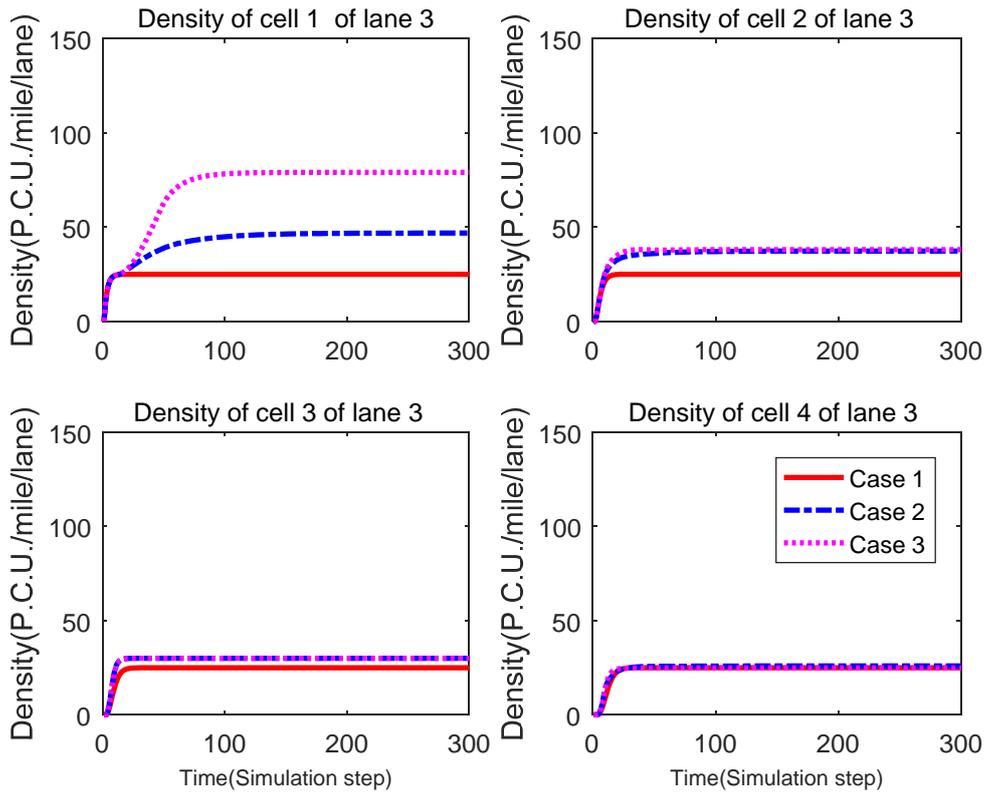


Figure 4.20. Comparison of traffic densities of the three cases for lane 3

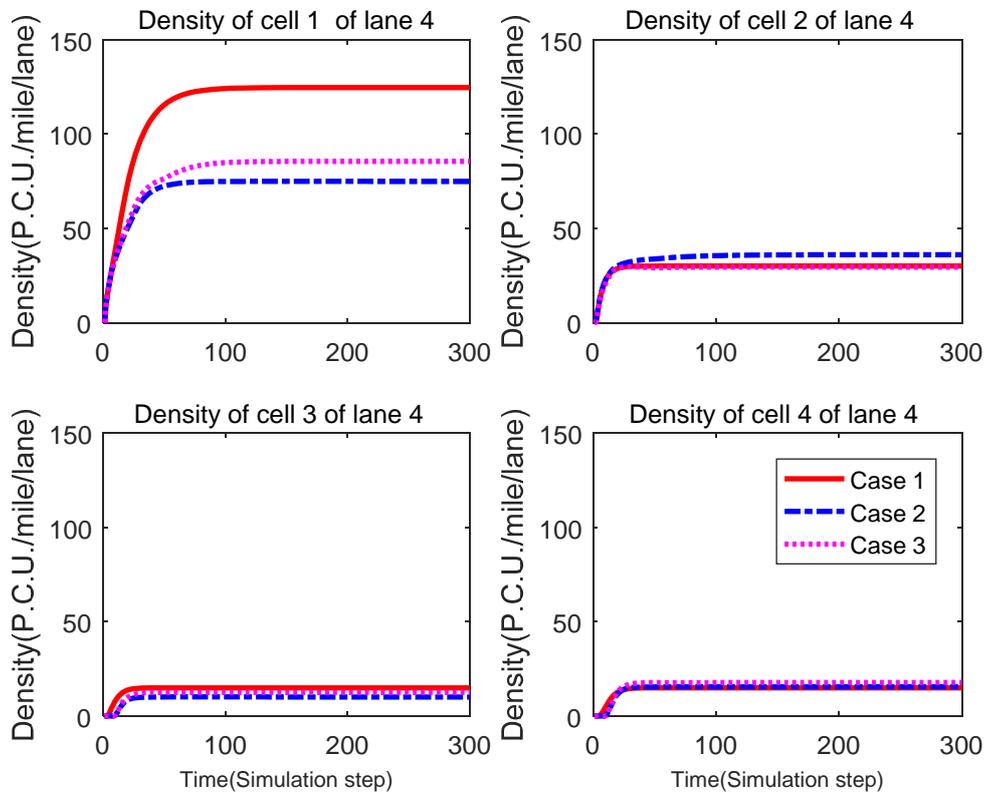


Figure 4.21. Comparison of traffic densities of the three cases for lane 4

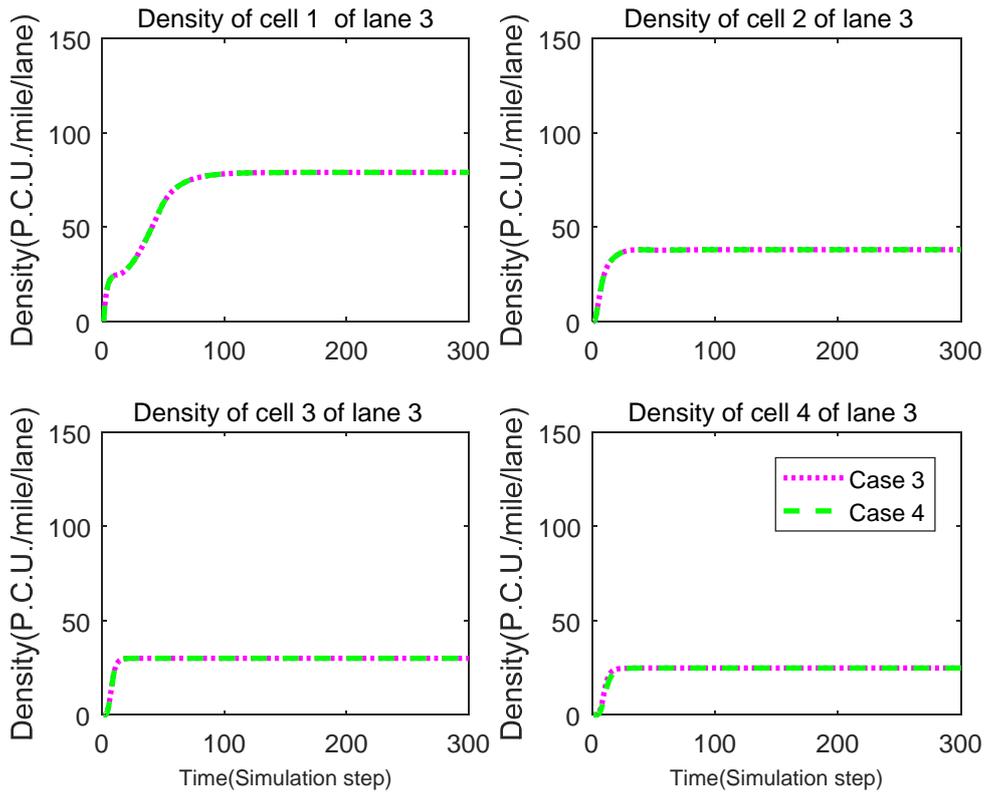


Figure 4.22. Test of different cell lengths

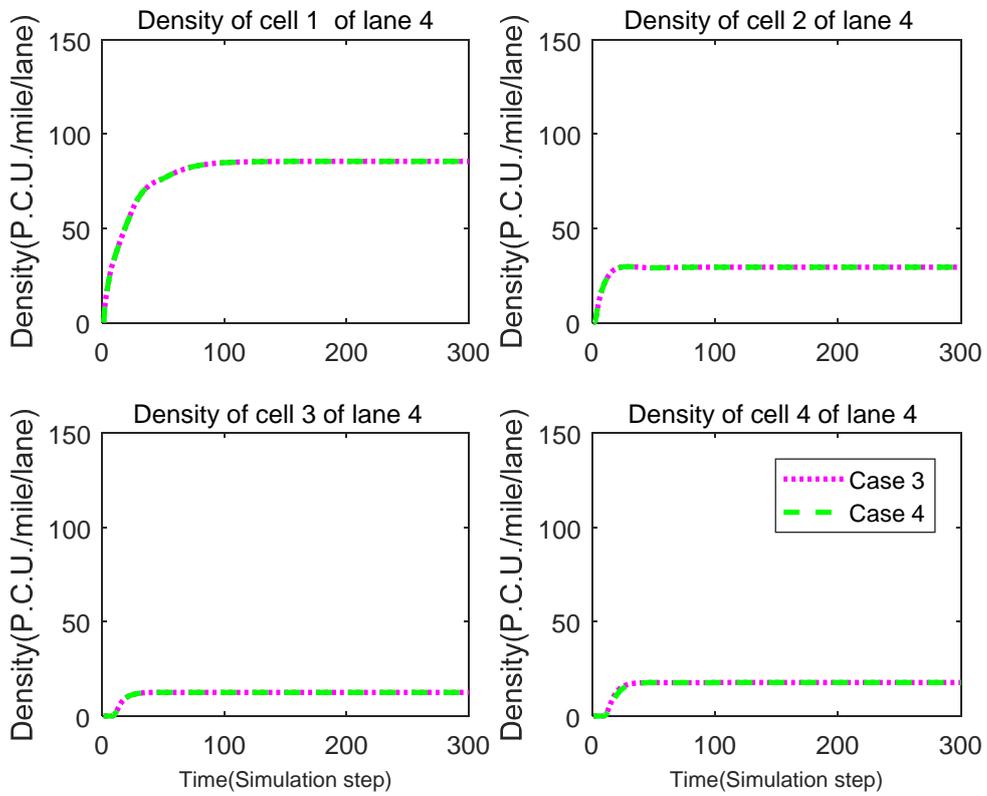


Figure 4.23. Test of different cell lengths

#### 4.8.4 The effects of stochastic elements

It is argued by Jabari and Liu (2012), Jabari and Liu (2013) that the mean process of any stochastic model of traffic flow should describe queue build-up and dissipation be consistent with well established (deterministic) traffic flow principles. However, this property is not (explicitly) required in a rich body of literature on stochastic traffic flow modeling in the fields of physics, mathematics as well as transportation research (see e.g. Hoogendoorn and Bovy, 2001, Mahnke and Kaupu, 2001, Mahnke et al., 2005, Wang and Papageorgiou, 2005, Wang et al., 2007, Wang et al., 2009, Khoshyaran and Lebacque, 2009, Sumalee et al., 2011 and the references therein). Jabari and Liu (2012) and Jabari and Liu (2013) criticized that missing this property is one of the major drawbacks of the existing stochastic traffic flow models. It will be illustrated in this test that whether or not a stochastic traffic flow model possesses such a consistency property is closely related to the way that the uncertainty is incorporated.

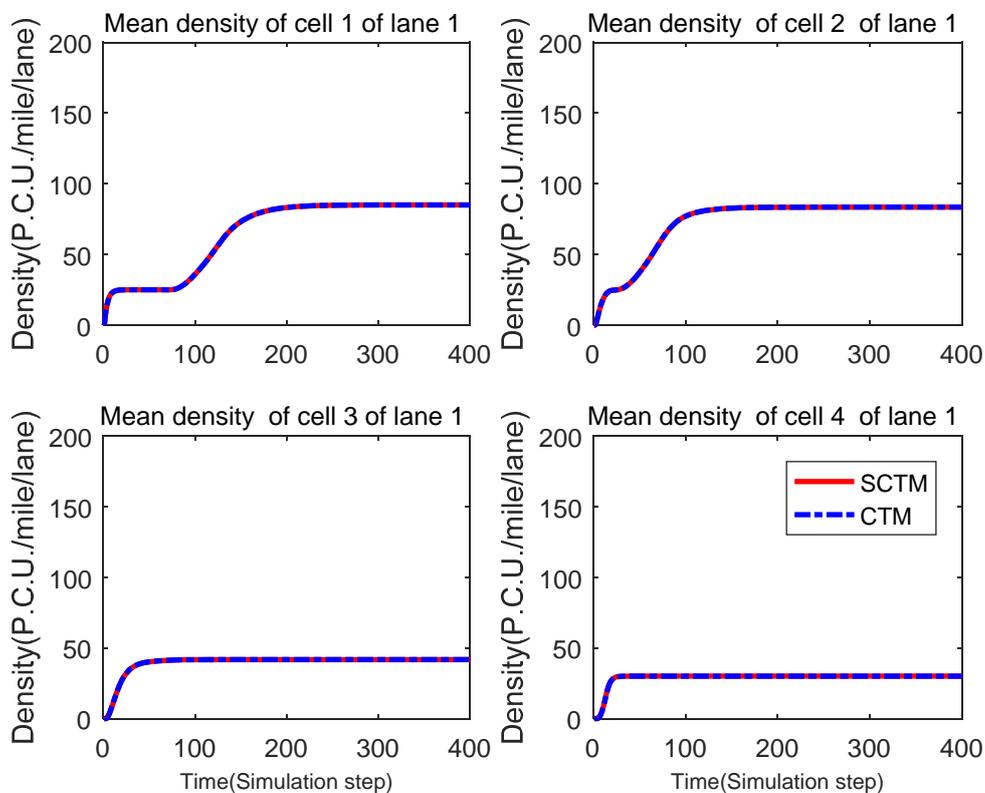


Figure 4.24. The mean process of the SCTM when the uncertainty is trivial and the deterministic dynamics of CTM for Scenario 1

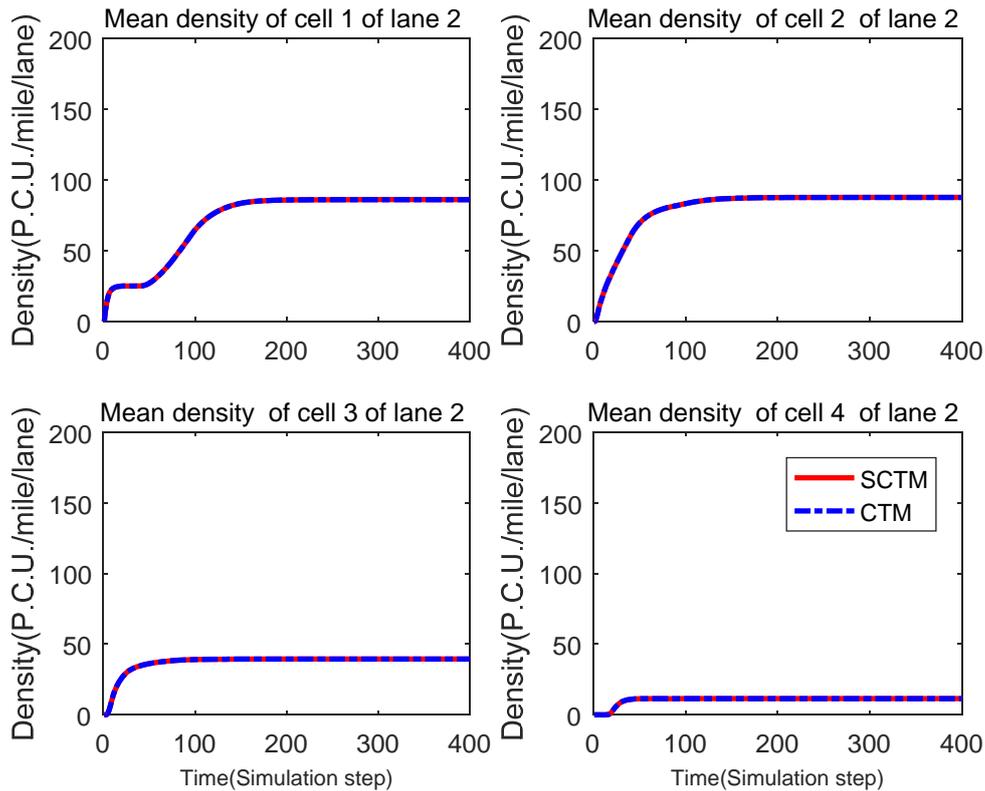


Figure 4.25. The mean process of the SCTM when the uncertainty is trivial and the deterministic dynamics of CTM for Scenario 1

In Jabari and Liu (2012), Jabari and Liu (2013), a stochastic refinement of the fluid process is directly added to the nominal (or the mean) traffic dynamics wherein the stochastic refinement is obtained by amplifying the deviation between a scaled queuing process and the deterministic fluid process. In the sense that the random term is directly added to the state process (traffic density process), it may be claimed that this approach incorporates the uncertainty as an internal or endogenous factor. To this end, the mapping from the random term to the system state is linear. Then the authors can show the perturbed traffic process weakly converge to the mean process (which in turn implies the consistency property) by certain functional central limit theorem (when the sample size and the amplification factor tend to infinity).

In traffic state estimation literature (see e.g., Wang and Papageorgiou, 2005, Wang et al., 2007, Wang et al., 2009, Khoshyaran and Lebacque, 2009, Sumalee et al., 2011,

Mihaylova et al., 2012), the uncertainty is taken as exogenous factor to the traffic dynamics, i.e., random traffic state is a result of demand and supply uncertainties (Wang et al., 2011, Wang et al., 2013).

To model random traffic states, the demand and supply (parameters of the fundamental diagram) functions are assumed to be polluted by certain noise terms. The noise terms affect traffic states through the lens of demand and supply functions to the traffic dynamics equations (but in contrast the noise term is directly added to the mean process in Jabari and Liu (2012), Jabari and Liu (2013)). In this sense, the mapping from noise terms to traffic state is generally nonlinear due to the nonlinearity of the fundamental diagram and dynamic equations, e.g., the multiplicative demand and supply uncertainties in Sumalee et al. (2011). As a consequence, the mean dynamics usually does not coincide with the original deterministic dynamics. The original deterministic dynamics under this approach can be regarded as a realization of the stochastic traffic state given a realization of the set of random demand and supply functions while it is similar to the (Monte Carlo) mean process in Jabari and Liu (2012), Jabari and Liu (2013) (when the sample size and the amplification factor tend to infinity). The mean process of the exogenous uncertainty modeling approaches would be equal to the original deterministic dynamics if the uncertainty is trivial or the realization of the set of random demand and supply functions is the nominal one as shown in Figures 4.24-4.25.

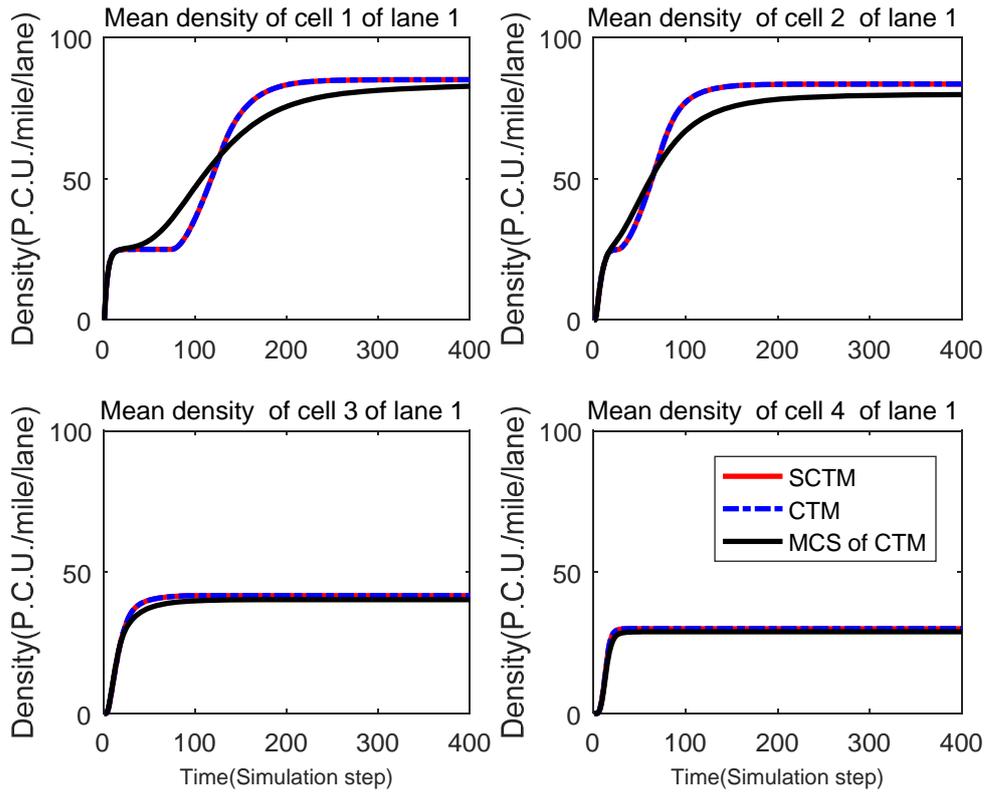


Figure 4.26. Comparison of mean processes produced by the SCTM, CTM, and MCS of CTM for Scenario 1

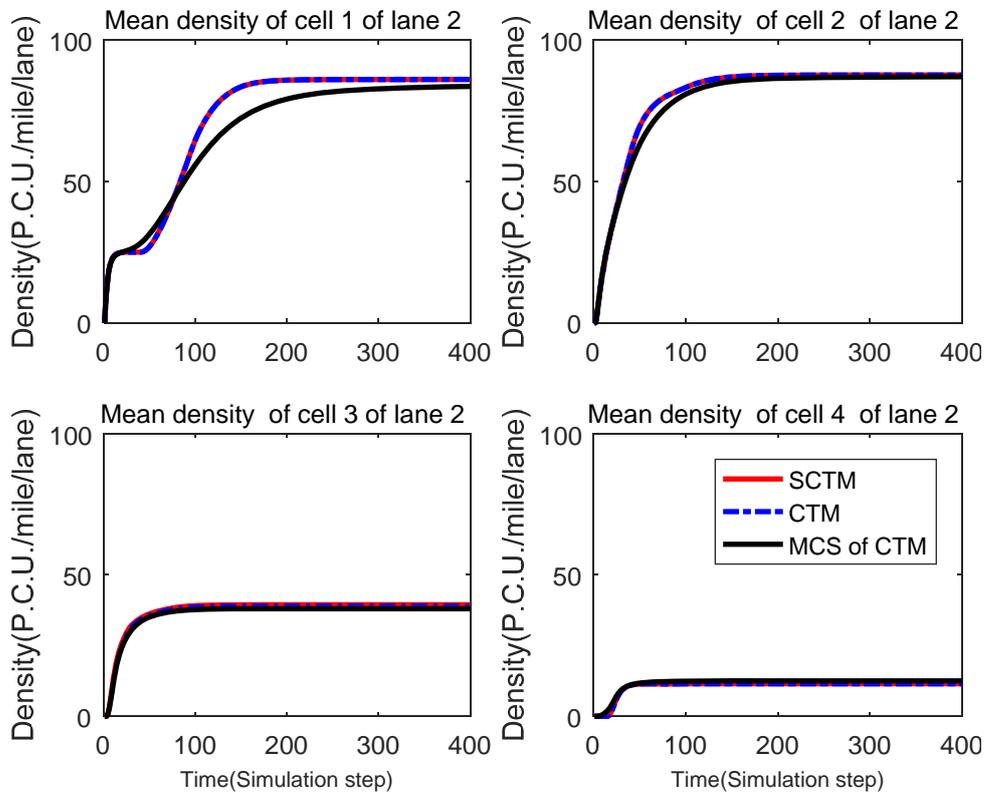


Figure 4.27. Comparison of mean processes produced by the SCTM, CTM, and MCS of CTM for Scenario 1

Indeed, under the bilinear formulation (also known as linear systems with Markovian jumps and multiplicative noise) of the SCTM proposed by Sumalee et al. (2011), the noises from demand and supply sides are modeled as intermediate multiplicative noises. The mapping from these intermediate multiplicative noises to the system state is linear. The mean process of the SCTM maybe not coincide with the original deterministic dynamics due to the following operations: the multiplicative noise from original Gaussian noises may not be Gaussian and the finite mixture of the possible scenarios. However, simulation results reveal that the SCTM can produce a mean process that is very close to the deterministic CTM (see Figures 4.26-4.27). The Monte Carlo Simulation (MCS)<sup>11</sup> produce a result that is quite different from both the SCTM and CTM because of the nonlinearity as previously explained.

Comparisons of the stochastic traffic states produced by the SCTM and the MCS of CTM with lane-changing for Scenario 1 are depicted in Figures 4.28-4.29. The MCS of CTM produces a mean process with slower transient dynamics than those produced by the SCTM and CTM. The reason may be due to the nonlinearity of the CTM and the raw sampling process adopted in the simulation. Nevertheless, the trends of the standard deviations (SDs) of traffic densities in both cases are similar. The variance produced by the MCS of CTM is also generally more significant than that of SCTM while they tend to have more and more close values when the traffic states approach their steady states. It is noteworthy that the MCS is subject to the sampling error, which usually over-approximates the variance. Various techniques for variance-reduction sampling have thus been proposed in the literature. Choosing more appropriate sampling process may reduce the variance and increase the accuracy of the transient dynamics produced by the MCS of CTM.

---

<sup>11</sup> The CTM with 5000 trials were tried out in this MCS. The stochastic elements are assumed to be 10% of their nominal values.

As it can be observed from the figures, the one standard deviation confidence interval produced by the SCTM includes a large part of the mean process of the MCS of CTM. When the traffic is in the free-flow steady state (i.e., flow moving downstream), the SD of traffic density propagates downstream. On the other hand, it can be observed with the propagation of the uncertainty in the reverse direction of the traffic flow under the congested condition. This verifies that both the SCTM and the MCS of SCTM capture dynamic queue and wavefront propagations in terms of mean and variance. Finally, the mixture distribution nature of the SCTM was indirectly proven by showing the emergence of mixture distributions in the solution of LWR model with random terms analytically by Blandin et al. (2012).

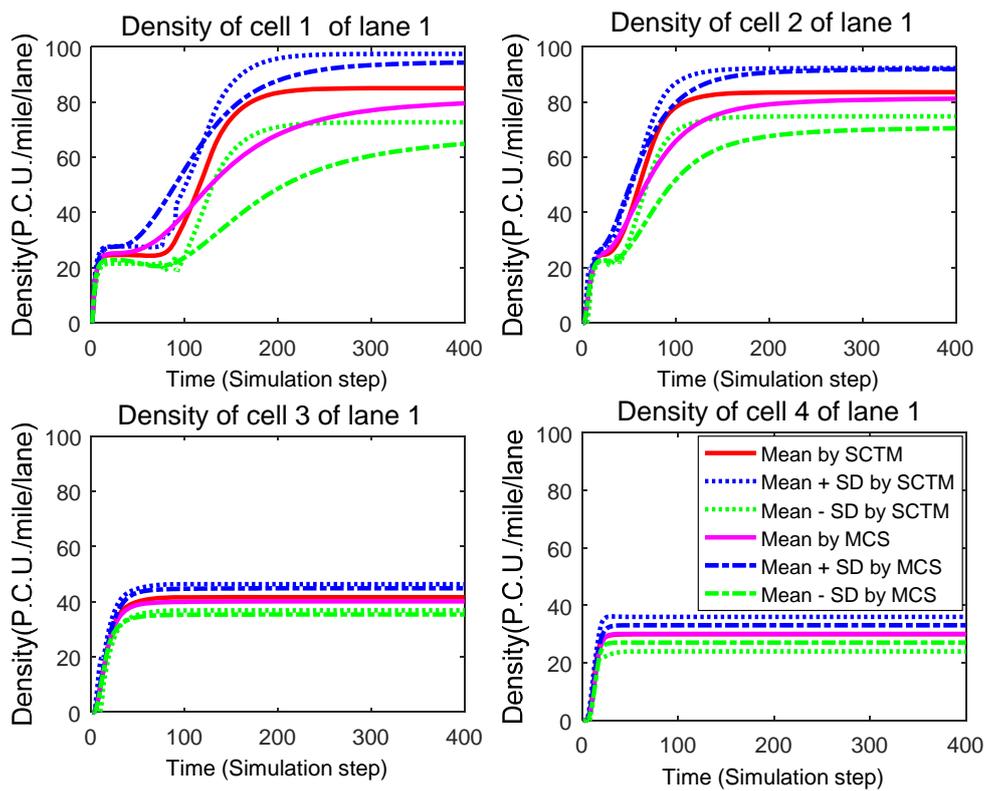


Figure 4.28. Comparison of stochastic traffic states produced by the SCTM and MCS of CTM for Scenario 1

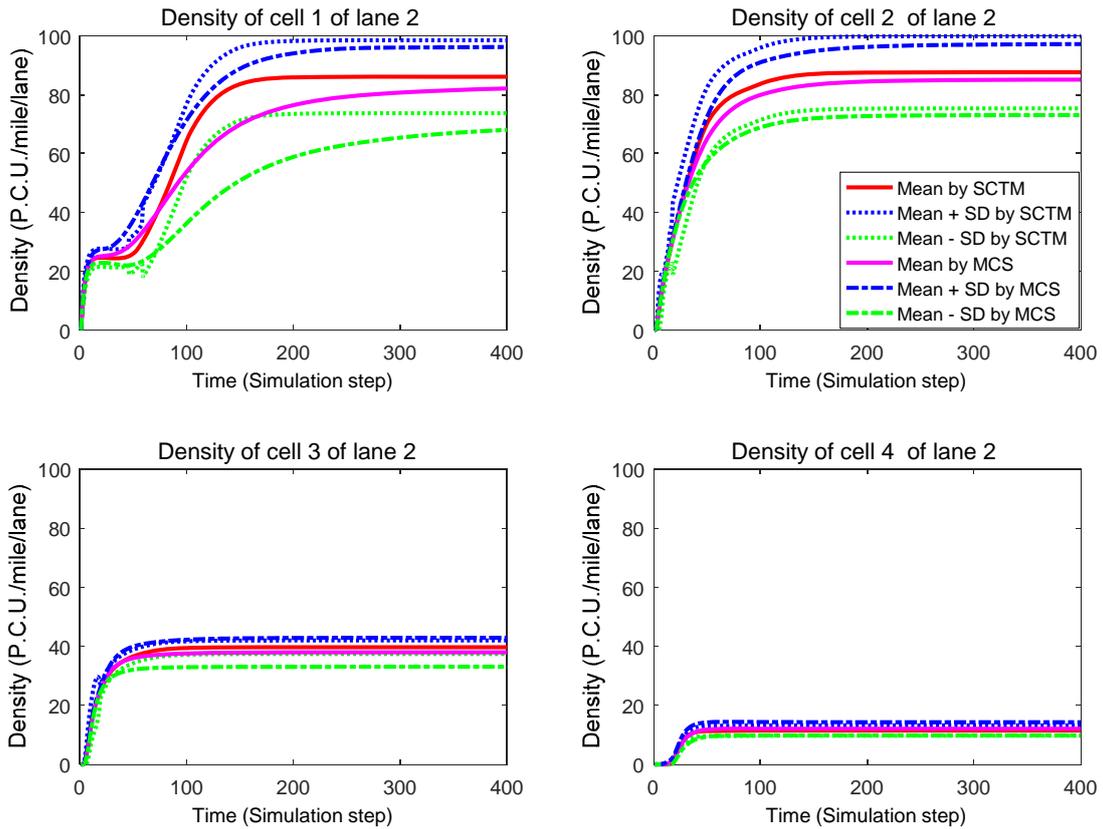


Figure 4.29: Comparison of stochastic traffic states produced by the SCTM and MCS of CTM for Scenario 1

## 4.9 Conclusions

This chapter extended the multilane hybrid (MH) theory or the multilane hybrid CTM by incorporating the lane-based fundamental diagram and certain speed-density relation to assimilate cell-lane traffic speeds. To avoid the disadvantages of triangular fundamental diagram (as it may not fit for investigating lane-changing maneuvers under free-flowing traffic conditions), this study has adopted the speed-density relationship proposed by Del Castillo and Bentez (1995a) and modified by Jin (2010a) for Lighthill-Whitham-Richard (LWR) model with lane-changing for speed data assimilation. DLC probability ratios are defined according to lane speed heterogeneity. Link (cell)-node junction models were introduced to propagate traffic dynamics with lane changes. The lane-changing ratios define virtual node splitting ratios which are propagated by the IT

and PIT principles. Based on these extensions, this chapter extended the SCTM to simulate the effects of vehicle lane-changing on traffic dynamics with demand and supply uncertainties. Flow propagations of both MLC and DLC maneuvers are calculated by demand-supply reaction laws based on the stochastic extensions of the IT and PIT principles in line with Zhong et al. (2013). In this manner, the lane-changing flows can be well defined by using the measurements of boundary variables and the average execution time of lane-changing only. Compared with the original SCTM, the new model is of the following features. First, the traffic states are given in terms of cell and lane based. Second, cell-lane-changing probabilities are augmented as additional states that define node splitting ratios. Third, an additional process is adopted to resolve cell-lane traffic speed from traffic density estimation. However, the link (cell)-node junction formulation integrates the multilane SCTM and the original SCTM into a unified framework. On the practical side, the proposed models do not require high-resolution traffic data which is a significant improvement over the existing models. Research into empirical justification, model calibration and validation of the proposed multilane SCTM is an emerging future task.

