Congestion Management in Urban Traffic Networks based on Macroscopic Fundamental Diagram

YE LI

Doctor of Philosophy

THE UNIVERSITY OF
SYDNEY

Supervisor: Mohsen Ramezani
Associate Supervisor: David Levinson

A thesis submitted in fulfilment of
the requirements for the degree of
Doctor of Philosophy

School of Civil Engineering
Faculty of Engineering
The University of Sydney
Australia

16 April 2022

Statement of originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Name: Ye Li

Signature:

Abstract

The number of vehicles is growing rapidly with the development of urbanization and the economy, and the traffic congestion problem is becoming increasingly serious all around the world. To alleviate congestion, the scale and structure of major cities have undergone notable changes. Previously, many infrastructures were built to distribute excessive traffic demands, but this approach neither cost-effective nor sustainable in the long term. With the expeditious development of urban modernization, traffic congestion, air pollution, and road safety problems are necessary to be paid attention to. Therefore, the crucial research field of Intelligent Transportation System (ITS) is heavily employed and explored to monitor traffic congestion in real-time and apply responsive control schemes.

However, effective control and management approaches for large-scale urban traffic networks still remain a big challenge, due to the complicated modeling of traffic dynamics in large networks (e.g. large number of road links, signalized intersections) and unpredictable traveler patterns (such as route and mode choices, and departure times). Unlike the microscopic methods that employ disaggregate models, e.g. lane changing and car following models that consider detailed behavior of each vehicle, this dissertation proposes innovative network-level traffic management approaches utilizing the Macroscopic Fundamental Diagram (MFD) to improve mobility and decrease vehicles total time spent in the urban network.

Chapter 1 of this thesis introduces the background and motivation in studying network-level modeling and control strategies based on MFD of large-scale urban networks. Recent studies demonstrate that MFD can provide a well-defined and low-scattered relationship between network vehicle accumulation and mean weighted flow, while it merely applied to the networks with uniformly distributed congestion. Accordingly, the heterogeneous urban network should be partitioned into multiple regions to further ensure the development of various network-level management strategies. In this thesis, two major control approaches are studied, i.e. perimeter flow control and congestion pricing. Perimeter flow controllers operate on the border among

regions that manipulate the percentages of inter-region transfer flows to maximize the number of vehicles reaching their destinations through cooperative traffic signal control. Congestion pricing is another effective way to improve network performance by surcharging travelers of private vehicles in peak hours. The background corresponding to the current development of these two control strategies and the motivation for further improvement based on the existing methods will be introduced in due course.

Chapter 2 develops a perimeter flow control strategy employing Sliding Mode Control (SMC) theory to diminish the influence of heterogeneous inter-region transfer flows and cordon queues on traffic congestion of the two-region network. In this chapter, two different MFD models are utilized, i.e. accumulation-based and trip-based MFDs. The former is more aggregated that assumes all vehicles with the same trip lengths while the latter considers detailed vehicles information, e.g. remaining trip lengths and travel times. Besides, it is customary in the accumulation-based MFD model to assume the transfer outflows are equal to the ratio of the instantaneous number of vehicles based on their destinations. This presumption is relaxed by introducing an uncertain variable. Thus, the accumulation-based MFD model with uncertainties is applied as an operation model to design a robust control strategy while the trip-based MFD model represents the plant (i.e. real network) to investigate the proposed controller performance in lessening vehicles' total time spent in the whole network.

Chapter 3 extends the perimeter control strategy and MFD modeling to a multi-region heterogeneous network to provide a more complex control framework considering location-varying region boundaries. The network is assumed to be partitioned into multiple homogeneous subregions that are grouped into two large heterogeneous regions. Accordingly, two aggregated models, i.e. region- and subregion-based MFD models, are developed to explore the dynamics of traffic propagation in the network. The traffic control problem is determined through Linear Quadratic Regulator (LQR) approach to minimize vehicles' total time spent in the overall network. Besides, a cordon selection algorithm is constructed to separate congested and uncongested regions according to the real-time congestion level of each subregion. To examine the effectiveness of the proposed control strategy, several control policies are investigated.

Chapter 4 introduces a congestion pricing scheme grounded on Model Predictive Control (MPC) for a multi-region urban network as an elegant control approach to simultaneously achieve

reducing vehicles' total time spent as well as maintaining revenue neutrality of the whole network. Two MFD models in different aggregated levels, i.e. prediction and simulation models, are respectively introduced. The former is utilized to design the controller, and the latter is employed to investigate the efficiency of the proposed pricing scheme. To ease the controller computational efforts, the prediction model ignores detailed regional origin-destination information and the receiving capacity of the destination region while these are involved in the simulation model considering current-best route choices which serves as the plant (i.e. reality). The optimal time-varying toll in each region impacts travelers' route selections, such that they can choose the route with the lowest travel costs. Besides, MPC optimization depends on an accurate dynamic model that can capture the intricate nonlinear traffic dynamics. The Long-Short Term Memory (LSTM) neural network is adopted to reap an accurate estimation of inter-region transfer flow that is integrated into the prediction model of the controller.

Keywords: *Network Fundamental Diagram, Congestion Management, Perimeter Control, Congestion Pricing, Accumulation-based Model, Trip-based Model, Cordon Queue, Traffic Flow, Switching System, Sliding Mode Control, Linear Quadratic Regulator, Model Predictive Control, Deep Learning, Long-Short Term Memory Network*

Authorship attribution statement

Chapters in this thesis published as papers:

Chapter 2 of this thesis is published as [Y. Li, M. Yildirimoglu, and M. Ramezani, "Robust perimeter control with cordon queues and heterogeneous transfer flows", *Transportation Research Part C: Emerging Technologies*, May 2021, 126, 103043].

I designed Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review and editing.

Chapter 3 of this thesis is published as [Y. Li, R. Mohajerpoor, and M. Ramezani, "Perimeter control with real-time location-varying cordon", *Transportation Research Part B: Methodological*, 150, pp.101-120].

I designed Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review and editing.

Chapter 4 of this thesis is submitted to [*Transportation Research Part B: Methodological*, and it is under peer-review.

I designed Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review and editing.

Attesting authorship attribution statement

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Student Name: Ye Li
Signature:
Date: 28/12/2021

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Supervisor Name: Mohsen Ramezani
Signature:
Date: 28/12/2021

Acknowledgements

To write this "thank you" is to say farewell to my school-days. There is no doubt that I am fortunate to be a member of TransportLab at the University of Sydney. First of all, I would like to say thank you to my supervisor, Dr. Mohsen Ramezani, for his strict requirements, systematic guidance, and non-stop encouragement to me all through my doctoral study. From subject matter resolution to thesis writing, this paper embodies Mohsen's endless painstaking efforts. His recommendation and encouragements are necessary for me to make each progress. His profound knowledge, meticulous educational attitude, eager educational vision, profound insight, and honest fashion have delivered an inexhaustible wealth to my life. A few phrases are now not ample to specific my gratitude, I would like to use this thesis to exhibit my most honest thanks to Mohsen.

Then, my sincere thanks goes to Prof. David Levinson and Dr. Emily Moylan. I recognize their earnest help and conducive hints when I encountered troubles with my research and presentations. I would like to specific my honest gratitude to Dr. Reza Mohajerpoor and Dr. Mehmet Yildirimoglu, with whom I had opportunities to have shut collaboration on quite a few components of this thesis. I would like to thank my colleagues and pals in Sydney, Amir Valadkhani, Dong Zhao, Dr. Mengying Cui, Behman Lahoorpoor, Ang Ji, Dr. Hao Wu, Yadi Wang, and Lifang Zhang. It is pleasure to be aware of you and to be capable to work together. These have been incredible and memorable days in my life.

Last however no longer least, my different and deepest thanks go to my parents, Yanbin Li and Hua Zhu. Without their non-stop aid and care, I should in no way develop up in such joyous and comfy surroundings nor have the braveness to confront any barriers on my way to success. And importantly, I would like to thank Dr. Yuyu Guo. It is your understanding, support, and encouragement that make this thesis possible. My heart swells with gratitude to you all.

Sydney, Australia. December 28, 2021

Ye Li

Contents

Statement of originality ii

Abstract iii

Authorship attribution statement vi

Attesting authorship attribution statement vii

Acknowledgements viii

Contents ix

List of Figures xii

List of Tables xvii

Chapter 1 Introduction 1

1.1 Research motivation and background 2

1.2 Thesis objectives 6

1.3 Thesis contributions 8

1.4 Thesis outline 10

Chapter 2 Robust Perimeter Control with Cordon Queues and Heterogeneous Transfer Flows 13

2.1 Introduction 14

2.2 Macroscopic traf c modeling 17

2.2.1 Trip-based MFD model: accounting for perimeter control and cordon queue dynamics 18

2.2.2 Accumulation-based MFD model: accounting perimeter control and heterogeneous transfer ow 25

2.3 Controller design 27
2.3.1 Sliding mode control 27
2.3.2 Improved Bang-Bang controller 30
2.4 Numerical experiments 31
2.4.1 Models consistency and comparison 31
2.4.2 Controller performance 34
2.5 Summary 39

Chapter 3 Perimeter Control with Real-time Location-varying Cordon 42
3.1 Introduction 42
3.2 Network traf c ow modelling 46
3.2.1 Subregion-level traf c ow model 47
3.2.2 Region-level traf c ow model 49
3.2.3 Relationship between subregion-level measurements and region-level variables 52
3.3 Controller design 54
3.3.1 Cordon switching algorithm 55
3.3.2 LQR control framework 58
3.3.2.1 The steady-state traf c condition 59
3.3.2.2 Linearized dynamic equations 60
3.3.2.3 Controller synthesis 62
3.4 Numerical experiments 63
3.4.1 Settings 63
3.4.2 Results and discussion 64
3.5 Summary 71

Chapter 4 Real-time quasi revenue-neutral congestion pricing in cities: crediting drivers to avoid CBDs 74
4.1 Introduction 74
4.1.1 Background and motivation 74
4.1.2 Related works 75
4.1.3 Contributions and chapter structure 78

4.2 Optimization model 79
4.2.1 MFD traf c dynamics 80
4.2.2 LSTM for inter-region transfer ow estimation 82
4.3 Model predictive control 84
4.3.1 Simulation plant 85
4.3.2 The MPC formulation 87
4.4 Numerical experiments 90
4.4.1 Simulation settings 90
4.4.2 Results and discussion 92
4.5 Summary 97

Chapter 5 Conclusions and Future Research 98
5.1 Robust perimeter control with cordon queues and heterogeneous transfer ows 98
5.2 Perimeter control with real-time location-varying cordon 99
5.3 Real-time quasi revenue-neutral congestion pricing in cities: crediting drivers to avoid CBD 100

Appendix A Appendix for Chapter 3: Perimeter Control with Real-time Location-varying Cordon 102
A1 Nomenclature 102
A2 Descriptions of linearized matrices A_s and B_s 104
A3 Estimating Region accumulation STD 105

Appendix B Appendix for Chapter 4: Real-time quasi revenue-neutral congestion pricing in cities 108
B1 The LSTM neural network 108
B1.1 The LSTM training procedure and validation 110

Bibliography 111

List of Figures

2.1 (a) Production MFD used in Accumulation-based MFD models, (b) Speed MFD used in Trip-based MFD models, (c) the two-region network structure with cordon queues, and (d) different trip length distributions (sampled of ine and used in the trip-based MFD model) 18

2.2 The control architecture of the proposed perimeter control method based on SMC. The accumulation-based MFD model is employed to design the controller. The trip-based MFD model is considered as the plant. The solid arrows represent online procedures. 29

2.3 (A) Slow varying two-staircase demand pro le; Accumulation evolution when trip lengths follow exponential distribution: (B) accumulation-based MFD model, (C) trip-based MFD model. 33

2.4 Accumulation evolution when trip lengths are randomly distributed: (A) accumulation-based MFD model, (B) trip-based MFD model; Evolution of internal and external ows: (C) accumulation-based MFD model, (D) trip-based MFD model; Evolution of (t): (E) accumulation-based MFD model, (F) trip-based MFD model. 35

