On Robust Distributed Control of Transportation Networks

Nilsson, Gustav

2019

Document Version:
Publisher's PDF, also known as Version of record

On Robust Distributed Control of Transportation Networks

Gustav Nilsson

LUND
UNIVERSITY

Department of Automatic Control

Ph.D. Thesis TFRT-1123
ISBN 978-91-7753-958-2 (print)
ISBN 978-91-7753-959-9 (web)
ISSN 0280–5316

Department of Automatic Control
Lund University
Box 118
SE-221 00 LUND
Sweden

Printed in Sweden by MediaTryck.
Lund 2019

Abstract

With the ever-growing traffic demands, the transportation networks are getting more and more congested. While expanding these networks with more roads is both costly and in many cities not even feasible, the rapid development of new sensing and communication techniques has made it possible to perform control of transportation networks in real-time. With the right usage of such technologies, existing transportation networks' capacities can be utilized better in order to lower the congestion levels. However, the control has to be done robustly, since real-time control and close to maximal utilization also make the networks more fragile and if not, even a small perturbation can have a tremendous impact on the traffic network. In this thesis, a few solutions that lead to better transportation network utilization are presented, designed with said robustness requirements in mind.

In the first part of the thesis, a decentralized control strategy for traffic signals is presented. The proposed policy, which we call Generalized Proportional Allocation (GPA), is inspired by the proportional fairness allocation for communication networks. The original proportional fairness controller does not explicitly take the overhead time needed to shift between different activation phases into account. We, therefore, enhance the proportional fairness so that it adapts its cycle length to the current demand. When the demand is higher, one wants longer signal cycles not to waste too much of the time overhead, while for lower demands, the cycle lengths should be shorter, so that the drivers do not have to wait for a long time. Stability for an averaged version of this control strategy is proved together with throughput-optimality of the controller. This means that no other control strategy can handle larger exogenous inflows to the network than the GPA-controller. Since the traffic signal controllers such as the GPA may allocate service to an empty line, due to the fact that several lanes can receive green light simultaneously, a model that handles this issue is proposed. For this model, the well-posedness of the dynamical system is shown when the traffic signal controller is Lipschitz continuous.

The GPA controller's performance is also evaluated in a microscopic traffic simulator. In the microsimulations, it is shown how the proposed feedback controller outperforms the standard fixed-time controller for a scenario based on all traffic over the duration of one full day in Luxembourg. The controller's performance is also

compared to another decentralized controller for traffic signals, the MaxPressure controller, for an artificial Manhattan-like network. From these simulations, it can be concluded that the GPA performs better than MaxPressure during low demands, but the MaxPressure performs better when the demand is high. The fact that the GPA does not require any information about the network, apart from the current queue lengths, makes it robust to perturbations. In other words, the control strategy does not have to be updated when the demand or topology of the network changes.

The second part of the thesis is devoted to routing problems. First, the problem of routing a fleet of vehicles in an optimal way for the whole fleet is considered. The objective is then to achieve a minimum delay in average for the entire fleet. The routing algorithm takes into account the presence of regular drivers that are trying to optimize their own traveling time in the network. Conditions are posted for when such a routing assignment exists, and two algorithms to compute it are shown.

At last, a type of dynamic routing policies for multicommodity flows is studied. The routing policies are designed with the objective to avoid congested routes. It has previously been shown that if only one class of vehicles are present, the network is robust to perturbations with these routing policies. A model for multicommodity flows is proposed, and it is shown that the robustness properties for the single-commodity case do not necessarily hold in the multicommodity case.

Acknowledgments

First of all, I would thank my supervisor Giacomo Como. Having the opportunity to work with him for over five years, and learn some of his ways to think about problems and research in general, has been invaluable to me. I also would like to thank my co-supervisor Anders Rantzer. During my first time as a Ph.D. student, Enrico Lovisari patiently guided me into the world of research. Also, I have had many good discussions with Michelle Chong. Thanks to all of you!

Last winter, I got the opportunity to do an internship at Mitsubishi Electric Research Laboratories (MERL), which was a truly inspiring time. My hosts, Uros Kalabic and Piyush Grover, gave me many new valuable insights about research.

I would also express my gratitude to Institute for Pure and Applied Mathematics at UCLA for inviting me to their long-term program titled "New Directions in Mathematical Approaches for Traffic Flow Management". The program and the following workshops have been a great source of inspiration for me.

During the about five years long journey of Ph.D. studies, I have been lucky to get to know a lot of people, who have helped me not just to improve as a researcher, but also a person in general. Thanks to all of you, you know who you are!

Although this acknowledgment could have ended with the previous paragraph, a few people deserve an explicit mention of making this thesis possible. I want to thank, Erik Bylow, Johan Fredriksson, Martina Maggio, Claudio Mandrioli, Richard Pates, Anders Robertsson, and Christian Rosdahl, for their comments on the draft on this thesis. Bo Bernhardsson should also get kudos for his careful review of my publications when the plan for this thesis was evaluated. A large part of this thesis was formed during my stays in Turin. I would, therefore, like to thank all the people at Dipartimento di Scienze Matematiche "Giuseppe Luigi Lagrange", Politecnico di Torino, for making my stay there enjoyable. Also, the cyclists in "SQUADRA CORSE Ciclofficina Artigiana" should be acknowledged for making my spare time in Turin way more fun by letting me join their amazing rides. As a Ph.D. student in Lund, I have received excellent administrative support with various issues. Thanks, Cecilia, Eva, Ingrid, Lizette, Mika, and Monika! Anders and Anders have always been helpful with various computer issues and Pontus with practical issues, which has made my work much easier. Leif's LaTeX knowledge has been invaluable for this

thesis. If LaTeXing was an Olympic discipline, I am sure that Leif would be on the podium.