# **Chapter 5 Optimal control strategies for freeway traffic mixed with regular human-piloted vehicles and connected automated vehicles**

## **5.1 Introduction**

This chapter proposes an integrated freeway traffic flow control framework that aims to improve flow efficiency for traffic mixed with a given penetration rate of CAVs and RHVs. The proposed control framework seeks to maximizing the throughput (thus minimizing the total delay, emission cost, electricity cost, and reducing the fluctuation of control). The objective of the proposed framework is to devise an integrated action of several control strategies such as variable speed limit (VSL) and lane changing (LC) measures for dual-two lane freeways. The CAVs are assumed to follow full compliance with the variable speed limit control (VSLC) and lane changing control (LCC) through the VACS. By contrast, the RHVs that are not connected to the VACS would make decisions in response to the control disseminated by the VMS. Advice, such as mandatory lane changing recommendation (LCR) and variable speed limit recommendation (VSLR), are disseminated upstream such that the drivers can make a response to a situation, which is still imperceptible to them, to allow tactical maneuvers. For example, the advisory system would advise drivers choosing proper lanes (change lanes to left or right or keep lane) from a distance to the incident locations or their target off-ramps for RHVs such that these drivers can respond to a situation that they cannot yet perceive themselves that allows tactical maneuvers. However, only part of the drivers will gradually follow such lane control instruction.

The content of this chapter is organized as follows:

- 1) Section 5.2 describes the system architecture of the proposed framework.

2). A multiclass dynamic macroscopic model is developed in Section 5.3, which enables:

- \*determining the variation in the fundamental diagram with respect to the varying proportion of CAVs and the implemented speed limit on the freeway segment;

- \*identifying minimum headway acceptance criteria for lane changing maneuvers proposed by CAVs and RHVs with different motivations;

- \*evaluating sending flows with an advanced Priority Incremental-Transfer (PIT) principle;

- \*propagating the lane-cell specific multiclass traffic flow and density.

3). Optimal control is formulated in Section 5.4:

- \*using the multiclass traffic model developed in Section 5.3 as network loading model;

- \*devising optimal control strategies such as VSLC, LCC to CAVs, and VSLR, LCR to RHVs, and minimum gap controls.

4). Due to the complexity of the multiclass multilane traffic flow model, the optimization problem is with the non-differentiable complicated functional structure of the control vector, as well as the constraint set that would induce many local optima. Gradient-based policy search (or iteration) approaches may not be suitable choices for the proposed optimal control problem. A gradient-free approach, i.e., a cross-entropy method (CEM) based on reinforcement learning policy search algorithm, is developed to solve the optimal control problem in Section 5.5. Rather than searching the value function iteration, this algorithm parameterizes the control (or policy) and searches for the optimal parameters that lead to maximal returns (or to minimizing the objective function). The dynamic proportion of CAVs on a short freeway segment would significantly affect the freeway's throughput and traffic speed. The effect of the penetration rate of CAVs is investigated by sensitivity analysis.

5). Numerical simulations are conducted in Section 5.6 to:

- \*illustrate the efficiency of lane changing control policy using CAVs;
- \* assess the performance of integrated control strategies; and
- \*analyze the impact of penetration rate and congestion level on the control performance.

6). Finally, Section 5.7 concludes the chapter and depicts the future works.

## **5.2 System architecture**

This chapter proposes an optimal control framework for improving the efficiency of freeway traffic with mixed RHVs and CAVs. As depicted in the framework shown in Figure 5.1, some of the vehicles, i.e., CAVs, traveling on the freeway are equipped with VACS, which enable traffic data transmission from vehicle-to-infrastructure (V2I) and the delivery of control instructions from infrastructure-to-vehicle (I2V). The infrastructures are all connected to a traffic management center. For V2I communication, the onboard devices can be used as sensors to provide traffic data for the assessment of real-time traffic states. For I2V communication, the infrastructures release time-dependent control instructions, such as LCC and VSLC, to the CAVs. Meanwhile, VMS gantries which are sparsely installed at fixed locations disseminate LCR and VSLR information to RHVs along the freeway.

A traffic control policy (possibly an integration of several control measures) is devised based on the perceived traffic states of the freeway (or environment). Then, it is implemented to control the traffic dynamics on the freeway via control actuators, which causes the traffic state to transition into a new state. After implementation, the resulting traffic state is measured by real-time observations. The reward is then evaluated from the objective function, which reflects the quality of the applied traffic control policy. Naturally, only a certain number of drivers will comply with the control instructions provided by the conventional traffic advisory system, e.g., the VMS gantries. The difference between the predicted performance and the observed reward (real performance)

is used to adjust the control policy for the next time step (in a model-based reinforcement learning manner). The adjustment is designed to reduce the difference between the predicted performance and the observed reward and to remedy potential control failures in real time.

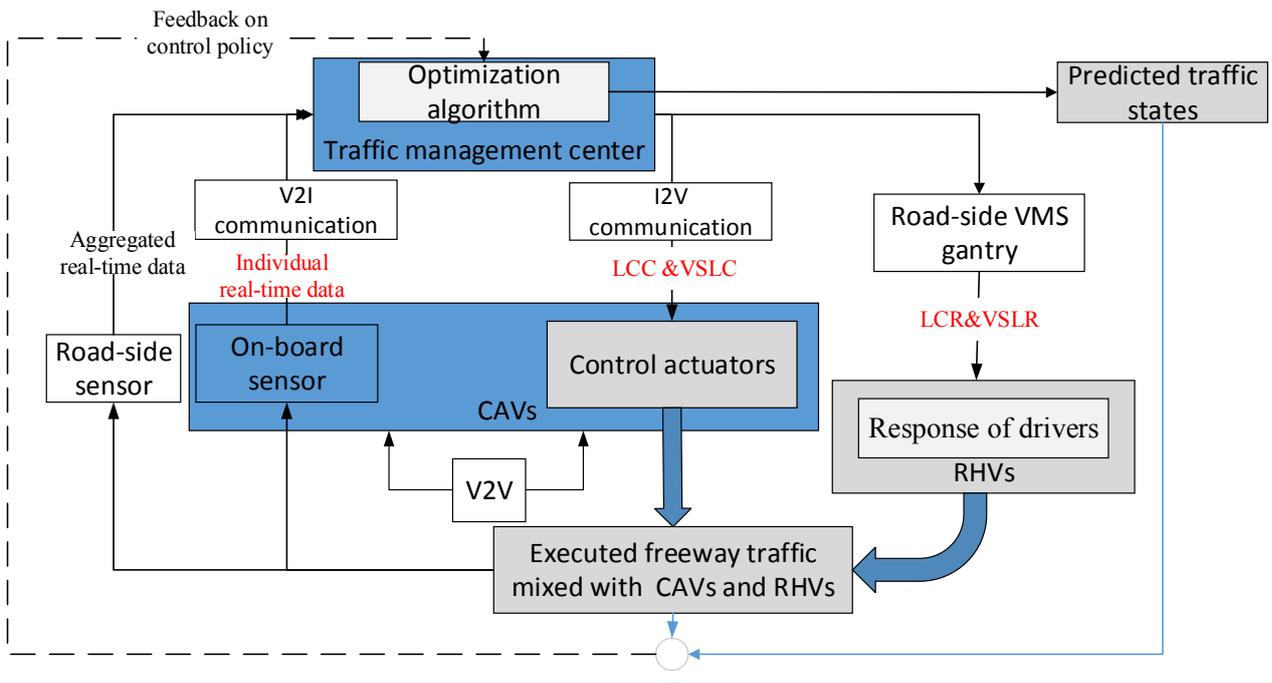


Figure 5.1. The system architecture of the proposed integrated control framework

### 5.3 Multiclass traffic flow characteristics mixed with CAVs and RHVs

As reviewed in Chapter 1 and Chapter 2, several factors, e.g., the uncertain penetration rate of CAVs, human drivers' behavior, and lane control policy, affect traffic flow with mixed RHVs and CAVs. In this section, the multilane cell transmission model (CTM) is extended from a single vehicle class to multiple vehicle classes to consider the factors mentioned above. The proposed multiclass traffic model will serve as the network loading model for formulating the optimal control problem which will be developed in Sections 5.2.

### 5.3.1 Headway distribution and fundamental diagram

Headway, defined as the time/space between the same positions of two consecutive vehicles, is an important measure of traffic flow characteristics, and thus it is essential for studying traffic flow. Because CAVs have significantly different operating characteristics compared to RHVs, it is deemed that CAVs can significantly reduce the headway between vehicles and hence increase the roadway capacity (Levin and Boyles, 2016). A recent experiment at the California Partners for Advanced Transportation Technology (PATH) showed that CAVs in platoons can maintain a time gap as small as 0.6 sec, compared to 1.5 sec for RHVs (Chen et al., 2017). In view of this, it is necessary to define a headway distribution law to model the potential capacity enhancement with the introduction of CAVs and to enable minimum headway control for mixed traffic.

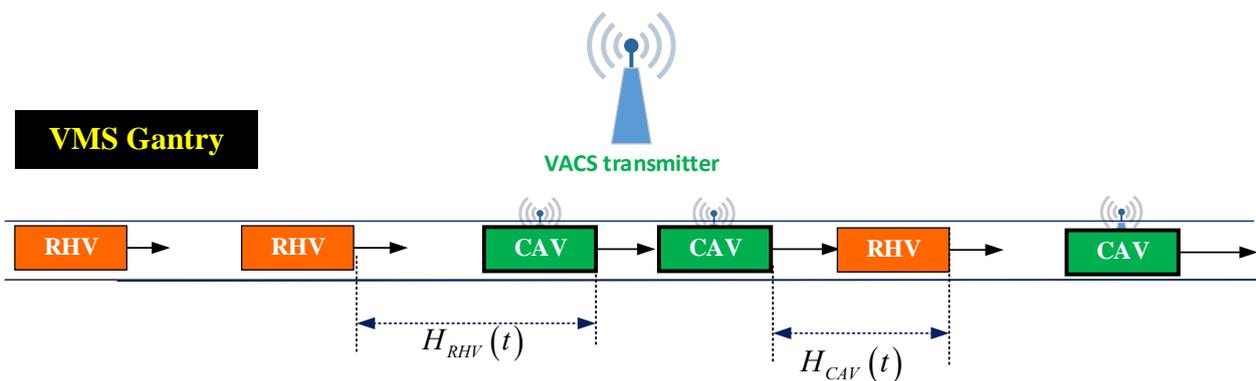


Figure 5.2. Illustration of space headway for mixed freeway traffic.

In line with Levin and Boyles (2016) and Chen et al. (2017), the effect of CAVs on roadway capacity was investigated by considering a single-lane freeway segment, as depicted in Figure 5.2. The space headway is defined as the distance from the head of the leading vehicle to the head of the following vehicle. Suppose a platoon of vehicles is traveling along a freeway section in the same lane and that the traffic flow is stable without interruption from on-ramp/off-ramp or traffic incidents. Based on the rear-end collision avoidance principle (Jepsen, 1998; Levin and Boyles, 2016), for vehicles

traveling at speed  $v(t)$  (mile/hour), the space headway criterion  $H_d(t)$  (mile) for a vehicle of class  $d$  is defined as:

$$H_d(t) \geq v(t)\Delta T_d + l + C, \quad \text{for } d = CAV, RHV \quad (5.1)$$

where  $\Delta T_{CAV}$  and  $\Delta T_{RHV}$  (hour) denote the response times of CAVs and RHVs, respectively,  $l$  (mile) denotes the vehicle length of the leading vehicle, and  $C$  is the safety gap (Jepsen, 1998) or minimum safe constant gap (Hidas, 2005) when all the vehicles are at a standstill. Compared with RHVs, CAVs can tolerate a much smaller space headway in front of it because of the smaller response time. In this study, it is assumed that the leading vehicle does not affect the minimum headway. Suppose at time  $t$  the traffic on this unit length freeway segment (otherwise, multiply both sides of (5.2) by the segment length) is of a proportion  $P(t)$  of CAVs and a proportion  $1 - P(t)$  of RHVs, the relationship between space headway and traffic density is described as

$$\rho(t) \left( P(t)H_{CAV}(t) + (1 - P(t))H_{RHV}(t) \right) = 1 \quad (5.2)$$

where  $H_{CAV}(t)$  and  $H_{RHV}(t)$  denote the space headway of CAVs and RHVs, respectively, as illustrated in Figure 5.2.

Based on Equation (5.1), the following equations are obtained:

$$\begin{aligned} \rho(t)P(t)H_{CAV}(t) &\geq \rho(t)P(t)(v(t)\Delta T_{CAV} + l + C) \\ \rho(t)(1 - P(t))H_{RHV}(t) &\geq \rho(t)(1 - P(t))(v(t)\Delta T_{RHV} + l + C) \end{aligned}$$

Summing the left and right sides, respectively,

$$\begin{aligned} \rho(t) \left( P(t)H_{CAV}(t) + (1 - P(t))H_{RHV}(t) \right) \\ \geq \rho(t)P(t)(v(t)\Delta T_{CAV} + l + C) + \rho(t)(1 - P(t))(v(t)\Delta T_{RHV} + l + C) \end{aligned}$$

This in conjunction with Equation (5.2) provides

$$1 \geq \rho(t)P(t)(v(t)\Delta T_{CAV} + l + C) + \rho(t)(1 - P(t))(v(t)\Delta T_{RHV} + l + C)$$

Therefore, the maximum speed  $\tilde{v}_s(t)$  that can be evaluated by specific traffic density and proportion of CAVs is evaluated as follows:

$$\tilde{v}_s(t) = \frac{(1-l \cdot \rho(t) - C \cdot \rho(t))}{\rho(t)} \cdot \frac{1}{(P(t) \Delta T_{CAV} + (1-P(t)) \Delta T_{RHV})} \quad (5.3)$$

without considering the speed limit or vehicles' mechanical capability. As reported in the literature, the average speed of RHVs is roughly equal to the average speed of CAVs in a traffic stream (Bekiaris-Liberis et al., 2016; Fountoulakis et al., 2017). To this end, it is assumed that this maximum speed  $\tilde{v}_s(t)$  will be **spontaneously** followed by both CAVs and RHVs as it guarantees the minimum space gap for avoiding collisions in accordance with different levels of congestion specified by traffic density. Therefore, in this chapter, the variable  $\tilde{v}_s(t)$  is named the **spontaneous speed limit (SSL)**, which might be adopted as the maximal speed by drivers and CAV controllers for ensuring safety. However, when the traffic density tends to zero (i.e., no vehicle is traveling on the freeway), the SSL approaches an infinitely large value. To remedy this, one can set an upper bound to this SSL, called a vehicle's **maximum mechanical speed**. For example, 170 miles/hour is a common upper bound of the speedometer of private vehicles. Generally speaking, the speed is actually restricted by the posted **permanent compulsory (upper bound) speed limit** for freeway traffic management purpose. In the United States under normal conditions, the posted permanent compulsory speed limit ranges from 90 miles/hour in rural areas to 40 miles/hour in urban areas (Highway Capacity Manual, 2010). Additionally, a temporary VSL can be issued as a control strategy for traffic incident management or congestion resolution if needed. As it can be expected, **maximum mechanical speed > permanent compulsory speed limit > temporary VSL**; therefore, the implemented speed limit (ISL) is given as either the permanent compulsory speed limit or the temporary VSL.

In line with the VSL control literature, e.g., Hegyi et al. (2005a, b), the turning point of ISL and SSL is located at the critical density  $\rho_c(t)$  (noting that the critical density is also affected by the implemented speed limited control itself), as shown in Figure 5.3. From the above analysis, the traffic flow speed  $v(t)$  is finally defined as a function of ISL (the

lower of the posted compulsory speed limit and the VSL), traffic density, and the proportion of CAVs.

$$v(t) = \begin{cases} \tilde{v}_l(t) & \text{if } \rho(t) \leq \rho_c(t) \\ \tilde{v}_s(t) = \frac{(1-l \cdot \rho(t) - C \cdot \rho(t))}{\rho(t)} \cdot \frac{1}{(P(t) \Delta T_{CAV} + (1-P(t)) \Delta T_{RHV})} & \text{if } \rho(t) > \rho_c(t) \end{cases} \quad (5.4)$$

where  $\tilde{v}_l(t)$  denotes the ISL (which is solved from the optimal control problem as formulated in Section 5.4 or the permanent compulsory speed limit posted by freeway management center) and  $\rho_c(t)$  is a function of the penetration rate of CAVs and ISL as follows:

$$\rho_c(t) = \frac{1}{\tilde{v}_l(t) (P(t) \Delta T_{CAV} + (1-P(t)) \Delta T_{RHV}) + l + C} \quad (5.5)$$

In traffic flow theory, the flow (rate) is the product of speed,  $v(t)$ , and density,  $\rho(t)$ , whereas the capacity is the maximum traffic flow (rate) observed at critical density  $\rho_c(t)$ . By  $Q_m(t) = \rho_c(t) \tilde{v}_l(t)$ , the roadway capacity can be defined below as a function of the penetration rate of CAVs and ISL:

$$Q_m(t) = \frac{\tilde{v}_l(t)}{\tilde{v}_l(t) (P(t) \Delta T_{CAV} + (1-P(t)) \Delta T_{RHV}) + l + C} \quad (5.6)$$

The back-wave speed is evaluated by:

$$w_c(t) = \frac{l + C}{(P(t) \Delta T_{CAV} + (1-P(t)) \Delta T_{RHV})} \quad (5.7)$$

On the other hand, the jam density is determined by the average vehicle length and the minimum safe constant gap:

$$\rho_J(t) = \frac{1}{l + C} \quad (5.8)$$

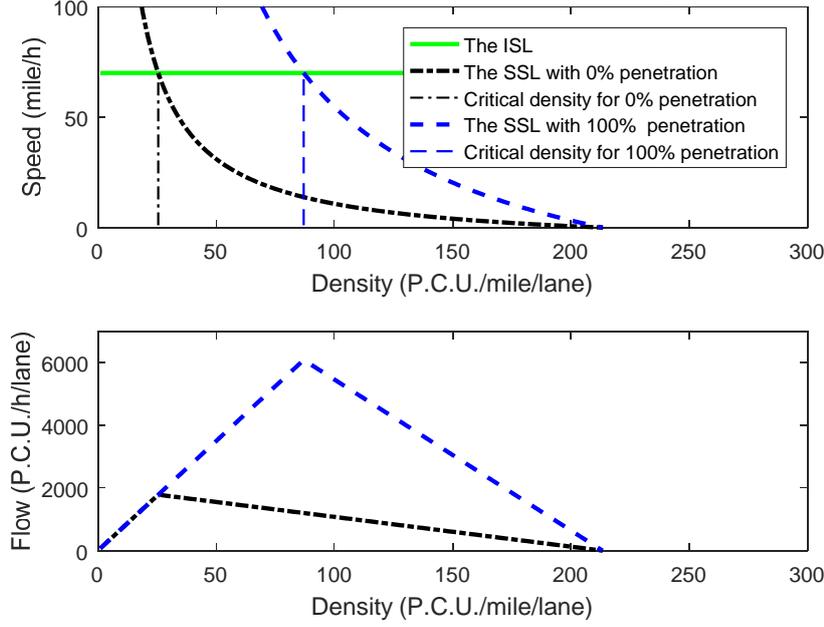


Figure 5.3. Lane-specific fundamental diagram as a function of the penetration rate of CAVs.

To ensure safety, the minimum space headway criteria  $\tilde{H}_d(t)$  for vehicle class  $d = CAV$  and  $RHV$  are evaluated respectively as

$$\begin{aligned}\tilde{H}_{CAV}(t) &= v(t)\Delta T_{CAV} + l + C \\ \tilde{H}_{RHV}(t) &= v(t)\Delta T_{RHV} + l + C\end{aligned}\quad (5.9)$$

The total space  $\tilde{O}(t)$  reserved for the minimum-space headway of all vehicles traveling on the freeway segment is calculated according to the proportions of the two vehicle classes.

$$\tilde{O}(t) = \rho(t)P(t)\tilde{H}_{CAV}(t) + \rho(t)(1-P(t))\tilde{H}_{RHV}(t)\quad (5.10)$$

When  $\rho(t) \geq \rho_c(t)$ , the speed is determined by SSL  $\tilde{v}_s(t)$ , according to Equation (5.10), the total space  $\tilde{O}(t)$  reserved for the minimum-space headways fills the total spaces. That is, the relationship between the traffic speed and density attains a critical value, as illustrated in Figure 5.3. The flow speed cannot be increased so as to maintain safety. However, when  $\rho(t) < \rho_c(t)$ , ISL  $\tilde{v}_i(t)$  can be lower than SSL  $\tilde{v}_s(t)$ , thus allowing extra space.

$$\tilde{O}(t) \begin{cases} < 1 & \text{if } \rho(t) < \rho_c(t) \\ = 1 & \text{if } \rho(t) \geq \rho_c(t) \end{cases}$$

That is to say, if there is no ISL  $\tilde{v}_i(t)$ , vehicles can travel faster under light traffic conditions. The free space  $1-\tilde{O}(t)$  not occupied by vehicles for safe traveling can be randomly distributed among them. For simplification, the safe space headway criteria are assumed to be magnified by  $\frac{1}{\tilde{O}(t)}$ , i.e.,

$$\begin{aligned} H_{CAV}(t) &= \tilde{H}_{CAV}(t) / \tilde{O}(t) \\ H_{RHV}(t) &= \tilde{H}_{RHV}(t) / \tilde{O}(t) \end{aligned} \quad \text{if } \rho(t) < \rho_c(t)$$

To sum up, the (average) space headway distributions are thus calculated:

$$\begin{aligned} H_{CAV}(t) &= \\ & \begin{cases} \frac{v(t)\Delta T_{CAV} + l + C}{\rho(t) \cdot (l + C) + \rho(t) \cdot v(t) \cdot (P(t)\Delta T_{CAV} + (1 - P(t))\Delta T_{RHV})} & \text{if } \rho(t) < \rho_c(t) \\ v(t) \cdot \Delta T_{CAV} + l + C & \text{if } \rho(t) \geq \rho_c(t) \end{cases} \\ H_{RHV}(t) &= \\ & \begin{cases} \frac{v(t)\Delta T_{RHV} + l + C}{\rho(t) \cdot (l + C) + \rho(t) \cdot v(t) \cdot (P(t)\Delta T_{CAV} + (1 - P(t))\Delta T_{RHV})} & \text{if } \rho(t) < \rho_c(t) \\ v(t) \cdot \Delta T_{RHV} + l + C & \text{if } \rho(t) \geq \rho_c(t) \end{cases} \end{aligned} \quad (5.11)$$

As stated in Green (2000), the brake response time of a human driver is composed of three components: mental processing, muscle movement, and brake engagement time. On average, the mental processing takes about 1.3 sec for unexpected occasions, the average muscle movement takes 0.2 sec (Wierwille and Casali, 1983; Lerner et al., 1995; Green 2000), and the brake engagement time takes 0.35 sec under emergency conditions. Therefore, the braking response time of RHVs to an unexpected occasion is 1.85 sec. However, CAVs do not have mental reactions nor muscle movement; therefore, the response time of CAVs is considered to be 0.35 sec for the brake engagement process. This value is consistent with the 0.6 sec time headway between successive CAVs (Chen et al., 2017).

Figure 5.4 demonstrates the impact of the penetration rate of CAVs<sup>23</sup> and the compulsory speed limit on road capacity. In this example, all vehicles are considered passenger car equivalent (P.C.E.) or passenger car unit (P.C.U.) vehicles 20 feet in length. The safe constant gap is 6.5 feet (when the related vehicles are at a standstill), and the compulsory speed limit is 70 miles/hour. As stated above, the response times of RHVs and CAVs are set to be 1.85 sec and 0.35 sec, respectively. As shown in Figure 5.4, the capacity monotonically increased with the increasing penetration rate of CAVs and the ISL.

The solid black line quantifies the variation of capacity with respect to the penetration rate of CAVs varying from 0% to 100% by fixing the compulsory speed limit at 70 miles/hour. As demonstrated in this figure, the capacity varies significantly from 1719 P.C.U./hour/lane (with 100% RHVs) to 6055 P.C.U./hour/lane (with 100% CAVs). On the other hand, by typical macroscopic traffic flow theory, the increase in free-flow speed (or the compulsory speed limit in our case) would introduce an increase in capacity. The red dotted line presents an example for this using a fixed 66.7% penetration rate of CAVs.

---

<sup>23</sup> In this case, the penetration rate is equivalent to proportion as cell-lane capacity is concerned.

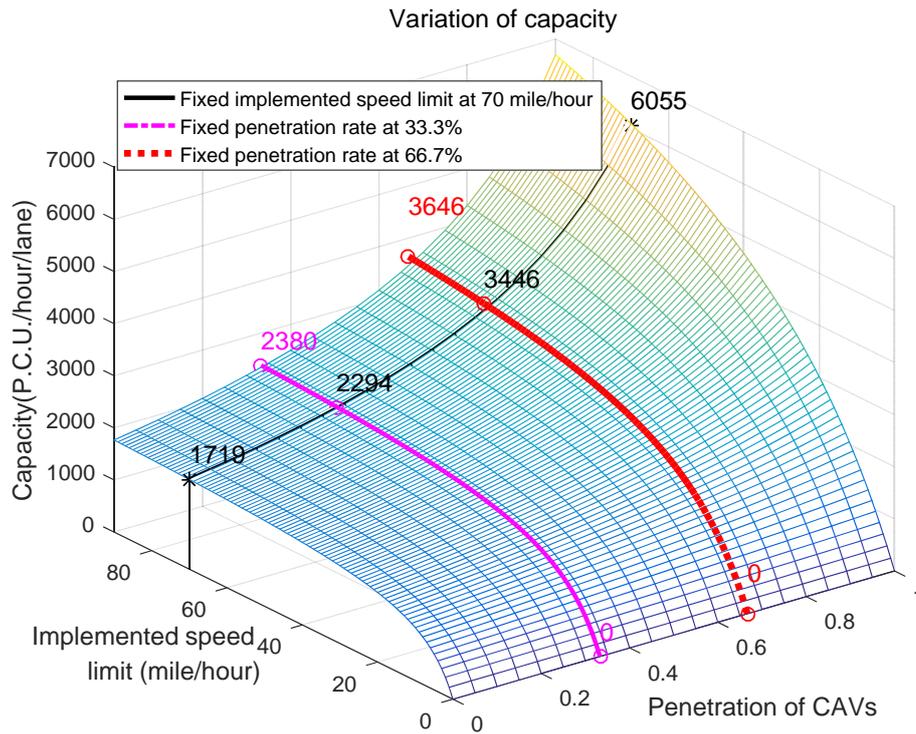


Figure 5.4. Capacity as a function of the penetration rate of CAVs and the implemented speed limit.

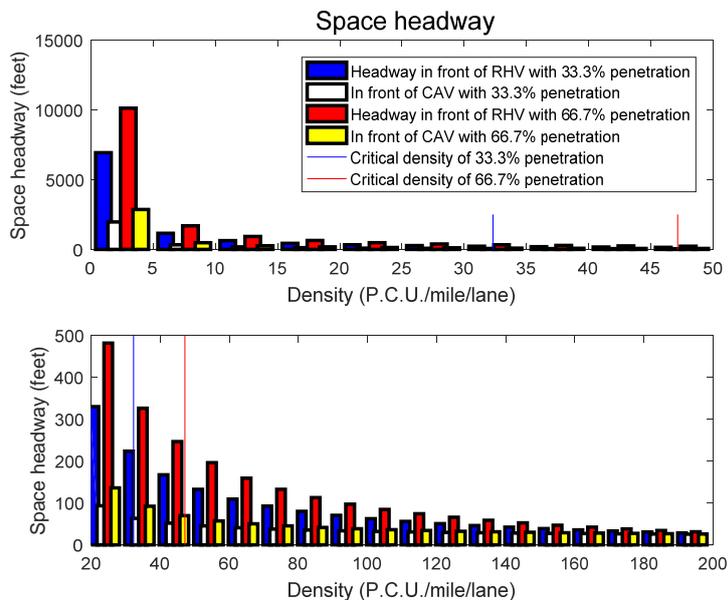


Figure 5.5. The distribution of space headway as a function of penetration rate.