2.5 (A) Peak-hour demands; Accumulation evolution: (B) no control, (C) improved bang-bang control, (D) proposed sliding mode control; Perimeter control actions: (E) improved bang-bang control, (F) proposed sliding mode control; (G) Production-MFD of Region 1, (H) Production-MFD of Region 2. The hysteresis loop in MFDs and the drop in average region speed are consequences of cordon queues. Note that the controllers update interval is 60 seconds. 38

2.6 The histogram of vehicles' travel times under three control methods (23202 vehicles in total): (A) no control; (B) I-BB control, and (C) the proposed SMC. Note that the controllers update interval is 60 seconds. 40

3.1 A large-scale two-region heterogeneous network comprising 19 homogeneous subregions. The borderline separating the two regions is highlighted using a solid

green and a dashed red line, representing gating ow exchange from Region 2 to 1 and from Region 1 to 2, respectively. (A) The initial boundary of the two regions, Region 1 encompasses Subregions 1 to 12, and Region 2 is formed by Subregions 13 to 19, where Subregion 19 is not directly in uenced by the perimeter control traf c signals. (B) First cordon switching when Subregion 17 becomes uncongested and is allocated to Region 1. (C) Second cordon switching where Subregion 8 becomes congested and is allocated to Region 2. Note that the proposed controller can possibly shrink Region 2 to only include one subregion, e.g., Subregion 19. Details of modeling are given in Section 3.2. 46

3.2 The closed-loop architecture of the proposed dynamic cordon perimeter control method based on LQR. The thick black arrows indicate the triggered operations at the switching time-steps,. The blue arrows indicate the continuous ow of information. Note that the subregion-level model is employed as the traf c simulator to examine the controller performance. The region-level model is adopted to design the controller. (For interpretation of the colours in this gure, the reader is referred to the web version of the manuscript). 56

3.3 (A) Regional demands in the network during the simulation period for the static cordon scenarios. The values of the (regional) demands would alter once the region boundaries change. However, the OD demands in the subregion-level are identical across all control scenarios. (B) (t) and (C) (t) parameters variations with time when applying the LQR + static cordon control algorithm. 65

3.4 Results of the no-control scenario ($U_{12}(t) = U_{21}(t) = 0:9$): (A) subregional and (B) regional accumulations for the whole simulation period. The accumulations of several subregions (i.e. Subregions 13-19) reach to (or very close to) the jam accumulation $n^{jam} = 10000$ and face gridlock during a long period, and accordingly Region 2 reaches to the hypercongested regime. 65

3.5 Subregions' accumulations with: (A) BB + static cordon; (B) LQR + static cordon; (C) BB + dynamic cordon (= 5); and (D) LQR + dynamic cordon (= 5). The proposed LQR + dynamic cordon method results in a more homogeneous accumulation

distribution in the whole network as well as shorter network clearance time compared to the other control strategies. 67

3.6 Snapshots of subregions' accumulations over time when implementing the LQR controller with dynamic cordon algorithm (= 5). Each snapshot corresponds to a cordon switching instant during the simulation. The border of the regions and the magnitude of perimeter control outputs in each snapshot is demonstrated by a dashed line ($\psi_{12}(t)$) and a solid line ($\psi_{21}(t)$). The color of the dashed and solid lines respectively indicate the relative magnitude, with red and green corresponding to U^{min} and U^{max}, respectively. The colorbar indicates the level of accumulation in each subregion. (For interpretation of the colors in this gure, the reader is referred to the web version of the manuscript). 69

3.7 Perimeter control outputs during the simulation period when applying the following control strategies: (A) BB + static cordon; (B) LQR + static cordon; (C) BB + dynamic cordon (= 5); and (D) LQR + dynamic cordon (= 5). 70

3.8 Regional accumulations during the simulation period when applying the following control strategies: (A) BB + static cordon; (B) LQR + static cordon; (C) BB + dynamic cordon (= 5); and (D) LQR + dynamic cordon (= 5). In the dynamic cordon scenarios, the accumulation in Region 2 reduces around time 5000 [s] because of the location-varying cordon that only Subregion 19 is in Region 2. Then the accumulation increases since more subregions are added to Region 2, see Fig. 3.6. 72

3.9 Production MFDs obtained from simulations when applying the following control strategies: (A) BB + static cordon; (B) LQR + static cordon; (C) BB + dynamic cordon (= 5); and (D) LQR + dynamic cordon (= 5). Note that P_1 and P_2 represent actual production MFDs obtained from the subregion-level model, and they are compared against those estimated from Eq. (3.5). Note that a subregion might transfer to a different region over time due to the cordon switching, resulting in jumps in MFDs with dynamic cordon. 73

4.1 (a) Network structure. A large-scale heterogeneous network comprising 7 homogeneous regions. Region 7 is the CBD of the network. (b) Three possible routes from Region 1 to Region 4. Green arrows mean less travel cost routes (greater travel distance but

less toll) while the blue arrow represents higher travel cost route (shorter route, more congested route, more toll). (c) Well-de ned MFD of each region. The dashed line indicates the critical accumulation with respect to the maximum production MFD. 80

4.2 LSTM neural networks development and the application in MPC. 83

4.3 The control architecture of the proposed congestion pricing management approach based on MPC. The aggregated optimization MFD model including the LSTM predictor for the percentage of inter-region transfer ows is employed to design the controller. The detailed MFD model (simulation plant) is considered to represent real multi-region traf c network. 84

4.4 Relationship between different time intervals and time steps of MPC formulation. 88

4.5 Regional accumulation evolution for the whole studied period obtained from simulations when applying the following pricing strategies: (a) No toll; (b) PI with $^{min} = 0$; (c) PI with $^{min} =$ tt VOT; (d) MPC with $^{min} = 0$ and $= 0$; (e) MPC with $^{min} =$ tt VOT and $= 0$; (f) MPC with $^{min} =$ tt VOT and $= 0{:}5$. In the No toll scenario, several regions (i.e. Regions 1, 3, 4, 5, 7) become congested, while Regions 2 and 6 are uncongested. Utilizing MPC revenue-neutral pricing method contributes to a more homogeneous accumulation distribution in the whole network as well as shorter network clearance time compared to the PI control. 94

4.6 Toll in each region during the simulation period when applying the following pricing strategies: (a) PI with $^{min} = 0$; (b) PI with $^{min} =$ tt VOT; (c) MPC with $^{min} = 0$ and $= 0$; (d) MPC with $^{min} =$ tt VOT and $= 0$; (e) MPC with $^{min} =$ tt VOT and $= 0{:}5$. For all scenarios, $^{max} = 5$. The average toll paid per vehicle for the ve scenarios are \$1.41, \$1.15, \$1.57, \$0.92, and \$0.09, respectively. 95

A.1 Equiout ow lines in the average occupancy vs. STD occupancy plane. Each solid line represents the variation of the STD of density with respect to the average density for a network with out ow constrained to a certain bin. In this gure, represents the total out ow of the network. The data for this gure is obtained as Figure 4 in Ramezani et al. 2015. 107

B.1 Illustrations of the LSTM neural network with repeating module that allows information to be passed from one step to the next. 109

List of Tables

2.1 The Improved Bang-Bang Control Values for $[U_{12}(t); U_{21}(t)]$ 30

2.2 Network-level statistics (averaged over 10 runs) for the three scenarios. The numbers in parenthesis indicate the improvement in performance compare with the no control case. Cycle time denotes the controller updating intervals. 39

3.1 Total time spent (expressed in 10^3 [veh s]) in the network when implementing various control strategies. Note that the no-control strategy has resulted in the network gridlock. The numbers in parenthesis indicate the relative reduction in TTS against the no-control scenario. 68

4.1 Network-level statistics (averaged over 10 replications) for the studied scenarios. The numbers in parenthesis indicate the improvement in performance compared with the No toll scenario. The total number of vehicles is 66064 [veh]. Note that "$toll^{min} < 0$" means the minimum toll in related scenario is negative (i.e. $toll^{min} = -tt \cdot VOT$). 93

CHAPTER 1

Introduction

With the rapid economic development and improvement of personal living standards, private vehicles have become a major mode of transport for people. However, the ensuing traf c congestion, air pollution, and traf c safety are prevalent phenomenons in cities that affect people's daily lives. With the growth of travel demand, the problem of traf c congestion is becoming more obvious and imminent, which restricts the quality of life in cities, that necessitates effective investigations for congestion alleviation. Therefore, it is a vital research direction to promptly detect traf c congestion in metropolitan cities and take effective control measures to enhance traf c mobility and accessibility. This dissertation develops state-of-the-art traf c congestion mitigation approaches based on network modeling and control for large-scale urban transportation networks.

Recently, several cities have established intelligent traf c control systems to monitor traf c situations in real-time to improve the quality of travel. An ef cient way to solve urban traf c congestion is reasonable management and control methods to limit the redundant traf c demand entering the congested network. This requires establishing a complex relationship between traveler behaviors and traf c status. In addition, recent advances in machine learning tools offer an opportunity to more precisely predict traf c states (e.g. traf c ow, travel time and speed, etc.) to optimize urban traf c systems.

In this introductory chapter, Section 1.1 interprets thesis motivation and concisely introduces research background in current literature, while the corresponding detailed literature review will be provided at each chapter in due course. Section 1.2 highlights the objectives that the thesis follows to achieve. Section 1.3 elaborates the main contributions of the thesis in each chapter. Finally, Section 1.4 depicts the outline of the thesis.

1.1 Research motivation and background

Traf c congestion occurs when the traf c demand exceeds the road or network capacity that leading to traf c delay. Traf c congestion imposes dramatically negative effects on people and society, such as excessive travel times, high economic losses, safety issues, etc. According to a congestion impact study from the consulting rms INRIX and Centre for Economics and Business Research (Cebr) around four countries (i.e. UK, USA, France, and Germany) in 2014, the annual cost of traf c congestion in the UK will surge 63 percent by 2030 to £21 billion. The annual cost includes direct costs (e.g. the effect on the price of goods caused by the additional time wasted, and the fuel consumption) and indirect costs (e.g. the loss of labor productivity), and it was estimated that the congestion caused 22.5 billion dollars loss in France, and for 124 billion dollars loss in the USA in 2013. The fuel wasted in congestion generates an additional 1.9 megatons of carbon dioxide emissions in France and 8.6 megatons in the US. Besides, British drivers spend approximately 124 hours stuck in gridlock annually, and this is set to rise to 136 hours in 2030 (CEBR and INRIX 2014). In addition, a report of HILDA Survey indicates that road users in mainland state capitals (i.e. Sydney, Melbourne, Brisbane, Adelaide, Perth) had consistently longer commute times, commonly spent around 66 minutes travelling to and from work per day in 2017 which is a 20% increase from 55 minutes in 2002. With the growing number of people expected to live in the cities, these impacts will be more pronounced. Therefore, the key to solving traf c problems is to promote the coordinated development of transport science and strategies, especially attaching great importance of transportation systems within smart cities.

Traf c congestion alleviation has been taken into account by numerous researchers and engineers who developed various control approaches corresponding to transportation demand management (TDM). These approaches are mainly focused on increasing overall network performance by encouraging people shift traf c modes (e.g. from private vehicles to public transport or active modes), traveling out of rush hours, or choosing uncongested (but long-distance) routes. Detailed strategy of such policies are as follows: (i) traf c supply and demand management, e.g. promoting transit use and carpooling (Jamal et al. 2017; Shen et al. 2021; Kumar and Khani 2021), parking management (Wang et al. 2021; Bahrami et al. 2021); (ii) land use and urban design, e.g. improving access to frequent transit service (Zhang et al. 2020; Smith et al. 2021), or modifying

road infrastructure such as the usage of dedicated high occupancy vehicles (HOV) lanes (Zang et al. 2020; Zhong et al. 2020) and the extension of bus lane operating times (Szarata et al. 2021). (iii) traf c control strategies, for instance, coordinated intersection management (Lei et al. 2019; Haddad and Mirkin 2020), freeway traf c control (e.g. ramp metering) (Han et al. 2020; Tajdari et al. 2020), route guidance (Sirmatel and Geroliminis 2017; Yildirimoglu et al. 2018), and congestion pricing (Simoni et al. 2019; Lehe 2019). Detailed overview of traf c control policies can refer to Papageorgiou et al. 2003.