At last, I would like to thank my family and friends for their support.

Financial Support

Part of this work has been supported by the Swedish Research Council (VR). The author is a member of the LCCC Linnaeus Center and the ELLIIT Excellence Center at Lund University.

Contents

List of Abbreviations

CTM	Cell Transmission Model
FIFO	First-In-First-Out
GPA	Generalized Proportional Allocation
MPC	Model Predictive Control
PDE	Partial Differential Equation
TTT	Total Travel Time

Nomenclature

$[a]_+$	The positive part of each entry in $a \in \mathbb{R}^n$, i.e., $[a]_+ = \max(a, 0)$ where the max is applied entry-wise
$[a]_-$	The negative part of each entry in $a \in \mathbb{R}^n$, i.e., $[a]_+ = \max(-a, 0)$ where the max is applied entry-wise
$\delta_i^{(j)}$	The indicator function, $\delta_i^{(j)} = 1$ if $i = j$, 0 otherwise
$\text{dist}(x, \mathcal{A})$	The shortest distance between a vector $x \in \mathbb{R}^n$, and a set $\mathcal{A} \subset \mathbb{R}^n$
$\hat{\lambda}$	The exogenous inflows to the junctions, used for traffic assignment purposes $\hat{\lambda} \in \mathbb{R}_+^{\mathcal{V}}$
$\hat{\mu}$	The exogenous outflows from the junctions, used for traffic assignment purposes $\hat{\mu} \in \mathbb{R}_+^{\mathcal{V}}$
$\text{int}(\mathcal{A})$	The interior of a set $\mathcal{A}$
λ	The exogenous inflows to the cells, $\lambda \in \mathbb{R}_+^{\mathcal{E}}$
$\mathcal{C}_T$	A vector of continuous functions on the interval $[0, T]$
$\mathcal{E}$	The set of cells, i.e., where traffic is stored

$\mathcal{E}^k$	The set of cells that vehicles of class $k \in \mathcal{K}$ have access to, $\mathcal{E}^k \subseteq \mathcal{E}$
$\mathcal{E}_v$	The set of in-coming cells to a junction $v \in \mathcal{V}$
$\mathcal{E}_v^+$	The set of out-going cells from a junction $v \in \mathcal{V}$
$\mathcal{E}_v^{k+}$	The set of outgoing cells from a junction $v \in \mathcal{V}$ that vehicles of class $k \in \mathcal{K}$ have access to, $\mathcal{E}_v^{k+} \subseteq \mathcal{E}_v^+$
$\mathcal{K}$	The different classes of traffic
$\mathcal{P}^{(v)}$	The set of phases associated with a junction v
$\mathcal{T}^{(v)}$	The signal program for a junction v
$\mathcal{U}^{(v)}$	The set of possible phase activations for a junction v
$\mathcal{V}$	The set of junctions, i.e., points where the cells are connected
$\mathcal{X}$	The space of traffic volumes $\mathcal{X} = \mathbb{R}_+^{\mathcal{E}}$
$\|\|x\|\|$	A general norm
$\tau_i(z_i)$	The delay function for a cell $i \in \mathcal{E}$
B	The node-link incidence matrix
C	The matrix with the outflow capacities on the diagonal, $C = \text{diag}(c)$
c	The vector of all outflow capacities $c \in \mathbb{R}_+^{\mathcal{E}}$
$d_i(x_i)$	The demand function for a cell $i \in \mathcal{E}$
i^+	The head of a cell $i \in \mathcal{E}$ such that $i^+ \in \mathcal{V}$
i^-	The tail of a cell $i \in \mathcal{E}$ such that $i^- \in \mathcal{V}$
R	The routing matrix, $R \in \mathbb{R}_+^{\mathcal{E}\times\mathcal{E}}$
$s_i(x_i)$	The supply function for a cell $i \in \mathcal{E}$
x	The traffic volumes in all cells
x^k	The traffic volumes of class $k \in \mathcal{K}$ in all cells
$x^{(v)}$	The traffic volumes in the incoming cells to a junction v
$x^{(v+)}$	The traffic volumes in the outgoing cells to a junction v
z	The outflow from each cell, $z \in \mathbb{R}_+^{\mathcal{E}}$

1

Introduction

1.1 Why Control of Transportation Networks?

While transportation networks have been around in our society for a while, they are now becoming more and more congested and are often not dimensioned to handle today's traffic demands. Increasing the capacity of the traffic networks to match the growing demands may be difficult. In many cities, there is no more space to build more roads. Even if autonomous vehicles and the shared economy will be more present in the future, it does not mean that the traffic demand will necessarily decrease. Simulation studies have shown that if people use transportation-on-demand services instead of their cars, the total number of vehicles needed will decrease, but the entire load in the traffic network will increase due to the need of re-balancing the vehicles [Spieser et al., 2014].

The ultimate goal of controlling traffic networks is to reduce congestions. Congestions in the US is today estimated to have a cost of about 300 billion USD [Cookson, 2018], which is about 1.5 % of the country's GDP [Organisation for Economic Co-operation and Development, 2018] when taking both the additional fuel consumption and hours lost in congestions into account. The increased fuel consumption has of course negative impact on the environment regarding pollution. By reducing congestion, transportation networks can also become more reliable. Reliability is important since predicable transportations are becoming more critical to our society. One example of this is factories that instead of storing material in the factories use just-in-time deliveries.