	Scenario: 33.3% Penetration rate			Scenario: 66.7% Penetration rate		
Density (P.C.U./mile/ lane)	Speed (mile/hour)	$H_{CAV}(t)$ (feet)	$H_{RHV}(t)$ (feet)	Speed (mile/hour)	$H_{CAV}(t)$ (feet)	$H_{RHV}(t)$ (feet)
10	70	196	694	70	286	1012
32	70	61	217	70	89	316
47	49	47	145	70	61	215
100	14	32	63	20	36	86
200	0.8	25	27	1	25	28

Table 5.1. Space headway as a function of density and penetration rate

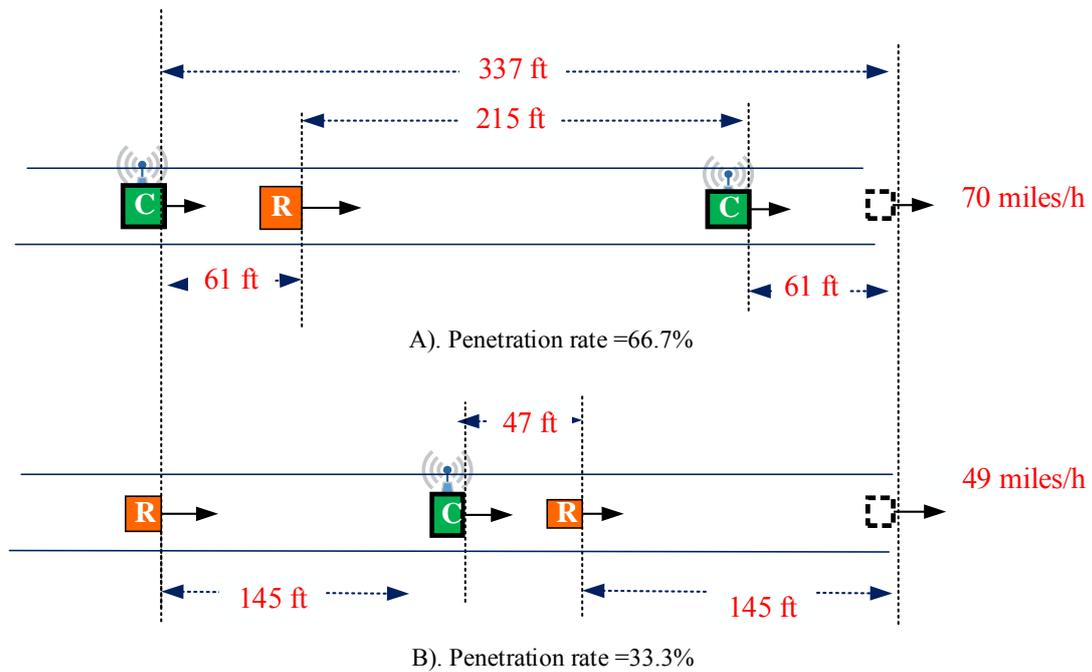


Figure 5.6. The variation of space headway with respect to penetration rate.

Figure 5.5 and Table 5.1 present the space headways of RHVs and CAVs with respect to traffic density and penetration of CAVs. The space headway distributions with respect to different penetration rates depicted in Figure 5.6 were obtained by setting the traffic

density to be 47 P.C.U./mile/lane, i.e., 3 P.C.U./337feets/lane, and assuming 66.7% and 33.3% penetration rates of CAVs, respectively. As it can be inferred from Figure 5.5, Figure 5.6 and Table 5.1 that:

- 1) The space headway in front of an RHV is larger than that in front of a CAV under the same traffic conditions, and this difference is more significant under free-flowing conditions (the results related to free-flowing conditions use bold font in Table 5.1).
- 2) The increment of traffic density leads to smaller space headways for both RHVs and CAVs.
- 3) The penetration rate of CAVs affects the space headway distribution. In particular, the space headway in front of RHV increases 50% when the penetration rate of CAVs raises from 33.3% to 66.7%.
- 4). With the same traffic density, the traffic with a higher penetration rate of CAVs might maintain a higher speed. As demonstrated in Figure 5.6, with the same traffic density 47 P.C.U./mile/lane, the traffic is 21 miles/hour faster when the penetration rate of CAVs raises from 33.3% to 66.7%.

As it can be inferred from the above observations, the introduction of CAVs can change the gap distribution of the target lanes, while this change makes the lane changing maneuvers easier to be accepted by the target lanes.

### **5.3.2 The minimum gap acceptance criteria**

Apart from the traffic flow characteristics associated with traffic flow on a single lane, such as the fundamental diagram and space headway distribution introduced in Section 5.3.1, the penetration of CAVs also has an impact on lane-changing maneuvers when traveling on multilane freeways. In this subsection, the lane-changing acceptance/rejection criteria for multiclass traffic is proposed by extending the counterparts from Chapter 3.

Suppose that the subject vehicle traveling on lane  $m$  intends to switch to lane  $\beta$  via the space gap between the leading vehicle and the following vehicle in the target lane at time  $t$ , as demonstrated in Figure 5.7. It is assumed that the subject vehicle is traveling at speed  $v_m(t)$  (mile/hour), whereas the leading vehicle and following vehicle in the target lane are both traveling at the same speed  $v_\beta(t)$  (mile/hour). Without loss of generality, the assumption  $v_m(t) < v_\beta(t)$  is proposed in this section, i.e., the lane-changing maneuver is executed to gain better driving conditions. Based on the speed difference assumption  $v_m(t) < v_\beta(t)$ , it is necessary for the subject vehicle to accelerate to avoid a collision with the following vehicle in the target lane. During the acceleration process, the space gap between the leading vehicle and the following vehicle must always fulfill the collision avoidance criteria as depicted previously.

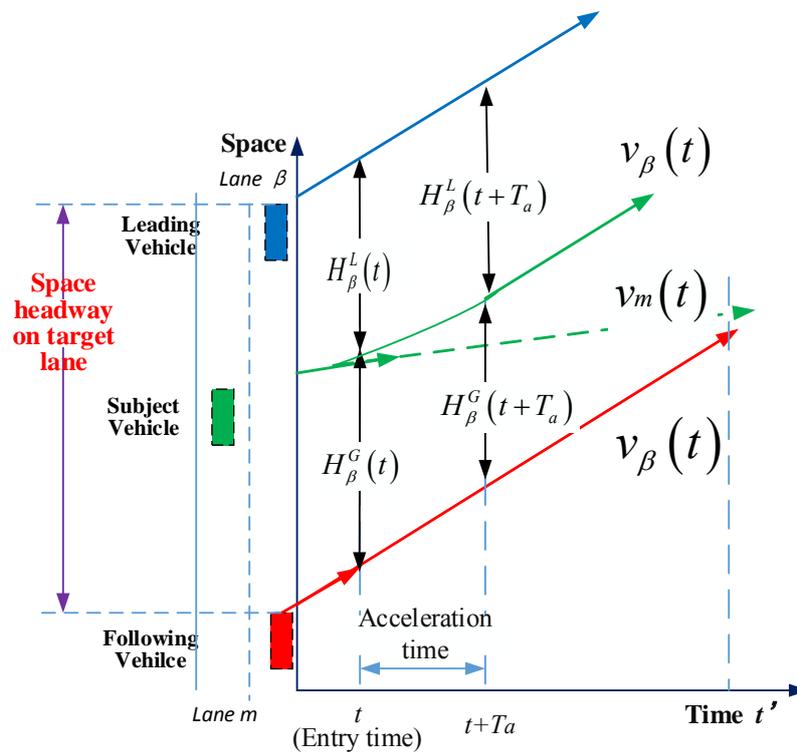


Figure 5.7. The headway acceptance analysis.

Denoting  $t' = t + T_a$  (hour) as the moment when the lane-changing maneuver is executed, as shown in Figure 5.7, the subject vehicle is supposed to travel at the same speed as the vehicles in the target lane. By assumption, this speed is  $v_m(t + T_a) = v_\beta(t + T_a) = v_\beta(t)$ , with  $T_a = \frac{v_\beta(t) - v_m(t)}{a}$ , where  $a$  (mile/hour<sup>2</sup>) denotes the acceleration of the subject vehicle. Considering the speed difference between the related vehicles during the lane-changing maneuver, the lead space headway<sup>24</sup>  $H_\beta^L(t')$  (mile) (as shown in Figure 5.7) will gradually increase whereas the lag space headway  $H_\beta^G(t')$  (mile) will decrease due to the acceleration and lane-changing maneuver of the subject vehicle, that is,

$$\begin{cases} H_\beta^L(t') = H_\beta^L(t) + (v_\beta(t) - v_m(t)) \cdot (t' - t) - \frac{1}{2} a \cdot (t' - t)^2 \\ H_\beta^G(t') = H_\beta^G(t) - (v_\beta(t) - v_m(t)) \cdot (t' - t) + \frac{1}{2} a \cdot (t' - t)^2 \end{cases} \quad \text{for } t \leq \forall t' \leq t + T_a$$

and

$$\begin{cases} H_\beta^L(t') = H_\beta^L(t + T_a) \\ H_\beta^G(t') = H_\beta^G(t + T_a) \end{cases} \quad \text{for } t' > t + T_a \quad (5.12)$$

According to the collision avoidance principle described in Section 5.3.1, the following conditions need to be fulfilled:

$$\begin{cases} H_\beta^L(t') \geq (v_m(t) + a \cdot (t' - t)) \Delta T_{d,s} + l + C \\ H_\beta^G(t') \geq v_\beta(t) \Delta T_{d,f} + l + C \end{cases} \quad \text{for } t \leq \forall t' \leq t + T_a \quad (5.13)$$

where  $\Delta T_{d,s}$  and  $\Delta T_{d,f}$  denote the response times of the subject and the following vehicles, respectively.

The space headway followed by vehicle  $f$  on the target lane  $\beta$ , i.e.,

---

<sup>24</sup> The lead space headway denotes the distance from the head of the lead vehicle to the head of the subject vehicle. The lag headway denotes the distance from the head of the subject vehicle to the head of the following vehicle.

$H_{s,m}^{f,\beta}(t) = H_{\beta}^L(t) + H_{\beta}^G(t)$ , is large enough for accepting a lane-changing demand required by subject vehicle  $s$  from lane  $m$  is determined as follows:

$$H_{s,m}^{f,\beta}(t) \geq \tilde{H}_{MLC,s,m}^{f,\beta}(t) = \begin{cases} \left[ \underbrace{v_{\beta}(t)\Delta T_{d,f}}_{item\ 1} + \underbrace{v_m(t)\Delta T_{d,s}}_{item\ 2} + \underbrace{\frac{(v_{\beta}(t)-v_m(t))^2}{2a}}_{item\ 3} \right] + \underbrace{(2l+2C)}_{item\ 4} & Case\ 1 \\ \left[ \underbrace{v_{\beta}(t)\Delta T_{d,f}}_{item\ 1} + \underbrace{v_m(t)\Delta T_{d,s}}_{item\ 2} + \underbrace{\frac{(v_{\beta}(t)-v_m(t))^2}{2a}}_{item\ 3} \right] \frac{x-x_c}{x_r-x_c} + \underbrace{(2l+2C)}_{item\ 4} & Case\ 2 \\ \underbrace{(2l+2C)}_{item\ 4} & Case\ 3 \end{cases} \quad (5.14)$$

The minimum space headway criterion  $\tilde{H}_{lc,s,m}^{f,\beta}(t)$  is the lower bound of  $H_{s,m}^{f,\beta}(t)$ , with lane changing intention  $lc = MLC$  or  $DLC$ .

For MLC intention, according to the level of urgency,  $\tilde{H}_{MLC,s,m}^{f,\beta}(t)$  is divided into three cases. Case 1 refers to the non-urgent scenario associated with a lane-changing demand proposed at a remote state with the remaining distance<sup>25</sup>  $x > x_r$ . In this case, drivers usually prefer a relatively large gap at the beginning of a lane-changing maneuver due to a risk-adverse attitude.  $\tilde{H}_{MLC,s,m}^{f,\beta}(t)$  decreases linearly with respect to the remaining

---

<sup>25</sup> The remaining distance  $x$  is defined as the distance from the current position of the subject vehicle to its target turning point.  $x_c$  and  $x_r$  denote the remaining distances by which the test section is partitioned as remote, median, and close sections, respectively, with each section corresponding to a specific level of mandatory lane-changing (MLC) urgency and minimum acceptance criterion. Taking a vehicle intending to execute an MLC as an example, the target turning point is considered to be remote as long as the remaining distance  $x > x_r$  and close if  $x < x_c$ . In Pan et al. (2016),  $x_c$  and  $x_r$  were calibrated to be 0.1 mile and 1 mile, respectively, in their empirical study.

distance when  $x$  is within the range  $x_c \leq x \leq x_r$  according to Case 2 of Equation (5.14). Both Case 1 and Case 2 were devised to model the minimum gap acceptance criterion of RHVs in line with Section 3.2. Under Case 1 and Case 2, because RHVs cannot identify whether the following vehicle on the target lane is an RHV or a CAV,  $\Delta T_{d,f} = \Delta T_{RHV}$  is adopted from the point of view of safety, if the subject vehicle is RHV, saying  $\Delta T_{d,s} = \Delta T_{RHV}$ . However, if the subject vehicle is CAV, the vehicle class of following vehicle on target lane can be identified, then both  $\Delta T_{d,f} = \Delta T_{RHV}$  and  $\Delta T_{d,f} = \Delta T_{CAV}$  are possible. Case 3 involves the urgent MLC demand, i.e.,  $x < x_c$ , proposed by RHVs or CAVs. For DLC intention which only can be proposed by RHVs, the  $H_{DLC,s,m}^{f,\beta}(t)$  is simplified as below:

$$H_{DLC,s,m}^{f,\beta}(t) \geq \tilde{H}_{DLC,s,m}^{f,\beta}(t) = \left[ \underbrace{v_\beta(t) \Delta T_{RHV,f}}_{item\ 1} + \underbrace{v_m(t) \Delta T_{RHV,s}}_{item\ 2} + \underbrace{\frac{(v_\beta(t) - v_m(t))^2}{2a}}_{item\ 3} \right] + \underbrace{(2l + 2C)}_{item\ 4}$$

As an RHV cannot identify the vehicle class of the following vehicle.

Whether the target lane could provide sufficient space headway to fulfill the minimum space headway criteria is a prerequisite of a successful lane-changing maneuver. However, the final execution of a lane-changing maneuver still needs to fulfill other supplementary conditions apart from this minimum space headway criteria. This issue will be further discussed in the next section.

In summary, compared with traffic of pure RHVs, traffic flow mixed with CAVs would introduce the following three advantages: 1) the higher penetration rate of CAVs induces a higher capacity for the same freeway; 2) the minimum space headway gap required by the CAVs is smaller than that required by RHVs, for both MLC and DLC lane changing intentions; 3) the space headway in front of RHVs can be significantly enhanced if more

CAVs are traveling on the segment under the same traffic conditions because CAVs tolerate a smaller headway, as illustrated in Figure 5.6. As it can be inferred from advantage 2) and 3), lane changing will be easier and safer.

## **5.4 Optimal control strategies**

This section formulates the optimal control problem for freeway traffic with a given penetration rate of CAVs. As explained in Section 5.2, the control architecture considers a central traffic manager that devises and evaluates the optimal control problem and disseminates the control strategies to both CAVs and RHVs. It is assumed that the traffic manager has complete knowledge of the traffic state, which is estimated by the multiclass multilane traffic flow model. The optimal control is for the aggregated traffic flow rather than individual vehicles. In line with the literature, it is assumed that the penetration rate of CAVs is sufficient and that the control variables considered in the optimal control problem can be implemented via appropriate actuators. Then there is a need to assess the effect of the penetration rate of CAVs by sensitivity analysis.

### **5.4.1 Optimal control problem formulation**

The objective of the optimal control problem is to improve the traffic efficiency through VSLC and LCC of CAVs, and VSLR and LCR for RHVs while maintaining minimum gap acceptance criteria for traffic safety. State-of-the-practice VSLR uses VMS gantries, and the control effect depends on the compliance of drivers. With CAVs, VSLC can be performed by CAVs without any intervention by the driver. Under the mixed traffic flow case, both VSLC and VSLR should not be updated too frequently, and the amplitude should not be changed too sharply to avoid excessive nuisance to drivers and passengers. Such issues are considered by discretization of speed limit control and penalizing the space-time differences of LC ratios, i.e., Equation (5.14).

In the present work, the VSL is specified by the traffic manager for each segment-lane such that all CAVs traveling on the segment-lane will follow the respective speed or speed limit. As reported in the literature, the average speed of RHVs is roughly equal to the average speed of connected vehicles in a traffic stream (Bekiaris-Liberis et al., 2016; Fountoulakis et al., 2017). Therefore, it will be sufficient to impose the speed limit, even on RHVs, given an adequate penetration of equipped vehicles, e.g., there is at least one CAV within each segment-lane. Considering the nuisance issue as previously discussed, the segment-lane would consist of several cells. The optimal lateral flow can be obtained by solving the optimization problem. However, the implementation of this control action is more challenging and uncertain unless all vehicles are CAVs that will fully follow the control. For RHVs, traditional actuators, i.e., road-side VMS gantries, are employed in parallel, but only a certain percentage of drivers will comply with the recommendation delivered by the VMS gantries. This type of driver compliance will be considered in the multiclass multilane traffic flow model. It is assumed that the characteristics of different drivers, e.g., the origin-destination of drivers, is given as external input to the traffic flow model. Otherwise, a complicated estimation algorithm is necessary to infer the potential mandatory lane-changing (MLC) demand (see Pan et al., 2016).

An abstract formulation of the optimal control problem can be thus defined.

$$\min J_p(\mathbf{x}(k_N)) + \sum_{k=1}^{k_N} J(\mathbf{x}(k), \mathbf{q}(k)) \quad (5.15)$$

subject to traffic flow conservation

$$\mathbf{x}(k+1) = \text{fun}(\mathbf{x}(k), \mathbf{q}(k), \mathbf{c}(K)), \quad 1 < k < k_N, \text{ and } K = \lceil k \cdot T_s / T_c \rceil$$

with definition constraints to **decision variables** (control variables)

$$\mathbf{0} \leq \mathbf{c}(K) \leq \mathbf{C}, \text{ for } 1 < K < \lceil k_N T_s / T_c \rceil.$$

and physical definition constraints to state variables and flow variables

$$\begin{cases} \mathbf{A}_1 \cdot \mathbf{q}(k) = \mathbf{B}_1 \\ \mathbf{A}_2 \cdot \mathbf{q}(k) \leq \mathbf{B}_2 \\ \mathbf{0} \leq \mathbf{x}(k) \leq \mathbf{X}(k) \\ \mathbf{0} \leq \mathbf{q}(k) \leq \mathbf{Q}(k) \end{cases} \quad \text{for } 1 < k < k_N$$

where the objective function will be specified in (5.16) with  $k = 1 \dots k_N$  as the simulation time steps of the control horizon.  $T_s$  denotes the time duration of a simulation step.  $\lceil x \rceil$  is the **ceiling** function that outputs the least **integer** greater than or equal to  $x$ . Table 5.2 (a) demonstrates the relationship between the simulation step, control cycle, and simulation horizon.

The term  $\text{fun}()$  is a compact representation of the underlying traffic flow model that describes the multiclass multilane traffic dynamics, which will be further elaborated in Sections 5.4.2–5.4.4.  $\mathbf{x}(k)$  is the vector of state variable during the period  $[kT_s, (k+1)T_s)$  and  $\mathbf{q}(k)$  is the vector of sending and receiving functions of traffic demands and flow that govern flow propagation in the multiclass multilane traffic flow model during  $[kT_s, (k+1)T_s)$ . The propagation of  $\mathbf{x}(k)$  and evaluation of  $\mathbf{q}(k)$  are described in Section 5.4.2.

$\mathbf{c}(K)$  is the vector of control variables (which are the decision variables). If the optimal control problem is solvable, the optimal control strategy  $\mathbf{c}^*(K)$  will be implemented during the period  $[KT_c, (K+1)T_c)$ , where  $T_c$  denotes the duration of each control cycle and the integer  $T_c/T_s$  is the number of simulation steps within each control cycle. The notations involved in defining the optimal control problem, and the constraints  $\mathbf{X}(k)$  and  $\mathbf{Q}(k)$  are defined and summarized in Table 5.3.

The objective function includes the total time spent, fuel consumption and emission cost of RHV, electricity cost of CAV, and the penalty on residual extra-queue within the pre-

defined period, as given in Equation (5.16).

$$\begin{aligned}
J(\mathbf{x}(k), \mathbf{q}(k)) = & \sum_{i=1}^N \sum_{x=m, \beta} \sum_{d=CAV, RHV} V_i \cdot \underbrace{(\rho_{i,x,D}(k) \cdot l_i + W_{x,D}(k))}_{\text{total time spent+Waiting time of extra-queue}} \cdot T_s \\
& + \underbrace{\sum_{i=1}^N \sum_{x=m, \beta} \sum_{Z=1,2,3} c_f \cdot F_{Z,i,x,RHV}(k) \cdot e^{(V_{Z,i,x}(k) \cdot K_f \cdot A_{Z,i,x}(k))^T}}_{\text{Fuel consumption of RHV}} \cdot T_s \\
& + \underbrace{\sum_{e=CO,HC,NO} \sum_{i=1}^N \sum_{x=m, \beta} \sum_{Z=1,2,3} c_e \cdot F_{Z,i,x,RHV}(k) \cdot e^{(V_{Z,i,x}(k) \cdot K_e \cdot A_{Z,i,x}(k))^T}}_{\text{Emission cost of RHV}} \cdot T_s \\
& + \underbrace{\sum_{i=1}^N \sum_{x=m, \beta} \sum_{Z=1,2,3} c_c \cdot F_{Z,x,CAV}(k) \cdot e^{(V_{Z,i,x}(k) \cdot K_c \cdot A_{Z,i,x}(k))^T}}_{\text{Electricity cost of CAV}} \cdot T_s
\end{aligned} \tag{5.16}$$

$$J_p(\mathbf{x}(k_N)) = \varphi \cdot V_i \cdot \underbrace{(W_{m,CAV}(k_N) + W_{m,CAV}(k_N))}_{\text{Penalty on the residual extra-queue}}$$

where  $N$  is the number of cells,  $k_N$  is the total simulation steps, and  $K_N$  is the total control cycles.

As depicted in Equation (5.16), the objective function is composed of the following parts:

- a) The first linear term represents the total travel time (TTT) spent on the modeling freeway stretch and the time spent in an extra queue upstream to the entry point if any.
- b) The second to the fourth linear terms represent the fuel consumption and emission cost associated with RHVs, and electricity cost of CAVs. These terms are closely related to speed, inter-cell, and cross-cell acceleration/deceleration. Traffic flows associated with cell  $(i, x)$ , where  $x$  stands for lane  $m$  or lane  $\beta$ , are categorized into three items ( $Z=1, 2, 3$ ) according to the moving directions. To be specific, the straightforward flow  $F_{1,i,x,d}(k)$  which enters cell  $(i+1, x)$  from cell  $(i, x)$  during time interval  $[kT_s, (k+1)T_s)$  with  $d$  stands for *RHV* and *CAV*, respectively; the lane-changing flow  $F_{2,i,x,d}(k)$  that switches lane from cell  $(i, x)$  to cell  $(i+1, \bar{x})$ , where  $\bar{x}$  denotes the lanes adjacent to lane  $x$ ; and  $F_{3,i,x,d}(k)$  the flow that retains its current cell-lane.  $c_f$  (USD/liter);  $c_e$

(USD/gram) and  $c_c$  (USD/watt) are coefficients that convert the fuel, emission and electricity costs into the monetary counterpart. The vector of speed  $V_{Z,i,x}(k) = [1, v_{Z,i,x}(k), v_{Z,i,x}^2(k), v_{Z,i,x}^3(k)]$  and the vector of acceleration/deceleration  $A_{Z,i,x}(k) = [1, a_{Z,i,x}(k), a_{Z,i,x}^2(k), a_{Z,i,x}^3(k)]$  are necessary when calculating the fuel consumption rate (liter/sec), emission rate (gram/sec) and electricity consumption rate (watt/sec). The elements of  $V_{Z,i,x}(k)$  and  $A_{Z,i,x}(k)$  are listed in Table 5.2 (b) (Liu et al., 2017).  $K_f, K_e$  and  $K_c$  are the model coefficient matrices (Tang et al., 2017; Ahn et al., 2007), which will be summarized in Table 5.6.

c) The last term  $J_p(\mathbf{x}(k_N))$  penalizes the extra-queue at the end of the simulation. The factor  $\varphi$  is set to be a large value to eliminate residual extra-queue.

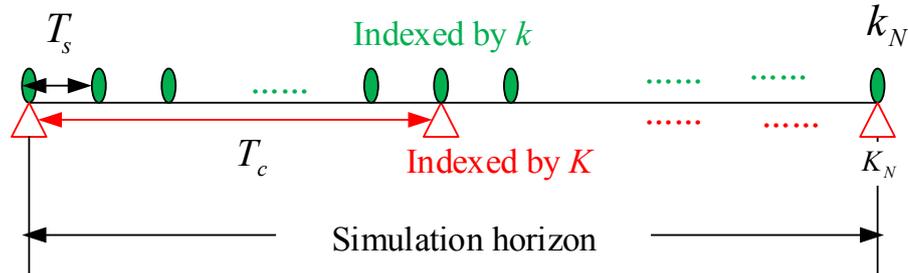


Table 5.2(a). Relationship between different time horizons of the optimal control problem

Z	$F_{Z,i,x,d}(k)$	$v_{Z,i,x}(k)$	$a_{Z,i,x}(k)$
1	$q_{st,i,x,d}(k) \cdot T_s$	$v_{i,x}(k) \rightarrow v_{i+1,x}(k+1)$	$\frac{v_{i+1,x}(k+1) - v_{i,x}(k)}{T_s}$
2	$(q_{m,i,x,d}^{i+1,\bar{x}}(k) + q_{d,i,x,d}^{i+1,\bar{x}}(k)) \cdot T_s$	$v_{i,x}(k) \rightarrow v_{i+1,\bar{x}}(k+1)$	$\frac{v_{i+1,\bar{x}}(k+1) - v_{i,x}(k)}{T_s}$
3	$S_{i,x,d}(k) \cdot T_s - \sum_{Z=1,2} F_{Z,i,x,d}(k)$	$v_{i,x}(k) \rightarrow v_{i,x}(k+1)$	$\frac{v_{i,x}(k+1) - v_{i,x}(k)}{T_s}$

Table 5.2(b). Speed and acceleration in determination of fuel consumption, emission and electricity cost during time interval  $[kT_s, (k+1)T_s)$

Vector	Elements	Notation description	Duration	Constraints
$\mathbf{x}(k)$	$\rho_{i,x,CAV}(k)$	Density of CAVs on cell $(i, x)$ during step $k$	$T_s$	$\left[0, \frac{1}{l+C}\right]$
	$\rho_{i,x,RHV}(k)$	Density of RHVs on cell $(i, x)$ during step $k$		$\left[0, \frac{1}{l+C}\right]$
	$P_{i,x,CAV}(k)$	Proportion/Penetration rate of CAVs on cell $(i, x)$ during step $k$		$[0, 1]$
	$v_{i,x}(k)$	Traffic speed on cell $(i, x)$ during step $k$		$\left[0, \hat{v}_{i,x,l}(K)\right]$
	$W_{x,CAV}(k)$	Virtual extra-queue state variables at the upstream boundary of cell $(1, x)$ with respect to vehicle class CAV or RHV during step $k$		----
	$W_{x,RHV}(k)$			----
	$S_{i,x,CAV}^{i+1,\bar{x}}(k)$	Sending function of CAVs issued by LCC from cell $(i, x)$ to cell $(i+1, \bar{x})$ during step $k$		
	$S_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$	Sending function of RHVs with MLC or DLC demand from cell $(i, x)$ to cell $(i+1, \bar{x})$ during step $k$		$\left[0, Q_{i,m}(k) \cdot P_{i,x,RHV}(k)\right]$
	$S_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$			
	$S_{st,i,x,CAV}(k)$	Sending function of flow with vehicle class CAV or RHV that intend to leave cell $(i, x)$ straightforwardly and be received by downstream cell $(i+1, x)$ during step $k$		$\left[0, Q_{i,m}(k)P_{i,m,CAV}(k)\right]$
	$S_{st,i,x,RHV}(k)$			$\left[0, Q_{i,m}(k)P_{i,m,RHV}(k)\right]$
	$\tilde{H}_{i,x,CAV}^{i+1,\bar{x}}(k)$	Minimum space headway criteria of $S_{i,x,CAV}^{i+1,\bar{x}}(k)$		----
	$\tilde{H}_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$	Minimum space headway criteria of $S_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$ and $S_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$		----
	$\tilde{H}_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$			----
	$q_{i,x,CAV}^{i+1,\bar{x}}(k)$	The amount of $S_{i,x,CAV}^{i+1,\bar{x}}(k)$ received by the target cell (or executed LCC)		$\left[0, S_{i,x,CAV}^{i+1,\bar{x}}(k)\right]$
$q_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$	The amount of $S_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$ and	$\left[0, S_{M,i,x,RHV}^{i+1,\bar{x}}(k)\right]$		

	$q_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$	$s_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$ received by the target cells, respectively			$[0, s_{D,i,x,RHV}^{i+1,\beta}(k)]$
	$q_{st,i,x,CAV}(k)$	The amount of $s_{st,i,x,CAV}(k)$ and $s_{st,i,x,RHV}(k)$ received by the target cells			$[0, s_{st,i,x,CAV}(k)]$
	$q_{st,i,x,RHV}(k)$				$[0, s_{st,i,x,RHV}(k)]$
	$R_{i,x}(k)$	Receiving function of cell $(i, x)$ during step $k$			$[0, Q_{i,x}(k)]$
$\mathbf{c}(K)$	$\hat{v}_{i,x,l}(K)$	Implemented speed limit on cell $(i, x)$ during control cycle $K$	$T_c$	$\mathbf{C}$	$[0, v_f]$
	$\hat{p}_{i,x,CAV}^{i+1,\bar{x}}(K)$	The initial ratio of CAVs that are issued to switch from cell $(i, x)$ to cell $(i+1, \bar{x})$ during control cycle $K$			$[0, 1]$
	$\hat{B}_{i,x}^{\bar{x}}(K)$	Variable denoting the whether the lane changing recommendation from cell $(i, x)$ to lane $\bar{x}$ is released during control cycle $K$			0 (not released) or 1 (released)

Table 5.3. Key variables of the optimal control problem

Finally, the constraints  $\begin{cases} \mathbf{A}_1 \cdot \mathbf{q}(k) = \mathbf{B}_1 \\ \mathbf{A}_2 \cdot \mathbf{q}(k) \leq \mathbf{B}_2 \end{cases}$  will be depicted in Section 5.4.2.