Although some progress has been made in urban traf c congestion management, ef cient management and control of large-scale heterogeneous networks still remain a big challenge because of the complicated and varying traf c dynamics. Current control approaches are effective but one-sided in view that (i) local adaptive traf c signal jurisdictions take care of a small part of the whole network; (ii) prede ned xed-time or immovable-location control does not respond to day-to-day uctuations in traf c patterns; and (iii) the network-level control methods, such as coordinated signal control and congestion pricing, requires convoluted and high computational efforts. Therefore, the principal objective of this thesis is to design effective real-time management approaches for large-scale urban networks to ameliorate traf c conditions and prevent gridlock phenomena in cities. The thesis proposes two major control methods, i.e. perimeter ow control (Chapters 2 and 3) and congestion pricing (Chapter 4), to address traf c congestion. Note that the thesis objectives based on the research background are presented in due course, while the particular literature review corresponding to the related research topic is introduced in each chapter.

The scope of this thesis is large-scale heterogeneous traf c modeling and management. For macroscopic traf c modeling, Geroliminis and Daganzo 2008 observed a well-de ned low-scattered macroscopic fundamental diagram (MFD) from eld data in Yokohama, Japan. MFD provides a phenomenological relationship between mean weighted ow and vehicle accumulation in the network. This relationship, approved by empirical data from multiple cities over the world, has supported the investigate of disparate traf c control schemes at the network level. Godfrey 1969 brought the idea of MFD and its mathematical formulation was introduced by Daganzo 2007 for a single region, while MFD existence was demonstrated in Geroliminis and Daganzo

2008 with the following properties: (i) a neighborhood-sized network exhibits an MFD when the congestion distribution is roughly homogeneous; (ii) there is almost a linear relationship between the regional average flow and its out flow (i.e. trip completion rate is the rate of vehicles finish their trips), which means the regional average travel length has almost a time-constant value; (iii) an MFD exists in the large urban network that is less-sensitive to the demand.

These findings of MFD offer opportunities to control traffic flow to improve network efficiency. Property (i) provides a network-level aggregated model to depict the congestion level in cities without considering detailed (road) link information. Property (ii) brings convenience to measure regional flow from different kinds of detectors, e.g. loop detectors, video sensors, etc, while region out flow observation is burdensome. Property (iii) illustrates the detailed origin-destination demand matrix is not required to design traffic control strategies. To improve the mobility of cities, numerous control approaches (e.g. Aboudolas and Geroliminis 2013; Mahmassani et al. 2013; Aalipour et al. 2018) have been developed utilizing the MFD concept. The core of these approaches is to adjust the traffic in flow to regions based on the congestion level which is represented by the MFD.

A well-defined MFD is not a universal law since congestion is not uniformly distributed in the real traffic network, which is the essential factor that affects the shape and scatter of the MFD (see Geroliminis and Sun 2011; Daganzo et al. 2011). A homogeneous network is more desirable compare to a heterogeneous one because the regional average flow is greater when the link density variance is small. One way to reduce the heterogeneousness of the network is to divide the whole network into several homogeneous regions with small link density variances (Geroliminis and Sun 2011; Ji and Geroliminis 2012). It should be mentioned that each region has a distinct time to reach the congested regime (decreasing part of MFD). The partitioning further enables development of elegant control strategies as the MFD can be utilized once a heterogeneous network is split into a number of homogeneous regions. Note that partitioning does not indicate the region boundaries are invariant over time. It is demonstrated in this thesis that the location-varying region boundaries can improve the efficiency (in terms of vehicles' total time spent) of the network. Therefore, the proposed control methods offer advanced technologies to solve traffic problems while the practical implementations require significant efforts. Furthermore,

the traditional MFD, which can refer to accumulation-based MFD, with a rough assumption that all vehicles with the similar average trip length in a region. Recently, a few studies (e.g. Arnott 2013; Fosgerau 2015; Mariotte et al. 2017) explored trip-based MFD reformulation to address the in uence of divergent travel lengths on congestion propagation in a single region. The trip-based MFD model considers individual vehicle travel length by presuming a well-de ned relation between mean speed and vehicle accumulation of a network. Up to date, different control strategies have mostly been designed based on accumulation-based MFD models considering similar trip lengths per region (see Bichiou et al. 2020; Guo and Ban 2020). Compared to previous research, Ramezani et al. 2015 and Sirmatel and Geroliminis 2017 studied variable average travel distance in the implementation of MFD models for traf c control purposes. It is veri ed that regional mean travel length can be affected by congestion level if the network is non-uniformly congested. However, little attention has been paid to the studies for the proper settings of trip lengths in MFD-based applications. To highlight the signi cance of various trip lengths in macroscopic traf c models, this thesis offers a framework to simultaneously utilize the parsimonious accumulation-based MFD for optimization procedures and detailed characteristics of the trip-based MFD to describe a real transportation network. Besides, to ll the gap that the trip-based MFD model developed grounded on a single region, this thesis proposes a multi-region trip-based MFD model that considers cordon queues and inter-region transfer ows as well as various traveler behaviors in terms of route choices and departure time that create additional complexity.

Macroscopic (network-level) models aim at depicting the transportation network through aggregated traf c states such as mean speed, density, etc. Several effective traf c management approaches have been developed and applied at the network level to alleviate local pockets of congestion, such as perimeter ow control (Li et al. 2021b; Li et al. 2021a; Sirmatel et al. 2021; Sirmatel and Geroliminis 2021), congestion pricing (Daganzo and Lehe 2015; Liu and Geroliminis 2017; Chen et al. 2021; Mansourianfar et al. 2021), route guidance (Yildirimoglu et al. 2015; Yildirimoglu et al. 2018), and ride-sourcing operation control (Ramezani and Nourinejad 2018). Perimeter ow control has been demonstrated its ef ciency in enhancing network mobility in numerous research while there is a lack of works corresponding to congestion pricing based on MFD, especially for multi-region networks. To ll the gap in the literature, this thesis

integrates multiple control objectives (i.e. reducing vehicles' total time spent and maintaining revenue neutrality) in the optimization framework and develops a congestion pricing scheme that can track the congestion propagation in the multi-region network and consequently enlarge the system performance. This is crucial since for decades practitioners and city managers have recommended congestion pricing (i.e. special charges for traveling on congested roadways or networks) as a way to encourage more effective utilization of the transport system, as well as lessen congestion and air pollution, providing net bene ts to society and economy. Recently, various forms of congestion pricing have been implemented in many cities, such as London, Singapore, Stockholm, etc, while fairness problems should be carefully addressed. This thesis provides theoretical support for equity concerns in implementing congestion pricing by taking into account maintaining revenue-neutrality of the network.

1.2 Thesis objectives

The goal of this thesis is to develop real-time traf c management strategies for metropolitan cities to enhance the system mobility and effectiveness. To this end, the research methodology copes with two major traf c control strategies, i.e. perimeter ow control and congestion pricing, for large-scale heterogeneously congested transportation networks.

In light of the prementioned research background and motivation, the particular objectives of this thesis are divided into two categories, (i) to propose a network-level coordinated traf c signal control strategy employing advanced control theory and (ii) to develop a congestion pricing scheme integrating multiple control objectives. The detailed objectives of this thesis corresponding to each chapter are listed in the following:

Perimeter ow control

Chapter 2 principal objective is to develop a real-time, robust perimeter controller based on the Sliding Mode Control (SMC) theory to address modeling uncertainties associated with internal and external out ow estimations in accumulation-based MFD models. The modeling basis of the method should be consistent with MFD traf c dynamics. To this end, the studied network of this chapter is assumed to be partitioned

into two regions, and the perimeter controllers are located between the region boundary. The second objective is to introduce a two-region trip-based MFD model considering cordon queue dynamics and to test the proposed robust perimeter control on the trip-based MFD model to investigate its effectiveness. The controller is applied to the trip-based MFD model that accounts for the effects of cordon queues which are considered as one aggregated queue consistent with the physics of MFD.

The key objective of Chapter 3 is to introduce a perimeter flow control method with location-varying cordon to reduce congestion in a heterogeneous large-scale urban network. To further investigate the effectiveness of perimeter control method developed in Chapter 2, the network is assumed to be divided into 2 concentric regions that initially consist of 12 (in periphery region) and 7 (in the centered (protected) region) homogeneous subregions. The location-varying cordon selection algorithm designed in this chapter has enabled effective management of spatial and temporal heterogeneity of demand and propagation of congestion by recurrently allocating uncongested subregions to the peripheral region and concentrating congested subregions into the protected region. Moreover, the perimeter control strategy is solved through the Linear Quadratic Regulator (LQR), while the optimization model and the plant (i.e. testbed transportation network) are in different levels, i.e. regional and subregional. That is, the regional model is used to design the controller while the subregion-level model is the model that performs the traffic simulator. Additionally, the second objective of this chapter is to design a more realistic model by permitting vehicles to cross the region boundaries multiple times.

Congestion pricing

The main objective of Chapter 4 is to construct a novel real-time congestion pricing scheme established on the Model Predictive Control (MPC) concept to simultaneously achieve dual goals of improving network performance and keeping revenue neutral (i.e. total toll collected be close to zero) of a multi-region network. The network is assumed to be partitioned into 7 regions each with a well-defined MFD. To simplify the

optimization design procedure, this chapter introduces two models with dissimilar aggregated levels, i.e. optimization and simulation models. Note that the MPC framework requires accurate estimation of inter-region transfer flows in the optimization model. Therefore, another objective of this chapter is to incorporate the deep-learning method (e.g. Long-short Term Memory (LSTM) network) into MPC to precisely estimate the transfer flows. Besides, a more detailed simulation model is introduced by considering a dynamic number of current-best shortest paths based on which region the vehicle is located.

The consecutive section describes contributions of each chapter related to the aforementioned objectives.

1.3 Thesis contributions

On the basis of thesis motivation and objectives, the contributions of this thesis at each chapter are as follows.

Chapter 2: Robust perimeter control with cordon queues and heterogeneous transfer flows

The chapter develops a network-level control strategy for urban heterogeneous transportation networks named perimeter flow control. The network is assumed to be divided into two regions each with a well-defined MFD. The perimeter controller is implemented utilizing coordination among traffic signals installed at boundaries between the regions to regulate the rate of transferring vehicles from one region to another to prevent hypercongestion of the network. The traffic modeling is consistent with the parsimonious accumulation-based MFD model that offers a relationship between regional trip completion rate and accumulation, and accordingly, the two-region MFD model is formulated on the basis of vehicle conservation equations. An uncertain variable is introduced to release the hypothesis in the percentage of internal and external transfer flows in the accumulation-based MFD model. Another outstanding contribution in this chapter is that a multi-region trip-based MFD model is proposed integrating more traffic states,

such as traveling and queued vehicles. The trip-based MFD connects regional mean speed and vehicle accumulation that allows us to build a disaggregated yet city-scaled model where commuters detailed travel information can be taken into account. In this chapter, the accumulation-based MFD model is employed to design the controller while the trip-based MFD model is adopted as the plant (i.e. real network) to explore the effectiveness of the proposed control strategy. The control problem is tackled through SMC since it is robust to deal with uncertainties in the model.

Chapter 3: Perimeter control with real-time location-varying cordon

This chapter moves a further step towards the development of the perimeter ow control framework for multi-region urban networks. The modeling is an extension of the MFD modeling in Chapter 2, while the network is partitioned into two regions and each with a number of subregions. We assume a non-linear, low-scattered MFD exists for each subregion. To take into consideration the effects of link density heterogeneity on the MFD, two distinctive aggregation models are applied. The region-level model incorporates variant trip lengths in each region, considers the heterogeneity of spatial congestion distribution, and embeds a simple route choice model because vehicles can pass region boundaries multiple times, while the subregion-level model describes detailed traf c dynamics of each subregion, integrates a current-best route choice model, and considers the effect of the boundary and receiving capacities. In addition, the optimal control problem is solved by the LQR scheme by considering the regional MFD model for controller synthesis and the subregional MFD model as traf c simulators (plant). Another essential contribution in terms of network-level control is to propose a novel traf c responsive cordon switching algorithm integrated within a location-varying perimeter control method. That is splitting the congested and uncongested regions recurrently in real-time. Moreover, the regional model is more realistic by permitting vehicles to pass region boundaries multiple times.