Luckily, the recent development in sensing and communication technology provides tools that can, when correctly used, help to improve the reliability and performance of transportation networks. Even if not each car is connected yet, most cars have one or several connected devices within it, such as a cell phone or a GPS routing device. Those kinds of connections are now used both to obtain traffic state information and to provide route guidance to drivers. But it is not just information from each car that is easier to obtain nowadays. Development of cheap cameras and computer vision have now made it easy to get real-time traffic information. While this traffic information was before obtained by loop-detectors, that had to be installed

Figure 1.1 Example of a smart phone application that gives the driver a recommended route to follow.

in the ground and only gave an estimate of the queue lengths, cameras which are easier to install can now provide accurate information about the queue lengths in real-time [Citilog, 2018].

However, utilizing the traffic networks close to capacity can also make them more fragile to perturbations. For instance, if everyone is using the same route guidance services, and it for some reason provides wrong information, the consequences can be much worse compared to when every driver made her own path decision. However, this is just one way new technologies put higher requirements on robust solutions. While traffic signals before were tuned based on a nominal behavior of drivers, this nominal behavior can now change rapidly when the drivers get new routing directions in real-time. Hence the traffic signals need to adapt to the new traffic pattern quickly.

There are several methods to incorporate feedback control into traffic networks. While some of the controllers in every single vehicle, such as adaptive cruise control, may have some impact on the macroscopic traffic flow in the network, we will in this thesis focus on control actions that have a more direct and more significant effect on the overall network state. The three most significant actuators in transportations networks are:

Traffic Signal Control The control of traffic signals has a significant impact on the urban and arterial traffic networks. Well-tuned traffic signal controllers can improve both the maximum throughput a traffic network can handle, and the smoothness of the traffic flows, in the sense of how many times the drivers have to stop for red signals during their journeys.

Routing With more real-time traffic data available, many drivers nowadays are using route guidance applications, such as shown in Figure 1.1 to avoid congestions

(a) Ramp-metering. Photo by Benpaul12 at English Wikipedia (Public domain), via Wikimedia Commons.

(b) Variable Speed Limit. Photo by Oregon Department of Transportation (CC BY 2.0).

Figure 1.2 Examples of ramp-metering, where a traffic light limits the inflow of vehicles to a highway, and variable speed limit signs. In this example, the variable speed limit is used to improve the safety, but it can be utilized to improve the traffic conditions as well.

and minimize their traveling time. Also, the introduction of tolls or route recommendations through road signs can affect the drivers' route choices.

Ramp-metering and Variable Speed Limits By limiting the inflow to a highway or reducing the speed on the highway, as shown examples of in Figure 1.2, one can avoid congestion effects. Fundamentals of traffic flow theory state that the traffic flows better when it is not too congested, so by limiting the upstream flow on a highway strip by ramp-metering and variable speed limit signs, one can achieve shorter travel time by avoiding the traffic entering into a congested state.

1.2 Related Work

In this section, we make a review of previous work. We start by looking at different models for traffic flows in networks, then discuss the three main control problems in traffic networks: traffic signal control, routing, and ramp-metering and variable speed limits.

Modeling

One of the most classical models for traffic flow on a highway is the LWR (Lighthill-Whitham-Richards) model [Lighthill and Whitham, 1955; Richards, 1956]. It is a macroscopic model, i.e., describes the vehicles as a continuous quantity and models the aggregate traffic flow and not each vehicle's movement.

In [Daganzo, 1994], the author proposes a discrete time model for traffic flows, where the highway is split into cells, where the flow between two succeeding cells is

determined as the minimum between the upstream cells demand and the downstream cell's supply. This model is referred to as the cell transmission model (CTM) is a discretization of the LWR-model. In [Daganzo, 1995] the cell transmission model is extended from a highway segment to a network, where several cells can merge into one cell, and one cell can split into several.

In the model presented in [Daganzo, 1994; Daganzo, 1995], the demand, and supply functions are both affine, which implies that the whole system dynamics is a switched linear system. In [Munoz et al., 2003] the authors present this fact, together with observability and controllability results, when the actuators are ramp meters. A formal stability analysis of this switched linear system is done in [Pisarski and Canudas-de-Wit, 2012].

While the previous references study the CTM when the demand and supply functions are affine, extensions to non-linear functions can be made as well. In [Lovisari et al., 2014] the stability of systems with general demand and supply functions are analyzed under the assumption that the vehicles can overtake each other, i.e., the flows progress in a non-first-in-first-out (non-FIFO) manner. The stability analysis utilizes monotonicity properties of the dynamical system, which is further explained in [Como, 2017]. In [Coogan and Arcak, 2015; Coogan and Arcak, 2016] the authors show that the monotonicity may be lost when the flows split up according to the FIFO-rule, i.e., when vehicles can not pass by each other, and the loss of monotonicity is beneficial for some control actions.

While the cell transmission model is a first-order model, second-order models exist as well, such as the METANET model [Messmer and Papageorgiou, 1990; Kotsialos et al., 2002]. Since these models are of higher order, they can capture more phenomena observed in traffic flows. However, more complex models are also more challenging to analyze, and often first-order dynamical models are considered good enough.

Traffic Signal Control

The first traffic lights in the early 20th century were controlled manually by a police officer in every junction [McShane, 1999] . Technological development made it soon possible to both control the traffic lights automatically and centralized.

With a centralized approach to traffic signal control, it is possible to coordinate the cycles in the traffic signals, so that they allow traffic on the main corridors in a city to progress smoothly, sometimes referred to as "green-waves". One early computer implementation of an algorithm that computes an optimal traffic signal control is TRANSYT [Robertson, 1969], which compute a static signal program. Later, other approaches to compute the optimal offset in signal timing has been developed, for instance, [Gomes, 2015; Coogan et al., 2017b; Mehr et al., 2018].