### 5.4.2 Traffic-state propagation

In this section, a multilane cell transmission model is proposed to simulate the effects of penetration rates of CAVs, human drivers' behavior, speed limit and lane control policies. The multiclass multilane CTM serves as the network loading model in the optimal control problem. To enable the simulation of lane-changing maneuvers for traffic flow mixed with RHVs and CAVs, the proposed model incorporates several new features. First, it considers the effect of time-dependent proportion of CAVs on the minimum safe gap

acceptance and the cell-lane fundamental diagram. Second, it defines the new reaction laws of sending and receiving functions that govern the flow propagation with respect to the multiclass cell-lane fundamental diagram. And third, it enables lane-changing priorities with respect to different vehicle classes and the lane-changing motivations.

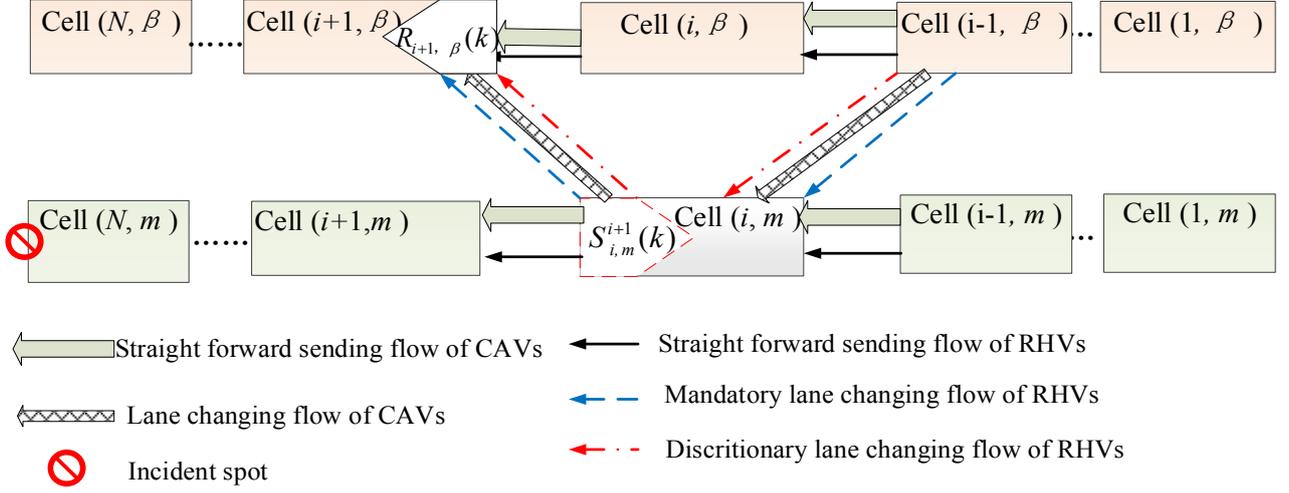


Figure 5.8. Merging and diverging induced by lane-changing of CAVs and RHVs.

Consider a dual-lane freeway segment divided into several cell packages, from cell package 1 to package  $N$ , along the longitudinal dimension, as shown in Figure (5.8). Each cell package includes two cells located on lane  $m$  and lane  $\beta$ , denoted as cell  $(i, m)$  and cell  $(i, \beta)$ , respectively. It is assumed that there is only one bottleneck (caused by lane drop or traffic accident) on this freeway segment, which is located at the end of cell  $(N, m)$ . Take cell  $(i, m)$  as shown in Figure 5.8 as an example, based on flow conservation, the cell densities  $\rho_{i, m, CAV}(k+1)$  (P.C.U./mile/lane) of CAVs and  $\rho_{i, \beta, RHV}(k+1)$  (P.C.U./mile/lane) of RHVs, are estimated as Equation (5.17):

$$\left\{ \begin{array}{l} \rho_{i, m, CAV}(k+1) = \rho_{i, m, CAV}(k) + \frac{T_s}{l_i} \left( q_{st, i-1, m, CAV}^{i, m}(k) + q_{i-1, \beta, CAV}^{i, m}(k) \right) - \frac{T_s}{l_i} \left( q_{st, i, m, CAV}^{i+1, m}(k) + q_{i, m, CAV}^{i+1, \beta}(k) \right) \\ \rho_{i, m, RHV}(k+1) = \rho_{i, m, RHV}(k) + \frac{T_s}{l_i} \left( q_{st, i-1, m, RHV}^{i, m}(k) + q_{M, i-1, \beta, RHV}^{i, m}(k) + q_{D, i-1, \beta, RHV}^{i, m}(k) \right) \\ \quad - \frac{T_s}{l_i} \left( q_{st, i, m, RHV}^{i+1, m}(k) + q_{M, i, m, RHV}^{i+1, \beta}(k) + q_{D, i, m, RHV}^{i+1, \beta}(k) \right) \end{array} \right. \quad (5.17)$$

respectively. The estimation of density on cell  $(i, \beta)$  can be similarly defined.

The cell density is a summation of the two vehicle classes, *i.e.*,

$$\rho_{i,x}(k+1) = \rho_{i,x,CAV}(k+1) + \rho_{i,x,RHV}(k+1), \text{ for } x = m \text{ or } \beta \quad (5.18)$$

whereas the proportion of CAVs in a particular cell-lane unit is given by

$$P_{i,x,CAV}(k+1) = \rho_{i,x,CAV}(k+1) / \rho_{i,x}(k+1), \text{ if } \rho_{i,x}(k+1) > 0 \quad (5.19)$$

According to Equation (5.3), the cell-lane speed is given by

$$v_{i,x}(k+1) = \begin{cases} \hat{v}_{i,x,l}(K), & \text{if } \rho_{i,x}(k+1) \leq \rho_{c,i,x}(k) \\ \frac{(1-l \cdot \rho_{i,x}(k+1) - C \cdot \rho_{i,x}(k+1))}{\rho_{i,x}(k+1) (P_{i,x,CAV}(k+1) \Delta T_{CAV} + (1 - P_{i,x,CAV}(k+1)) \Delta T_{RHV})}, & \text{if } \rho_{i,x}(k+1) > \rho_{c,i,x}(k) \end{cases} \quad (5.20)$$

depending on whether the speed limit control or the permanent compulsory speed limit is implemented. The above flow conservation and speed equations can be regarded as equality constraints of the optimization problem

With  $x$  denoting either lane  $m$  or lane  $\beta$  and  $\bar{x}$  denoting lane  $\beta$  or lane  $m$  accordingly, the constraints on the control variables are listed in Equation (5.21). To ensure the feasibility of speed limit control and to maintain traffic flow stability,  $\hat{v}_{i,x,l}(K)$  is constrained on the lateral (the 1<sup>st</sup> equation), longitudinal (the 2<sup>nd</sup> equation and the 3<sup>rd</sup> inequality), and temporal dimensions (the 4<sup>th</sup> inequality). And it is discretized into several intervals with sufficient increments (rather than changing continuously, the 5<sup>th</sup> constraint). Besides,  $\hat{B}_{i,x}^{\bar{x}}(K)$ , denoting whether the RHVs traveling on cell  $(i, x)$  is recommended to switch to the adjacent lane  $\bar{x}$ , is also constrained on the lateral (the 6<sup>th</sup> equation), longitudinal (the 7<sup>th</sup> equation), and their changes are restricted to be less than two times (the 8<sup>th</sup> inequality). The LC ratio  $\hat{P}_{i,x,CAV}^{j+1,\bar{x}}(K)$  of each cell should be within 0%-100% (the 9<sup>th</sup> constraint).

- 1).  $\hat{v}_{i,x,l}(K) - \hat{v}_{i,\bar{x},l}(K) = 0,$
- 2).  $\hat{v}_{i,x,l}(K) - \hat{v}_{i-1,x,l}(K) = 0,$  if  $i \in \mathbf{G}$  and  $i \geq 2$
- 3).  $|\hat{v}_{i,x,l}(K) - \hat{v}_{i-1,x,l}(K)| \leq \Delta v_c,$  if  $i \in \tilde{\mathbf{G}}$  and  $i \geq 2$
- 4).  $|\hat{v}_{i,x,l}(K) - \hat{v}_{i,x,l}(K-1)| \leq \Delta v_c, \quad \forall K \geq 2$
- 5).  $\hat{v}_{i,x,l}(K) \in V_L,$
- 6).  $\hat{B}_{i,x}^{\bar{x}}(K) \cdot \hat{B}_{i,\bar{x}}^x(K) = 0,$  (5.21)
- 7).  $\hat{B}_{i,x}^{\bar{x}}(K) - \hat{B}_{i-1,x}^{\bar{x}}(K) = 0, \quad \forall i \geq 2$
- 8).  $\sum_{K=2}^K |\hat{B}_{i,x}^{\bar{x}}(K) - \hat{B}_{i,x}^{\bar{x}}(K-1)| \leq 2, \quad \forall i \geq 2$
- 9).  $0 \leq \hat{p}_{i,x,CAV}^{i+1,\bar{x}}(K) \leq 1, \quad \forall i \leq N-1$
- 10).  $\hat{p}_{i,x,CAV}^{i+1,\bar{x}}(K) \cdot \hat{p}_{i,\bar{x},CAV}^{i+1,x}(K) = 0.$

First, on the lateral dimension, vehicles (both RHVs and CAVs) in parallel lanes are compelled to follow the same speed limit at the same longitudinal location, as stated by the 1<sup>st</sup> equation of Equation (5.21). This is because implementing a lane-specific speed limit might encourage DLC maneuvers, which would further reduce the freeway capacity and induce instability in traffic flow, especially in the case of a traffic incident or over-saturated traffic conditions. Furthermore, different speed limits issued by a single VMS might confuse the human drivers.

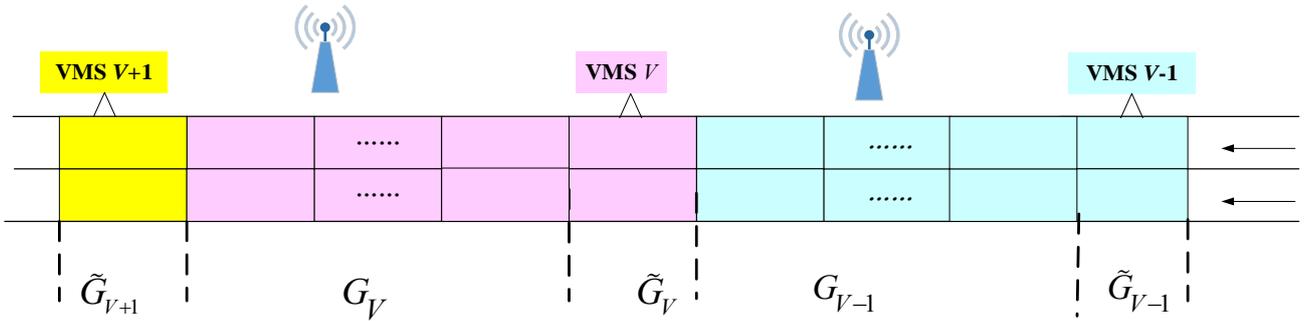


Figure 5.9. Illustration of freeway traffic control by VMS gantries.

As depicted in Figure 5.9,  $G_V$  denotes a freeway segment governed by VMS  $V$ , whereas  $\tilde{G}_V$  denotes the boundary cell of this segment connecting the segment governed by VMS  $V-1$ . Let set  $\mathbf{G}$  to be the collection of all  $G_V$ , for  $V= 1, 2, \dots, N_V$ , where  $N_V$  is the

number of VMS gantries installed along the study site. Similarly,  $\tilde{\mathbf{G}}$  denotes the combination of all boundaries  $\tilde{G}_V$ . The 2<sup>nd</sup> and 3<sup>rd</sup> items of formula (5.21) constrain the speed limit fluctuation on the longitudinal dimension to take into account the operational feasibility and comfort in driving. First of all, the speed limit should be the same within the segment governed by a gantry. Moreover, the speed limit should not admit sharp variations from cell to cell for two reasons. Because the cell length is not long, e.g., 200 meters, varying the speed limit sharply from cell to cell would trigger oscillations in the traffic flow that would further induce instability. Furthermore, the VMS gantries cannot be installed so densely as to disseminate cell-to-cell VSL control to RHDVs. Even if gantries are installed for each cell, human drivers cannot adapt to such quickly varying speed limit control. The 3<sup>rd</sup> inequality suggests that at the same time the VSLR released by two adjacent VMS gantries should not admit sharp variation. The term  $\Delta v_c$  is the maximal tolerable fluctuation of speed limit between two adjacent boundary cells of two successive VMS segments on the spatial dimension, such as between  $\tilde{G}_V$  and  $G_{V-1}$ , as well as the fluctuation between control cycles on the temporal dimension. In this study, the value of  $\Delta v_c$  is restricted to ensure that the new speed limit could be achieved within a simulation step, e.g.,  $\Delta v_c = 20$  mile/hour can be achieved in 10 seconds by acceleration/deceleration. The same fluctuation restriction  $\Delta v_c$  is also applied to constrain the temporal variation of VSL on VMS, as expressed in the 4<sup>th</sup> inequality of formula (5.21).

Also, the increment (or decrement) (also termed as quantization in Zhang and Ioannou (2017)) of VSL must be practical for human drivers. For example, a change of 1 mile/hour in the VSL might be imperceptible to human drivers, whereas a change of 5 miles/hour is quite perceptible to drivers and thus may be more suitable. To this end, in this chapter, the speed limit control is discretized as the finite value set  $L_V = \{10, 15, \dots, v_f - 5, v_f\}$ , with increments of 5 miles/hour, saying the lowest VSL issued by the system is 10 miles/hour. The five items as introduced above constrain the variables

associated with the VSLR released by the VMS while the VSLC issued by the VACS also follows the same constraints.

Some of the constraints have been included in different forms in the literature. For example, Zhang and Ioannou (2017) considered the discretization, quantization of VSL and saturation of speed limit variations. The constraint set proposed here is more general than that in Zhang and Ioannou (2017) since both spatial-temporal saturation of speed limit variations meanwhile asynchronous simulation steps for traffic flow model and the cycle for control implementation are considered in this chapter.

The 6<sup>th</sup> formula of (5.21) suggests that the opposite lane changing orders cannot be released on the same VMS segment during the same control cycle. The 7<sup>th</sup> equation of (5.21) assumes the VMS upstream to the incident spot release the same LCR. The 8<sup>th</sup> inequality prevents the case that the RHVs are suggested to keep changing their lanes more than twice as frequent lane-changing maneuvers would disrupt traffic stability. When traffic on lane  $x$  is suggested to change to the adjacent lane  $\bar{x}$ , the traffic on lane  $\bar{x}$  must not be suggested to change to lane  $x$  (10<sup>th</sup> equation).

### 5.4.3 Sending and receiving functions for multiclass multilane traffic flow propagation

The evaluation of flow vector  $\mathbf{q}(k)$  is vital for the propagating the multiclass multilane traffic flow for determining the state vector  $\mathbf{x}(k)$ . Apart from the state variable  $\mathbf{x}(k)$  and the vector of the control variable  $\mathbf{c}(K)$ , the external input  $\mathbf{u}(k)$  in conjunction with the fundamental diagram  $\mathbf{f}(k)$  is also necessary for the evaluation of  $\mathbf{q}(k)$ . To this end,  $\mathbf{q}(k)$  is represented as a mapping of these vectors as follows:

$$\mathbf{q}(k) = \text{fun}(\mathbf{x}(k), \mathbf{c}(K), \mathbf{u}(k), \mathbf{f}(k)), \quad \text{for } 1 < k < k_N, \text{ and } K = \lceil k \cdot T_s / T_c \rceil$$

Vector	Elements	Notation description	Duration
$\mathbf{f}(k)$	$\hat{v}_{i,x,l}(K)$	Speed limit control on cell $(i, x)$ during cycle $K$	$T_c$
	$\rho_{c,i,x}(k)$	Critical density on cell $(i, x)$ at simulation step $k$	$T_s$
	$Q_{i,x}(k)$	Capacity on cell $(i, x)$ at simulation step $k$	$T_s$
	$w_{i,x}(k)$	Wave-back speed on cell $(i, x)$ at simulation step $k$	$T_s$
	$\rho_{i,x,j}(k)$	Jams density on cell $(i, x)$ at simulation step $k$	$T_s$
$\mathbf{u}(k)$	$S_{0,x}(k)$	Inflow demand of lane $x$ at simulation step $k$	$T_s$
	$P_{0,x,CAV}(k)$	Proportion of CAVs associated with $S_{0,x}(k)$	$T_s$
	$R_{E,x}(k)$	Available space of the downstream section at time step $k$	$T_s$

Table 5.4. Variables in traffic flow propagation

The sending function  $s_{i,x,d}^{i+1}(k)$  (P.C.U./hour/lane), which quantifies the traffic flow that intends to leave cell  $(i, x)$  for vehicle class  $D$ , is defined as

$$s_{i,x,d}^{i+1}(k) = \begin{cases} \hat{v}_{i,x}(K) \cdot \rho_{i,x}(k) \cdot P_{i,x,d}(k) & \text{if } \rho_{i,x}(k) < \rho_{i,x,c}(k) \quad \text{for } i = 1, 2, \dots, N-1 \\ Q_{i,x}(k) \cdot P_{i,x,d}(k) & \text{if } \rho_{i,x}(k) \geq \rho_{i,x,c}(k) \quad \text{for } i = 1, 2, \dots, N-1 \\ S_{0,x}(k) \cdot P_{0,x,d}(k) + W_{x,d}(k) & \text{for } i = 0 \end{cases} \quad (5.23)$$

where  $P_{i,x,d}(k)$  denotes the proportions of CAVs and RHVs in cell  $(i, x)$ , with  $d = CAV$  and  $RHV$ , respectively. Where

$$P_{i,x,RHV}(k) = 1 - P_{i,x,CAV}(k) \quad (5.24)$$

where  $P_{0,x,d}(k)$  is the proportion of vehicle class  $d$  arriving at the upstream boundary,  $S_{0,x}(k)$  is the inflow demand towards lane  $x$  and  $W_{x,d}(k)$  is the number of vehicles of class  $d$  queuing at the upstream boundary to cell  $(1, x)$  at simulation time step  $k$ .

The sending function  $s_{i,x,d}^{i+1}(k)$  determines the flow intended to leave cell  $(i, x)$  during

time step  $k$  and moves toward the downstream cell package ( $i + 1$ ) involving all possible lanes, i.e., cell  $(i+1, \bar{x})$  in the adjacent lane and cell  $(i+1, x)$  in the current lane. The lane changing demands can be further categorized into three possible types: a) CAVs following the lane-changing instructions issued by VACS, which is denoted by  $s_{i,x,CAV}^{i+1,\bar{x}}(k)$ ; b) RHVs making MLC decisions based on lane-changing suggestions, which is denoted by  $s_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$ ; and c) RHVs making DLC decisions to increase speed, which is denoted by  $s_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$ . These lane-changing flows can be evaluated as

$$\begin{aligned}
s_{i,x,CAV}^{i+1,\bar{x}}(k) &= \left\lfloor s_{i,x,CAV}^{i+1}(k) \cdot \hat{p}_{i,x,CAV}^{i+1,\bar{x}}(K) \cdot T_s \right\rfloor / T_s \\
s_{MLC,i,x,RHV}^{i+1,\bar{x}}(k) &= s_{i,x,RHV}^{i+1}(k) \cdot p_{MLC,i,x,RHV}^{i+1,\bar{x}}(k) \cdot \hat{B}_{i,x}^{\bar{x}}(K) \\
s_{DLC,i,x,RHV}^{i+1,\bar{x}}(k) &= s_{i,x,RHV}^{i+1}(k) \cdot p_{DLC,i,x,RHV}^{i+1,\bar{x}}(k) \cdot \left(1 - \hat{B}_{i,x}^{\bar{x}}(K)\right)
\end{aligned} \tag{5.25}$$

$\lfloor x \rfloor$  represents the **floor** function that outputs the largest integer smaller than or equal to  $x$ , as the LCC orders cannot be issued to non-integer number of CAVs. The other two proportions  $s_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$  and  $s_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$  are associated with RHVs based on human drivers' individual decisions. The determination of mandatory lane changing demand proportion of RHVs  $s_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$  follows the compliance rule of RHVs to lane changing recommendation issued by VMS gantries. The evaluation depends on the following elements: the remaining distance  $d_i$  from the current cell package  $i$  to the incident spot; the cell density on target lane  $\rho_{i,\bar{x}}(k)$ ; the total MLC demand that intends to switch from lane  $x$  to lane  $\bar{x}$ ,  $S_{MLC,x,RHV}^{\bar{x}}(k)$ ; and the accumulation of executed MLC flow from cell  $(1, x)$  to cell  $(i, x)$ .

$$p_{MLC,i,x,RHV}^{i+1,\bar{x}}(k) = \left\{ \left[ S_{MLC,x,RHV}^{\bar{x}}(k) \cdot e^{-(d_i - d_c)^2 / (\alpha_1 + \alpha_2 \cdot \rho_{i,\bar{x}}(k))^2} - \sum_{j=1}^{i-1} q_{MLC,j,x,RHV}^{j+1,\bar{x}}(k) \right] / s_{i,x,RHV}^{i+1}(k) \right\} \tag{5.26}$$

The detailed discussion on mandatory lane changing and calibration of parameters  $\alpha_1$ ,  $\alpha_2$  and critical density  $d_c$  can be referred to Chapter 3 (Yang and Koutsopoulos, 1996).

DLC is motivated by the speed advantage of the adjacent lane if no LCR is issued to the current segment on the current lane, i.e.,  $\hat{B}_{i,x}^{\bar{x}}(K) = 0$ .

$$p_{DLC,i,x,RHV}^{i+1,\bar{x}}(k) = \frac{\max(0, v_{i,\bar{x}}(k) - v_{i,x}(k))}{v_f \tau} \quad (5.27)$$

The term  $\tau$  can be interpreted as the average time a driver takes to decide and execute a lane change when the original lane is stopped, and the target lane is free-flowing (Laval and Daganzo, 2006).

Finally, the straightforward flow intended to leave cell  $(i, x)$  and enter cell  $(i+1, x)$  is

$$\begin{aligned} s_{st,i,x,CAV}^{i+1,x}(k) &= s_{i,x,CAV}^{i+1}(k) - s_{i,x,CAV}^{i+1,\bar{x}}(k) \\ s_{st,i,x,RHV}^{i+1,x}(k) &= s_{i,x,RHV}^{i+1}(k) - s_{MLC,i,x,RHV}^{i+1,\bar{x}}(k) - s_{DLC,i,x,RHV}^{i+1,\bar{x}}(k) \end{aligned} \quad (5.28)$$

However, the final execution of all the lane changing demands and straightforward demands still depends on two external factors: the fulfillment to minimum gap acceptance criteria by target lane, and the accommodation of several merging demands considering specific priorities.

#### 5.4.4. Allocation of available space for merging flows

The available space of the target cell  $(i+1, \bar{x})$   $R_{i+1,\bar{x}}(k)$  (P.C.U./hour/lane) can be evaluated according to the fundamental diagram  $\mathbf{f}(k)$ .

$$R_{i+1,\bar{x}}(k) = \begin{cases} w_{i+1,\bar{x}}(k) (\rho_{J,i+1,\bar{x}}(k) - \rho_{i+1,\bar{x}}(k)) & \text{if } \rho_{i+1,\bar{x}}(k) < \rho_{c,i+1,\bar{x}}(k) \text{ for } i = 0, 2, \dots, N-1 \\ Q_{i+1,\bar{x}}(k) & \text{if } \rho_{i+1,\bar{x}}(k) \geq \rho_{c,i+1,\bar{x}}(k) \text{ for } i = 0, 2, \dots, N-1 \\ R_{E,\bar{x}}(k) & \text{for } i = N \end{cases} \quad (5.29)$$

where  $E$  denotes the downstream sink connected to the last cell  $N$ .

Toward the target cell  $(i+1, \bar{x})$ , there are five sending flows, i.e.,  $s_{st,i,x,CAV}^{i+1,x}(k)$ ,  $s_{st,i,x,RHV}^{i+1,x}(k)$ ,  $s_{i,x,CAV}^{i+1,\bar{x}}(k)$ ,  $s_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$ , and  $s_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$ . The last three items assess

the feasibility in merging to target lane via the three minimum space headway criteria  $\tilde{H}_{i,x,CAV}^{i+1,\bar{x}}(k)$ ,  $\tilde{H}_{MLC,i,x,RHV}^{i+1,\bar{x}}(k)$ , and  $\tilde{H}_{DLC,i,x,RHV}^{i+1,\bar{x}}(k)$ , respectively. Considering the limited available space provided by the target cell, the flows that can be received by the target cell are calculated by Equations (5.30) and (5.31).

$$q_{lc,i,x,d}^{i+1,\bar{x}}(k) = \begin{cases} s_{lc,i,x,d}^{i+1,\bar{x}}(k) & \text{if } U_i^{i+1,\bar{x}}(k) \leq R_{i+1,\bar{x}}(k) \text{ and } \tilde{H}_{lc,i,x,d}^{i+1,\bar{x}}(k) \leq \bar{H}_{i+1,\bar{x}}(k) \\ \frac{s_{lc,i,x,d}^{i+1,\bar{x}}(k)}{U_i^{i+1,\bar{x}}(k)} R_{i+1,\bar{x}}(k) & \text{if } U_i^{i+1,\bar{x}}(k) > R_{i+1,\bar{x}}(k) \text{ and } \tilde{H}_{lc,i,x,d}^{i+1,\bar{x}}(k) \leq \bar{H}_{i+1,\bar{x}}(k) \\ 0 & \text{if } \tilde{H}_{lc,i,x,d}^{i+1,\bar{x}}(k) > \bar{H}_{i+1,\bar{x}}(k) \end{cases} \quad (5.30)$$

$$q_{st,i,\bar{x},d}^{i+1,\bar{x}}(k) = \begin{cases} s_{st,i,\bar{x},d}^{i+1,\bar{x}}(k) & \text{if } U_i^{i+1,\bar{x}}(k) \leq R_{i+1,\bar{x}}(k) \\ \frac{s_{st,i,\bar{x},d}^{i+1,\bar{x}}(k)}{U_i^{i+1,\bar{x}}(k)} R_{i+1,\bar{x}}(k) & \text{if } U_i^{i+1,\bar{x}}(k) > R_{i+1,\bar{x}}(k) \end{cases} \quad (5.31)$$

The subscript  $lc$  in  $q_{lc,i,x,d}^{i+1,\bar{x}}(k)$  (P.C.U./hour/lane) refers to the lane-changing motivations and  $d$  refers to vehicle types (see Table 5.3).  $\bar{H}_{i+1,\bar{x}}(k)$  denotes the average gap length on the target cell ( $i+1, \bar{x}$ ):

$$\bar{H}_{i+1,\bar{x}}(k) = P_{i+1,\bar{x},CAV}(k) \cdot H_{i+1,\bar{x},CAV}(k) + (1 - P_{i+1,\bar{x},CAV}(k)) \cdot H_{i+1,\bar{x},RHV}(k)$$

$U_i^{i+1,\bar{x}}(k)$  (P.C.U./hour/lane) denotes the total space required by the lane-changing demand (by fulfilling the minimum space headway criterion in the target lane) toward cell ( $i+1, \bar{x}$ ).

$$U_i^{i+1,\bar{x}}(k) = s_{st,i,\bar{x},CAV}^{i+1,\bar{x}}(k) + s_{st,i,\bar{x},RHV}^{i+1,\bar{x}}(k) + \sum_{lc=MLC}^{DLC} \sum_{d=CAV}^{RHV} s_{lc,i,x,d}^{i+1,\bar{x}}(k) \frac{\tilde{H}_{lc,i,x,d}^{i+1,\bar{x}}(k)}{l}, \quad \text{for } \forall \tilde{H}_{lc,i,x,d}^{i+1,\bar{x}}(k) < \bar{H}_{i+1,\bar{x},RHV}(k) \quad (5.32)$$

Finally, the number of vehicles (including CAVs and RHVs) queuing upstream at the boundary cell ( $1, x$ ) is evaluated as follows:

$$W_{x,CAV}(k+1) = W_{x,CAV}(k) + (S_{0,x}(k)P_{0,x,CAV}(k) - q_{st,0,x,CAV}^{1,x}(k))T_s \quad (5.33)$$

$$W_{x,RHV}(k+1) = W_{x,RHV}(k) + (S_{0,x}(k)P_{0,x,RHV}(k) - q_{st,0,x,RHV}^{1,x}(k))T_s$$

A flowchart of the proposed optimal control framework is summarized in Figure 5.10.

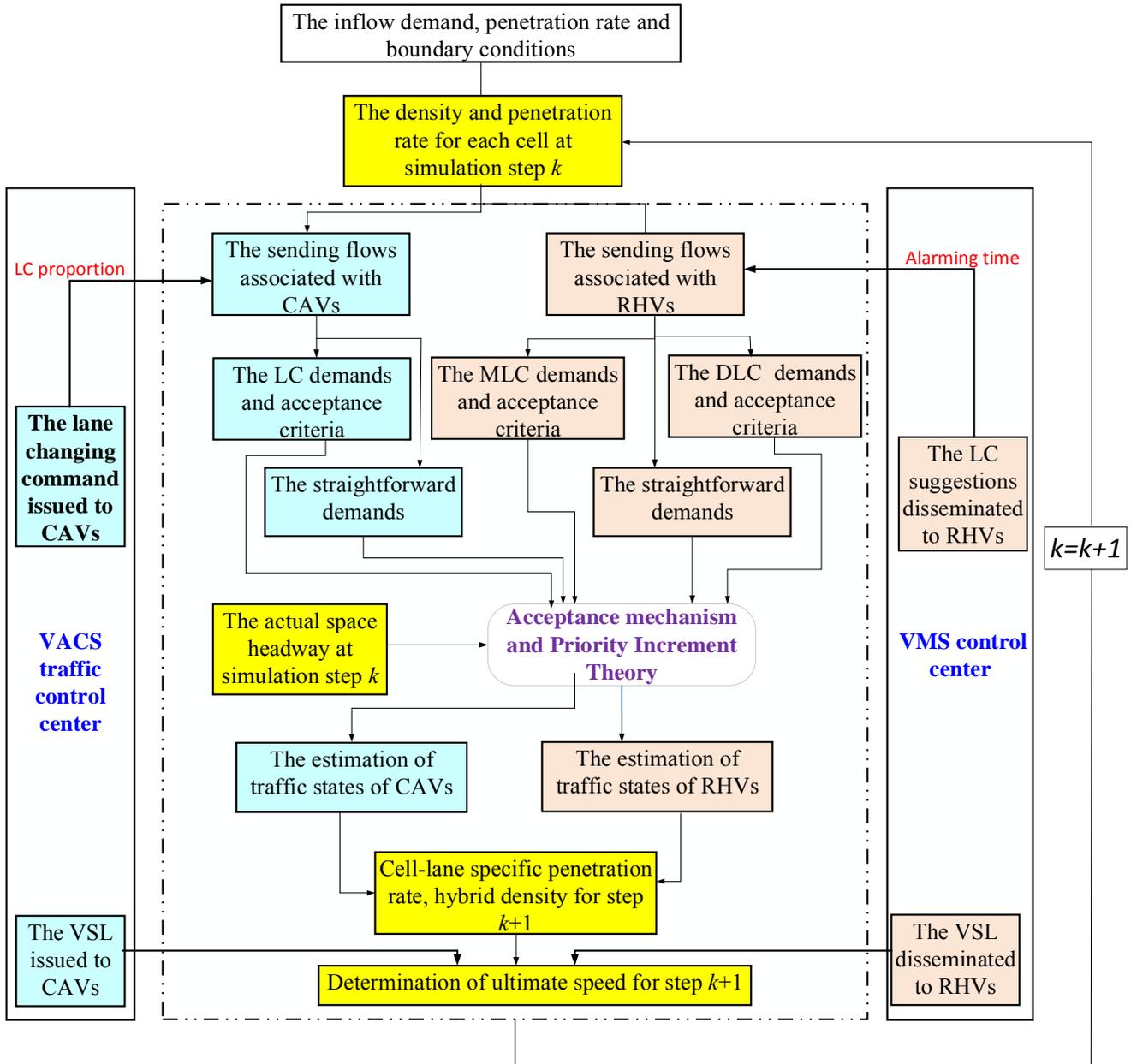


Figure 5.10. Flowchart of multiclass traffic model incorporating control strategies.