Chapter 4: Real-time quasi revenue-neutral congestion pricing in cities: crediting drivers to avoid CBD

The control problem of a large-scale traf c network, comprising seven homogeneous regions is formulated. The congestion pricing scheme as a means of providing congestion control has been studied in considerable research, while its development based on MFD is carefully addressed, especially for multi-region networks. The tolls in each region vary grounded on the current congestion level that is obtained through the MPC controller. To take into account the complexity of traf c dynamics as well as equity issues of pricing, the proposed MPC controller integrates dual objectives, i.e. reducing vehicles' total time spent and maintaining revenue neutral of the whole network. The urban region traf c modeling is based on the MFD while in different levels, i.e. a more aggregated model for controller optimization while a detailed model is utilized for the traf c simulation. However, the MPC controller requires an accurate optimization model to track the convoluted and varying traf c dynamics, such that a deep-learning LSTM network is incorporated to precisely estimate the percentage of inter-region transfer ows in the optimization model. Besides, the simulation model along with a simple route choice procedure in which the number of current-best paths is various based on the location of vehicles.

1.4 Thesis outline

This thesis contains 5 chapters that are concisely presented in the following paragraphs. The principal 3 chapters (excluding Chapter 1, Introduction, and Chapter 5, Conclusion and future research) are organized based on different network-level control strategies. Chapter 2 and Chapter 3 corresponding to perimeter ow control that covers the research toward the application of various novel control theories as well as the development of large-scale urban traf c modeling based on the MFD. Chapter 4 proposes a real-time congestion pricing scheme for heterogeneous urban networks based on the MFD framework. It should be noted that each chapter is a complete stand-alone research article comprising abstract, introduction, methodology, results, and conclusions with its related (mathematical) nomenclatures.

Chapter 2 is the pioneer to develop a multi-region trip-based MFD model considering disaggregated and detailed traf c states, such as traveling and queued vehicles. The studied network is assumed to be divided into two regions. The perimeter controller can be realized by a coordinated traf c signal controller that is located at the region boundary between regions to manipulate the inter-region transfer ows in order to prevent the occurrence of hypercongestion. The control problem is solved through the SMC approach. The parsimonious accumulation-based MFD model is utilized to develop the controller while the trip-based MFD model is regarded as the plant. The results of the method are presented in:

> Y. Li, M. Yildirimoglu, and M. Ramezani, "Perimeter traf c ow control: a robust control approach based on trip-based and accumulation-based Macroscopic Fundamental Diagrams", *TRANSW*, September 2019.
>
> Y. Li, M. Yildirimoglu, and M. Ramezani, "Integrating the dynamics of boundary queues in two-region trip-based Macroscopic Fundamental Diagram models", *BTER* online conference, January 2020.

Chapter 2 is structured as a self stand-alone article published as:

> Y. Li, M. Yildirimoglu, and M. Ramezani, "Robust perimeter control with cordon queues and heterogeneous transfer ows", *Transportation Research Part C: Emerging Technologies*, May 2021, 126, 103043.

Chapter 3 is the pioneer to propose the location-varying perimeter ow control strategy as a city-scale traf c management approach for heterogeneous transportation networks based on the MFD modeling. The studied network is assumed to be divided into two regions each with a few subregions. A novel switching cordon algorithm is proposed to capture the spatial and temporal traf c demand uctuations by reallocating the congested and uncongested subregions so as to improve the network performance. The optimal inter-region transfer ow rates are determined through LQR theory. Two different models are employed, regional and subregional models, to design the controller and to examine its effectiveness, respectively. The preliminary results of the method are presented in:

Y. Li, R. Mohajerpoor, and M. Ramezani, "Perimeter flow control with time-varying cordon based on macroscopic fundamental diagram", TRANSW, November 2018.

Y. Li, R. Mohajerpoor, and M. Ramezani, "Perimeter flow control with time varying cordon based on Macroscopic Fundamental Diagram", TRISTAN, June 2019.

Chapter 3 is organized as a self stand-alone article published as:

Y. Li, R. Mohajerpoor, and M. Ramezani, "Perimeter control with real-time location-varying cordon", Transportation Research Part B: Methodological, vol. 150, pp.101-120.

Chapter 4 proposes a real-time network-level congestion pricing scheme based on the MFD modeling. The large-scale heterogeneous network is partitioned into seven regions. The optimal toll in each region is modified by the MPC approach considering bi-objective, i.e. keeping revenue-neutral and diminishing vehicles' total time spent. The modeling significance of this chapter is to incorporate (deep-learning) LSTM network into the MPC framework to accurately estimate inter-region transfer flows of the optimization model. Besides, a more detailed MFD model is constructed taking care of a varying number of current-best shortest paths. The preliminary results of this chapter are presented in:

Y. Li and M. Ramezani, "Optimal real-time pricing in cities: crediting drivers to avoid CBDs", TRANSW, December 2021.

Chapter 4 is a self stand-alone article that is submitted to Transportation Research Part B: Methodological, and is currently under review.

Finally, Chapter 5 closes this thesis by summarizing key contributions of each chapter, explaining the possible applications of the proposed control management strategies, and introducing the future research directions.

CHAPTER 2

Robust Perimeter Control with Cordon Queues and Heterogeneous Transfer Flows

---

Accumulation-based Macroscopic Fundamental Diagram (MFD) model is widely employed to design perimeter control methods to improve traf c operation in urban networks. While the accumulation-based MFD assumes a low-scatter, non-linear relationship between region production and accumulation, the out ow relationship in formulating dynamics of multi-region networks requires simplifying assumptions. The existing perimeter control methods are grounded on accumulation-based MFD models where the number of transferring vehicles is approximated by the ratio of the instantaneous number of vehicles based on their destinations. Moreover, perimeter control may lead to more vehicles queuing at the region boundary (i.e. cordon queues) which add local impediments on traveling vehicles and impact the accuracy of well-de ned MFDs. To address these shortcomings under time-varying conditions, this chapter develops a robust perimeter control method based on the Sliding Mode Control to minimize total travel time in the entire network. To test the performance of the proposed control method, a trip-based MFD model is developed that accounts for cordon queues and various trip lengths of individual travelers. In this chapter, two-region accumulation-based and trip-based MFD models are compared through numerical experiments. The results pinpoint the proposed robust perimeter control method can effectively alleviate congestion and improve network ef ciency during traf c rush hours.

2.1 Introduction

Excessive traf c congestion has become a quotidian problem in urban cities with the growing transport demand. To alleviate congestion and improve the network ef ciency, several management approaches for large-scale transport systems have been developed, such as perimeter ow control (e.g. Geroliminis et al. 2012; Fu et al. 2020; Ingole et al. 2020), congestion pricing (e.g. Zheng and Geroliminis 2020), and coordinated ramp metering (e.g. Han et al. 2020). Perimeter ow control was found to be an effective approach to manipulate traf c ows between regions to maintain vehicle accumulation in regions under a prede ned value (i.e. critical accumulation). The perimeter control is implemented using coordination among traf c signals installed at boundaries between the regions to regulate the rate of transferring vehicles from one region to another to prevent hypercongestion.

Recently, sizable research studied perimeter control approaches, see Haddad and Shraiber 2014; Ramezani et al. 2015; Yang et al. 2018; Lei et al. 2019; Haddad and Mirkin 2020 among others. These approaches utilize accumulation-based MFD model grounded on a parsimonious, low-scattered, and non-linear relationship between vehicle accumulation [veh] and production [veh m/s] of the network, see Fig. 1(a). While the congestion is homogeneously distributed in the network, the aggregated accumulation-based MFD can be employed to model the traf c dynamics (see Geroliminis and Daganzo 2008). The MFD model was originally proposed by Godfrey 1969. The theoretical concepts of MFD were rst established in Daganzo 2007, and its existence was demonstrated in Geroliminis and Daganzo 2008 utilizing eld data from downtown Yokohama, Japan. The MFD dynamic foundation of a single region was established in Daganzo 2007, where the evolution of traveling vehicle accumulation is modeled by tracking the incoming traf c demand and endogenous traf c out ow. In networks without uniformly distributed vehicle densities, MFD exhibits a hysteresis loop. To cope with this issue, one recognized approach is to partition a heterogeneously congested network into multiple regions with homogeneous congestion distributions (see Saeedmanesh and Geroliminis 2016; Saedi et al. 2020). The partitioning further enables to manipulate the transferring ows between regions by the traf c signals located at region boundaries. The perimeter ow control is a coordinated action among these traf c signals.

Perimeter control grounded on accumulation-based MFD has been introduced for single-region (e.g. Daganzo 2007; Keyvan-Ekbatani et al. 2012) and multi-region networks (e.g. Ramezani et al. 2015; Kouvelas et al. 2017; Csikós et al. 2017; Yang et al. 2018; Yang et al. 2019). The perimeter control settles at the peripheries between regions to manipulate the percentages of transfer ows. However, restricting transferring vehicles might result in queuing vehicles at region boundaries. The queuing vehicles affect the traf c ow dynamics and consequently might in uence the out ow of the region and accuracy of well-de ned MFD Haddad 2017. So the cordon queue impacts should be carefully addressed. Ni and Cassidy 2019 proposed to model time-varying queue impacts on network production by re-scaling MFD at any time step, where the MFD is scaled downward or upward respectively when the cordon queue size grows or shrinks. In this chapter, we develop a robust perimeter controller that tackles the effect of cordon queues on congestion propagation.

Many studies have constructed perimeter control framework using techniques such as Bang-Bang control (see Daganzo 2007; Aalipour et al. 2018), Proportional-Integral (PI) control (see Keyvan-Ekbatani et al. 2015a; Mohajerpoor et al. 2020; Ingole et al. 2020), Model Predictive Control (MPC) (see Geroliminis et al. 2012; Fu et al. 2017; Kim et al. 2019), and Linear Quadratic Regulator (LQR) (see Haddad and Shraiber 2014; Haddad 2015). The model-based perimeter controllers for multi-region networks require an assumption for the transferring ows in the accumulation-based MFD model. In most cases, the ratio between the internal and external out ows is assumed to be equal to the ratio of the instantaneous number of vehicles based on their destinations, (e.g. in Ramezani et al. 2015; Zhong et al. 2018; Yang et al. 2019). This assumption imposes modeling inaccuracy and would lessen the performance of model-based controllers, speci cally the ones based on model predictions such as MPC.

this chapter proposes a robust perimeter controller based on Sliding Mode Control (SMC) method to address the inherent mismatch between the traf c ow model (i.e. MFD) and real world traf c dynamics. This is crucial because MFD is inherently a spatially and temporally aggregated model and neglects several (micro) aspects of traf c. Speci cally, the proposed SMC perimeter controller is designed to tackle the following issues in multi-region MFD modeling; (i) the effect of cordon queues on congestion heterogeneity and MFD and (ii) the split ratio between

internal and external out ows. The SMC is a variable structure control method that accounts for uncertainties in modeling formulation and measurements. The SMC enables discontinuous and nonlinear control gains to stabilize the system by manipulating the state of the system to prede ned (sliding) surfaces in nite time and maintain the states on them thereafter (e.g. Aalipour et al. 2017; Bichiou et al. 2020).