By using magnetic loop detectors to measure the traffic flow, several solutions have been proposed on how to retune the signal programs, for example, SCAT [Sims and Dobinson, 1980], SCOOT [Robertson and Bretherton, 1991], UTOPIA [Mauro

and Di Taranto, 1990]. While those retuning strategies take several practical aspects into account, they do not have any formal performance guarantees, such as stability of the dynamical system or that they will achieve throughput optimality.

One feedback solution for traffic signal control is the MaxPressure controller [Varaiya, 2013b; Varaiya, 2013a]. It determines how the traffic light should be controlled in real-time, and it has formal guarantees about the stability of the traffic network and throughput optimality. The MaxPressure controller is utilizing ideas from the BackPressure controller, a controller developed for communication networks. The BackPressure controller was originally proposed in [Tassiulas and Ephremides, 1992] and has then been further developed in, e.g., [Neely, 2003]. The idea of adapting the BackPressure controller to traffic signal control has also independently been proposed in [Wongpiromsarn et al., 2012]. While the BackPressure controller both controls which queue that should be served and how the packets should be routed, only service control can usually be done in traffic networks. This because the drivers themselves do the routing. For the traffic setting, the controller then needs information on how the drivers propagate through the network. A variant of the MaxPressure controller when the controller does not have the right information about how the vehicles turn has been proposed in [Gregoire et al., 2014]. In that paper, the authors show that if the routing information to the controller is not correct, the maximum throughput in the traffic network will be less. Due to the need for accurate routing information in the MaxPressure controller, approaches have been taking to estimate this information from loop detector information [Coogan et al., 2017a]. Also, the idea of utilizing the routing suggestions from the BackPressure controller and variants thereof has been proposed in [Zaidi et al., 2016; Gregoire et al., 2016; Le et al., 2017]. While the original MaxPressure controller does not take into account the practical limitation on finite storage capacities on the lanes, approaches to solving this without full guarantees about staying below capacity have been taken in [Gregoire et al., 2015].

If the routing information is known to the controller, Model Predictive Control-like (MPC) solutions for signal control is possible as well, as proposed in [Grandinetti et al., 2018], [Bianchin and Pasqualetti, 2018] and [Hao et al., 2018a; Hao et al., 2018b].

Another feedback scheduling policy for communication networks is the proportional fairness policy, proposed in [Kelly, 1997]. Its stability properties have been analyzed for stochastic networks in [Massoulié, 2007] and [Walton, 2014]. The feedback controller for traffic signal control in this thesis is developed based on the ideas of proportional fairness. The idea of using proportional fairness for signal control has also been proposed in [Kovacs et al., 2016]. While the implementation in [Kovacs et al., 2016] adjusts the cycle lengths on a longer time scale compared to the phase activation times, the solution in this thesis adjust the cycle length for every cycle by computing the phase activation and cycle length simultaneously.

Signalized intersections can also be utilized to balance the traffic load among different regions in a city, commonly referred to as perimeter control. In [Geroliminis

et al., 2013] the authors propose an MPC-approach for doing this balancing and in [Mehr et al., 2017] the authors show a solution that does perimeter and signal control jointly inside the region.

Routing

The basic principles of traffic equilibria with respect to route choices were stated in [Wardrop, 1952]. Wardrop introduced two different equilibria, the user optimal and the system optimal. The user optimal equilibrium is the one when each driver tries to take their time-optimal path, i.e., at equilibrium the driver cannot change her route choice to shorten her total travel time. On the other hand, in the system optimal equilibrium, the average journey time is optimal. Since the equilibria usually do not coincide, one usually refers the loss of efficiency in a user optimal equilibria as the price of anarchy and bounds on this price has been presented [Roughgarden and Tardos, 2002]. For a comprehensive overview of the static traffic assignment problem, see [Patriksson, 2015].

The dynamic traffic assignment (DTA) problem was introduced in [Merchant and Nemhauser, 1978a; Merchant and Nemhauser, 1978b]. At this time, on-line guidance was certainly not available to the drivers, so the main interest of this was for planning purposes. Later a linear programming solution to compute the system optimal routing assignment was proposed in [Ziliaskopoulos, 2000]. In [Como et al., 2016] the authors show that the dynamic traffic assignment can be computed through convex relaxation, and in [Ba and Savla, 2016; Rosdahl et al., 2018] it is shown how this convex relaxation could be solved in a distributed way.

While many of the DTA solutions assume that it is possible to control all drivers' route choices, [Samaranayake et al., 2018] proposes a solution when only a fraction of vehicles' route choices can be controlled.

In [Como et al., 2013a; Como et al., 2013b] the authors study a locally responsive routing policy, where each driver adjusts her route depending on the traffic volume in the cells. Under the assumption that the drivers are avoiding congestion, i.e., when the traffic volume increases in one cell, it is not more likely that the driver will choose that cell, the authors show both stability of the system. They also show that under this kind of routing policies, cascade effects will not occur in the network. This means that the network is resilient to perturbations. While the results in [Como et al., 2013a; Como et al., 2013b] only focus on local behaviors in the network, the work in [Como et al., 2013c] investigates local responsive routing policies together with a global routing dynamics, where the drivers converge towards a Wardrop equilibrium. In [Yazicioglu et al., 2018] the authors combines the locally responsive routing policies with a variable speed limit control to improve the transportation network's resilience. Another model for how drivers update their route choices dynamically is presented in [Shah and Shin, 2010], where the drivers' probability to reevaluate their route choices depends on the current congestion level.