## 5.5 Solution algorithm based on cross-entropy method and control parameterization

The optimal control problem for general nonlinear systems usually adopts gradient-based

approaches to devise solution algorithms based on several assumptions, such as the system dynamics (governing the state space) is at least continuously differentiable whereas the control is continuous. It can be seen from the development of the multiclass multilane traffic flow model that the underlying system dynamics of our optimal control problem is not differentiable, and it is too complicated to evaluate the gradient of the objective function along the state trajectory and the constraint set. Even its simplest case, i.e., the CTM case, is non-differentiable (Zhong et al., 2014, Zhong et al., 2016b). On the other hand, as explained in the previous section, the control is also subject to several constraints for practical considerations, e.g., the VSL is discretized, and the control is implemented in a cycle-to-cycle manner (2 min interval while the simulation time step is 10 sec). To this end, the optimal control problem will not be solved directly but some heuristic optimization algorithms will be investigated for the equivalent reformulation of the optimal control problem concerned.

### 5.5.1 Control parameterization

To begin with, the discretization of the control (or the actions) is first presented. As previously discussed, the VSL control is discretized into finite values in 5 mile/hour increments, e.g., [5, 10, 15, 20, ..., 60, 65, 70] mile/hour for several practical issues. Indeed the lane flow distribution ratio needs to be discretized similarly. Taking a cell with a length of 1/8 miles as an example, if the density is 160 P.C.U./mile/lane (very congested traffic), the number of vehicles in the cell is 20. Although the lane flow distribution ratio can be any real number not greater than 1, a lane flow distribution ratio smaller than 5% implies a flow volume of less than one vehicle, which lacks any physical meaning. Therefore, for this case, it is wise to discretize the lane flow distribution ratio into finite values in 5% increments in this example. All the elements from the discretization of an originally larger (e.g., continuous) action space  $U$  consist of the new action space of the problem. It is assumed that this set contains  $M$  distinct actions and is denoted as  $U_d = \{u_1, \dots, u_M\}$ . It is further assumed that the state space  $X$  comprises  $D$  variables.

The control (or policy) parameterization technique is a popular numerical scheme for solving the optimal control problem. It involves approximating the control function (as a vector in certain functional space) by a linear combination of the basis functions. By doing so, the coefficients in the linear combination are decision variables to be determined instead of solving the optimal control problem directly. Applying the control parameterization technique yields a low dimensional approximation of the original optimal control problem.

For the current purpose, to enable a random search by CEM and to cover the effects missed by the control discretization, a policy parameterization is presented in line with that of Busoniu et al. (2010). In this study, the  $\aleph$  basis functions (BFs) are used below, i.e.,  $\varphi_i(x; \xi) : X \rightarrow \mathbb{R}, i = 1, \dots, \aleph$ , defined over the state space and parameterized by a vector  $\xi$  that gives the locations and shapes of the BFs. The BFs are associated with (control) actions by a many-to-one mapping, which can be represented as a vector  $\vartheta \in \{1, \dots, M\}^{\aleph}$  that associates each BF  $\varphi_i$  to a discrete (control) action index  $\vartheta_i$ , or equivalently to a discrete action  $u_{\vartheta_i}$ .

As reported by Busoniu et al. (2010), the number  $\aleph$  of BFs, in conjunction with the type of BFs, determine the accuracy of the control (or policy) approximation. Given the type of BFs, a good value of  $\aleph$  for a given problem cannot be determined theoretically in general but could be found empirically. In the literature, the use of a control parameterization technique for optimal control of a relatively small number of BFs is often sufficient to provide a good policy approximation. If prior knowledge about the complexity of optimal policy is available, one can choose a reasonable type of BF and the value of  $\aleph$  beforehand such as in our case.

There are many types of BFs, such as splines, polynomials, and radial basis functions (RBFs), which could be used. In this chapter, the Gaussian radial basis functions are

chosen to adapt to the Gaussian assumption that has been adopted in the CEM sampling process. The Gaussian RBFs are defined by

$$\varphi_i(x; \xi) = \exp\left(-\sum_{d=1}^D \left(\frac{(x_d - c_{i,d})^2}{b_{i,d}^2}\right)\right) \quad (5.34)$$

where  $D$  is the number of state variables,  $c_i = [c_{i,1}, \dots, c_{i,D}]^T$  is the  $D$ -dimensional center of the  $i^{\text{th}}$  RBF, and  $b_i = [b_{i,1}, \dots, b_{i,D}]^T$  is its width. Denoting the vector of centers by  $\mathbf{c} = [c_1^T, c_2^T, \dots, c_{\aleph}^T]^T$  and the vector of widths by  $\mathbf{b} = [b_1^T, b_2^T, \dots, b_{\aleph}^T]^T$ , the BFs parameter vector is  $\xi = [\mathbf{c}^T, \mathbf{b}^T]^T$ . In our case, all the elements of the vectors of center and width are non-negative. Using the VSL control as an example, a schematic representation of this parameterization is given in Figure 5.11. The spike of the RBF is determined by the parameters of the basis function. For example, because 20 miles/hour lies at the boundary of the speed limit control, one may use a spiked RBF to parameterize this control, whereas one may use an RBF with a high standard deviation to parameterize the VSL controls in the middle (which may yield a many-to-one mapping). For any state  $x$ , the control associated with a BF that takes the largest value at  $x$  will be chosen (Busoniu et al., 2010), i.e.,

$$h(x; \xi, \vartheta) = u_{\vartheta_{i^*}}, \text{ where } i^* = \underset{i}{\operatorname{argmax}} \varphi_i(x; \xi).$$

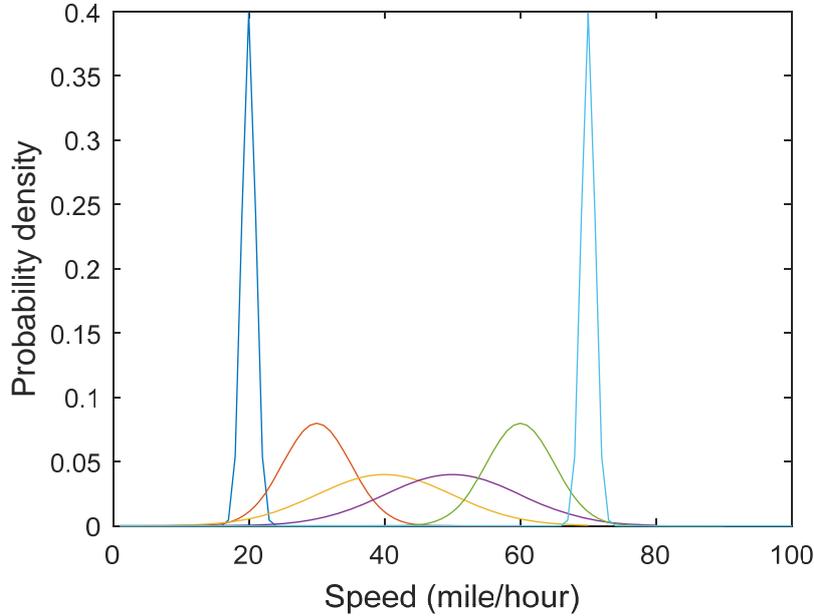


Figure 5.11. **Illustration of the policy parameterization by radial basis functions.**

### 5.5.2 Recapitulation of the cross-entropy method for optimization

Due to the complexity of the multiclass multilane traffic flow model as previously discussed, the optimization problem is with a non-differentiable complicated functional structure of the control vector and the constraint set, which would induce many local optima. Gradient-based policy search (or iteration) approaches may not be suitable choices for the optimal control problem proposed in this chapter. A gradient-free approach, i.e., a Cross-Entropy-Method-based (CEM-based) optimal control search algorithm adopted from the reinforcement learning community for learning the game of Tetris (Szita and Lorincz, 2006), is adopted for solving the optimal control problem concerned. Rather than searching the optimal control directly, this algorithm parameterizes the control (or policy) and searches for the optimal parameters that lead to maximal returns (or minimizing the objective function). Thiery and Scherrer (2009) optimized the weights with the CEM that led to one of the best publicly known controllers to the game of Tetris (Gabillon et al., 2013). Note that, after the control parameterization, the decision variables (parameters) are continuous. The CEM-based algorithm proposed in this chapter is an extension of that of Zhong et al. (2016a) for continuous optimization problem in transportation.

The Kullback–Leibler (K–L) distance/divergence (also known as cross/relative entropy) is often used to measure the dissimilarity of two probability distributions, e.g.,  $p(x)$  and  $q(x)$ , in information theory and machine learning communities. When using a distribution  $q(x)$  to approximate an unknown distribution  $p(x)$ , the K–L divergence of  $q$  from  $p$ , denoted as  $D_{KL}(p||q)$ , measures loss/change of information when  $q(x)$  is used to approximate  $p(x)$ . For discrete probability distributions  $p_k$  and  $q_k$ , the K–L divergence of  $q$  from  $p$  is defined to be

$$D_{KL}(p \parallel q) = \sum_{k=1}^K p_k \ln \frac{p_k}{q_k}.$$

For continuous distributions  $p(x)$  and  $q(x)$ , the K–L divergence is defined to be

$$D_{\text{KL}}(p \parallel q) = \int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} dx.$$

The K–L distance has two attractive features as a metric for model fidelity (Majda and Wang, 2006):

1. It is always non-negative,

$$D_{\text{KL}}(p \parallel q) \geq 0,$$

with  $D_{\text{KL}}(p \parallel q) = 0$  if and only if  $p(x) = q(x)$  almost everywhere.

2. It is invariant under general (invertible) nonlinear changes of variables.

To begin with, a recapitulation of the CEM is firstly presented for the general optimization problem (Rubinstein and Kroese, 2004). Consider a general minimization problem of the following form:

$$\gamma^* = \min_{x \in \chi} S(x), \quad (5.35)$$

where  $\gamma^*$  represents the (global) minimum of  $S(x)$  and  $x$  is defined in a certain function space  $\chi$ . Obtaining an (global) optimum solution for a general optimization problem can be regarded as a rare event, especially when the problem is nondifferentiable and nonconvex. To be specific, a rare event such as  $S(x) \leq \gamma$  is defined for the minimization of  $S(x)$  by defining the minimum  $\gamma^*$  as a threshold (or some  $\gamma \geq \gamma^*$  but sufficiently close to  $\gamma^*$ ). To this end, CEM was used to formulate a family of probability density functions (PDFs) distributed in  $\chi$ , denoted by  $f(x; v)$ , parameterized by  $v$  to apply the Monte Carlo approach to solve above rare event probability. By the above rare event analogy, it can be defined that

$$\ell(\gamma) = P_u(S(X) \leq \gamma) = E_u(I_{\{S(X) \leq \gamma\}}), \quad (5.36)$$

where  $X = (X_1, X_2, \dots, X_n)$  is a random vector generated by PDF with the parameter  $v$  set to  $u$  (i.e.,  $u$  is a realization of  $v$ ) in  $f(x; v)$ .  $P_u$  denotes the evaluation of probability.  $E_u$  denotes the expectation.  $I(\cdot)$  is the indicator function, i.e.,  $I_{\{S(X) \leq \gamma\}} = 1$ , if and only if  $S(X) \leq \gamma$  is true, 0 otherwise. The original optimization problem is then converted into a rare event probability estimation problem by CEM, and the

objective is to maximize  $\gamma$  such that  $\ell(\gamma)$  approaches 0. In other words,  $\gamma$  is the minimum of  $S(x)$  in a probabilistic sense because  $\gamma$  gives the largest value such that  $S(X)$  is greater than  $\gamma$ , i.e.,  $S(x) > \gamma$  with very high probability if  $\ell(\gamma)$  approaches 0. For a given  $\gamma$ , a practical way to estimate  $\ell(\gamma)$  is to generate some samples from  $\chi$  and perform Monte Carlo simulations. To this end,  $N$  samples are generated from  $f(x; u)$  and  $\ell(\gamma)$  is estimated as

$$\hat{\ell} = \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{x}_i) \leq \gamma\}},$$

where  $\mathbf{x}_1, \dots, \mathbf{x}_N$  denotes a random sample. This crude Monte Carlo simulation idea is computationally intensive in that it requires a huge number of samples to accurately estimate  $\ell(\gamma)$  when  $S(X) \leq \gamma$  is a rare event. To overcome the computation problem, the CEM exploits the power of the importance sampling technique, which uses a different probability density function,  $\kappa(x; \epsilon)$ , on  $\chi$  and computes the estimation of  $\ell(\gamma)$  as  $\hat{\ell}(\gamma)$  by

$$\hat{\ell}(\gamma) = \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{x}_i) \leq \gamma\}} \frac{f(\mathbf{x}_i; u)}{\kappa(\mathbf{x}_i; \epsilon)}.$$

Defining

$$\kappa^*(x, \epsilon) = \frac{I_{\{S(x) \leq \gamma\}} f(x; u)}{\ell(\gamma)},$$

and replacing  $\kappa$  by  $\kappa^*$ , results in

$$\hat{\ell}(\gamma) = \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{x}_i) \leq \gamma\}} \frac{f(\mathbf{x}_i; u)}{\kappa^*(\mathbf{x}_i; \epsilon)} = \ell(\gamma).$$

The problem now turns to the estimation of  $\kappa^*$  because  $\ell(\gamma)$  is unknown. The CEM defines the distance between two PDFs,  $\kappa(x; \epsilon)$  and  $f(x; v)$ , using K–L divergence, which is also known as cross-entropy and is defined as follows:

$$D_{KL}(\kappa||f) = E_{\kappa} \left( \ln \frac{\kappa(x; \epsilon)}{f(x; v)} \right) = \int \kappa(x; \epsilon) \ln \kappa(x; \epsilon) dx - \int \kappa(x; \epsilon) \ln f(x; v) dx,$$

where  $E_{\kappa}$  means the expectation under the probability density function  $\kappa$ , and  $D_{KL}(\kappa||f) \geq 0$  with  $D_{KL}(\kappa||f) = 0$  if and only if  $f(x) = \kappa(x)$ . Minimizing the K–L distance to approach  $\kappa^*$  is equivalent to

$$\max_v \int \kappa^*(x, \epsilon) \ln f(x; v) dx,$$

by selecting  $v$ .

Substituting the definition of  $\kappa^*$  results in

$$\max_v \int \frac{I_{\{S(x) \leq \gamma\}} f(x; u)}{\ell(\gamma)} \ln f(x; v) dx.$$

Using the definition of  $\ell(\gamma)$ , an equivalent optimization problem is

$$v^* = \operatorname{argmax}_v E_u I_{\{S(X) \leq \gamma\}} \ln f(X; v).$$

Note that the nominal PDF  $f$  is assumed to be parameterized by a finite-dimensional vector  $u$ , i.e.,  $f(x) = f(x; u)$ , and that the importance sampling PDF is  $f(\cdot; v)$  for some parameter  $v$ . Again, using importance sampling, with a change of measure  $f(\cdot; w)$  the above optimization problem is rewritten as

$$v^* = \operatorname{argmax}_v E_w I_{\{S(X) \leq \gamma\}} W(X; u, w) \ln f(X; v),$$

for any reference parameter  $w$ , where

$$W(x; u, w) = \frac{f(x; u)}{f(x; w)},$$

is the likelihood ratio at  $x$  between  $f(\cdot; u)$  and  $f(\cdot; w)$  (Rubinstein et al., 2004).

Finally,  $v^*$  is estimated by solving the following stochastic program:

$$\hat{v}^* = \operatorname{argmax}_v \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{x}_i) \leq \gamma\}} W(\mathbf{x}_i; u, w) \ln f(\mathbf{x}_i; v) \doteq \operatorname{argmax}_v \hat{D}(v),$$

where  $\mathbf{x}_1, \dots, \mathbf{x}_N$  is a random sample from  $f(\cdot; w)$ . In typical applications the function  $\hat{D}$  is convex and differentiable with respect to  $v$ . The solution may be readily obtained by solving the following system of equations (with respect to  $v$ ).

$$\frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{x}_i) \leq \gamma\}} W(\mathbf{x}_i; u, w) \nabla_v \ln f(\mathbf{x}_i; v) = 0,$$

where  $\nabla_v$  indicates the gradient with respect to  $v$ .

To sum up the above description briefly, the CEM comprises two key steps:

1. Generate trial decision variable sets randomly according to the chosen distributions, which are usually specified as Bernoulli or Uniform distributions (for discrete variable)

or Gaussian distribution (for continuous variable).

2. Update the probability distribution used to generate the random trial sets according to the principle of “importance sampling.”

The CEM algorithm for optimization is summarized as follows:

1. Set  $\hat{v}_0$  to  $u$  and initialize the iteration counter  $l$  to 1.
2. Generate a set of random samples  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$  from  $p(x; \hat{v}_{l-1})$ .

Calculate the objective function  $S(\mathbf{x}_i)$  for each sample and then determine  $\hat{\gamma}_l$  by the following equality:

$$\hat{\gamma}_l = S_{(\lceil (1-\rho)N \rceil)}, \quad (5.37)$$

where  $\rho$  is a small real number.

3. Use the same samples  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$  to estimate parameter vector  $\hat{v}_l^{new}$  by solving the equality below.

$$\hat{v}_l^{new} = \operatorname{argmax}_v \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{x}_i) \leq \hat{\gamma}_l\}} W(\mathbf{x}_i; u, \hat{v}_{l-1}) \ln f(\mathbf{x}_i; v). \quad (5.38)$$

4. Let  $0 \leq \beta \leq 1$  be a smoothing parameter, the parameter vector is updated according to the smoothed updating law.

$$\hat{v}_l = \beta \hat{v}_l^{new} + (1 - \beta) \hat{v}_{l-1}$$

5. If the standard deviation of samples generated by  $p(x; \hat{v}_l)$  is lower than  $\varepsilon$ , stop the algorithm and determine the minimum of  $S(x)$ .

$$\gamma^* = \hat{\gamma}_l. \quad (5.39)$$

Otherwise, increase the iteration counter  $l$  by 1, and return to Step 2.

### 5.5.3 Optimal parameter search based on the CEM

Through the control parameterization depicted in Section 5.1, the optimal control problem is converted to a parameter optimization problem. From the CEM optimization process for parameter optimization as described in the above section, there is a need to search the mean and variance for each decision variable. Therefore, the density of each

center  $c_{i,d}$  is parameterized by its mean  $\mu_{c_{i,d}}$  and its standard deviation  $\delta_{c_{i,d}}$ , whereas the probability density function for a width  $b_{i,d}$  is likewise parameterized by  $\mu_{b_{i,d}}$  and  $\delta_{b_{i,d}}$ . Similar to the centers and widths themselves, the vectors of the probability density function for the radial basis function (RBF) parameters comprises all of these vectors, i.e.,

$$\mathbf{v}_\xi = [(\boldsymbol{\mu}_c)^T, (\boldsymbol{\delta}_c)^T, (\boldsymbol{\mu}_b)^T, (\boldsymbol{\delta}_b)^T]^T.$$

Here, Gaussian distribution is used to generate samples for  $\mathbf{v}_\xi$  in line with Zhong et al. (2016a).

If the centers are chosen beforehand, e.g., the centers of the RBFs associated with the VSL are specified as [10, 15, 20, ..., 60, 65, 70] miles/hour, then the vector of the centers is not a decision one and

$$\mathbf{v}_\xi = [(\boldsymbol{\mu}_b)^T, (\boldsymbol{\delta}_b)^T]^T$$

The dimension of the problem can be then be reduced but at the price that the control law may be a Pareto optimum. According to Busoniu et al. (2010), the vector  $\vartheta$ , which contains the assignments of discrete actions to the BFs, is represented in binary code. Each element  $\vartheta_i$  is represented using  $N^{\text{bin}} = \lceil \log_2 M \rceil$  bits, so that  $\vartheta$  has  $\aleph N^{\text{bin}}$  bits. Every bit is drawn from a Bernoulli distribution parameterized by its mean  $\mu^{\text{bin}} \in [0, 1]$  ( $\mu^{\text{bin}}$  gives the probability of selecting 1; the probability of selecting 0 is  $1 - \mu^{\text{bin}}$ ).

The original dynamics are now approximated by incorporating the parameterized control as input.

$$x_{k+1} \approx F(x_k, h(x; \xi, \vartheta)) = \hat{F}(x_k, \xi, \vartheta).$$

The function  $F(x_k, \cdot)$  is only a compact representation of the multiclass multilane traffic flow model, whereas  $\hat{F}(x_k, \xi, \vartheta)$  is its approximation by incorporating the parameterized control as input. The original cost function is approximated by

$$J(x_k) \approx \hat{J}(x_k, \xi, \vartheta)$$

This form is used to highlight that a) the dynamics originally driven by the control input are now determined by the parameters of the parameterized control so as the cost function; b) finding an optimal control is now about finding optimal parameters that minimize the approximated cost function. For a dynamic system, the state trajectory depends on both the initial state and the control input. Therefore, the optimal control problem is discussed given an initial state or a set of initial states depending on whether the underlying process is deterministic or stochastic. Unlike the study by Busoniu et al. (2010), our multiclass multilane traffic flow model is a deterministic process; it suffices to simulate a single trajectory initialized by  $x_0$  in  $X_0$  to find the optimal control sequence associated with  $x_0$ . To this end, the score function (objective function to be minimized by the CEM) is defined as

$$\hat{J}(x_0, \xi, \vartheta) = \sum_{k=1}^{k_N} \hat{J}(x_k, \xi, \vartheta)$$

Minimizing  $\hat{J}(x_0, \xi, \vartheta)$  is a mixed-integer optimization problem to find optimal parameters  $\xi$  (which is the vector of parameters of the basis function) and  $\vartheta$  (which is the index). Now, it is appropriate to describe the CEM for policy approximation as summarized in Algorithm 1 below.

<b>Symbol</b>	<b>Meaning</b>
$N$	Number of samples
$\hat{J}(x_0, \xi, \vartheta)$	The objective function
$\xi, \vartheta$	The decision parameters
$N(;v)$	The normal distribution of model parameters with parameters $v$
$v_{\xi_0} = [\mu_{\xi_0}, \delta_{\xi_0}]$	The parameters of density with mean and standard deviation
$\rho$	The percentage of elite samples on all samples
$\alpha$	The smoothing parameter
$\varepsilon$	The prescribed threshold
$X_{im}, X_{un}$	The important/unimportant parameter combinations
$\delta$	The bandwidth
$w$	The flag of removing unimportant parameter combinations
$\varepsilon_1$	The threshold of adding unimportant parameter combinations

Table 5.5. Notation list of the CEM for parameter optimization

---

## ALGORITHM 1

The cross-entropy method for parameter optimization

---

**Input:** Scoring function  $\hat{f}(x_0, \xi, \vartheta)$ , density function  $\mathbb{N}(\cdot; \nu)$ , other parameters

1: initialize  $x_0$  and density parameters  $\nu_{\xi_0} = [\mu_{\xi_0}, \delta_{\xi_0}]$ , and  $\nu_{\vartheta_0} = [\mu_{\vartheta_0}, \delta_{\vartheta_0}]$

2.1: generate original samples  $\xi_0 = [\xi_1, \dots, \xi_N]$  from  $\mathbb{N}(\xi; \nu_{\xi_0})$  and  $\vartheta_0 = [\vartheta_1, \dots, \vartheta_N]$  from  $\mathbb{N}(\vartheta; \nu_{\vartheta_0})$ , and  $\text{Sam}_0 = [\xi_0, \vartheta_0]$

2.2: fix one parameter  $\xi^i$  or  $\vartheta^j$  to produce derived samples  $\text{Sam}_0 | \xi^i$  or  $\vartheta^j$

with  $i = 1, \dots, d$

2.3: compute  $S(x) = \hat{f}(x_0, \xi, \vartheta)$ , and use kernel density method to estimate original PDF and derived PDF:

$$\hat{g}(S(x)) = \frac{1}{N\delta} \sum_{i \in I} \frac{1}{2\pi} \left( -\frac{1}{2} \left( \frac{S(x) - \bar{S}(x_i)}{\delta} \right)^2 \right)$$

2.4: calculate the cross-entropy distance between original PDF  $f(x)$  and derived PDF  $g(x | X_i)$ , and determine the important parameter combinations  $X_{im}$

and unimportant parameter combinations  $X_{un}$

3:  $t \leftarrow 0$  and  $w \leftarrow 0$ , fix  $X_{un}$  to its constant value, remove  $\nu_{im}$  from  $\nu_{\xi_0}$

and  $\nu_{\vartheta_0}$  respectively, and initialize  $x_0$

4: repeat

5:  $t \leftarrow t + 1$

---

- 
- 6: if  $\max\{v_{\xi_{t-1}}^{\varepsilon}, v_{\vartheta_{t-1}}\} \leq \varepsilon_1$  and  $w=0$
- 7: add  $v_{un,0}$  into  $v_{\xi_0}$  and  $v_{\vartheta_0}$  respectively, and set  $w=1$
- 8.1: generate samples  $\xi_1, \dots, \xi_N$  from  $\mathbb{N}(\xi; v_{\xi_{t-1}})$  and  $\vartheta_1, \dots, \vartheta_N$   
from  $\mathbb{N}(\mathcal{G}; v_{\vartheta_{t-1}})$
- 8.2: compute  $\hat{J}(x_0, \xi_i, \vartheta_i)$  with  $i = 1, \dots, N$
- 8.3: reorder and reindex s. t.  $\hat{J}_0 \leq \dots \leq \hat{J}_N$
- 8.4:  $\gamma_t \leftarrow \hat{J}_{[(1-\rho)N]}$
- 8.5:  $\hat{v}_{\xi_t} \leftarrow [\mu_{\xi_t}, \delta_{\xi_t}^2]$  where  $u_{\xi_t} = \frac{1}{\rho N} \sum_{j=1}^{\rho N} \xi_{[j]}$ ,  $\delta_{\xi_t}^2 = \frac{1}{\rho N} \sum_{j=1}^{\rho N} (\xi_{[j]} - \mu_{\xi_t})^2$   
 $\hat{v}_{\vartheta_t} \leftarrow [\mu_{\vartheta_t}, \delta_{\vartheta_t}^2]$  where  $u_{\vartheta_t} = \frac{1}{\rho N} \sum_{j=1}^{\rho N} \vartheta_{[j]}$ ,  $\delta_{\vartheta_t}^2 = \frac{1}{\rho N} \sum_{j=1}^{\rho N} (\vartheta_{[j]} - \mu_{\vartheta_t})^2$
- 8.6:  $v_{\xi_t} = \alpha \hat{v}_{\xi_t} + (1-\alpha)v_{\xi_{t-1}}$  and  $v_{\vartheta_t} = \alpha \hat{v}_{\vartheta_t} + (1-\alpha)v_{\vartheta_{t-1}}$
- 9: until  $\max\{\delta_{\xi_t}, \delta_{\vartheta_t}\} \leq \varepsilon$  or  $t = t_{\max}$
- Output:**  $\mu_{\xi_t}$  and  $\mu_{\vartheta_t}$ , the best sample; and  $\gamma_t$ , the best score
-

## 5.6 Numerical example

### 5.6.1 Description of the “test site”

A numerical study was conducted on a virtual 2.4-mile-long freeway segment with two lanes, as shown in Figure 5.12. It is assumed that this freeway segment is far from on/off-ramps or other physical bottlenecks. Therefore, under normal traffic conditions, the only lateral movement on this segment is the DLC for achieving a speed advantage. The communication infrastructures that transmit real-time traffic data and issue lane-changing controls (LCCs) and variable speed limit controls (VSLCs) to CAVs are installed along the freeway section. It is also assumed that there is one VMS gantry installed on the upstream boundary of the segment that broadcasts lane-changing recommendations (LCRs) and variable speed limit recommendations (VSLRs) to RHVs. Assuming that the initial time for this simulation is 0, five minutes after that, cell (6, 1) is temporally closed due to a traffic incident. Only cell (6, 2) is available for traffic and a temporal bottleneck is thus formed. The lane closure lasts 20 min until the incident is removed. The simulation continues another 20 min to dissolve the queue afterward.

To meet the numerical stability requirement, i.e., the Courant–Friedrichs–Lewy (CFL) condition of the CTM, the cell length should be greater than or equal to the distance traveled during a simulation time step at the free-flow speed. Therefore, this 2.4 miles freeway link is divided into six cell packages with equal lengths of 0.4 miles. The driving behavior of RHVs is affected by the information on the VMS only until they pass the gantry. As expected, the compliance to these recommendations will gradually affect the upstream segments with the movement of RHVs.