One notable characteristic of employing accumulation-based MFD model is the numerical tractability that offers optimization and control opportunities. To this end, several studies assume an average trip length representing all vehicles within the network (e.g. Loder et al. 2019; Huang et al. 2020). Recent studies highlight the effect of various trip lengths on accumulation-based MFD model (Mariotte et al. 2017; Batista and Leclercq 2019). In particular, Mariotte et al. 2020b demonstrated that the calibration of trip lengths is crucial for accuracy of the MFD models. To take into consideration the different trip lengths of vehicles, the trip-based MFD was rst proposed by Arnott 2013, and applied in later research such as Lamotte and Geroliminis 2018; Yildirimoglu and Ramezani 2020; Jin 2020. Trip-based MFD assumes a well-de ned relation between vehicle accumulation and mean speed of the network, see Fig. 1(b). Mariotte et al. 2017 employed a one-region network to compare the accumulation-based MFD and trip-based MFD models, and pinpointed incoherence during transient phases. When the system operates under steady state, the two models are consistent with each other (i.e. analogous accumulation evolution), while in the loading and unloading phases, they exhibit substantial differences. Similar results are observed in multi-region networks Mariotte and Leclercq 2019. A multi-region trip-based MFD model requires more complicated states to describe traf c dynamics, e.g. vehicle accumulations towards internal or external destinations and their related trip lengths. this chapter introduces a two-region trip-based model that considers cordon queues to be employed as a testbed to investigate the proposed robust perimeter controller.

The contributions of this chapter are threefold: (i) designing a real-time, robust perimeter control based on Sliding Mode Control (SMC) method to tackle the inherent uncertainties in multi-region accumulation-based MFD model, (ii) developing a two-region trip-based MFD model considering cordon queue dynamics, and (iii) testing the proposed robust perimeter control on the trip-based MFD model to investigate its effectiveness. In this chapter, it is assumed the urban network is

divided into two regions. The perimeter control regulates the ows between the two regions, and vehicles waiting to cross the boundary represent the cordon queues, as delineated in Fig. 1(c). An accumulation-based MFD model with uncertain parameters is established to develop a robust traf c control method. The controller is applied to the trip-based MFD model that accounts for the effects of cordon queues which are considered as one aggregated queue consistent with the physics of MFD. To further investigate the SMC ef ciency, an improved Bang-Bang controller is introduced for comparison.

The remainder of this chapter is organized as follows. Section 2.2 introduces the two-region accumulation-based MFD model and the trip-based MFD model accounting for cordon queues. Section 2.3 develops an SMC scheme grounded on the accumulation-based MFD model to minimize the network total time spent. Also, an improved Bang-Bang controller is introduced for comparison. Section 2.4 employs numerical experiments to investigate the two models as well as to evaluate the controller performance, under two disparate demand pro les. Finally, in Section 2.5, the paper closes with the summary and future research directions.

2.2 Macroscopic traf c modeling

In this section, the two-region trip-based and accumulation-based MFD models are introduced. The former is considered as the plant that accounts for the perimeter control, cordon queues, and heterogeneous transfer ows. The trip-based MFD model, which is the more detailed model, tracks the number of vehicles grouped as queued and traveling with internal or external destinations. On the other hand, the proposed parsimonious accumulation-based MFD model is employed to derive the robust perimeter control based on SMC while neglecting the dynamics of cordon queues.

(A) (B)

(C)

(D)

FIGURE 2.1. (a) Production MFD used in Accumulation-based MFD models, (b) Speed MFD used in Trip-based MFD models, (c) the two-region network structure with cordon queues, and (d) different trip length distributions (sampled of ine and used in the trip-based MFD model)

.

2.2.1 Trip-based MFD model: accounting for perimeter control and cordon queue dynamics

The trip-based MFD model is a disaggregated yet city-wide model that takes into account individual travel details such as trip length and departure time (see Mariotte et al. 2017). In this section, we introduce the event-based implementation of the trip-based MFD model (Mariotte et al. 2017; Yildirimoglu and Ramezani 2020) for a network divided into two regions with perimeter control and considering the cordon queue dynamics. The events contain departure (from the origin) and arrival (to the destination), joining and leaving cordon queues for vehicles with origin in one region and destination in another region, and changes in perimeter control values. Every event directly affects the accumulations in regions and hence regions mean speeds.

The network is assumed to be partitioned into two regions, denoted by Region $I, J = \{1, 2\}$. The exogenous travel demands entering the network are associated to different trip lengths and departure times (from their origins). Let $N_{IJ}^{Q}(t)$ [veh] denote the number of queuing vehicles at the boundary in Region I with the destination in Region J at time t. Note that the queuing vehicles only consist of vehicles traveling between regions, and the vehicles suffer the queuing process at most once during their whole trips. The traveling accumulations in Region I, $N_I^T(t)$ [veh], includes the number of traveling vehicles heading to internal and external destinations, that is $N_I^T(t) = N_{II}^T(t) + N_{IJ}^T(t)$. As a result, the total number of vehicles in Region I at time t is, $N_I(t) = N_I^T(t) + N_{IJ}^Q(t)$. We consider the Production-MFD of Region I, $P_I(N_I(t))$ [veh m/s], is re-scaled with respect to the number of traveling and queuing vehicles at time t (Ni and Cassidy 2019), as shown in the following,

$$\begin{aligned} \tilde{P}_I(N_I^T(t); N_{IJ}^Q(t)) &= (1 - \frac{N_{IJ}^Q(t)}{N_I^{jam}}) \; P_I(\frac{N_I^T(t)}{1 - N_I^Q(t)/N_I^{jam}}) \\ &= a \; (N_I^T(t))^3 (\frac{N_I^{jam}}{N_I^{jam} - N_{IJ}^Q(t)})^2 + b \; (N_I^T(t))^2 (\frac{N_I^{jam}}{N_I^{jam} - N_{IJ}^Q(t)}) + c \; N_I^T(t); \end{aligned} \tag{2.1}$$

where N_I^{jam} [veh] is the jam accumulation of Region I, $\tilde{P}_I(N_I^T(t); N_{IJ}^Q(t))$ is the re-scaled Production-MFD, and a, b, c are the estimated parameters for the upper-envelope Production-MFD, i.e. $P_I(N_I(t)) = a \; (N_I(t))^3 + b \; (N_I(t))^2 + c \; N_I(t)$.

The re-scaled mean speed of Region I at time t, $\tilde{V}_I(N_I^T(t); N_{IJ}^Q(t))$ [m/s], used in trip-based MFD model[1] based on the traveling and queued vehicle accumulations is

$$\tilde{V}_I(N_I^T(t); N_{IJ}^Q(t)) = \frac{\tilde{P}_I(N_I^T(t); N_{IJ}^Q(t))}{N_I^T(t)}. \tag{2.2}$$

This indicates that the speed changes when the number of traveling and/or queued vehicles changes. Based on the network speed, the traveled distance between events is calculated, and the vehicles exit the network once their assigned trip lengths are completed. The model gets updated when an event occurs, i.e. the time instance when a vehicle enters or exits a region, or

[1] The difference between $\tilde{P}_I(N_I^T(t); N_{IJ}^Q(t))/N_I^T(t)$ and $\tilde{P}_I(N_I^T(t); N_{IJ}^Q(t))/N_I(t)$ is that the former represents the mean speed of traveling vehicles and the latter represents the mean speed of all vehicles in Region I. The proposed trip-based model tracks the travel distance of vehicles; this only constitutes traveling vehicles, thus the mean speed of traveling vehicles should be used.

joins or leaves a cordon queue, or perimeter control values changes. In which case, the entire time duration might be unevenly separated between events. The fundamental principle of individual vehicle i's trip length towards an internal destination, i.e., $L_{II,i}$ [m], should satisfy,

$$L_{II,i} = \int_{t_i^{dep}}^{t_i^{arr}} \bar{v}_I(N_I^T(t), N_{IJ}^Q(t))dt. \tag{2.3}$$

And for vehicle i's trip length towards an external destination, i.e., $L_{IJ,i}$ [m], should satisfy,

$$L_{IJ,i} = \int_{t_i^{dep}}^{t_i^{Q_I}} \bar{v}_I(N_I^T(t), N_{IJ}^Q(t))dt + \int_{t_i^{LQ_I}}^{t_i^{arr}} \bar{v}_J(N_J^T(t), N_{JI}^Q(t))dt \tag{2.4}$$

where t_i^{dep} and t_i^{arr} [s] are departure and arrival times of vehicle i, i.e., $t_i^{Q_I}$ and $t_i^{LQ_I}$ [s] are the time that vehicle i joins and leaves the cordon queue in Region I, and $\bar{v}_I(N_I^T(t), N_{IJ}^Q(t))$ and $\bar{v}_J(N_J^T(t), N_{JI}^Q(t))$ [m/s] are the re-scaled regional mean speed of Region I and J at time t, respectively. Note that $t_i^{Q_I}$, $t_i^{LQ_I}$, and t_i^{arr} are function of t_i^{dep}, see Eq. (2.6) and Eq. (2.7). Also, it should be noted that for vehicles traveling between regions, i.e., Eq. (2.4), $L_{IJ,i}$ contains two terms associated with travel distance in Region I, $L_{I,i}$, and in Region J, $L_{J,i}$ [m], such that $L_{IJ,i} = L_{I,i} + L_{J,i}$, as shown in Fig. 1(c). The travel lengths can be estimated using a real network data. Eq. (2.3) and Eq. (2.4) aim to realize the traveled distance of vehicle i, while the trip lengths and departure times are model inputs for both internal and external trips. Without loss of generality, we assume the travel lengths are the Euclidean distance between the origins and the destinations considering a plane network. With sampling numerous pairs of origins and destinations in both regions, the trip length distributions of L_{11}, L_{12}, L_{21}, and L_{22} are sampled offline, shown in Fig. 1(d). In the trip-based MFD model, the trip length for each vehicle is randomly assigned from the predefined distributions at the beginning of their trip.

The trip-based MFD model is served as the plant to replicate the actual number of transferring flows, such that the outputs of the model are the arrival times of vehicles with internal and external trips, and the time when the vehicles join and leave cordon queues. Note that we consider the cordon queues as one aggregated queue, one in each region, consistent with the physics of MFD; and the effect of growing or shrinking cordon queue size is captured in re-scaled Speed-MFD. Since the perimeter control restricts transfer flows between the regions, the model should take into account the entry flow capacity of regions. This is a joint function of the border capacity

between the regions and the available space in the destination region. Similar with Ramezani et al. 2015, we consider the entry flow capacity $C_{IJ}(N_J(t))$ [veh/s] as a piecewise two-segment function; first a constant value reflecting border capacity value $\bar{C}_{IJ}$, followed by a decreasing positive function of accumulation of Region J

$$C_{IJ}(N_J(t)) = \begin{cases} \bar{C}_{IJ} & \text{if } 0 \quad N_J(t) < \quad N_J^{jam}, \\ \bar{C}_{IJ} \left(1 \quad \left(\frac{N_J(t)}{N_J^{jam}}\right)^{!}\right) & \text{if } \quad N_J^{jam} \quad N_J(t) \quad N_J^{jam}, \end{cases} \tag{2.5}$$

where $0 < \quad < \quad 1$ defines the deflection accumulation in Region J when the entry flow capacity starts to decline; and $\bar{C}_{IJ}$ [veh/s] is the boundary capacity between Region I and J that denotes the possible maximum number of queued vehicles that can pass the boundary. $\bar{C}_{IJ}$ can be estimated in real life as an aggregation of the capacity of the streets upstream of perimeter traffic signals. Boundary capacity is the maximum number of vehicles per unit time that can pass through the intersections under disparate roadway, traffic, and signalization conditions, while the capacity can be maximized through several methods (see Amirgholy et al. 2020). Note that, other entry flow and exit flow models (e.g. Mariotte and Leclercq 2019; Mariotte et al. 2020a) can be readily integrated within the two-region trip-based model while validation with real data is a research priority. Furthermore, the calibration of is a challenging task especially in multi-region settings since it is needed to capture the complicated multi-path entry flow dynamics.