Ramp-Metering and Variable Speed Limits

The fact that limiting the inflow to a highway with ramp-metering will reduce the overall total traveling time has been shown in [Gomes et al., 2008], although the idea of ramp-metering has existed long before this observation. One of the earliest feedback control strategies for ramp-metering is ALINEA [Papageorgiou et al., 1991], which is a linear proportional control strategy. Later on, the ALINEA controller was extended with integral action [Wang et al., 2010]. Proportional-integral control has also been proposed for variable speed limit control [Carlson et al., 2011].

Model-predictive-control solutions have also been proposed for ramp-metering and variable speed limit control, for example, in [Hegyi et al., 2005; Gomes and Horowitz, 2006; Muralidharan and Horowitz, 2012]. As the traffic flow model is non-linear, those control strategies are computationally heavy, and approaches have been taken to decrease the number of computations needed, for instance by event-trigged approaches [Ferrara et al., 2015b], or make the computations distributed [Ferrara et al., 2015a]. Also, with MPC-techniques it is fairly easy to extend the models, in the hope to achieve better accuracy. For an example of a MPC-solution, see [Pasquale et al., 2015]. While the MPC-techniques are tractable, it is in [Schmitt et al., 2017] shown that in some cases, a non-predictive control will achieve as good performance, due to the monotonicity properties of traffic dynamics.

In [Pisarski and Canudas-de-Wit, 2016] the authors take a game-theoretic approach to ramp metering control, which results in a distributed solution.

1.3 Outline of the Thesis and Related Papers

This thesis consists of 8 chapters. Chapter 2 serves as an introduction to the models for traffic networks used in the thesis. Also an overview of the control problems in transportation networks is given in this chapter. In Chapters 3–5 different aspects of feedback based traffic signal control are discussed, while Chapters 6 and 7 discuss a few routing problems. The thesis is concluded in Chapter 8 together with some suggestions for further research.

Chapter 2 - Preliminaries on Dynamical Flow Networks

The goal of chapter 2 is to introduce concepts and notation that will be used in the following chapters. Hence, this chapter serves as survey and no original material is presented. We present the model for dynamical flow networks that we will use throughout the thesis. The model is a generalization of existent dynamical models for networks flows. Therefore, the chapter also shows how a couple of models used in the literature fit into the presented modeling framework. Those models are the cell transmission model, which is a well-used model to capture traffic flows on a macroscopic level with congestion effects, and a fluid model for point queues, i.e., a queuing model where spacial distribution of vehicles in a queue is disregarded. The

chapter also includes a description of how the control action from different physical actuators in traffic networks enters the model.

Chapter 3 - On the Well-Posedness of a Feedback Controlled Point Queue Model

In this chapter, we discuss the technical issue of existence and uniqueness of solutions to the dynamic model for point queues presented in Chapter 2. In short, this issue arises when an empty lane receives a green light in a signalized junction. While previous work only studied this problem when a predefined controller controls the traffic signal, we show that a unique solution to the dynamical point queue model exists when the control action is determined by a feedback controller.

This chapter is based on the following article:

Nilsson, G. and G. Como (2019). "On well-posedness of feedback-controlled outflows in dynamical flow networks". In preparation.

Authors' contributions: GN developed the theory together with GC. GC also gave valuable suggestion in how to improve the presentation. GN and GC wrote the manuscript together.

Chapter 4 - Decentralized Throughput Optimal Traffic Signal Control

This chapter, presents a fully decentralized feedback-based traffic signal controller. To avoid collisions, constraints are often present that limit which lanes that can receive green light simultaneously, commonly referred to as phases. The controller determines how to split the activation between different phases. To do this, the controller only needs information about the queue lengths on the incoming lanes to the signalized junction. We also present a fundamental bound on the maximal throughput that any traffic signal controller can achieve for a given set of phases. We show that the proposed controller achieves this bound, despite the little information the controller requires.

This chapter is based on the following article:

Nilsson, G. and G. Como (2019). "Generalized proportional allocation policies for robust control of dynamical flow networks". In preparation.

Authors' contributions: GN developed most of the results, with feedback from GC. GC also contributed with ideas how to improve some of the proofs. GN wrote most of the manuscript, with valuable comments and suggestions how to improve the clarity of the presentation from GC.

The paper above is a journal version where results from the following paper are included:

Nilsson, G. and G. Como (2017). "On generalized proportional allocation policies for traffic signal control". *IFAC-PapersOnLine* **50**:1. 20th IFAC World Congress, pp. 9643–9648. DOI: 10.1016/j.ifacol.2017.08.1728.

Authors' contributions: GN developed most of the results, in discussion with GC. GN wrote most of the paper, with valuable comments from GC on the manuscript.

A result for a simplified setting where only one lane in each junction can receive green-light simultaneously was presented in:

Nilsson, G., P. Hosseini, G. Como, and K. Savla (2015). "Entropy-like Lyapunov functions for the stability analysis of adaptive traffic signal controls". In: *2015 54th IEEE Conference on Decision and Control (CDC)*, pp. 2193–2198. DOI: 10.1109/CDC.2015.7402532.

Authors' contributions: The idea of using proportional fairness to control traffic signals was proposed by GC and KS. GN and PH developed the results independently in discussions with GC and KS respectively. GN and PH wrote the paper together, with input from GC and KS.