Please note that the CAVs are always with full automatic control. **Otherwise they are regarded as RHVs.** Therefore, when “**without control**” is used, it implies that the optimal control **is only applied to the CAVs rather than to coordinate the movements of the RHVs and CAVs simultaneously.**

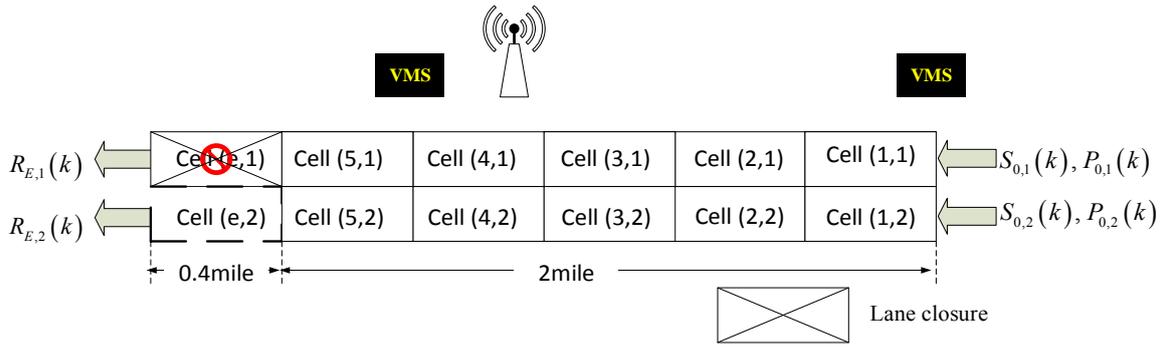


Figure 5.12. Topology of the “test site”

After the incident is detected, the proposed optimal control framework devises control strategies for traffic management. Then, the VMS gantries start to disseminate LCRs and VSLRs to RHVs. As discussed in Sections 5.4, considering the applicability, the VMS information should not vary too frequently, nor too sharp on temporal and spatial domains; otherwise, drivers will have difficulty in making decisions and taking action. In this example, the control cycle  $T_c$  is set to be 120 sec, which consists of 12 simulation time steps of 10 sec duration for each simulation time step. All vehicles are assumed to be P.C.U. 20 feet (6 meters) long, and the safe constant gap is 6.5 feet (2 meters; see the empirical study in Chapter 3). For determining the time-dependent fundamental diagram and minimum space headway acceptance criteria, the response time is set to 1.85 sec for RHVs and 0.35 sec for CAVs. Note that the response time of CAVs is assumed to vary from 0.25 sec to 1.5 sec in Levin and Boyles (2016a). The minimum headway required by CAVs is about 0.6 sec if the response time is chosen as 0.35 sec. Since the minimum headway required by CAVs would heavily affect the capacity of the freeway with mixed traffic of RHVs and CAVs, a sensitivity analysis on several typical values of the response time of CAVs will be conducted later in this section. It is assumed that the compulsory speed limit is 70 miles/hour under normal conditions. The parameters for evaluating the distribution of MLC and DLC demands are listed in Table 5.6. These parameters were calibrated and validated in Chapter 3. In this numerical study, the total delay is converted to monetary cost so that the fuel price, emission cost, and electricity cost can be included as well.

Longitudinal cumulative distribution function of MLC demand	$\alpha_1$	671
	$\alpha_2$ Weighting associated with the density of target lane	33.7
	$x_c$ Critical remaining distance	0.05 miles
DLC demand determination function	$\tau$ Average lane-changing reaction time	3 sec
Objective function	Value of travel time	24 USD/hour
	Cost of fuel (Petrol price)	0.64 USD/liter
	Emission cost of CO and HC	0.07 USD/mg
	Emission cost of NO	0.007 USD/mg
	Cost of electricity	0.12 USD/kWh
Fuel consumption matrix $K_f$ (L/P.C.U./s)	$\begin{bmatrix} -0.679439 & 0.135273 & 0.015946 & -0.001189 \\ 0.029665 & 0.004808 & -0.000020535 & 5.5409285E-8 \\ -0.000276 & 0.000083329 & 0.000000937 & -2.479644E-8 \\ 0.000001487 & -0.000061321 & 0.000000304 & -4.467234E-9 \end{bmatrix}$	
CO emission cost $K_{CO}$ (mg/P.C.U./s)	$\begin{bmatrix} 0.887447 & 0.148841 & 0.030550 & -0.001348 \\ 0.029665 & 0.004808 & -0.000020535 & 5.5409285E-8 \\ -0.000276 & 0.000083329 & 0.000000937 & -2.479644E-8 \\ 0.000001487 & -0.000061321 & 0.000000304 & -4.467234E-9 \end{bmatrix}$	
HC emission cost $K_{HC}$ (mg/P.C.U./s)	$\begin{bmatrix} -0.728042 & 0.012211 & 0.023371 & -0.000093243 \\ 0.024950 & 0.010145 & -0.000103 & 0.000000618 \\ -0.000205 & -0.000549 & 0.000037592 & -0.000000213 \\ 0.000001949 & -0.000061321 & 0.000003310 & -1.739372E-8 \end{bmatrix}$	
NO emission cost $K_{NO}$ (mg/P.C.U./s)	$\begin{bmatrix} -1.067682 & 0.254363 & 0.008866 & -0.000951 \\ 0.046423 & 0.015482 & 0.000000569 & 0.000000328 \\ -0.000173 & 0.002876 & -0.00005866 & 0.00000024 \\ 0.000000569 & -0.000321 & 0.000001943 & -1.257413E-8 \end{bmatrix}$	
Electricity cost $K_c$ (w/P.C.U./s)	$\begin{bmatrix} 0.02 & 74.4 & 3.5E-4 & 1.3 \\ 0.3426 & 1266 & 0 & 0 \end{bmatrix}$	

Table 5.6. Parameters in the traffic model and the optimization problem

## 5.6.2 Simulation results

In this numerical simulation, unless otherwise specified, optimal control was conducted on a dual-lane freeway segment with traffic flow mixed with 60% CAVs. To identify the contribution of each control strategy, i.e., LCRs and VSLRs for RHVs and LCCs and

VSLCs for CAVs, different combinations of these control strategies were investigated for alleviating the incident impact under heavy and light traffic conditions.

### 5.6.2.1 Optimal control for heavy traffic conditions

The inflows that intend to enter cell (1,1) and cell (1,2) are both set to 1500 P.C.U./hour/lane. Simulations are conducted for the following four cases:

- **Case 1: Baseline (without control).** No recommendations are disseminated to RHVs.

The CAVs traveling in the incident lane would follow the non-optimized LCCs, such as remaining in the current lane until the remaining distance to the incident spot is less than 0.2 miles.

- **Case 2:** Optimal VSLRs and VSLCs are disseminated to RHVs and CAVs, respectively.

Considering the full and immediate response of CAVs to VACS commands, the RHVs are assumed to fully comply to the VSLC. Also, non-optimized LCCs are still issued to CAVs, which are getting very close to the lane closure location.

- **Case 3:** Broadcast LCRs as long as the incident is identified until the incident is removed. Meanwhile, optimal LCCs are disseminated to the CAVs.

- **Case 4:** Finally, the full combination of optimal control strategies is integrated for implementation on both RHVs and CAVs.

Figure 5.13 demonstrates the spatial-temporal distribution of traffic density and speed after implementing the control strategy proposed in Case 1. The  $z$  value in Figure 5.13 presents lane-specific cell density; the color of the surface denotes the corresponding speed. As can be observed, the queue on lane 1 quickly spills upstream after the lane closure caused by the incident at 5 min. In view of the speed advantage in the adjacent lane, the RHVs intend to switch to it via MLC maneuvers (as they do not know about the incident until they reach the incident spot or join the queue), whereas the CAVs comply with the LCCs in the MLC mode. The lane-changing traffic of RHVs is traveling at a low

speed during the congested period. The lane-changing maneuvers under this congested situation dramatically reduce the speed in lane 2 and increase its density, causing congestion in lane 2. This speed reduction caused by the lane-changing maneuvers further induces the **capacity drop**, i.e., the capacity at the bottleneck is much lower than the theoretical value,<sup>26</sup> especially when the lane-changing flow ratio is large.

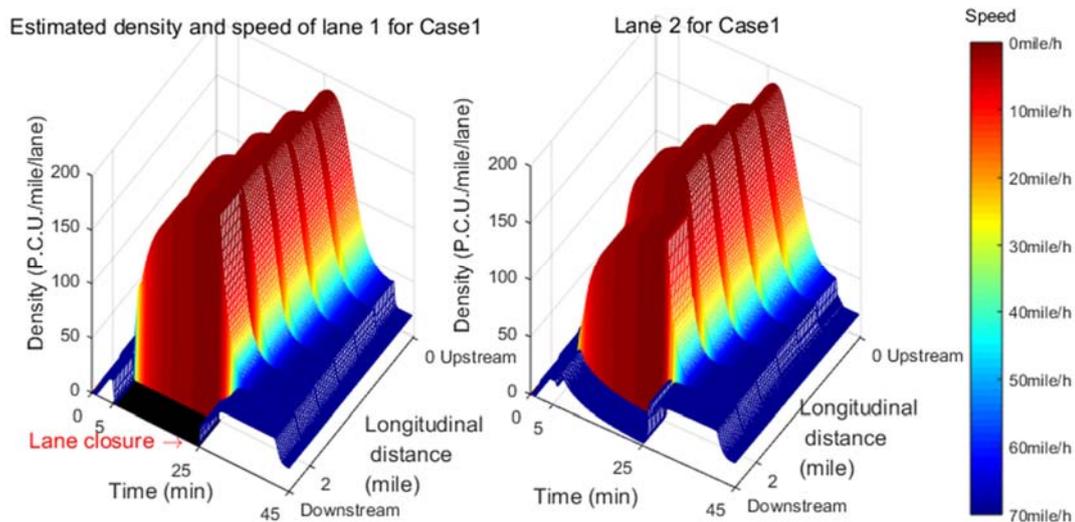


Figure 5.13 The estimated lane-specific density and speed of Case 1.

Vehicles would have to slow down considerably to seek lane-changing opportunities if they approach the incident location or the end of the queue without being aware that their lane is blocked. Indeed, this would not only cause a capacity drop but also **stop-to-go oscillation in traffic speed**. To see this, a simulation was conducted by assuming pure RHVs (the inflow rate is assumed to be 1500 P.C.U./hour/lane for both lanes) traveling on the freeway segment. Without information on the traffic incident, vehicles traveling in the incident lane would change lanes only when they reach the incident spot or arrive at the end of the queue. They have to slow down considerably, in either case, to change

<sup>26</sup> The capacity for 60% penetrated mixed traffic is as high as 3400 P.C.U./hour/lane, as estimated in Section 5.3. Ideally, if there is no interruption from lane-changing flow, the single lane capacity is enough for the total 3000 P.C.U./hour inflow demand of the two lanes.

to the open lane. When the low-speed vehicle changes to the open lane, vehicles traveling in the open lane must slow down to adapt to the speed of the lane-changing vehicle to guarantee traffic safety, as shown in Figure 5.14. The speed would increase when the lane-changing vehicles increase their speed. This equilibrium would be broken again when there are new lane changes from low-speed vehicles. On the other hand, there is the so-called “friction effect” from the empirical observation that drivers’ fear of moving fast in open lanes when an incident or slowly moving vehicles exist in adjacent lanes.

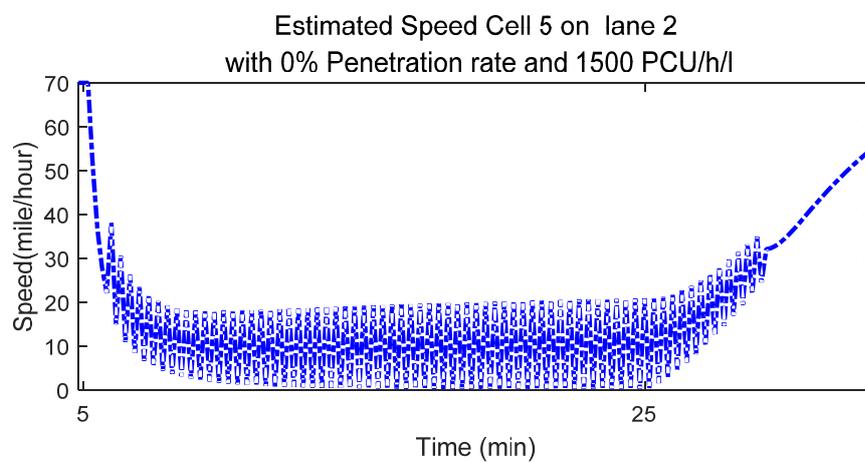


Figure 5.14. Stop-to-go oscillation in traffic speed of Case 1.

The congestion might be alleviated via two methods: 1) Reduce the traffic arriving at the bottleneck, i.e., via VSL, or 2) adjust the longitudinal distribution of the lane-changing demand along the horizontal distance of the freeway segment, i.e., via LCCs and LCRs. Table 5.7 presents the improvement of the implementation by using the strategies mentioned above obtained from the proposed optimal control framework. As shown in Table 5.7, with the control scheme outlined in Case 2, the total delay is slightly reduced compared with Case 1. The improvement is not significant because no lane-changing control is implemented. The RHVs would still travel to the end of the queue to seek a possible lane change. Such lane-changing maneuvers at low speed would reduce the traffic speed and capacity of the adjacent lane. On the other hand, it can be observed from the left figure of Figure 5.15 that the optimal value of the VSL hits a boundary (i.e., 10

miles/hour) during the congestion period. This is because the over-saturated traffic, in conjunction with the significant low speed (forced) lane-changing traffic, would indeed reduce the traffic speed while causing stop-to-go oscillation. To prevent this, the optimal control coordinates the speed to that of the RHVs. This indicates that under over-saturated traffic conditions, using VSL control only cannot prevent traffic breakdowns in a satisfactory manner. A similar result was observed in Zhang and Ioannou (2017), in which forced lane changes near the incident were the major cause of the capacity drop. Without combining the lane-changing control, once the low speed (forced) lane-changing maneuvers bring down the speed of traffic in adjacent lanes, the capacity drop takes place consequently. As such, the VSL control will have limited or no effect in improving travel time under this circumstance. This is because there is unlike for the VSL control technique to eliminate the capacity drop caused by the low-speed lane-changing traffic as observed in the right figure of Figure 5.15.

Case	Inflow					
	1500 P.C.U./hour/lane					
	Total value of travel time associated with lane 1 (USD)	Total value of travel time associated with lane 2 (USD)	Fuel consumption and emission cost of (40%) RHVs (USD)	Electricity cost of (60%) CAVs (USD)	Objective function (USD) and percentage of improvement	Total throughput (P.C.U.)
1. Baseline	2962	2758	292	148	<b>6154 (0%)</b>	2217
2. Optimal VSLR +VSLC	2850	2647	278	118	<b>5893 (-4%)</b>	2217
3. LCR+ optimal LCC	1658	2285	222	146	<b>4311 (-30%)</b>	2217
4. LCR+ optimal LCC+ VSLR+ VSLC	1346	2406	212	147	<b>4070 (-33.9%)</b>	2217

Table 5.7 Optimization results for heavy traffic demand

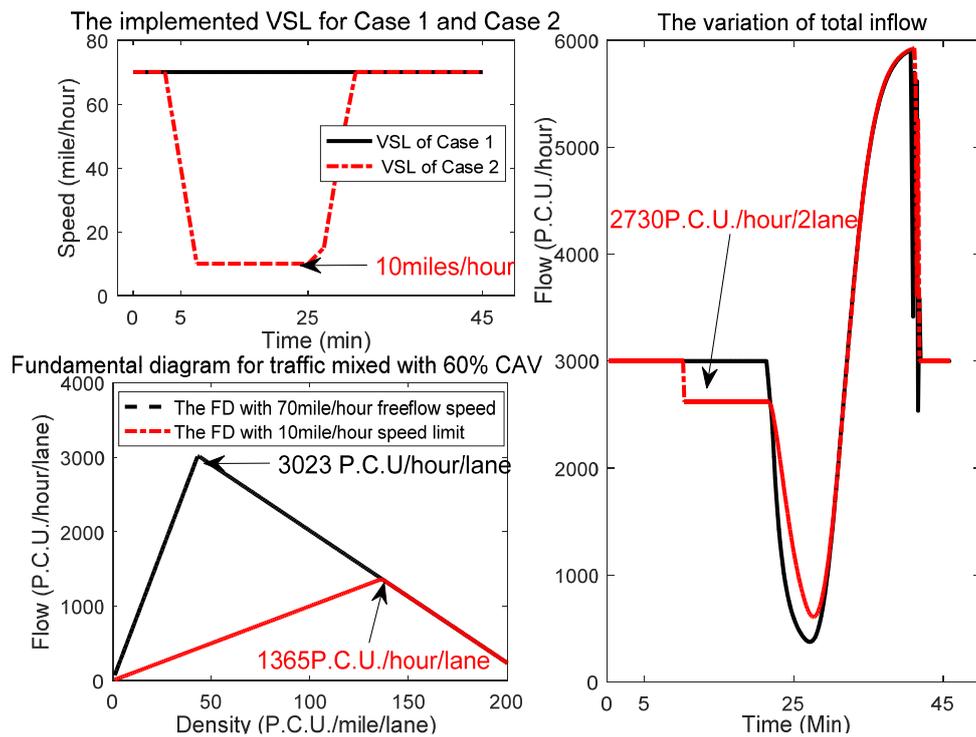


Figure 5.15. VSL and actual inflow rate.

From the observations in Case 2, Case 3 implemented lane-changing control to alleviate the incident impact. In this case, RHVs receive LCRs to change lanes from the location

where the VMS gantry is installed; meanwhile, CAVs receive and execute LCCs from the VACS. According to the results reported in Table 5.7, the total delay was significantly reduced by this control strategy. By disseminating appropriate LCRs and LCCs to upstream RHVs and CAVs, respectively, most of the lane-changing maneuvers are executed far away from the incident location where the traffic speed is still high. **Therefore, the capacity drop is dramatically reduced.** These results are also consistent with those reported in the literature (see, e.g., Zhang and Ioannou (2017) and the references therein). Compared with the total delay in lane 1, the total delay in lane 2 is larger, because most of the vehicles changed to lane 2 at locations far away from the incident spot and then traveled along that lane. In contrast, in Cases 1 and 2, vehicles travel until they reach the end of the queue in lane 1 and then slow down looking for a chance to execute a (forced) lane change. Both the volume and the average travel time in lane 1 of Cases 1 and 2 are large, which obviously results in a large total delay. This (and also the empirical study in Chapter 3 and the numerical simulations) proves that lane changing could smooth out differences between the lanes if properly controlled, which makes traffic dynamics continuous from a mathematical point of view.

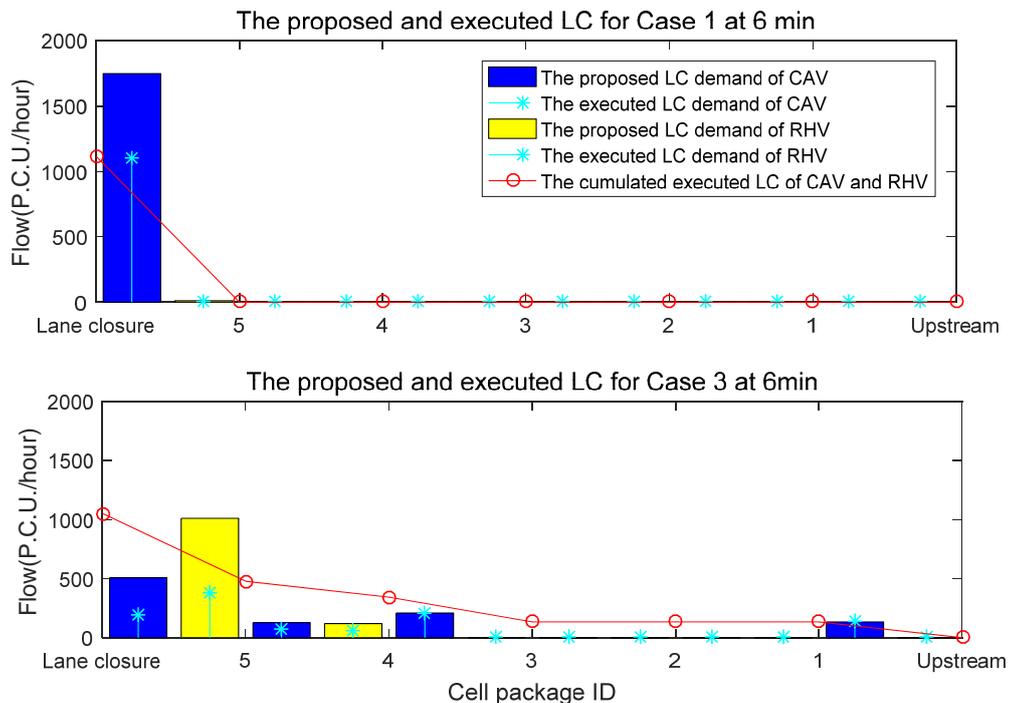


Figure 5.16. The proposed and executed lane changes associated with RHVs and CAVs.

Figure 5.16 demonstrates the proposed and executed lane changes from lane 1 to lane 2 six minutes immediately after the traffic incident is identified by the VACS system. Under Case 1, the baseline case, only CAVs are instructed to switch to lane 2 at the downstream boundary of Cell 5 by the VACS system, whereas the RHVs remain traveling in lane 1 until they reach the incident spot. In Case 3, in contrast, the lane-changing maneuvers are frequent, even quite a way upstream of the segment, for both RHVs and CAVs. The RHVs trigger the MLC maneuvers, as suggested by the VMS. However, there are still several lane changes close to the incident spot. This is because the incident just happened and the queue had not yet formed at the sixth minute.

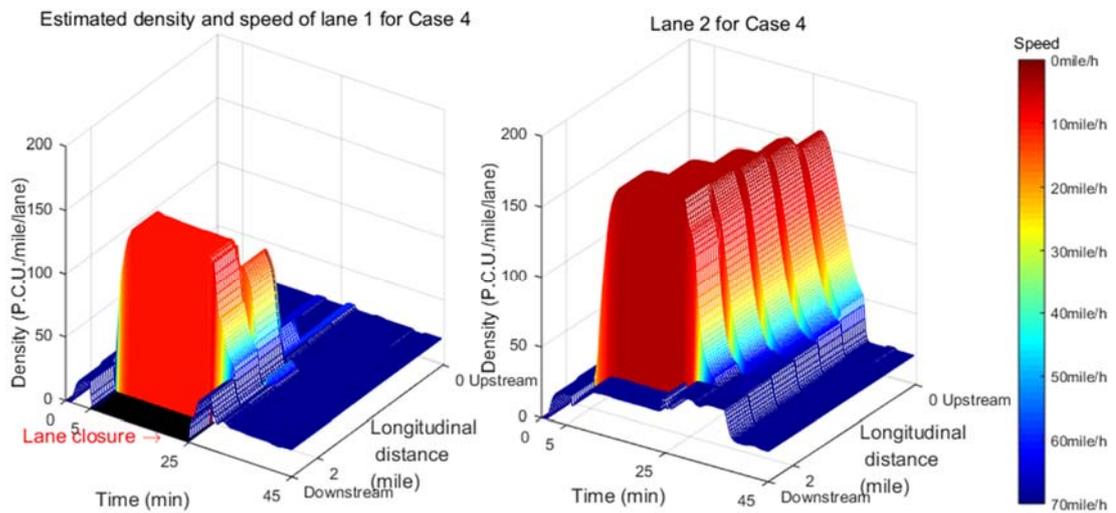


Figure 5.17. The estimated lane specific density and speed of Case 4.

Figure 5.17 presents the simulation results of traffic states implementing the integrated optimal control strategies, i.e., Case 4. Compared with Case 1, the traffic flow in lane 1 in Case 4 has a much lower level of congestion regarding both the congested longitudinal distance and the time span that the congestion lasts. This is because, similar to the results in Case 3, most vehicles change their lane far away from the incident spot because of the LCRs/LCCs implemented. Compared with Case 3, the performance further improves. These observations prove that LCRs/LCCs can **reduce the number of vehicles queuing at the bottleneck** (the point at which they are forced to change lanes at low speed) and

decrease the traffic density. From a mathematical point of view, **LCRs/LCCs make the system continuous and easier for the VSL controller to stabilize**. For traffic safety, the integration of VSL and lane-changing controls **can dramatically decrease the average number of stops**, therefore drastically reducing the instances of stop-and-go traffic, **smoothing the traffic flow, and damping the shockwaves**.

### 5.6.3 Sensitivity analysis

As previously mentioned, the penetration rate of the CAVs significantly affects the freeway capacity and throughput. Sensitivity analysis is conducted to demonstrate the performance of the optimal control against the penetration rate of CAVs, the level of congestion and the response time of CAVs.

In the first test, the performance of the integrated optimal control is investigated under both heavy traffic and light traffic scenarios with respect to different penetration rates of the CAVs. As shown in Section 5.3, the capacity is positively related with the penetration rate of CAVs, and thus the definitions of heavy traffic and light traffic conditions are differentiated with respect to different penetration rates of the CAVs, as indicated in Table 5.7. Under a pure RHV case, the integrated optimal control for this case, i.e., **VSLR +LCR**, reduced the cost for each vehicle under light traffic and heavy traffic conditions by 23.8% and 27.9%, respectively. The percentage improvement achieved under heavy traffic conditions was higher than that under light traffic conditions. This is because traffic is more disordered under heavy traffic conditions and the capacity drop is more likely to happen under this scenario. Implementing proper control can alleviate the congestion and prevent the capacity drop. From the absolute value of improvement per vehicle, the improvement under light traffic conditions was 0.524 USD, whereas it was 0.88 USD under heavy traffic conditions. Similar results were observed for the case of a 60% penetration rate of CAVs. Detailed analysis of this case is given in the previous section.

It is worthwhile to point out that, even though the inflow demand is the same, the cost incurred per vehicle under this setting, i.e., 1.39 USD, was much lower than that of the pure RHVs, i.e., 2.27 USD. When the penetration rate increased to 80%, the percentage improvement seems less significant than that achieved by the 60% penetration rate. The reason for this is that when the penetration rate increased to 80%, there were more CAVs, which are basically robots that fully comply with the optimal control. The improvement is achieved by coordinating the movements of RHVs and CAVs. When there are fewer RHVs, the relative gain from this coordination is also relatively smaller. However, the cost per vehicle under an 80% penetration rate is 1.46 USD (1500 P.C.U./hour/lane) and 1.62 USD (2000 P.C.U./hour/lane) under light traffic and heavy traffic conditions, respectively. Note that the costs under a 60% penetration rate are 1.39 USD (900 P.C.U./hour/lane) and 1.84 USD (1500 P.C.U./hour/lane) under light traffic and heavy traffic conditions, respectively. Comparing these two cases, it can be seen that CAVs can significantly improve traffic efficiency, even under denser traffic conditions. The freeway throughput is thus increased with respect to the penetration rate.

Penetration rate	0% (All RHVs)		60%		80%	
Inflow rate (P.C.U./hour/lane)	500	900	900	<b>1500</b>	1500	2000
Total cost (USD) / Total throughput (P.C.U.)	Baseline (no control)		LCC only		LCC only	
	1608/ 731	4135 / 1312	2923/ 1330	<b>6154/ 2217</b>	4433/ 2217	6348/ 2948
Total cost (USD)/ Total throughput (P.C.U.)	VSLR +LCR only		Integrated control strategies		Integrated control strategies	
	1225/ 731	2980/ 1312	1850/ 1330	<b>4070/ 2217</b>	3252/ 2217	4779/ 2948
Improvement for individual vehicle	23.8%	27.9%	20.6%	33.9%	25.0%	24.7%

Table 5.8 Sensitivity analysis on penetration rate and inflow demand (level of congestion).

Next, the average cost against the penetration rate is depicted in Figure 5.18 by assuming the inflow rate to be 900 P.C.U./hour/lane. It is found that the cost of each vehicle is

around 3.2 USD (the red circle) for the baseline case when the traffic consists of RHVs only. By applying the integrated traffic control, the cost decreases to 2.27 USD. This cost is a monotonic decrease with respect to the penetration rate of the CAVs. The minimum cost, i.e., 0.77 USD, is achieved when all vehicles are CAVs.

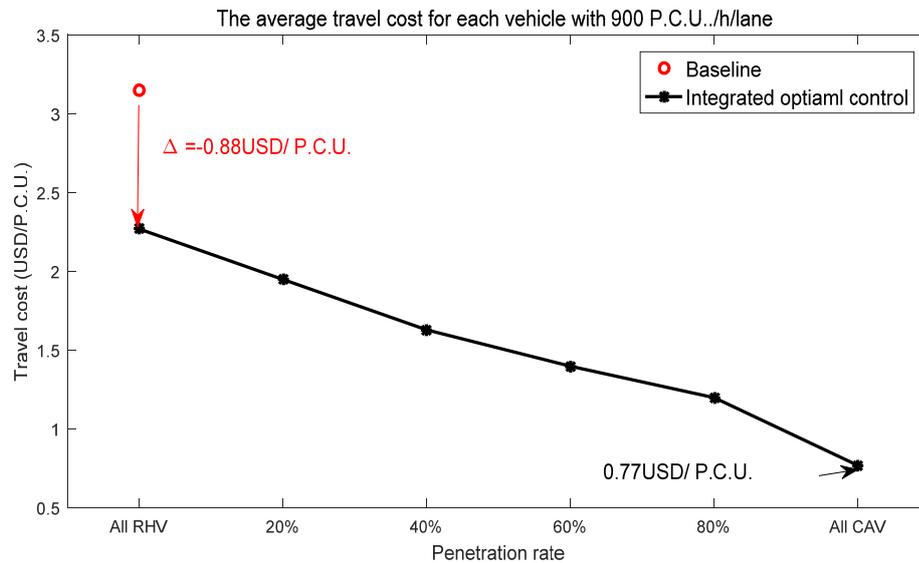


Figure 5.18. The average traveling cost with respect to the penetration rate of CAVs.

The response time of CAVs is also an important factor that influences the fundamental diagram. Thus, a sensitivity analysis on the response time of CAVs was conducted to see how it affected the performance of the integrated optimal control. This analysis was tested by assuming the traffic was mixed with 60% CAVs; inflow demand was set to be 900 P.C.U./hour/lane. As indicated by Table 5.9, the total cost increased with respect to the response time of the CAVs. This is not surprising because the response time determines the minimum headway and thus the capacity of the single-lane traffic. A higher response time of CAVs implies a larger minimum headway and thus a smaller capacity. The fundamental diagram would be more similar to that of pure RHVs when the response time of CAVs is getting closer to the response time of RHVs. Similar conclusions can be drawn for the freeway performance.

Response time of CAVs	0.35 sec	0.50 sec	1 sec
Total cost (USD) of the integrated optimal control/total throughput (P.C.U.)	4069/ 2217	4365/ 2217	5240/ 2087

Table 5.9 Sensitivity analysis with respect to the response time of CAVs

## 5.7 Conclusions

This chapter proposed an integrated optimal control framework for improving the performance efficiency of freeway traffic mixed with CAVs and RHVs. To furnish the controller design, a new multiclass multilane traffic flow model was proposed to simulate traffic flow dynamics mixed with CAVs and RHVs. This new multiclass multilane CTM-based model explicitly considered the effect of variations in cell capacity and backward wave speed in response to vehicle-class proportions within each cell. The impact of the implemented speed limit control on the fundamental diagram was also explicitly considered. Different safety gap acceptance criteria were proposed to consider various lane-changing maneuvers adopted for different vehicle types (RHVs or CAVs) with different lane-changing intentions to guarantee traffic safety and lane-changing priorities.