We define the time-varying remaining trip length as $L_{I;i}(t)$ and $L_{IJ;i}(t)$ [m]. They are updated at each time instance when a new event occurs such that if vehicle i travels inside Region I, $L_{II;i}(t') = L_{II;i}(t) \quad (t' \quad t) \quad v_I(N_I^T(t); N_{IJ}^Q(t))$, if vehicle i is in Region I and travels from Region I towards Region J, $L_{I;i}(t') = L_{I;i}(t) \quad (t' \quad t) \quad v_I(N_I^T(t); N_{IJ}^Q(t))$, and if vehicle i is in Region J and traveled from Region I towards Region J, $L_{J;i}(t') = L_{J;i}(t) \quad (t' \quad t) \quad v_J(N_J^T(t); N_{JI}^Q(t))$, where t [s] is the time of current event and t' [s] is the time of next event. Note that similar to the total travel distance $L_{IJ;i}$, the remaining trip length of vehicle i traveling between regions is $L_{IJ;i}(t) = L_{I;i}(t) + L_{J;i}(t)$. With this in mind, the estimation of arrival time of vehicle i at time t

traveling with an internal destination in Region I is,

$$t_i^{arr}(t) = t + \frac{L_{II;i}(t)}{\bar{v}_I(N_I^T(t); N_{IJ}^Q(t))}.^2 \tag{2.6}$$

If vehicle i travels from Region I to Region J, the estimated join queue time, leave queue time, and arrival time are,

$$t_i^{Q_I} = t + \frac{L_{I;i}(t)}{\bar{v}_I(N_I^T(t); N_{IJ}^Q(t))} \tag{2.7a}$$

$$t_i^{LQ_I} = t_i^{Q_I} + \frac{N_{IJ}^Q(t) + 1}{C_{IJ}(N_J(t)) \quad U_{IJ}(t)} \tag{2.7b}$$

$$t_i^{arr} = t_i^{LQ_I} + \frac{L_{J;i}(t)}{\bar{v}_J(N_J^T(t); N_{JI}^Q(t))} \tag{2.7c}$$

where $U_{IJ}(t)$ denotes the perimeter control manipulating the transfer flows from Region I to Region J at time t, and $0 \quad U_{min} \quad U_{IJ}(t) \quad U_{max} \quad 1$. Note that Eq. (2.6) and Eq. (2.7) are applied to all individual vehicles in the network after each event such that all potential event times are re-estimated and the next imminent event is identified. Eq. (2.7) are not only applied as potential estimates of event times but also employed for the vehicles already started their trips. Hence the current status of the vehicles (e.g. traveling or queuing), and their current locations (e.g. in Region I or Region J) should be taken into account. That is, if the size of cordon queue changes or in case the perimeter control value changes, then Eq. (2.7b) switches to, $t_i^{LQ_I} = t + z_i^{Q_I}(t)/(C_{IJ}(N_J(t)) \quad U_{IJ}(t))$. Here, $z_i^{Q_I}(t)$ is the position of vehicle i in the cordon queue in Region I at time t. Ultimately, when vehicle i passes the boundary and starts traveling in Region J, Eq. (2.7c) becomes $t_i^{arr} = t + L_{J;i}(t)/\bar{v}_J(N_J^T(t); N_{JI}^Q(t))$.

The pseudo-code of the two-region trip-based MFD model is detailed below. For the sake of brevity, region subscripts of t_i^Q, t_i^{LQ}, and t_i^{arr} are omitted.

Algorithm 1 Two-region trip-based MFD model pseudo-code

Initialize event_list=[ ]

Initialize previous_time = t_{init}

---

[2] Accordingly, the estimated arrival time at the time of departure reads $t_i^{arr}(t_i^{dep}) = t_i^{dep} + L_{II;i}/\bar{v}_I(N_I^T(t_i^{dep}); N_{IJ}^Q(t_i^{dep}))$.

for All vehicles (i) do
  Add t_i^{dep} to event_list
  Determine the initial t_i^{Q}, t_i^{LQ}, t_i^{arr} based on Eq. (2.6) or Eq. (2.7) considering initial mean region speed
  Add t_i^{Q}, t_i^{LQ}, and t_i^{arr} to event_list
end for
Sort event_list based on event time

while there are events in the event_list do
  current_time ← the time of the first event in event_list
  for All current traveling vehicles in the network do
    if vehicle i destination is inside Region I then
      $RL_{I,i} \leftarrow RL_{I,i} - \nabla_I(N_I^T(t), N_{IJ}^Q(t))$ (current_time - previous_time)
    else if vehicle i travels from Region I to J then
      if vehicle i currently in Region I then
        $RL_{I,i} \leftarrow RL_{I,i} - \nabla_I(N_I^T(t), N_{IJ}^Q(t))$ (current_time - previous_time)
      else if vehicle i currently in Region J then
        $RL_{J,i} \leftarrow RL_{J,i} - \nabla_J(N_J^T(t), N_{JI}^Q(t))$ (current_time - previous_time)
      end if
    end if
  end for
  if The first event is vehicle departure then
    if The vehicle departures in Region I then
      $N_I^T(t) \leftarrow N_I^T(t)+1$
    else if The vehicle departures in Region J then
      $N_J^T(t) \leftarrow N_J^T(t)+1$
    end if
  else if The first event is vehicle arrival then
    if The vehicle arrives from Region I then
      $N_I^T(t) \leftarrow N_I^T(t)-1$

else if The vehicle arrives from Region J then

$N_J^T(t) \quad N_J^T(t)\text{-}1$

end if

else if The first event is vehicle join cordon queue then

if The vehicle joins the cordon queue in Region I then

$N_{IJ}^Q(t) \quad N_{IJ}^Q(t)\text{+}1, N_I^T(t) \quad N_I^T(t)\text{-}1$

else if The vehicle joins the cordon queue in Region J then

$N_{JI}^Q(t) \quad N_{JI}^Q(t)\text{+}1, N_J^T(t) \quad N_J^T(t)\text{-}1$

end if

else if The first event is vehicle leave cordon queue then

if The vehicle leaves the cordon queue in Region I then

$N_{IJ}^Q(t) \quad N_{IJ}^Q(t)\text{-}1, N_J^T(t) \quad N_J^T(t)\text{+}1$

else if The vehicle leaves the cordon queue in Region J then

$N_{JI}^Q(t) \quad N_{JI}^Q(t)\text{-}1, N_I^T(t) \quad N_I^T(t)\text{+}1$

end if

else if The first event is a new perimeter control value then

update t_i^{LQ} and t_i^{arr} for all vehicles with respect to Eq. (2.6) and Eq. (2.7), considering current vehicle statuses and locations

end if

update $P_I(N_I^T(t); N_{IJ}^Q(t))$ and $P_J(N_J^T(t); N_{JI}^Q(t))$ using Eq. (2.1)

update $\bar{v}_I(N_I^T(t); N_{IJ}^Q(t))$ and $\bar{v}_J(N_J^T(t); N_{JI}^Q(t))$ using Eq. (2.2)

for all vehicles in the network do

update join queue time, leave queue time, and arrival time based on Eq. (2.6) and Eq. (2.7) with the new regional mean speed $\bar{v}_I(N_I^T(t); N_{IJ}^Q(t))$ and $\bar{v}_J(N_J^T(t); N_{JI}^Q(t))$

end for

Remove the first event in the event_list

end while

Note that the trip-based MFD model terminates once all the vehicles exit the network. Note also that the trip-based MFD model is not necessarily a first in first out (FIFO) model since all vehicles in the network might travel with distinct trip lengths, see Yildirimoglu and Ramezani 2020.

2.2.2 Accumulation-based MFD model: accounting perimeter control and heterogeneous transfer flow

In this section, we present the two-region accumulation-based MFD model with perimeter control and a general (uncertain) representation of transfer flows. In majority of literature (e.g. Yildirimoglu et al. 2015; Ramezani and Nourinejad 2018; Yildirimoglu et al. 2018; Sirmatel and Geroliminis 2019), the accumulation-based MFD model did not differentiate between traveling and queued vehicles; while Haddad 2017 and Ni and Cassidy 2019 modeled these two groups of vehicles separately. Also, a FIFO-based exit function as in Paipuri and Leclercq 2020 may be utilized to estimate the number of queuing vehicles in accumulation-based MFD models. In this chapter, we resort to employ the parsimonious accumulation-based MFD models with no differentiation between the traveling and queuing vehicles to simplify the controller design. However, the proposed trip-based MFD model considers traveling and queued vehicles to offer a more detailed picture of traffic dynamics in real networks with perimeter control. The time-varying vehicle accumulations in Region I and J are represented by $N_I(t)$ and $N_J(t)$ [veh], respectively. Vehicle accumulations in Region I bound for internal and external destinations are denoted by $N_{II}(t)$ and $N_{IJ}(t)$ [veh], such that $N_I(t) = N_{II}(t) + N_{IJ}(t)$. Furthermore, the accumulation fraction, $N_{IJ}(t)/N_I(t)$, is assumed to approximately model the transfer flow between Region I and Region J. This assumption is relaxed in this chapter by introducing unknown time-varying (yet partially observable) parameters in the accumulation-based MFD model. We define $0 \leq \alpha_I(t) \leq 1$ corresponding to the internal part of the outflow of Region I (i.e. trips with internal destinations), and $1 - \alpha_I(t)$ associated with the external part of the outflow of Region I (transfer flows, i.e. trips with external destinations[3]).

At the network level, the accumulation-based MFD considers `Production-MFD' in Region I, $P_I(N_I(t))$ [veh m/s], and assumes that the network outflow $G_I(N_I(t)) = P_I(N_I(t))/L_I$, where L_I [m] denotes the average trip length in Region I. The exogenous travel demands are composed of Region I's demands heading to its neighbour Region J, $D_{IJ}(t)$, and those towards destinations inside Region I, $D_{II}(t)$ [veh/s]. The perimeter control located at the boundary between the two regions, $U_{IJ}(t)$ and $U_{JI}(t)$, manipulate transfer flows such that

[3] $\alpha_I(t)$ can be accurately enough estimated as $N_{II}(t)/N_I(t)$ in steady state conditions in MFD models.

$0 \quad U_{min} \quad U_{IJ}(t); U_{JI}(t) \quad U_{max} \quad 1$. The presented accumulation-based MFD model does not consider cordon queue dynamics in order to simplify the controller design. With this choice, the perimeter flow controller does not require the number of queued vehicles and the re-scaled Production-MFD. Therefore, we adopted a robust control approach using SMC method to overcome this simplification in modeling. The accumulation conservation equations are as:

$$\frac{dN_{II}(t)}{dt} = Q_{II}(t) \quad {}_{I}(t) \quad \frac{P_I(N_I(t))}{L_{II}(t)} + (1 \quad {}_{J}(t)) \quad U_{JI}(t) \quad \frac{P_J(N_J(t))}{L_{JI}(t)} \tag{2.8a}$$

$$\frac{dN_{IJ}(t)}{dt} = Q_{IJ}(t) \quad (1 \quad {}_{I}(t)) \quad U_{IJ}(t) \quad \frac{P_I(N_I(t))}{L_{IJ}(t)} \tag{2.8b}$$

where $L_{II}(t)$ [m] is the average travel distance of vehicles with internal destinations and $L_{IJ}(t)$ and $L_{JI}(t)$ [m] are the average travel distance of vehicles with external destinations. Eq. (2.8) associates the change in vehicles accumulations with the inflows and outflows inside a region or from one region to another. Note that the average trip lengths, $L_{II}(t)$ and $L_{IJ}(t)$, can be time-varying while for the sake of simplicity we adopt time-invariant values in the numerical tests.

It is worth mentioning that ${}_{I}(t)$ is partially observed in real-time to a certain degree of accuracy by measuring the number of vehicles passing through the perimeter control intersections. However, in the accumulation-based MFD model, ${}_{I}(t)$ is estimated as $N_{II}(t) = N_I(t)$, and it is more challenging to be predicted as it intrinsically required in MPC approaches. In addition, ${}_{I}(t)$ is a source of uncertainty in modeling while SMC is robust to deal with unobservable variations in model variables, and hence a robust perimeter control based on SMC theory is applied that take into account uncertainties in the MFD model. Note that the proposed SMC perimeter controller is designed grounded on the accumulation-based MFD model while applied on the trip-based MFD model, such that ${}_{I}(t)$ is not required in the proposed controller. From there, the consistency check of these two different models is necessary where the numerical experiments are presented in Section 2.4.1.