Chapter 5 - Evaluation of GPA Control in a Microscopic Traffic Simulator

In this chapter, we present a discretized version of the traffic light controller introduced in Chapter 4. The discretized version is implemented in a microscopic traffic simulator, and its performance is evaluated against both a fixed-time controller and another well-known feedback controller for traffic signals, the MaxPressure controller. The performance evaluation is done in an artificial Manhattan-like grid, to make the setting as ideal as possible. To both illustrate how easy it is to implement the controller and how the controller can improve a realistic traffic situation, the controller is also implemented in a realistic traffic scenario. The scenario models all traffic in the city of Luxembourg during a full day.

This chapter is based on the following article:

Nilsson, G. and G. Como (2018). "A micro-simulation study of the generalized proportional allocation traffic signal control". Submitted.

Authors' contributions: GN developed the experiments in discussions with GC. GN implemented the simulations and wrote most of the manuscript with feedback from GC.

The article above is partly based on theory presented in:

Nilsson, G. and G. Como (2018). "Evaluation of decentralized feedback traffic light control with dynamic cycle length". *IFAC-PapersOnLine* **51**:9. 15th IFAC Symposium on Control in Transportation Systems CTS 2018, pp. 464–469. DOI: 10.1016/j.ifacol.2018.07.076.

Authors' contributions: GN developed the experiments in discussions with GC. GN implemented the simulations and wrote most of the manuscript with feedback from GC.

Chapter 6 - Two-Tier Traffic Assignment

In this chapter, we consider the problem when drivers who are taking their user-optimal path have to share the traffic network with a fleet of (possibly autonomous) vehicles trying to select optimal paths for the whole fleet. We present results about existence and uniqueness of such an equilibrium, together with algorithms to compute it.

This chapter is based on the paper:

Nilsson, G., P. Grover, and U. Kalabic (2018). "Assignment and control of two-tiered vehicle traffic". In: *57th IEEE Conference on Decision and Control*, pp. 1023–1028.

Authors' contributions: GN developed the idea under discussions with UK and PG. GN developed most of the results, with input from UK. GN and UK wrote the manuscript together, with feedback from PG.

Chapter 7 - Resilience of Dynamically Routed Multicommodity Flows

In this chapter, we study the problem when several flows are routed through the network. At each node in the network, the flows can make a routing decision about which of the outgoing links to follow. The decision is based upon the current congestion state on the outgoing links, and in a way such that the flows are trying to avoid congested links. Situations like this can occur when, e.g., commuters that know the best route choices well and do not use a routing device, one day have to avoid an accident or a roadwork. We show that a flow equilibrium exists for this kind of routing behaviors, and we also show that introducing a heterogeneous routing behavior may make the network less resilient to perturbations compared to previous results on homogeneous flow behavior.

This chapter is based on the paper:

Nilsson, G., G. Como, and E. Lovisari (2014). "On resilience of multicommodity dynamical flow networks". In: *53rd IEEE Conference on Decision and Control*, pp. 5125–5130. DOI: 10.1109/CDC.2014.7040190.

Authors' contributions: GN developed the theory together with EL and GC. GN and EL wrote the manuscript, with feedback from GC.

1.4 Notation

Most of the notation in the thesis will be introduced when it is used and it is also summarized in the Nomenclature. Nevertheless, some basic notation that will be used throughout the thesis is introduced here. We let $\mathbb{R}_{(+)}$ denote the (non-negative) reals. For a finite set $\mathcal{A}$, we let $\mathbb{R}^{\mathcal{A}}$ denote the set of vectors indexed by the elements in the set $\mathcal{A}$. In the same manner, $\mathbb{R}^{\mathcal{A}\times\mathcal{B}}$ denotes the matrices whose rows are indexed by the elements in the finite set $\mathcal{A}$ and whose columns are indexed by elements in the finite set $\mathcal{B}$.

2

Preliminaries on Dynamical Flow Networks

In this chapter, we introduce a general dynamical model for flow networks, gathering concepts and notation that will be used in the rest of the thesis.

The topology of a transportation network will be modeled as a directed multigraph, so the chapter starts with an introduction to some graph-theoretic notation. With a topological description in place, we then talk about different ways to determine the routing in the network, i.e., the paths that the traffic propagate through the network. Later, we describe the dynamical system of traffic flows in networks, that will serve as the basis for the whole thesis. It will also be shown how these dynamical systems relate to some other well-known models for traffic flows. The chapter is concluded with a brief description of some control challenges in transportation networks, and how those enter into our dynamical model.

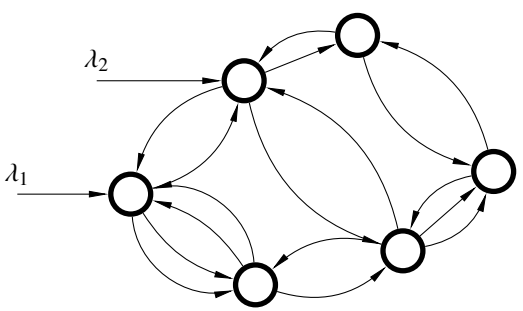


Figure 2.1 An example of a topology for a dynamical flow network.

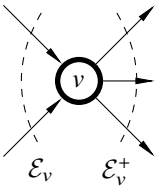


Figure 2.2 The set of in-going, $\mathcal{E}_v$, and out-going, $\mathcal{E}_v^+$, cells to a junction v.

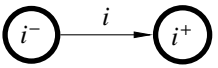


Figure 2.3 For a cell $i \in \mathcal{E}$, i^+ denotes the junction that the traffic from cell i flows out into and i^- denotes the junction that the traffic in the cell is coming from.

2.1 Traffic Networks As Graphs

We model the network topology as a directed multi-graph. We denote this multi-graph as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of *nodes* and $\mathcal{E}$ the multi-set of directed *links* between the nodes. In contrast to a graph, in a multi-graph there can be multiple edges between two nodes, as shown in Figure 2.1.