An optimal integrated control was then established using the proposed multiclass multilane traffic model as the network loading model. The control actions included VSL control, lane-changing control, and minimum safety gap control (implicitly considered in the flow model). The controls were actuated through the VACS for the CAVs and the en-route VMS for the RHVs. The CAVs were assumed to comply with the control commands through the VACS fully. In contrast, RHVs that were not connected to the VACS were expected to make decisions in response to the incident alarm and VSLRs given by the en-route VMSs).

Because of the complexity of the multiclass multilane traffic model, the optimization

problem is with the non-differentiable complicated functional structure of the control vector, as well as the constraint set that would induce many local optima. A gradient-free algorithm, which is a CEM-based reinforcement learning policy search approach, was developed to solve the above optimal control problem. Instead of searching for the value function by iteration, this gradient-free algorithm first parameterized the control (or policy) and then searched for the optimal parameters that would lead to the maximal returns (or minimize the objective function). It was found that the LCCs were able to reduce the number of vehicles queuing at the bottleneck and decrease the traffic density on the freeway. From a mathematical point of view, LCCs can make the freeway traffic system continuous and easier for the VSL controller to stabilize. In terms of road traffic safety, the integration of VSL and LCC can drastically reduce the instances of the stop-and-go traffic, smooth the traffic flow, and suppress the impact of the shockwaves on the freeway sections concerned. On the other hand, the dynamic proportion of CAVs on a short freeway segment would significantly affect its throughput and traffic speed. The effect of the penetration rate of CAVs was investigated by sensitivity analysis. The advantage of CAVs in reducing average cost is highlighted.

## **Chapter 6 Summary and recommendation**

### **6.1 Summary of the key findings**

One of the primary objectives of this dissertation was to develop multiclass multilane traffic flow models to support real-time simulation for various active traffic management (ATM) applications on freeway traffic control. The new models proposed in this dissertation capture several vital aspects of multilane traffic on expressways. Firstly, the discretionary lane-changing (DLC) and mandatory lane-changing (MLC) maneuvers were simultaneously modeled using lane-specific fundamental diagrams. Secondly, capacity drop and random elements caused by demand and supply uncertainties were taken into account. Thirdly, the emerging multiclass traffic mixed with regular human-piloted vehicles (RHVs) and connected automated vehicles (CAVs) were described. Integrated traffic control strategies were optimized in terms of various ATM schemes such as variable speed limit (VSL), lane-changing control (LCC) and minimum safety gap control for improving the freeway traffic operation in the presence of mixed RHVs and CAVs.

Freeway bottlenecks usually happen at the weaving sections and interchanges due to the complex traffic features mainly induced by vehicle lane-changing maneuvers. Such complex traffic features not only reduce the efficiency of the freeway but also affect the traffic safety particularly at the weaving sections and interchanges of the freeway. Without properly modeling the vehicle lane-changing maneuvers, it is difficult to optimize different traffic control strategies or ATM schemes to manage traffic flow on freeway for alleviating the bottleneck problems. In traffic flow theory, modeling lane-changing maneuvers is essential to capture several important characteristics of multilane traffic flow on the freeway, e.g., heterogeneous traffic flow distribution, capacity drop, and flow balancing effects. However, little attention has been given to

modeling DLC and MLC maneuvers simultaneously at a macroscopic level in conjunction with model validation from real-world data (Zheng, 2014).

Remarkably, most of the existing approaches for simulating vehicle lane-changing maneuvers typically differentiate DLC and MLC while different traffic models have been developed for considering these two lane-changing behaviors separately. Hence, a comprehensive traffic model was proposed in this dissertation for simulating vehicle lane-changing maneuvers simultaneously to assess the impacts of MLC and DLC maneuvers on the surrounding traffic explicitly. The proposed model could balance the trade-off between maximizing the model's predictive and explanatory power and minimizing the model's complexity for applications of various ATM applications simultaneously in practice.

To furnish an effective and efficient tool for optimizing different ATM schemes, a novel macroscopic multilane traffic model was proposed in Chapter 3 to simulate the effect of MLC and DLC maneuvers on freeways. In the proposed multilane traffic model, the lane-based fundamental diagrams were introduced to capture the relationship between traffic speed and flow per lane. Rules describing different lane-changing motivations and the corresponding levels of urgency were developed to identify the priority levels of different lane-changing maneuvers. A recursive lane-changing demand estimation algorithm that aims to minimize the differences of the lane flow distributions estimated by the proposed model and the measurement counterparts was devised to adjust the lane-changing demand on the basis of the traffic counts detected at boundaries of the freeway segments. Flow propagations of both the MLC and DLC maneuvers were then calculated by the demand-supply reaction laws based on the extended IT and PIT principles.

The main purpose of the proposed multilane traffic model is to provide a reliable and implementable simulation tool for active traffic management on the freeway to alleviate

the bottleneck problems at the freeway weaving sections nearby the interchanges. Therefore, it is essential that the proposed model can be calibrated and validated using traffic data available to most of the traffic management centers rather than requiring high-resolution traffic data collected by further expensive infrastructures.

To examine this, traffic data on a complex weaving section of the SR241 freeway in Orange County, California was downloaded from the PeMS data management system and was used to calibrate and validate the proposed macroscopic multilane cell transmission model. The results indicated that the proposed model can capture the impacts of different lane-changing maneuvers on the temporal and spatial traffic state, especially the lateral lane flow distribution and the queuing effects on the longitudinal dimension in conjunction with congestion spreading to adjacent lanes on the freeway segments concerned. With the use of the MLC demand estimation algorithm and the lane-changing probability distribution function on longitudinal dimension, it was found that the resulting estimations satisfactorily fit their measurement counterparts. **This promising finding implies that the proposed model can also be used to infer the turning ratios (e.g., off-ramp traffic demand) at ramps of the freeway, as well as the dynamic traffic demand by Origin-Destination (OD) for the freeway system.**

With the same data source, the proposed model outperforms the traditional CTM with respect to the estimation accuracy. Meanwhile, some important impacts of lane-changing such as capacity drop and flow balancing effect of DLC were also investigated in an empirical study. The empirical findings are essentially in agreement with the previous findings in the literature. As shown by the empirical study in this dissertation, the proposed models do not require high-resolution traffic data but only use the traffic data available to most of the freeway traffic management centers. Such a parsimonious data requirement can be regarded as an important extension of the existing models for applications in reality.

Another salient feature of the proposed model is that drivers are assumed to use the traffic condition spatially ahead that could be perceived by them to make their lane-changing decisions. The effect of the perceived traffic condition decreases exponentially with respect to the distance away from her/his current position. Such an exponential perception assumption results in an exponential probabilistic distribution function of the cumulative MLC. Unlike some of the existing models with the assumption that drivers do not perceive the downstream traffic condition to make their lane changing decisions, which is regarded as unrealistic by Shiomi et al. (2015), the proposed model assumes that drivers would perceive the traffic condition spatially ahead to make their lane changing decisions. In the empirical study of this dissertation, the exponential probabilistic distribution function of the cumulative MLC was calibrated and validated by using the real-world data.

In practice, freeways are always subject to traffic demand and supply uncertainties, particularly on congested segments with weaving maneuvers. The measured traffic data on multilane highway is polluted with noisy. Meanwhile, the driving styles and risk-taking capabilities vary with drivers especially under crash-prone conditions, over-saturated traffic, and adverse weather conditions. Developing a multilane traffic flow model to consider the random elements is of importance (Sumalee et al., 2011; Zhong et al., 2013). In Chapter 4, the multilane hybrid (MH) theory (also known as the multilane hybrid CTM) was extended to incorporate the lane-based fundamental diagram and the traffic speed-density relationship to assimilate lane traffic speeds and to take account the stochastic elements.

As it has been shown in the literature, the triangular fundamental diagram adopted in the conventional CTM does not perform well for traffic speed data assimilation with noisy, especially under free-flow traffic states. To avoid the disadvantages of the triangular fundamental diagram, the traffic speed-density relationship proposed by Del Castillo and Bentez (1995a) and modified by Jin (2010a) for Lighthill-Whitham-

Richard (LWR) model with consideration of the lane-changing effect was adopted for the traffic speed assimilation. The lane-changing ratios are defined according to the lane speed heterogeneity following the approach adopted by Laval and Daganzo (2005). The receiving functions are assumed to be polluted by the stochastic elements and thus are random variables. From the definition of discretionary lane-changing (DLC) demand, one needs to evaluate cell traffic speed to proceed the calculation of DLC demand for each cell. However, this may not be possible since it is not realistic to obtain all cell speeds from the conventional point detectors. In the approach proposed in this dissertation, for those cell speeds that are not directly available from measurement, one can interpolate the cell speeds by the fundamental traffic speed-density relationship.

Link (cell)-node junction models were introduced to propagate the traffic dynamics with lane changes in line with Zhong et al. (2013). The lane-changing ratios can be used to define the virtual node splitting ratios which are propagated by the IT and PIT principles. Following the operational modes in the SCTM, random events with different probabilities of occurrence are defined to govern the traffic demand-supply reactions. The “actual” flow received by the downstream lane-cell is then a finite mixture of these random events. Flow propagations of both mandatory and discretionary lane-changing maneuvers are calculated with the use of the traffic demand-supply reaction laws based on the stochastic extensions of the IT and PIT principles in line with Zhong et al. (2013). With this approach, the lane-changing flows can be defined by using the measurements of boundary variables and the average execution time for changing lane only.

The MLC model and lane-changing demand estimation algorithm developed in Chapter 3 is too complicated to be extended for considering the stochastic elements. Note from the literature and the empirical study in Chapter 3 that the capacity drop is one of the major consequences of MLC maneuvers. On the other hand, using capacity drop models is a common approach to simulate the impact of MLC on the freeway traffic flow, see Kontorinaki et al. (2016) for an overview. Therefore, the capacity drop model by

Leclercq (2011, 2016b) was adopted to simulate MLC maneuvers to simplify the stochastic lane-changing model. On the other hand, the conventional SCTM did capture the random fluctuations in the fundamental diagram with emphasis on the random capacity. Using the capacity drop model would facilitate the development of the SCTM framework proposed in this dissertation for simulating the impacts of MLC on the freeway traffic flow.

For practical applications, similar to the multilane traffic flow model developed in Chapter 3, the model proposed in Chapter 4 does not require high-resolution traffic data and even admits fewer parameters to be calibrated than the multilane traffic flow model proposed in Chapter 3. Distinguished from the conventional SCTM, the new model proposed in Chapter 4 features that: i) the traffic states are given in terms of cell and lane based; ii) cell-lane changing probabilities are augmented as additional states to define the node splitting ratios; iii) additional process is required to resolve the cell-lane traffic speed from traffic density estimation. In this dissertation, the link (cell)-node junction formulation has however been integrated with the multilane SCTM and the conventional SCTM into a unified framework. Numerical simulation results support that the proposed multilane SCTM can produce several essential properties of the lane-changing traffic flow on the freeway with uncertainties.

The models proposed in Chapter 3 and 4 are developed to model typical freeway traffic flows with cars only. For heterogeneous traffic flows with large vehicles, such as trucks, it is assumed that these larger vehicles can be converted to the equivalent passenger car unit (P.C.U.). Several important human factors such as imperfect driving, estimation error in spatial and temporal anticipation and perceptual threshold have yet been considered in the proposed models. Nevertheless, applications of the proposed models to generalized road traffic networks with different types of traffic, such as bikes, motorcycles, buses, and emergency vehicles, etc. are beyond the scope of this dissertation.

Given the advantages and a promising market prospect of the Vehicle Automation and Communication Systems (VACS) and Connected Automated Vehicles (CAVs) in the forthcoming decades, there has been an enormous effort by the automobile industry and research institutions towards the development, testing, and deployment of VACS. As it can be expected, the number of CAVs equipped with VACS will be rapidly increasing in the coming decade. Meanwhile, the regular human-piloted vehicles (RHVs) should still play the major role in the market in the short term (Levin et al., 2016a). Therefore, it will be very likely that the road network is to be shared by CAVs and RHVs in the near future. The penetration of CAVs and VACS may lead to improvements in freeway network performance and traffic flow efficiency.

To design the freeway traffic control strategies or ATM schemes for improving the performance efficiency of freeways with different market penetration of CAVs, a new multiclass multilane traffic flow model was proposed in Chapter 5 to simulate traffic flow dynamics mixed with CAVs and RHVs. This new multiclass multilane CTM-based model explicitly considers the effect of variations in cell capacity and backward wave speed in response to vehicle-class proportions within each cell by its fundamental diagram. The impact of the implemented speed limit control on the fundamental diagram is also explicitly considered in the new model proposed in Chapter 5. Meanwhile, different safety gap acceptance criteria were proposed to consider various lane-changing maneuvers adopted for the two different vehicle types (i.e., RHVs or CAVs) with different lane-changing intentions to guarantee the traffic safety and lane-changing priorities.

Based on the multiclass multilane traffic model proposed in Chapter 5, an optimal control framework was further devised for improving the efficiency of freeway traffic mixed with RHVs and CAVs using the VACS and the en-route Variable Message Signs (VMS). The objective this optimal control framework is to devise an integrated design of several traffic control strategies such as variable speed limit (VSL) and lane-changing control

(LCC) to minimize the generalized cost for efficiency and greenness on the dual-two lane freeway sections under various traffic conditions. The minimum gap control for safety is implicitly modeled by the traffic flow model which is regarded as a constraint in the optimization problem concerned. The CAVs are assumed to follow full compliance with the control commands through the VACS. In contrast, the RHVs that are not connected to the VACS would make decisions in response to the incident alarm and VSL given by the en-route Variable Message Signs (VMS). It was found that the LCC would balance the ratio of mandatory vs. discretionary lane-changing maneuvers while the VSL could improve the lane-changing efficiency and safety by adjusting the minimum acceptance gap. Advice, such as optimal mandatory lane-changing location and VSL, are disseminated upstream such that drivers can make response to a situation that they cannot yet perceive to allow tactical maneuvers.

It is very interesting and important to examine the trade-off between the model complexity and the computational efficiency for practical applications of the proposed models in reality. In this dissertation, this trade-off effect has been investigated in threefold as summarized below:

The first one is the issue of continuity. Due to the complexity of the multiclass multilane traffic model proposed in Chapter 5, the optimization problem concerned is with the non-differentiable complicated functional structure of the control vector as well as the constraint set that would induce many local optima. The optimal control problem for general nonlinear systems usually adopts gradient-based approaches to devise solution algorithms based on several assumptions, such as the system dynamics (governing the state space) is at least continuously differentiable whereas the control is continuous. It can be seen from the development of the multiclass multilane traffic flow model that the underlying system dynamics for the optimal control problem concerned is not differentiable, and it is too complicated to evaluate the gradient of the objective function along the state trajectory and the constraint set. The gradient-based policy search (or

iteration) approaches may not be a good choice for solving the proposed optimal control problem. To this end, a gradient-free algorithm, which is a cross-entropy method (CEM) based reinforcement learning policy search algorithm, was developed to solve the above optimal control problem. Instead of searching for the value function by iteration, this gradient-free algorithm first parameterizes the control (or policy) and then searches for the optimal parameters that lead to the maximal returns (or minimizing the objective function). As indicated in Thiery and Scherrer (2009), this CEM based algorithm led to one of the best publicly known controllers to the game of Tetris (Gabillon et al., 2013).

The second one is the cycle-to-cycle control implementation: The control is also subject to several constraints for practical considerations, e.g., the VSL is discretized, and the control is implemented in a cycle-to-cycle manner (2–5 min interval while the simulation time step is 10 s). Such cycle-to-cycle control implementation would reduce the dimension of the control variables to  $\frac{\text{simulation time step}}{\text{control cycle}}$  (e.g.,  $\frac{10}{300}$ ).

The third one is the control parameterization: Although the above cycle-to-cycle control implementation could reduce the dimension of the optimization problem. However, it would still be huge if the simulation (optimization) horizon is comparatively long. Conventional approaches based on dynamic programming or nonlinear programming would suffer from the curse of dimensionality. Rather than searching the optimal control parameters directly, this algorithm parameterizes the control (or policy) and searches for the optimal parameters that lead to maximal returns (or minimizing the objective function). The control (or policy) parameterization technique is a popular numerical scheme for solving the optimal control problem. It involves approximating the control function (as a vector in certain functional space) by a linear combination of the basis functions. The coefficients in the linear combination are decision variables to be determined (independent of the time discretization) instead of solving the optimal control problem directly (involving many time-dependent control variables). Control parameterization technique yields a low dimensional parameter optimization

approximation of the original high dimensional optimal control problem. Combining this with the aforementioned reinforcement learning technique could yield an efficient computational algorithm for the proposed optimal control problem.

The effectiveness and the computational feasibility of the optimal control framework were illustrated via numerical example for a variety of penetration rates of CAVs under various traffic conditions. It was found that the LCC would reduce the number of vehicles queuing at the bottleneck and decrease the traffic density on the freeway. From a mathematical point of view, it was shown that the LCC could make the freeway traffic system continuous and easier for the VSL controller to stabilize. For road traffic safety, the integration of VSL and LCC can drastically reduce the instances of the stop-and-go traffic, smooth the traffic flow and suppress the impact of the shockwaves on the freeway sections concerned. On the other hand, the dynamic proportion of CAVs on a short freeway segment would significantly affect its throughput and traffic speed. The effect of different penetration rate of CAVs was investigated by sensitivity tests in the numerical example.

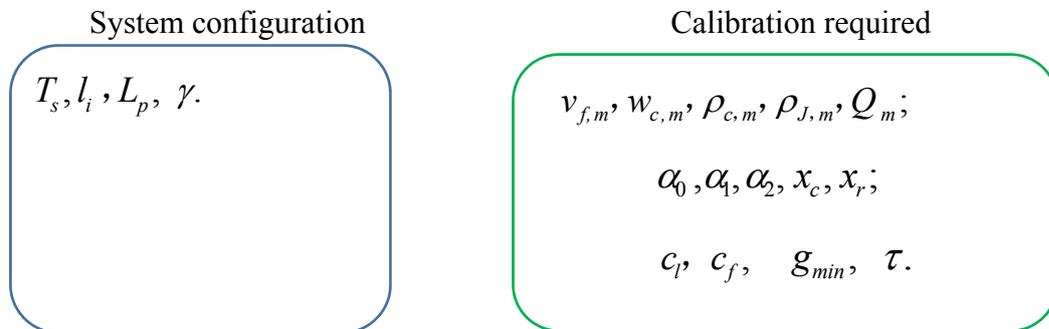
For the optimal control design, since there is no available data set for freeway traffic flow mixed with CAVs and RHVs, it was assumed that the proposed multiclass multilane traffic model can describe the characteristics of traffic mixed with CAVs and RHVs accurately. The connection established by V2V communication was assumed to be reliable to support the cooperative driving of CAVs under dynamic traffic environment regardless the time-varying communication topology. It was also assumed that the arrival pattern of the mixed **traffic is generalized by a known process**. Furthermore, large-scale network optimal control applications have not yet been included in this dissertation. It is interesting to look into the computational efficiency for large-scale applications in future.

## **6.2 Recommendation for future works**

## Multilane traffic flow model

A comprehensive calibration procedure should be required for both the multilane CTM model proposed in Chapter 3 and the multilane SCTM model proposed in Chapter 4. Take the multilane CTM model proposed in Chapter 3 as an example; there are 18 parameters for the whole system as summarized below.

Parameters:



Among the parameter set, four parameters such as sample time  $T_s$  are user-specific, i.e., input to the proposed model by the user directly. The remaining 14 parameters would require calibration for quantifying their values. These parameters can be calibrated in three separate groups which are not difficult to obtain in practice. The calibration procedure in Chapter 3 is an illustrative one in the sense that some typical calibrated values from the literature were adopted in this study. The reason is that these parameters are stable ones (i.e., not fast time-varying parameters, e.g., the average reaction time  $\tau$ ).

Parameters  $v_{f,m}, w_{c,m}, \rho_{c,m}, \rho_{J,m}, Q_m$  are from the lane-specific fundamental diagrams which are induced by the macroscopic traffic flow model. In Chapter 3, these parameters were calibrated by the conventional least squares method as explained in the empirical study. The parameters in the cumulative distribution function of mandatory lane-changing demand on longitudinal dimension, i.e.,  $\alpha_0, \alpha_1, \alpha_2, x_c, x_r$ , were calibrated based on the data collected by PeMS using the curve fitting method as

discussed in the section of Chapter 3 for the empirical study. A typical value of the average reaction time  $\tau$  is taken from the literature regarding its calibration using both NGSIM and PeMS data sets. The parameters regarding the minimum acceptance gap, i.e.  $c_l, c_f, g_{min}$ , do require trajectory data for recording the lane-changing maneuvers. In Chapter 3, the typical values for these parameters have been adopted from literature and were calibrated in Hidas (2005) whereas a small set of trajectory data obtained from traffic video was applied to calibrate these parameters. For a typical freeway segment, such data can be easily recorded, and the calibration can be achieved in the way as Hidas (2005) did.

Although the calibration method adopted is heuristic, the macroscopic multilane traffic model proposed in Chapter 3 can be used with only the data required by the conventional CTM. Although these parameters have not been fine-tuned, the proposed model can still produce quite accurate estimation results through this heuristic calibration procedure with very limited traffic data as indicated by the model validation results in this dissertation. However, it should be noted that this is mainly because these parameters are quite stable under recurrent traffic conditions. If there is any anomaly on the freeway, such heuristic calibration procedure may not be appropriate. Therefore, a robust comprehensive calibration method would be needed for non-recurrent traffic conditions. On the other hand, devising comprehensive calibration method is independent but with less attention. In literature for freeway traffic model development, very few comprehensive calibration methods can be found even for the well-known cell or link transmission models and car-following models. Moreover, the research on sensitivity analysis for model calibration is even rarer, see Zhong et al. (2015, 2016) and the references therein. In view of this, there is an emerging avenue of research for the development of such a robust comprehensive calibration method and its sensitivity analysis for future work.

Apart from the fact that these parameters are quite stable in general, another observation

from the empirical study in Chapter 3 would be that the model results are not very sensitive to these stable parameters. In Chapters 3 and 4, a one-at-a-time sensitivity analysis (SA) was conducted to test the sensitivity of the average reaction time  $\tau$ . The results indicated that the proposed lane-changing model is not sensitive to  $\tau$  for a small change of  $\tau$  around 3 sec since the simulation result is also related to the sample time  $T_s$ . The ratio  $\tau/T_s$  would have more impacts (or sensitivity value) on the accuracy of the simulation. Of course, more rigorous sensitivity analysis (SA) is needed to identify the critical parameters to reduce the calibration effort and data requirement. Also, further study is required for developing a comprehensive sensitivity analysis framework to assess the important parameters to improve the calibration process. Future research on empirical justification is also needed for calibration and validation of the proposed multilane SCTM. Similar to the SCTM case (Sumalee et al., 2011; Zhong et al., 2013), the heterogeneity and uncertainty in traffic flow can be better described by incorporating stochastic elements. Developing a comprehensive calibration procedure considering the stochastic elements, e.g., the data noise and model uncertainty, is also an interesting topic for future study. Nevertheless, extensions of the proposed models to generalized road traffic networks with different vehicle types require further research effort.

Unlike the conventional multilane CTM, with the VACS, it is possible to extend the macroscopic models developed in Chapter 5 to incorporate the position and speed information from CAVs and measurement from a minimum number of point detectors that specify the boundary conditions of the freeway sections under investigation. Although extensive research has been dedicated to traffic state estimation at the link level, e.g., the freeway corridor, the structural observability, an important but yet examined issue, is missing in the literature even for the simple cases such as freeways or urban expressways. Observability is a quantitative description of a dynamic system is inherently limited by the ability to infer the internal state of the system from its accessible outputs. A difficulty for this with conventional fixed-point detectors is the

structural observability depends on the traffic condition especially during the transient period, which would introduce too dense detectors or the installation of fixed detectors should adapt to real-time traffic condition (either of which is unlikely to happen). With the penetration of CAVs, it is possible to use the CAVs as floating detectors (but unlike GPS that cannot provide detailed and accurate vehicle information) to compensate the fixed detectors to determine the structural observability.

### **Handling uncertainties**

It was assumed in Chapter 5 that the proposed multiclass multilane traffic model could well describe the characteristics of traffic mixed with CAVs and RHVs. However, the mismatch between the dynamic process and the underlying prediction models is inevitable. Uncertainties such as penetration rate of CAVs in the near future, variations of the freeway supply functions such as capacity, a stochastic compliance rate of human drivers and dynamic traffic delays prevailing in the control loop should be systematically investigated. Regarding the integrated traffic control strategies on freeways, a new future research direction is to consider the robustness property of the proposed optimal control framework against the uncertainties for freeway traffic management in practice. In this dissertation, it was assumed that the nominal functions of these elements are known, and the control delays are negligible. In reality, uncertainties in the penetration rate of CAVs yield stochastic fundamental diagram and thus lead to variations of the freeway supply functions. Data from onboard sensors are prone to errors and lags exist in vehicle dynamics through information transmission. Due to these inherent uncertainties, optimal control based solutions become suboptimal, and the performance increasingly deteriorates with increasing disturbance prediction and model errors. Studying the impact of uncertainties on the controller performance is necessary. Time lags in the control loop may even render the system unstable. The proposed optimal control problem against the uncertainties and time lags concerned is an interesting and vital future work.

## **Hierarchical traffic control architecture**

Another promising research avenue is to develop a **hierarchical** control architecture for mixed CAV and RHV traffic flows for cooperative strategy. The top layer (global) optimizes the network performance while the lower level optimizes the link level (local) performance. The control measure developed in Chapter 5 can be regarded as a **local level control** that the integrated traffic control system minimizes the **total travel cost** along each road link for roadway traffic mixed with CAVs via VACS and RHVs via traffic signals and en-route Variable Message Signs (VMS). At the **global or network-wide level**, one promising approach is to use the macroscopic (network) fundamental diagram (MFD) framework. Under the notion of MFD, a city can be partitioned into multiple regions with each represented by an MFD model. The flow control to each region is actuated by regulating the critical intersections effectively to alleviate the level of congestion. The advantage of the MFD approach is that the complexity of the optimization problem (so as the dimension of the decision variables) can be reduced drastically such that a good balance between computational burden and performance can be achieved.

## **Data-driven control and reinforcement learning based control strategies**

Besides the above **hierarchical** control architecture, there is room for further improving the balance between computation burden and performance. Since there is no available data set for freeway traffic flow mixed with CAVs and RHVs, the reinforcement learning based control framework developed in Chapter 5 still relies on a reliable traffic model to estimate the traffic dynamics, i.e., a traffic model for network loading. When the VACS can support sufficient data, pure data-driven control framework can be devised to further reduce the computation burden by leveraging recent advances in deep reinforcement learning (RL) or the approximate dynamic programming (ADP). Under this RL solution approach, rather than searching the optimal control directly, the RL

algorithm parameterizes the control policy (and evaluate the reward for each action as well) and searches for the optimal parameters that lead to the maximum gains.

## References

- Ahmed, K., 1999. Modeling Drivers' Acceleration and Lane-changing Behavior. Ph.D. dissertation, Massachusetts Institute of Technology.
- Ahn, S., Cassidy, M., 2007. Freeway traffic oscillations and vehicle lane-change maneuvers. In: Allsop, R.E., Bell, M.G.H., Heydecker, B. (Eds.), Proceedings of the 17th International Symposium on Traffic and Transportation Theory, 691-710.
- Alecsandru, C., 2006. A stochastic mesoscopic cell-transmission model for operational analysis of large-scale transportation networks. PhD dissertation, Department of Civil and Environmental Engineering, Louisiana State University, USA.
- Asano, M., Kuwahara, M., Tanaka, S., 2007b. Multi-directional pedestrian flow model based on empirical data. Proceedings of 11th World Conference on Transport and Safety Research.
- Asano, M., Sumalee, A., Kuwahara, M., Tanaka, S., 2007. Dynamic cell transmission-based pedestrian model with multidirectional flows and strategic route choices. Transportation Research Record, No. 2039, 42-49.
- van Arem, B., van Driel, C., and R. Visser, 2006. The Impact of Cooperative Adaptive Cruise Control on Traffic-Flow Characteristics. IEEE Transactions on Intelligent Transportation Systems, 7(4), 429–436.
- Bekiaris-Liberis, N., Roncoli, C., and Papageorgiou, M., 2016. Highway traffic state estimation with mixed connected and conventional vehicles. IEEE Transactions on Intelligent Transportation Systems, 17(12), 3484-3497.
- Ben-Akiva, M., 1996. Development of a deployable real-time dynamic traffic assignment system, task D interim report: analytical developments for DTA system. MIT, Cambridge, MA.
- Ben-Akiva, M., Choudhury, C., Toledo, T., 2006. Lane changing models. Proceedings of the 2nd International Symposium of Transport Simulation, Lausanne, Switzerland, September.

- Ben-Akiva, M., Choudhury, C., Toledo, T., 2006. Lane-changing models. In: Proc. Int. Symp. Trans. Simu., Lausanne, Switzerland.
- Bertsekas, D., 2005. Dynamic Programming and Optimal Control. Athena Scientific, Belmont, MA.
- Bishop, C., 2006. Pattern Recognition and Machine Learning. Springer-Verlag New York.
- Blandin, S., Couque, A., Bayen, A., Work, D., 2012. On sequential data assimilation for scalar macroscopic traffic flow models. *Physica D*, 241 (17), 1421-1440.
- Busoniu, L., Babuska, R., De Schutter, B., Ernst, D., 2010. Reinforcement Learning and Dynamic Programming Using Function Approximators. CRC Press, Boca Raton, FL.
- Carey, M., Balijepalli, C., Waling, D., 2015. Extending the cell transmission model to multiple lanes and lane-changing. *Netw. Spat. Econ.*, 15 (3), 507-535.
- Carlson, R., Papamichail, I., Papageorgiou, M., 2011. Local feedback-based mainstream traffic flow control on motorways using variable speed limits. *IEEE Transactions on Intelligent Transportation Systems*, 12(4), 1261-1276.
- Carlson, R., Papamichail, I., Papageorgiou, M., Messmer, A., 2010. Optimal motorway traffic flow control involving variable speed limits and ramp metering. *Transportation Science*, 44 (2), 238-253.
- Carter, M., Rakha, H., Aerde, M., 1999. Variability of traffic-flow measures across freeway lanes. *Can. J. Civ. Eng.*, 26, 270-281.
- Cassidy, M., Bertini, R., 1999. Some traffic features at freeway bottlenecks. *Transportation Research Part B*, 33, 25-42.
- Cassidy, M., Rudjanakanoknad, J., 2005. Increasing the capacity of an isolated merge by metering its on-ramp. *Transportation Research Part B*, 39 (10), 896-913.
- Calvert, S., Taale, H., Snelder, M., Hoogendoorn, S., 2015. Vehicle specific behaviour in macroscopic traffic modelling through stochastic advection invariant. *Transportation Research Procedia* 10, 71 – 81.
- Chen, D., Ahn, S., Chitturi, M., and Noyce, D. 2017. Towards vehicle automation:

- Roadway capacity formulation for traffic mixed with regular and automated vehicles. *Transportation Research Part B*, 100, 196-221.
- Chen, C., 2003. Freeway performance measurement system (PeMS). PhD dissertation, University of California, Berkeley.
- Cheu, R., Martinez, J., Duran, C., 2009. A cell transmission model with lane changing and vehicle tracking for port of entry simulations. *Transportation Research Record*, No. 2124, 241-248.
- Chiabaut, N., Leclercq, L., Buisson, C., 2010. From heterogeneous drivers to macroscopic patterns in congestion. *Transportation Research Part B*, 44 (2), 299-308.
- Choudhury, C., Ben-Akiva, M., Rao, A., Lee, G., Toledo, T., 2007. State dependence in lane changing models. In *Transportation and Traffic Theory*, Allsop, R., Bell, M., and Heydecker, B., ed., Elsevier, 711-733.
- Choudhury, C., Ben-Akiva, M., Toledo, T., Lee, G., Rao, A., 2007. Modeling cooperative lane changing and forced merging behavior. In *Proceedings of the 86th Annual Meeting of the Transportation Research Board*, Washington, D.C.
- Coifman, B., Mishalani, R., Wang, C., Krishnamurthy, S., 2006. Impact of lane-change maneuvers on congested freeway segment delays: pilot study. *Transportation Research Record*, No. 1965, 152-159.
- Corthout, R., Flöteröd, G., Viti, F., Tampere, C., 2012. Non-unique flows in macroscopic first-order intersection models. *Transportation Research Part B*, 46(3), 343-359.
- Daamen, W., Loot, M., Hoogendoorn, S., 2010. Empirical analysis of merging behavior at freeway on-ramp. *Transportation Research Record*, No. 2188, 108-118.
- Daganzo, C., 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B*, 28 (4), 269-287.
- Daganzo, C., 1995. The cell transmission model: Network traffic. *Transportation Research Part B*, 29 (2), 79-93.