2.3 Controller design

2.3.1 Sliding mode control

Sliding Mode Control (SMC) is a robust control approach to provide a desired dynamic behavior regardless of uncertainties and disturbance in the model (see Slotine 1984; Canale et al. 2008; Spurgeon 2014). In this section, SMC is designed building on the accumulation-based MFD model in Eq.(2.8) to manipulate the system states reach prede ned sliding surfaces and maintain on the surfaces thereafter (Bichiou et al. 2020). This chapter applies a similar control framework as Aalipour et al. 2017.

We rearrange the state variables of the system as $X_1(t) = N_{11}(t) + N_{21}(t)$, $X_2(t) = N_{12}(t)$, $X_3(t) = N_{21}(t)$, and $X_4(t) = N_{12}(t) + N_{22}(t)$. The traf c dynamics of the two-region network with the new rearranged states are,

$$\dot{X}_1(t) = Q_{11}(t) + Q_{21}(t) \quad M^X_{11}(t) \tag{2.9a}$$

$$\dot{X}_2(t) = Q_{12}(t) \quad M^X_{12}(t)U_{12}(t) \tag{2.9b}$$

$$\dot{X}_3(t) = Q_{21}(t) \quad M^X_{21}(t)U_{21}(t) \tag{2.9c}$$

$$\dot{X}_4(t) = Q_{22}(t) + Q_{12}(t) \quad M^X_{22}(t) \tag{2.9d}$$

such that the external out ows and control signals are merely appeared in two traf c states (i.e. $X_2(t)$ and $X_3(t)$). The rearranged internal and external out ows are,

$$M^X_{11}(t) = \beta_1(t) \frac{P_1(X_1(t) \quad X_3(t) + X_2(t))}{L_{11}(t)} \tag{2.10a}$$

$$M^X_{12}(t) = (1 \quad \beta_1(t)) \frac{P_1(X_1(t) \quad X_3(t) + X_2(t))}{L_{12}(t)} \tag{2.10b}$$

$$M^X_{21}(t) = (1 \quad \beta_2(t)) \frac{P_2(X_4(t) \quad X_2(t) + X_3(t))}{L_{21}(t)} \tag{2.10c}$$

$$M^X_{22}(t) = \beta_2(t) \frac{P_2(X_4(t) \quad X_2(t) + X_3(t))}{L_{22}(t)}: \tag{2.10d}$$

Note that the rearranged out ows hold akin calculation logic with the out ows in Eq.(2.8), e.g. $M_{12}(t) = M^X_{12}(t) = (1 \quad \beta_1(t)) \quad P_1(N_1(t))/L_{12}(t)$.

The sliding surfaces are considered as $S_1(t) = X_4(t) - k_1 X_2(t)$ and $S_2(t) = X_1(t) - k_2 X_3(t)$, where positive parameters k_1 and k_2 are designed to ensure stability. The evolution of the states of the system over time are forced towards the sliding surfaces $S_1(t) = 0$ and $S_2(t) = 0$. The derivatives of sliding surfaces are,

$$\dot{S}_1(t) = Q_{22}(t) + (1 - k_1)Q_{12}(t) - M^X_{22}(t) + k_1 U_{12}(t) M^X_{12}(t) \tag{2.11a}$$

$$\dot{S}_2(t) = Q_{11}(t) + (1 - k_2)Q_{21}(t) - M^X_{11}(t) + k_2 U_{21}(t) M^X_{21}(t). \tag{2.11b}$$

To analyze the stability, we define the following inequalities,

$$\alpha_1(t) \geq \left| \frac{Q_{22}(t) + (1 - k_1)Q_{12}(t) - M^X_{22}(t)}{k_1 M^X_{12}(t)} \right| \tag{2.12a}$$

$$\alpha_2(t) \geq \left| \frac{Q_{11}(t) + (1 - k_2)Q_{21}(t) - M^X_{11}(t)}{k_2 M^X_{21}(t)} \right|; \tag{2.12b}$$

such that the above inequalities can be guaranteed by defining the variables $\alpha_1(t)$ and $\alpha_2(t)$ as,

$$\alpha_1(t) = \frac{Q_{22,max} + (k_1 - 1)Q_{12,max} + M^X_{22}(t)}{k_1 M^X_{12}(t)} \tag{2.13a}$$

$$\alpha_2(t) = \frac{Q_{11,max} + (k_2 - 1)Q_{21,max} + M^X_{11}(t)}{k_2 M^X_{21}(t)}; \tag{2.13b}$$

where $Q_{11,max}$, $Q_{12,max}$, $Q_{21,max}$, and $Q_{22,max}$ [veh/s] are the maximum values of $Q_{11}(t)$, $Q_{12}(t)$, $Q_{21}(t)$, and $Q_{22}(t)$, respectively. The Lyapunov candidate function is considered as $v(t) = (S_1(t)^2 + S_2(t)^2)/2$, to compel the trajectories moving towards sliding surfaces and to ensure the surfaces are stable. The derivative of $v(t)$ is,

$$\begin{aligned}
\dot{v}(t) &= S_1(t)\dot{S}_1(t) + S_2(t)\dot{S}_2(t) \\
&= S_1(t)(Q_{22}(t) + (1 - k_1)Q_{12}(t) - M^X_{22}(t) + k_1 M^X_{12}(t) U_{12}(t)) \\
&\quad + S_2(t)(Q_{11}(t) + (1 - k_2)Q_{21}(t) - M^X_{11}(t) + k_2 M^X_{21}(t) U_{21}(t)) \\
&\leq |S_1(t)| \alpha_1(t) k_1 M^X_{12}(t) + U_{12}(t) S_1(t) k_1 M^X_{12}(t) \\
&\quad + |S_2(t)| \alpha_2(t) k_2 M^X_{21}(t) + U_{21}(t) S_2(t) k_2 M^X_{21}(t).
\end{aligned} \tag{2.14}$$

To satisfy precedent conditions, we de ne the discontinuous perimeter control inputs as,

$$U_{12}(t) = \quad _1(t) \quad \mathrm{sgn}(S_1(t)) \tag{2.15a}$$

$$U_{21}(t) = \quad _2(t) \quad \mathrm{sgn}(S_2(t)) \tag{2.15b}$$

where $_1(t) > \ _1(t) + \ _0$ and $_2(t) > \ _2(t) + \ _0$, and $_0$ is a positive and relatively small constant. Substituting control inputs in Eq. (2.14), the inequalities (2.16) show the derivative of V is negative, such that the state trajectories will be captivated towards the sliding surfaces and stay on them.

$$\begin{aligned} \underline{v}(t) \quad & S_1(t) \ _1(t)k_1M_{12}^{X}(t) \quad (\ _1(t) + \ _0)k_1M_{12}^{X}(t)S_1(t)\mathrm{sgn}(S_1(t)) \\ & + S_2(t) \ _2(t)k_2M_{21}^{X}(t) \quad (\ _2(t) + \ _0)k_2M_{21}^{X}(t)S_2(t)\mathrm{sgn}(S_2(t)) \\ & \quad _0|S_1(t)|k_1M_{12}^{X}(t) \quad _0|S_2(t)|k_2M_{21}^{X}(t) \quad 0: \end{aligned} \tag{2.16}$$

A bounding lter is applied to ensure the control sequences obtained from Eq. (2.15) are between the lower and the upper bounds (i.e., U_{min} and U_{max}). For instance, if $U_{12}(t)$ is smaller (or greater) than U_{min} (or U_{max}), then $U_{12}(t) = U_{min}$ (or $U_{12}(t) = U_{max}$) will be adopted. Note that the proposed SMC perimeter control does not necessarily require $_1(t)$ and $_2(t)$ values and traf c state measurements can be used. The control architecture is depicted in Fig. 2.2.

FIGURE 2.2. The control architecture of the proposed perimeter control method based on SMC. The accumulation-based MFD model is employed to design the controller. The trip-based MFD model is considered as the plant. The solid arrows represent online procedures.

2.3.2 Improved Bang-Bang controller

To compare the performance of the proposed SMC, we consider an improved Bang-Bang (I-BB) controller. The Bang-Bang (BB) perimeter controller is a feedback control approach that aims to ensure the vehicular accumulation in a region remains under the critical value (i.e. the accumulation that generates the maximum production). The BB controller applied in the previous research (e.g. Geroliminis et al. 2012; Aalipour et al. 2018) was designed corresponding to fixed critical and jam accumulations of the region, i.e. N_i^{cr} and N_i^{jam} [veh]. The I-BB proposed here follows the same control policy yet with time-varying critical and jam accumulations,[4] i.e. $N_i^{cr}(t)$ and $N_i^{jam}(t)$ [veh], to capture the effect of cordon queues on the MFD.

The $N_i^{cr}(t)$ is determined via the derivative of the re-scaled Production-MFD with respect to the number of traveling vehicles in Region i at time t, $\frac{\partial P_i(N_i^T(t); N_{ij}^Q(t))}{\partial N_i^T(t)} = 0$. And the re-scaled jam accumulation is considered as $N_i^{jam}(t) = N_i^{jam} - N_{ij}^Q(t)$. The I-BB controller aims to protect the more congested region by restricting the number of vehicles to enter and allowing more vehicles to leave to the less congested region. Accordingly the control values ($U_{12}(t)$ and $U_{21}(t)$) operate either at the minimum or the maximum (U_{min} or U_{max}). Detailed analytical properties of BB control policy can be referred to Aalipour et al. 2018, whereas there is a difference that the proposed I-BB uses traveling accumulations ($N_i^T(t)$) rather than total accumulations ($N_i(t)$) of Region i. The I-BB control law is summarized in Table 2.1.

TABLE 2.1. The Improved Bang-Bang Control Values for $[U_{12}(t); U_{21}(t)]$

$N_1^T(t) \le N_1^{cr}(t)$, $N_2^T(t) \le N_2^{cr}(t)$	$N_1^T(t) \le N_1^{cr}(t)$, $N_2^T(t) > N_2^{cr}(t)$	$N_1^T(t) > N_1^{cr}(t)$, $N_2^T(t) \le N_2^{cr}(t)$	$N_1^T(t) > N_1^{cr}(t); N_2^T(t) > N_2^{cr}(t)$	
$[U_{max}; U_{max}]$	$[U_{min}; U_{max}]$	$[U_{max}; U_{min}]$	if $\frac{N_1^T(t)}{N_1^{jam}(t)} > \frac{N_2^T(t)}{N_2^{jam}(t)}$ $[U_{max}; U_{min}]$	if $\frac{N_1^T(t)}{N_1^{jam}(t)} < \frac{N_2^T(t)}{N_2^{jam}(t)}$ $[U_{min}; U_{max}]$

[4] The performance of BB in the numerical experiment was worse than the no control case.

2.4 Numerical experiments

In this section, the accumulation-based and trip-based MFD models and the proposed SMC perimeter control are scrutinized with two scenarios. (i) A slow-varying demand scenario to investigate the consistency between the two-region accumulation-based and trip-based MFD models considering cordon queues. This is essential because, first, the controller design and application are based on different models (i.e. accumulation-based and trip-based), second, the accumulation-based MFD model does not consider the boundary capacity and cordon queues. And (ii) a fast-varying peak-hour demand scenario to investigate the proposed SMC efficiency comparing with no control and I-BB control cases. The trip-based MFD model is applied as the plant for a typical peak-hour period in this scenario. The numerical settings under two types of demand scenarios are introduced first, and results will be analyzed and discussed in detail in the corresponding subsections. It is worth mentioning that the average trip length of each region applied in the accumulation-based MFD model is 2300 [m], and this value is the mean of the predefined distributions (as shown in Fig. 2.1(d)) which are employed to randomly assign trip length of each vehicle in the trip-based MFD model for both consistency check and controller performance experiments.