The topology of a multi-graph can be described through the node-link incidence matrix. For a given multi-graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the node-link incidence matrix $B \in \{-1, 0, 1\}^{\mathcal{V}\times\mathcal{E}}$ is defined as

$$B_{vi} = \begin{cases} 1 & \text{if the tail of link } i \text{ is node } v, \\ -1 & \text{if the head of link } i \text{ is node } v, \\ 0 & \text{otherwise.} \end{cases}$$

A *path* between link i and j is a sequence of cells $\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_l)$ such that $\gamma_0 = i$, $\gamma_l = j$ and $\gamma_n^+ = \gamma_{n+1}^-$ for all $0 \leq n < l$.

The links will further on be referred to as *cells*, and each cell contains traffic moving in one direction. While the nodes may correspond to *junctions*, where incoming traffic can proceed to several downstream cells, nodes can also be introduced to model the boundary between two consecutive cells. As we will see later on, this type of nodes may be useful for space discretization of specific traffic models. However, we will throughout the thesis assume that all the nodes correspond to an actual signalized junction unless otherwise stated. For a given junction $v \in \mathcal{V}$, we let $\mathcal{E}_v$ denote the set of incoming cells to that junction and $\mathcal{E}_v^+$ the set of cells going out from junction v. This is illustrated in Figure 2.2. For a cell $i \in \mathcal{E}$, we let i^+ denote the junction $v \in \mathcal{V}$ that the traffic from cell i will flow out into. In the same manner, i^- denotes the junction that the traffic in cell i is coming from, as shown in Figure 2.3.

To a non-empty subset of the cells, that we will refer to as *source cells*, there is

an exogenous possibly time-varying inflow $\lambda_i(t) \geq 0$. We will denote the set of the source cells as $\mathcal{S} \subset \mathcal{E}$. Moreover, we will denote the full vector of exogenous inflows $\lambda(t) \in \mathbb{R}_+^{\mathcal{E}}$, and let $\lambda_i(0) \equiv 0$ for all non-source cells $\mathcal{E} \setminus \mathcal{S}$. To model the dynamical flows within the network, we will throughout the thesis denote the outflow from each cell i with z_i, and $z \in \mathbb{R}_+^{\mathcal{E}}$ is the vector of outflows from all cells.

To describe how the traffic flow splits up in the junctions, we introduce a possibly both time-varying and state-dependent *routing matrix* $R \in \mathbb{R}_+^{\mathcal{E}\times\mathcal{E}}$, whose elements are all non-negative. Each element in the routing matrix is telling how large fraction of the outflow of one cell that will turn to another, and the elements will therefore be referred to as *turning ratios*. The routing matrix is assumed to have the following properties:

1. $R_{ij} > 0$ if cell i is connected to cell j through a node.
2. $R_{ij} \leq 1$ is the fraction of flow that proceeds from cell i to cell j.
3. For a given cell $i \in \mathcal{E}$, $1 - \sum_{j\in\mathcal{E}} R_{ij} \geq 0$ is the fraction of flow that leaves the network after have passed by cell i.

In practice, the routing matrix can either be obtained from traffic measurements or computed by an assignment problem. In the Section 2.2, we will show how such an assignment can be computed, both under the assumption that each driver takes her optimal path, and when the routing is controlled by a central authority.

With the exogenous inflow vector and the routing matrix in place, we can state a couple of connectivity properties. The first one is about if a cell is reachable by any exogenous inflow and the second one if traffic in a cell can leave the network by following some path in the network.

Definition 2.1
A pair, (λ, R), of an exogenous inflow vector and routing matrix is said to be *inflow-connected* if there for any given cell $i \in \mathcal{E}$ exists a path from a cell $j \in \mathcal{E}$ with $\lambda_j > 0$ to cell i. Moreover, a routing matrix is said to be *out-connected* if there from every cell i, exists a path to some cell j, denoted *sink-cell*, such that $\sum_{k\in\mathcal{E}} R_{jk} < 1$. □

We also note the following properties of an out-connected routing matrix, that we will make use of later:

Proposition 2.1
If R is out-connected, then

i) it has spectral radius strictly less than 1,
ii) there exists an induced norm $\|\cdot\|_\dagger$ such that $\left\|R^T\right\|_\dagger < 1$, and,
iii) for any subset $\mathcal{I} \subset \mathcal{E}$ the routing matrix describing the routing between the cells in $\mathcal{I}$, $R_{\mathcal{I}\mathcal{I}}$, i.e., the restriction of R on $\mathcal{I} \times \mathcal{I}$, is out-connected.
iv) the matrix $I - R^T$ is invertible. □

The proof of the proposition is given in Appendix A.1.

2.2 Static Flow Optimization in Transportation Networks

In the previous section, the routing matrix was introduced to describe the paths of the traffic through the transportation network. While the routing matrix can be estimated from traffic data, it can also be estimated from models of drivers' path preferences, referred to as traffic assignment.

To compute a traffic assignment, information about the exogenous inflows to the nodes and outflows from the nodes is needed. To distinguish the exogenous inflow to the cells from the exogenous inflows to the cells, we let $\hat{\lambda} \in \mathbb{R}_+^{\mathcal{V}}$ denote the vector of exogenous inflows to the nodes and $\hat{\mu} \in \mathbb{R}_+^{\mathcal{V}}$ denote the exogenous outflows from the cells. This modification does not make us lose generality with respect to the previously introduced model. An exogenous inflow $\hat{\lambda}_v$ to a node $v \in \mathcal{V}$ can be represented as a cell i connected to the node, with $\lambda_i = \hat{\lambda}_v$.