- Daganzo, C., 2005. A variational formulation of kinematic waves: Basic theory and complex boundary conditions. *Transportation Research Part B*, 39 (2), 187-196.
- Daganzo, C., Geroliminis, N., 2008. An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B*, 42(9), 771-781.
- Daganzo, C., Lin, W., del Castillo, J., 1997. A simple physical principle for the simulation of freeways with special lanes and priority vehicles. *Transportation Research Part B*, 31(2), 103-125.
- Del Castillo, (2012) Del Castillo, J., 2012. Three new models for the flow-density relationship: Derivation and testing for freeway and urban data. *Transportmetrica*, 8 (6), 443-465.
- Del Castillo, J., and Bentez, F., 1995a. On the functional form of the speed-density I): General theory. *Transportation Research Part B*, 29 (5), 373-389.
- Del Castillo, J., and Bentez, F., 1995b. On the functional form of the speed-density II): Empirical investigation. *Transportation Research Part B*, 29(5), 391-406.
- Diakaki, C., Papageorgiou, M., Papamichail, I., Nikolos, I., 2015. Overview and analysis of Vehicle Automation and Communication Systems from a motorway traffic management perspective. *Transportation Research Part A*, 75, 147-165.
- Duret, A., Ahn, S., Buisson, C., 2012. Lane flow distribution on a three-lane freeway: General features and the effects of traffic controls. *Transportation Research Part C*, 24 (1), 157-167.
- Flöteröd, G., Bierlaire, M., Nagel, K., 2011. Bayesian demand calibration for dynamic traffic simulations. *Transportation Science*, 45 (4), 541-561.
- Flöteröd, G., Rohde, J., (2011). Operational macroscopic modeling of complex urban road intersections. *Transportation Research Part B*, 45 (6), 903-922.
- Fountoulakis, M., Bekiaris-Liberis, N., Roncoli, C., Papamichail, I., Papageorgiou, M., 2017. Highway traffic state estimation with mixed connected and conventional vehicles: Microscopic simulation-based testing. *Transportation Research Part C*, 78, 13-33.

- Gabillon, V., Ghavamzadeh, M., Scherrer, B., 2013. Approximate Dynamic Programming Finally Performs Well in the Game of Tetris. *Neural Information Processing Systems (NIPS) 2013*, Dec 2013, South Lake Tahoe, United States.
- Gentile, G., Meschini, L., Papola, N., 2005. Macroscopic arc performance models with capacity constraints for within-day dynamic traffic assignment. *Transportation Research Part B*, 39 (10), 319-338.
- Geroliminis, N., Daganzo, C., 2008. Existence of urban-scale macroscopic fundamental diagrams: some experimental findings. *Transportation Research Part B*, 42(9), 759-770.
- Gipps, P., 1986. A model for the structure of lane-changing decisions. *Transportation Research Part B*, 20 (5), 403-414.
- Golob, T., Recker, W., Alvarez, V., 2004. Safety aspects of freeway weaving sections. *Transportation Research Part A*, 38 (1), 35-51.
- Goñi-Ros, B., Knoop, V., Takahashi, T., Sakata, I., van Arem, B., and Hoogendoorn, S., 2016. Optimization of traffic flow at freeway sags by controlling the acceleration of vehicles equipped with in-car systems. *Transportation Research Part C*, 71, 1–18.
- Green, M., 2000. “How long does it take to stop?” methodological analysis of driver perception -brake times. *Transportation human factors*, 2(3), 195-216.
- Gunay, B., 2007. Car following theory with lateral discomfort, *Transportation Research Part B*, 41 (7), 722-735.
- Jin, H. and Jin, W., 2015. Control of a lane-drop bottleneck through variable speed limits. *Transportation Research Part C: Emerging Technologies*, 58, 568–584.
- Han, Y., Chen D., and Ahn, S., 2017. Variable speed limit control at fixed freeway bottlenecks using connected vehicles. *Transportation Research Part B*, 98, 113-134.
- Han, Y. (2017). *Fast Model Predictive Control Approaches for Road Traffic Control*. TRAIL Ph.D. dissertation series. Delft University Press.
- He, Z., Zheng, L., Song, L., Zhu, N., 2017. A jam-absorption driving strategy for

- mitigating traffic oscillations. *IEEE Transactions on Intelligent Transportation Systems*. 18 (4), 802-813.
- Hegy, A., Schutter, B., Hellendoorn, H., 2005a. Model predictive control for optimal coordination of ramp metering and variable speed limits. *Transportation Research Part C*, 13(3), 185-209.
- Hegy, A., Schutter, B., Hellendoorn, H., 2005b. Optimal coordination of variable speed limits to suppress shock waves. *IEEE Transactions on Intelligent Transportation Systems*, 6(1), 102-112.
- Herrera, J., Bayen, A., 2010. Incorporation of Lagrangian measurements in freeway traffic state estimation. *Transportation Research Part B*, 44 (4), 269-287.
- Hidas, P., 2005. Modelling vehicle interactions in microscopic simulation of merging and weaving. *Transportation Research Part C*, 13(1), 37-62.
- Hoogendoorn, S.P., 1999. Multiclass Continuum Modelling of Multilane Traffic Flow. TRAIL Ph.D. dissertation series. Delft University Press.
- Hoogendoorn, S., Bovy, P., 2001a. Generic gas-kinetic traffic systems modeling with applications to vehicular traffic flow. *Transportation Research Part B*, 35, (4), 317-336.
- Hoogendoorn, S., Bovy, P., 2001b. Platoon-based multiclass modeling of multilane traffic flow. *Networks and Spatial Economics*, 1(1-2), 137-166.
- Hou, Y., Edara, P., Sun, C., 2015. Modelling mandatory lane-changing using Bayes classifier and decision trees. *IEEE Transactions on Intelligent Transportation Systems*, 15 (2), 647-655.
- Jabari, S., Liu, H., 2012. A stochastic model of traffic flow: Theoretical foundations. *Transportation Research Part B*, 46 (1), 156-174.
- Jepsen, M., 1998. On the Speed-Flow Relationships in Road Traffic: A Model of Driver Behaviour. In *Proceedings of the Third International Symposium on Highway Capacity*.
- Jia, D., Ngoduy, D., 2016. Platoon based cooperative driving model with consideration of realistic inter-vehicle communication. *Transportation Research Part C*, 68,

245–264.

- Jia, D., Ngoduy, D., 2016. Enhanced cooperative car-following traffic model with the combination of V2V and V2I communication. *Transportation Research Part B*, 90, 172–191.
- Jin, W., 2010a. A kinematic wave theory of lane-changing traffic flow. *Transportation Research Part B*, 448(9), 1001-1021.
- Jin, W., 2010b. Macroscopic characteristics of lane-changing traffic. *Transportation Research Record*, No. 2188, 55-63.
- Jin, W., Zhang, H., 2003. On the distribution schemes for determining flows through a merge. *Transportation Research Part B*, 37 (6), 521-540.
- Jin, W., Zhang, Y., Chu, L., 2006. Measuring first-in-first-out violation among vehicles. In *Proceedings of the 85th Annual Meeting of the Transportation Research Board*, Washington, D.C.
- Kesting, A., Treiber, M., Helbing, D., 2007. General lane-changing model mobile for car-following models. *Transportation Research Record*, No. 1999, 86-94.
- Keyvan-Ekbatani, M., Daamen, W, and Knoop, V.L., 2016. Categorization of the lane-change decision process on freeways. *Transportation Research part C*, 69, 515-526.
- Khondaker, B., Kattan, L., 2015. Variable speed limit: A microscopic analysis in a connected vehicle environment. *Transportation Research Part C*, 58, 146-159.
- Khoshyaran, M., Lebacque, J., 2009. A stochastic macroscopic traffic model devoid of diffusion. In: Appert-Rolland, C., Chevoir, F., Gondret, P., Lassarre, S., Lebacque, J., Schreckenberg, M. eds.), *Traffic and Granular Flow '07*. Springer, New York, pp. 139-150.
- Knoop, V., Hoogendoorn, S., Shiomi, Y., Buisson, C., 2012. Quantifying the number of lane changes in traffic: an empirical analysis. *Transportation Research Record*, No. 2278, 31-41.
- Knoop, V.L., Buisson, C., 2015. Calibration and validation of probabilistic discretionary lane-change models. *IEEE Transactions on Intelligent*

- Transportation Systems, 16 (2), 834 – 843.
- Kontorinaki, M., A. Spiliopoulou, C. Roncoli, and M. Papageorgiou, 2016. Capacity drop in first-order traffic flow models: Overview and real data Validation. In Processing of Transportation Research Board 95th Annual Meeting, 3541.
- Kurzhanskiy, A., Varaiya, P., 2010. Active traffic management on road networks: a macroscopic approach. Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences, Part A 368(1928), 4607–4626.
- Lauer, M., 2011. Grand Cooperative Driving Challenge 2011. Intelligent Transportation Systems Magazine, 3(3), 38–40.
- Laval, J., Cassidy, M., Daganzo, C., 2007. Impacts of lane changes at on-ramp bottlenecks: A theory and strategies to maximize capacity. (In: Kühne, R., Poeschl, T., Schadschneider, A., Schreckenberg, M., Wolf, D. Eds.), Traffic and Granular Flow '05'. Springer, Berlin, 577-586.
- Laval, J., Daganzo, C., 2006. Lane-changing in traffic streams. Transportation Research Part B, 40 (3), 251-264.
- Laval, J., Leclercq, L., 2008. Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model. Transportation Research Part B, 42 (6), 511-522.
- Laval, J., Leclercq, L., 2010. A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic. Philosophical Transactions of Royal Society A, 368 (1928), 4519-4541.
- Lebacque, J., Khoshyaran, M., 2002. First-order macroscopic traffic flow models for networks in the context of dynamic assignment. In: M. Patriksson and M. Labbé, eds. Transportation planning the state of the art. Dordrecht: Kluwer, 119-140.
- Lebacque, J., Khoshyaran, M., 2005. First-order macroscopic traffic flow models: Junction modeling, network modeling. In: H.S. Mahmassani, ed. Proceedings of the 16th international symposium on transportation and traffic theory, Amsterdam, Elsevier, 365-386.

- Leclercq, L., 2007. Hybrid approaches to the solutions of the "Lighthill-Whitham-Richards" model. *Transportation Research Part B*, 41 (7), 701-709.
- Leclercq, L., Chiabaut, N., Laval, J., Buisson, C., 2007. Relaxation phenomenon after changing lanes: Experimental validation with NGSIM dataset. *Transportation Research Record*, No. 1999, 79-85.
- Leclercq, L., Laval, J., Chiabaut, N., 2011. Capacity drops at merges: An endogenous model, *Transportation Research Part B*, 45 (10), 1302-1313.
- Leclercq, L., Moutari, S., 2007. Hybridization of a class of "second order" models of traffic flow. *Simulation Modelling Practice and Theory*, 15 (8) 918-934.
- Leclercq, L., Knoop, V.L., Marczak, F., and Hoogendoorn, S.P., 2016a. Capacity drops at merges: New analytical investigations. *Transportation Research Part C*, 62, 171-181.
- Leclercq, L., Marczak, F., Knoop, V.L., and Hoogendoorn, S.P., 2016b. Capacity drops at merges: analytical expressions for multilane freeways. *Transportation Research Record*, No. 2560, 1-9.
- Lee, J., Cassidy, M., 2009. An Empirical and Theoretical Study of Freeway Weave Bottlenecks. PATH Research Report, UCB-ITS-PRR-2009-13.
- Lerner, M, Huey, R., Mcgee, H., and Sullivan. A (1995). Older driver perception-reaction time for intersection sight distance and object detection, 1(1995), 33-40. Report FHWA-RD-93-168, Federal Highway Administration, US-DOT, Mclean, Virginia.
- Levin, M., and Boyles, S., 2016a. A multiclass cell transmission model for shared human and autonomous vehicle roads. *Transportation Research Part C*, 62, 103-106.
- Levin, M., and Boyles, S., 2016b. A cell transmission model for dynamic lane reversal with autonomous vehicles. *Transportation Research Part C*, 68, 126-143.
- Li, J., Chen, Q., Wang, H., Ni, D., 2012. Analysis of LWR model with fundamental diagram subject to uncertainties. *Transportmetrica*, 8 (6), 387-405.
- Li, Z., Liu, P., Wang, W. and Xu, C., 2014. Development of a control strategy of variable

- speed limits to reduce rear-end collision risks near freeway recurrent bottlenecks. *IEEE Transactions on Intelligent Transportation Systems*, 15(2), 866-877.
- Li, Z., Liu, P., Xu, C. and Wang, W., 2016. Optimal Mainline Variable Speed Limit Control to Improve Safety on Large-Scale Freeway Segments. *Computer-Aided Civil and Infrastructure Engineering*, 31, 366-380.
- Lighthill, M., Whitham, G., 1955. On kinematic waves. I: Flow movement in long rivers; II: A theory of traffic flow on long crowded roads. In: *Processing of Royal Society A*, 229, 281-345.
- Liu, Y., Chang, G.L., 2011. An arterial signal optimization model for intersections experiencing queue spillback and lane blockage. *Transportation Research Part C*, 19 (1), 130-144.
- Liu, S., Hellendoorn, H., and Schutter, B., 2017. Model Predictive Control for Freeway Networks Based on Multiclass Traffic Flow and Emission Models. *IEEE Intelligent Transportation System*, 18(2), 306-321.
- Luo, Y., Xiang, Y., Cao, K., Li, K., 2016. A dynamic automated lane change maneuver based on vehicle-to-vehicle communication. *Transportation Research Part C*, 62 87-102.
- Mauch, M., Cassidy, M., 2002. Freeway Traffic Oscillations: Observations and Predictions. *Transportation and Traffic Theory in the 21<sup>st</sup> Century*, 653-673.
- Menendez, M., 2006. An analysis of HOV lanes: Their impact on traffic. Ph.D. dissertation, Department of Civil Engineering, University of California, Berkeley, USA.
- Milanés, V., Shladover, S.E., 2014. Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. *Transportation Research Part C*, 48, 285–300.
- Milanes, V., Shladover, S.E., Spring, J., Nowakowski, C., Kawazoe, H., Nakamura, M., 2014. Cooperative adaptive cruise control in real traffic situations. *IEEE Transactions on Intelligent Transportation Systems*, 15, 296–305.
- Muralidharan, A., and R. Horowitz, 2015. Computationally efficient model predictive

- control of freeway networks. *Transportation Research Part C: Emerging Technologies*, 58, 532–553.
- Munjal, P., Pipes, L., 1971. Propagation of on-ramp density perturbations on unidirectional two- and three-lane freeways. *Transportation Research Part B*, 5(4), 241-255.
- Muñoz, L., Sun, X., Horowitz, R., Alvarez, L., 2003. Traffic density estimation with the cell transmission model. *Processing of American Control Conference*, Denver, Colorado, USA, 5, 3750-3755.
- Muñoz, L., X. Sun, R. Horowitz and L. Alvarez, 2006. A piecewise-linearized cell transmission model and parameter calibration methodology. *Transportation Research Records*, No. 1965, 183-191.
- Newell, G., 1993. A simplified theory of kinematic waves in highway traffic. I: general theory. II: queuing at freeway bottlenecks. III: multi-destination flows. *Transportation Research Part B*, 27, 281-313.
- Ng, M., Waller, S., 2010. A computationally efficient methodology to characterize travel time reliability using the fast Fourier transform. *Transportation Research Part B*, 44 (10), 1202-1219.
- Ni, D., 2011. Multiscale Modeling of Traffic Flow. *Mathematica Aeterna*, 1(1), 27-54.
- Nie, J., Zhang, J., Ding, W., Wan, X., Chen, X., Ran, B., 2017. Decentralized cooperative lane-changing decision-making for connected autonomous vehicles. *IEEE Access*, 4, 9413 - 9420.
- Ngoduy, D., Jia, D., 2017. Multi anticipative bidirectional macroscopic traffic model considering cooperative driving strategy. *Transportmetrica B*, 5 (1), 96-110.
- Ozbay, K., Kachroo, P., 1999. *Incident management in intelligent transportation systems*. Artech House, Boston.
- Pan, T., Lam, W.H.K., Sumalee, A., Zhong, R., 2016. Modelling the impacts of mandatory and discretionary lane-changing maneuvers. *Transportation Research Part C*, 68, 403-424.
- Pan, T., Sumalee, A., Zhong, R., Indra-payoong, N., 2013. Short-term traffic state

- prediction based on temporal-spatial correlation. *IEEE Transactions on Intelligent Transportation Systems*, 14 (3), 1242 – 1254.
- PATH, <http://www.path.berkeley.edu/>.
- Patire, A., Cassidy, M., 2011. Lane changing patterns of bane and benefit: Observations of an uphill expressway. *Transportation Research Part B*, 45 (4), 656-666.
- Pipes, L., 1967. Car following models and the fundamental diagram of road traffic. *Transportation Research*, 1(1), 21-29.
- Powell, W.B., 2011. *Approximate dynamic programming: solving the curses of dimensionality*, 2nd edition. Wiley, Hoboken, NJ.
- Hall, R., and C. Chin, 2005. Vehicle sorting for platoon formation: Impacts on highway entry and throughput. *Transportation Research Part C: Emerging Technologies*, 13(5-6), 405–420.
- Richards, P., 1956. Shockwaves on the highway. *Operation Research*, 4, 42-51.
- Roncoli, C., Papageorgiou, M., Papamichail, I., 2015a. Traffic flow optimisation in presence of vehicle automation and communication systems – Part I: a first-order multilane model for motorway traffic. *Transportation Research Part C* 57, 241–259.
- Roncoli, C., Papageorgiou, M., Papamichail, I., 2015b. Traffic flow optimisation in presence of vehicle automation and communication systems – Part II: optimal control for multilane motorways. *Transportation Research Part C* 57, 260–275
- Roncoli, C., Papamichail, I., Papageorgiou, M., 2016. Hierarchical model predictive control for multilane motorways in presence of Vehicle Automation and Communication Systems. *Transportation Research Part C* 62, 117-132.
- Rubinstein, R., and Kroese, D., 2004. *The cross-entropy method: a unified approach to combinatorial optimization, monte-carlo simulation and machine learning*. Springer.
- Tsugawa, S., and Kato, S., 2010. Energy ITS: another application of vehicular Communications. *Communications Magazine IEEE*, 48(11), 120–126.
- Sarvi, M., 2013. Freeway weaving phenomena observed during congested traffic.

- Transportmetrica, 9 (4), 299-315.
- Schakel, W., Knoop, V., and van Arem, B., 2012. LMRS: An integrated lane change model with relaxation and synchronization. In: Processing of Transportation Research Board, Washington, D.C.
- Schnetzler, B., Louis, X., Lebacque, J., 2012. A multilane junction model. Transportmetrica, 8 (4), 243-260.
- Sheu, J., 2004. A sequential detection approach to real-time freeway incident detection and characterization. European Journal of Operational Research, 157 (2), 471-485.
- Sheu, J., Ritchie, S., 2001. Stochastic modeling and real-time prediction of vehicular lane-changing behavior. Transportation Research Part B, 35 (7), 695-716.
- Shiomi, Y., Taniguchi, T., Uno, N., Shimatomo, H., 2015. Multilane first-order traffic flow model with endogenous representation of lane-flow equilibrium. Transportation Research Part C, 59, 198-215.
- Shvestsov, V., Helbing, D., 1999. Macroscopic dynamics of multilane traffic. Physical review E, 59 (6), 6328-6338.
- 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, 30, 161-177.
- Sumalee, A., Pan, T., Zhong, R., Uno, N., 2013. Dynamic stochastic journey time estimation and reliability analysis using stochastic cell transmission model: Algorithm and case studies. Transportation Research Part C, 35, 263–285.
- Sumalee, A., Zhong, R., Pan, T., and Szeto, W., 2011. Stochastic cell transmission model (SCTM): A stochastic dynamic traffic model for traffic state surveillance and assignment, Transportation Research Part B, 45 (3), 507-533.
- Sun, D., Elefteriadou, L., 2010. Research and implementation of lane-changing model based on driver behavior. Transportation Research Record, 2161, 1-10.
- Sun, D., Elefteriadou, L., 2011. Lane-changing behavior on urban streets: A focus group-based study. Applied Ergonomics, 42 (5) 682-691.

- Sun, D.J., Elefteriadou, L., 2010. Research and implementation of lane-changing model based on driver behavior. *Transportation Research Record*, No. 2161, 1-10.
- Sun, D.J., Kondyli, A., 2010. Modeling vehicle interactions during lane changes on arterial streets. *Computer-Aid Civil and Infrastructure Engineering*, 25(8), 557-571.
- Szeto, W.Y., 2008. The enhanced lagged cell transmission model for dynamic traffic assignment. *Transportation Research Record*, 2085, 76-85.
- Stern, R. E., Cui, S., Monache, M.L., Bhadani, R., Bunting, M., Churchill, M., Hamilton, N., Hauley, R., Pohlmann, H., Wu, F., Piccoli, B., Seibold, B., Sprinkle, J., Work, D. B., 2018. Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments. *Transportation Research Part C*, 89, 205–221.
- Szita, I., Lorincz, A., 2006. Learning Tetris Using the Noisy Cross-Entropy Method. *Neural Computation*, 18(12): 2936–2941.
- Talebpour, A., Mahmassani, H. S., 2016. Influence of connected and autonomous vehicles on traffic flow stability and throughput, *Transportation Research Part C*, 71, 143–163.
- Tampère, C., Corthout, R., Cattrysse, D., and Immers, L., 2011. A generic class of first order node models for dynamic macroscopic simulation of traffic flows. *Transportation Research Part B*, 45(1), 289-309.
- Tampère, C., Hoogendoorn, S., van Arem, B., 2009. Continuous traffic flow modeling of driver support systems in multiclass traffic with intervehicle communication and drivers in the loop. *IEEE Transactions on Intelligent Transportation Systems*, 10 (4), 649-657.
- Tang, T., Wong, S., Huang, H., Zhang, P., 2009. Macroscopic modeling of lane-changing for two-lane traffic flow. *Journal of Advanced Transportation*, 43(3), 245-273.
- Tang, T., Wang, T., Chen L., and Shang H., 2017. Impacts of energy consumption and emissions on the trip cost without late arrival at the equilibrium state. *Physica*

- A, 479,341–349.
- Thiemann, C., Treiber, M., Kesting, A., 2008. Estimating acceleration and lane-changing dynamics from next generation simulation trajectory data. *Transportation Research Record*, No. 2088, 90-101.
- Thiery, C., Scherrer, B., 2009. Improvements on Learning Tetris with Cross Entropy. *International Computer Games Association Journal*, 32 (1), 23-33.
- Toledo, T., Choudhury, C., Ben-Akiva, M., 2005. A lane-changing model with explicit target lane choice. *Transportation Research Record*, No. 1934, 157-165.
- Tsugawa, S., and Kato, S., 2010. Energy ITS: Another application of vehicular communications, *IEEE Communication Magazine*, 48 (11), 120–126.
- Varaiya, P., 1993. Smart cars on smart roads: problems of control. *IEEE Transactions on Automatic Control*, 38 (2), 195-207.
- van de Weg, G. S. (2017). *Efficient Algorithms for NetworkWide Road Traffic Control*. TRAIL Ph.D. dissertation series. Delft University Press.
- van Wageningen-Kessels, F., van Lint, H., Hoogendoorn, S., Vuik, K., 2010. Lagrangian formulation of multiclass kinematic wave model. *Transportation Research Record*, No. 2188, 29–36.
- Wang, C., Coifman, B., 2008. The effect of lane-change maneuvers on a simplified car-following theory. *IEEE Transactions on Intelligent Transportation Systems*, 9 (3), 523-535.
- Wang, H., Li, J., Chen, Q., Ni, D., 2011. Logistic modeling of the equilibrium speed-density relationship. *Transportation Research Part A*, 45 (6), 554-566.
- Wang, H., Ni, D., Chen, Q., Li, J., 2013. Stochastic modeling of equilibrium speed-density relationship. *Journal of Advanced Transportation*, 47 (1), 126-150.
- Wang, M., Daamen, W., Hoogendoorn, S.P., van Arem, B., 2016a. Connected variable speed limits control and car-following control with vehicle-infrastructure communication to resolve stop-and-go waves. *Journal of Intelligent Transportation Systems*, 20 (6), 559-572.
- Wang, M., Daamen, W., Hoogendoorn, S.P., van Arem, B., 2016b. Cooperative car-

- following control: Distributed algorithm and impact on moving jam features. *IEEE Transactions on Intelligent Transportation Systems*, 17 (5), 1459-1471.
- Wang, M., Daamen, W., Hoogendoorn, S.P., van Arem, B., Happee, R. 2015. Game theoretic approach for predictive lane-changing and car-following control. *Transportation Research Part C*, 58, 73-92.
- Wang, R., Li, Y., and Work, D. 2017. Comparing traffic state estimators for mixed human and automated traffic flows. *Transportation Research Part C*, 28(2017), 95-110.
- Wang, Y., Papageorgiou, M., 2005. Real-time freeway traffic state estimation based on extended Kalman filter: A general approach. *Transportation Research Part B*, 39 (2), 141-167.
- Wang, Y., Papageorgiou, M., Messmer, A., 2007. Real-time freeway traffic state estimation based on extended Kalman filter: A case study. *Transportation Science*, 41, 167-181.
- Wang, Y., Papageorgiou, M., Messmer, A., Coppola, P., Tzimitsi, A., Nuzzolo, A., 2009. An adaptive freeway traffic state estimator. *Automatica*, 45 (1), 10-24.
- Wierwille, W., and Casali, J., 1983. Driver steering reaction time to abrupt-onset. *Human factors*, 1983, 2(1), 103-116.
- Work, D., Blandin, S., Tossavainen, O., Piccoli, B., Bayen, A., 2010. A traffic model for velocity data assimilation. *Applied Mathematics Research eXpress*, 2010 (1), 1C35.
- Yang, Q., Koutsopoulos, H., 1996. A microscopic traffic simulator for evaluation of dynamic traffic management systems. *Transportation Research Part C*, 4 (3), 113-129.
- Yao, S., Knoop, V.L., Van Arem, B., 2017. Optimizing traffic flow efficiency by controlling lane changes: collective, group and user optima. *Transportation Research Record*, No. 2622, 96-104.
- Yuan, K., Knoop, V.L., and Hoogendoorn, S.P., 2017. A kinematic wave model in Lagrangian coordinates incorporating capacity drop: application to

- homogeneous road stretches and discontinuities. *Physica A*, 465, 472-485.
- Zhang, Y., and Ioannou, P., 2017. Combined Variable Speed Limit and Lane Change Control for Highway Traffic. *IEEE Transactions on Intelligent Transportation Systems*, inpress.
- Zheng, Z., 2014. Recent developments and research needs in modeling lane changing. *Transportation Research Part B*, 60, 16-32.
- Zheng, Z., Ahn, S., Chen, D., Laval, J., 2013. The effects of lane-changing on the immediate follower: Anticipation, relaxation, and change in driver characteristics. *Transportation Research Part C*, 26, 367-379.
- Zhong, R., Sumalee, A., Pan, T., Lam, W.H.K, 2014. Optimal and robust strategies for traffic management under demand and supply uncertainties: An overview and general theory. *Transportmetrica*, 10 (10), 849-877.
- Zhong, R., Chen, C., Andy, C., Pan, T., Yuan, F., He, Z., 2015. Automatic calibration of fundamental diagram for first-order macroscopic freeway traffic models. *Journal of Advanced Transportation*, 50 (3), 363-385.
- Zhong, R., Fu, K., Sumalee, A., Ngoduy, D., Lam, W.H.K., 2016. A cross-entropy method and probabilistic sensitivity analysis framework for calibrating microscopic traffic models. *Transportation Research Part C*, 63, 147-169.
- Zhong, R., Sumalee, A., Pan, T., Lam, W.H.K, 2014. Optimal and robust strategies for traffic management under demand and supply uncertainties: An overview and general theory. *Transportmetrica*, 10 (10), 849-877.
- Zhong, R., Sumalee, A., Pan, T., Lam, W.H.K., 2013. Stochastic cell transmission model for traffic network with demand and supply uncertainties. *Transportmetrica*, 9(7), 567-602.
- Zhong, R., Yuan, F., Pan, T., Chow, A., Chen, C., Yang, Z., 2016. Linear complementarity system approach to macroscopic freeway traffic modeling: Uniqueness and convexity. *Transportmetrica A: Transport Science*, 12 (2), 142–174.
- Zhou, M., Qu, X., Jin, S., 2017. On the impact of cooperative autonomous vehicles in

improving freeway merging: a modified intelligent driver model-based approach. *IEEE Transactions on Intelligent Transportation Systems*, 18 (6), 1422-1428.

Zhu, F., Ukkusuri, S., 2017. Modeling the proactive driving behavior of connected vehicles: A cell-based simulation approach. *Computer-Aided Civil and Infrastructure Engineering*, inpress.