2.4.1 Models consistency and comparison

In this section, the numerical tests are adopted to investigate the characteristics and consistency between the two-region accumulation-based and trip-based MFD models. We apply the trip-based MFD model for a two-region network where each region has the same MFD with the one observed in Yokohama, Japan in Geroliminis and Daganzo 2008, as $P_i(N_i(t)) = a\ (N_i(t))^3 + b\ (N_i(t))^2 + c\ N_i(t)$, where $a = 9.98\ 10^{-8}$, $b = -0.002$, and $c = 9.78$. The network schematic is shown in Fig. 2.1(c) with the average trip length of 2300 [m] for both regions. The boundary capacities are $\overline{C}_{12} = \overline{C}_{21} = 10$ [veh/s] and $= 0.75$. According to Lamotte and Geroliminis 2018, the two models should be equivalent when the demands vary slowly and the trip lengths are exponentially distributed, or the two models are consistent when the system reaches steady-state in spite of trip length distributions. With this in mind, we first select a slow-varying demand, as

shown in Fig. 2.3(a), and disparate type of trip length distributions (i.e. exponential, random), for consistency check. The trip lengths are randomly selected from different distributions in an of ine procedure. The initial accumulations in the onset of the simulation are assumed uncongested in Region 1 and congested in Region 2, where $N_1(0) = 3{,}000$ [veh] and $N_2(0) = 3{,}500$ [veh]. Note that the accumulation-based MFD model is FIFO and does not track cordon queue dynamics. Besides, the time step between events in the trip-based MFD model is uneven while the whole simulation period is determined by the time that the last vehicle arrives at its destination.

We check the consistency and conduct a comparison of the MFD models established in the previous section. The results are depicted in Fig. 2.3 and Fig. 2.4 with exponentially and randomly distributed trip lengths with average trip length 2300 [m]. Note that in the trip-based MFD model, each vehicle has a distinct unique trip length drawn from a distribution. For the consistency, we make sure the average trip lengths used in Eq. (2.8) are consistent with the mean of the distributions of the trip lengths of the vehicles. Note that the random trip lengths do not follow any speci c distributions while we ensure they have the same average trip lengths as of the accumulation-based MFD model. Also note that trip lengths are not constant; each vehicle has a unique trip length in the trip-based MFD model. Fig. 2.3(b) and Fig. 2.3(c) indicate the accumulation evolution of the accumulation-based and trip-based MFD models with trip lengths following exponential distribution, while those accumulations of accumulation-based MFD model are estimated grounded on the steady-state assumption of $n_i(t) = N_{II}(t)/N_I(t)$ using Eq.(2.8). It appears that the accumulation evolution of the two MFD models are alike in trend and magnitude under the condition of slow varying demands; in accordance with the ndings of Jin 2020. In Fig. 2.4(a) and Fig. 2.4(b), accumulation exhibits discrepancies at the beginning (before 2000 [s]) while remaining constant when the system reaches steady state (after 2200 [s]). This is in accordance with Lamotte and Geroliminis 2018 which stated the two models are similar at steady-state despite how trip lengths distributed.

The accumulations in Region 1 and Region 2 in the trip-based MFD model abruptly increase when there is a large number of vehicles generated around 800 -1800 [s]. However, the accumulation-based MFD model reacts moderately with the demand surge. Since the queuing vehicles for both regions primarily leave the cordon queues at 500 - 600 [s], they increase the traveling

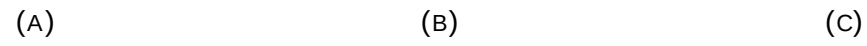

FIGURE 2.3. (A) Slow varying two-staircase demand pro le; Accumulation evolution when trip lengths follow exponential distribution: (B) accumulation-based MFD model, (C) trip-based MFD model.

accumulations in the other region. As the accumulation-based MFD does not consider the boundary capacity and cordon queues, it is expected to see it underestimates the accumulations, which is observed in Fig. 2.3 and Fig. 2.4.

Fig. 2.4(c) and Fig. 2.4(d) display the internal and external out ows of accumulation-based and trip-based MFD models with randomly selected trip lengths. The values of internal and external out ows are measured through trip-based MFD model by counting the number of vehicles which nish their trips or join the queues at each time step ρ(e.g. 2 [min]). Consequently, the sum of internal and external out ows of a region may exceed the maximum value of out ow-MFD (i.e. approximately 6.3 [veh/s]) since (i) each vehicle in the trip-based MFD model has a unique trip length which might temporarily lead to a surge in the out ow from the network. (ii) Because of the causality effect (the instantaneous change of region average speed with the change in accumulation) in transition periods the observed out ows are higher. (iii) Lastly, the queuing process at the regions boundary adds inherent delays in external out ows which propagates in accumulations and internal out ows.

The evolution of (t) of the two MFD models are depicted in Fig. 2.4(e) and Fig. 2.4(f). In accumulation-based MFD model, $_{,}(t)$ is continuous in time and instantaneous and is estimated as $N_{II}(t) = N_{I}(t)$. Yet it appears that this estimation is not always accurate. As in Fig. 2.4(e), $_{i}(t)$ should be very close to one at the end of the studied period as a result of extra time needed for transferred vehicles to nish their trips. Unlike the estimation of $_{I}(t)$ in accumulation-based MFD model, we consider to obtain it via out ows in the trip-based MFD model, i.e. $_{i}(t) =$

$M_{II}(t) = (M_{II}(t) + M_{IJ}(t))$. Note that considering $N_{IJ}(t) = N_I(t)$ equals to $_I(t)$ might be reliable under several conditions, e.g. exponential distributed trip lengths and steady state conditions, whereas it is less accurate under time-varying and complex situations. Thus a robust controller with respect to (t) uncertainty (which is a byproduct of cordon queues and heterogeneous distributions of destinations) is needed.

2.4.2 Controller performance

In this section, a scenario with typical peak-hour travel demands is presented to explore the performance of the proposed SMC perimeter controller. The two-region trip-based MFD model considering cordon queues is employed as the plant. In this scenario, the network is initially in uncongested conditions for both regions with initial accumulations $N_1(0) = 2{,}300$ [veh] and $N_2(0) = 2{,}500$ [veh] respectively. Note that the accumulation-based and trip-based MFD models rely on the same MFD functions but with different representations, i.e. Production-MFD and Speed-MFD. Perimeter control values are calculated using I-BB and SMC with the lower bound as $U_{min} = 0.1$ and the upper bound as $U_{max} = 0.9$. Both controllers change the perimeter control values every minute. In the no control case, $U_{12}(t) = U_{21}(t) = U_{max}$ are applied over the studied time duration. Furthermore, the parameters in sliding surface are selected as $k_1 = 2$ and $k_2 = 4$. Considering the upper-envelope MFD, the region critical and jam accumulations are 3,334 and 10,000 [veh], respectively. However, the re-scaled MFD alters the critical and jam accumulations, i.e. $N_I^{cr}(t)$ and $N_I^{jam}(t)$, over time such that the two accumulations will be adjusted in every iteration under I-BB control framework.

Three scenarios are compared to examine the proposed controller ef ciency: 1) no perimeter control, 2) I-BB perimeter control, and 3) SMC. We take the no control case as the base scenario and apply the I-BB control for comparison. The I-BB perimeter control is a state feedback control method considering the current number of traveling vehicles in the regions $N_I(t)$ and comparing it with re-scaled time-varying critical and jam accumulations in order to protect the more congested region. A fast varying rush-hour demand pro le is employed, as shown in Fig. 2.5(a). The network faces increasing traf c demands before 900 [s], while the demands gradually drop to zero to ensure the clearance of the network at the end of the simulation.

(A) (B)

(C) (D)

(E) (F)

FIGURE 2.4. Accumulation evolution when trip lengths are randomly distributed: (A) accumulation-based MFD model, (B) trip-based MFD model; Evolution of internal and external ows: (C) accumulation-based MFD model, (D) trip-based MFD model; Evolution of (t): (E) accumulation-based MFD model, (F) trip-based MFD model.

The accumulation evolution over studied period of trip-based MFD model corresponding to no control, I-BB control, and SMC are depicted in Fig. 2.5(b)-(d) respectively. With the absence of control, the control sequences always operate at the maximum such that a few queued vehicles (i.e. $N^{Q}_{12}(t) < 500$ [veh]) are detected and the cordon queues dissipate promptly, as shown in Fig. 2.5(b). Also, Region 2 is the more congested one and requires protection in the no control case. It can be observed that the accumulation difference between the two regions shrinks when either I-BB or SMC is employed, whereas both controllers lead to additional vehicles queuing at the cordons compare with the no control case. The control sequences of I-BB and SMC controllers are depicted in Fig. 2.5(e) and Fig. 2.5(f). Note that before 500 [s], there are no queued vehicles in I-BB control case because both $U_{12}(t)$ and $U_{21}(t)$ are the maximum value, but they are observed in SMC case since the perimeter control is active from the beginning to restrict the vehicles passing the boundary between Region 1 and 2. Afterwards, between 500 - 1800 [s], the perimeter control oscillates either at the maximum or the minimum values of the I-BB control case. The chattering behavior of control sequences is evident as I-BB tries to proportionally equalize the accumulations of both regions. In contrast, the SMC protects the more congested Region 2. As shown in Fig. 2.5(f), $U_{21}(t)$ operates at U_{max} between 0 and 2100 [s] while $U_{12}(t)$ is chattering to limit the vehicles in Region 1 to enter Region 2.

Perimeter control is a way to distribute vehicles more uniformly in the network yet could ease the network congestion to some extent. The number of cordon queued vehicles in both regions are higher in I-BB control case compare to no control and SMC, such that more vehicles suffer queuing delays because of the perimeter control. The cordon queues in Region 1 and Region 2, $N^{Q}_{12}(t)$ and $N^{Q}_{21}(t)$, are continuously expanding and concentrating at the cordons from 500 [s] to 1800 [s] with the peak values of 2255 [veh] and 1558 [veh] respectively. The vehicles queued near cordons reduces the average speed of traveling vehicles. However, the number of queuing vehicles corresponding to SMC is much smaller than I-BB control case. With SMC, the cordon queues are long-lived (appear sooner and diminish later) and moderate, which helps curb the negative effect of cordon queues on the speed of traveling vehicles. Furthermore, SMC results in chattering perimeter control actions that reduce the maximum size of cordon queues, with the peak value of 693 [veh] in SMC and 2255 [veh] in I-BB. Note that the control values of SMC are oscillating between U_{min} and U_{max} since the controllers are incorporated with a limiting filter. No

perimeter control values are observed in between as a consequence of the magnitude of sliding surface parameters, i.e. k_1 and k_2. Higher parameters can be applied whereas the controller performance might be deteriorated.

It assuredly alleviates the congestion in Region 2 when either controller applies; the maximum accumulation reduces from 6790 [veh] with no control to 4658 [veh] with I-BB control and 4207 [veh] with the proposed SMC. This comes at an expense of more vehicles in Region 1; the maximum accumulation extends from 4188 [veh] with no control to 4834 [veh] with I-BB control and 5198 [veh] with SMC. One observation worth mentioning is in the no control case, Region 2 is the more congested one that needs protection, but this order switches (Region 1 becomes more congested) in the SMC case.

Fig. 2.5(g) and Fig. 2.5(h) depicts the Production-MFD of Region 1 and Region 2 for all control cases where hysteresis loops are observed. Previous research (e.g. Mahmassani et al. 2013 suggested that the clockwise hysteresis loops are more likely to happen as a consequence of the heterogeneity of congestion distribution in a region. In this chapter, the trip-based MFD model considers the impact of cordon queues, therefore the congestion is not uniformly distributed especially with considerable cordon queues or with protracted time for queue dissipation. Results indicate that SMC leads to a more homogeneous network as the Production-MFDs for both regions display negligible hysteresis loops and the accumulation difference between regions is insigni cant. The Production-MFD in the trip-based MFD model is the re-scaled one, therefore it can be observed that the Production-MFD curves are lower than the upper-envelope Production-MFD.

Comparing the three control cases, (see Fig. 2.5, Fig. 2.6, and Table 2.2, we notice the advantages of SMC compared to the I-BB controller. The objective of the SMC is minimizing the total time spent that is equivalent to minimize the total number of vehicles in the network. Table 2.2 lists the average travel time [s/veh] and total time spent [veh] of I-BB perimeter control, the proposed SMC cases along with the no control case. The controller updating cycle times are 60 and 120 seconds, respectively. Tabulated values are obtained through the average results out of ten simulations. The values in parentheses are the change in the percentage of total network performance compare with the no control case. With a limited number of switching control