To model the propagation delay, every cell $i \in \mathcal{E}$ is equipped with a *delay function* $\tau_i : \mathbb{R}_+ \to \mathbb{R}_+$, that is dependent upon the traffic flow z_i in the cell. The delay functions are assumed to be non-negative, differentiable, and non-decreasing.

As mentioned in Chapter 1.2, two common equilibria for static traffic assignment is the system optimal assignment (SO) and the user optimal equilibrium (UE) [Ozdaglar and Menache, 2011].

Let us first start with considering a single-origin single-destination assignment with exogenous inflow $\alpha > 0$, origin node $o \in \mathcal{V}$, and destination node $d \in \mathcal{V}$. In other words, $\hat{\lambda} = \alpha\delta^{(o)}$ and $\hat{\mu} = \alpha\delta^{(d)}$, where $\delta^{(o)}$ and $\delta^{(d)}$ are vectors with all zero elements apart from a 1 element at index o and d respectively.

The system optimal assignment is computed by solving

$$\begin{aligned} \underset{z \in \mathbb{R}_+^{\mathcal{E}}}{\text{minimize}} \quad & \sum_{i \in \mathcal{E}} z_i \tau_i(z_i), \\ \text{subject to} \quad & Bz = \hat{\lambda} - \hat{\mu}\,. \end{aligned}$$

The system optimal assignment optimizes the total travel time for the overall transportation network. This may imply that some users can get a shorter travel time by deviating from their assigned path.

In the user optimal equilibrium (UE), also referred to as Wardrop equilibrium, each driver is trying to take the path with the shortest delay. This means that no driver can achieve a shorter delay by taking a different path. Let Γ denote the set of all paths in the network that start from a cell connected from node o and end with a cell connected to node d. Formally, a traffic assignment $z^* \in \mathbb{R}_+^{\mathcal{E}}$ is an user optimal assignment if

$$\begin{aligned} &\sum_{i \in p_1} \tau_i(z_i^*) = \sum_{j \in p_2} \tau_j(z_j^*) \quad \text{for all paths } p_1, p_2 \in \Gamma \text{ with } \sum_{i \in p_1} z_i^* > 0, \sum_{j \in p_2} z_j^* > 0 \text{ and} \\ &\sum_{i \in p_1} \tau_i(z_i^*) \geq \sum_{j \in p_2} \tau_j(z_j^*) \quad \text{for all paths } p_1, p_2 \in \Gamma \text{ with } \sum_{i \in p_1} z_i^* = 0, \sum_{j \in p_2} z_j^* > 0. \end{aligned}$$

It has been shown in [Beckman et al., 1956] that the user optimal equilibrium can be computed by solving the following optimization problem

$$\begin{aligned} \underset{z \in \mathbb{R}_+^{\mathcal{E}}}{\text{minimize}} \quad & \sum_{i \in \mathcal{E}} \int_0^{z_i} \tau_i(s) \mathrm{d}s \,, \\ \text{subject to} \quad & Bz = \hat{\lambda} - \hat{\mu} \,. \end{aligned}$$

While the user equilibrium determines the routing if each driver takes her optimal path, sometimes referred to as drivers behaving anarchistic, the system optimal assignment minimizes the total travel time for the whole system. Since the total travel time for a user optimal equilibrium is always greater than or equal to a system optimal assignment, there will be a loss of optimality with user optimal behaviors. This loss of system-optimality is often referred to as the *price of anarchy*, which is the ratio between the total travel time for the user optimal equilibrium and the system optimal assignment. In the case of affine delay functions, i.e., $\tau_i(z_i) = \alpha_i z_i + \beta_i$ where $\alpha_i > 0$ and $\beta_i \geq 0$, the price of anarchy is bounded by $4/3$, as shown in [Roughgarden and Tardos, 2002], while there exist other convex delay functions that will make the price of anarchy arbitrarily large.

One solution to control the user optimal equilibrium towards the system optimal, and hence reduce the price of anarchy, is to introduce tolls. Under the assumption that there is an equivalence between time and money, tolls denoted ω_i that act as extra delays can be computed as

$$\omega_i = z_i^* \tau_i'(z_i^*),$$

for all $i \in \mathcal{E}$ and where z_i^* are the flows at system optimum [Nisan et al., 2007]. With those tolls, the delay for each cell is now given by $d_i(z_i) + \omega_i$, and the user optimal equilibrium will coincide with system optimal assignment. Thus the price of anarchy is reduced to one.

When the assigned flows have been computed, which we will denote z^*, the routing matrix can be determined. In order to do so, for each junction $v \in \mathcal{V}$ with $\hat{\lambda}_v > 0$ we add an on-ramp cell $i' \in \mathcal{E}$ going into the junction with constant flow $\hat{\lambda}_v$. The routing matrix for all cells $i, j \in \mathcal{E}$ is then given by

$$R_{ji} = \begin{cases} \dfrac{z_i^*}{\sum_{\ell \in \mathcal{E}_{i^-}^+} z_\ell^*} & \text{if } \sum_{\ell \in \mathcal{E}_{i^-}^+} z_\ell^{k*} > 0 \text{ and } j^+ = i^- \,, \\ 0 & \text{otherwise.} \end{cases}$$

While we in this section until now only have considered the traffic assignment problem with one origin and one destination, the theory applies for multiple origin-destination pairs as well [Ozdaglar and Menache, 2011]. A natural relaxation of the problem with multiple origin-destination pairs is instead of describing the exogenous flows by origin-destination pairs, describe the exogenous flows as general vectors $\hat{\lambda}$ and $\hat{\mu}$ with the aggregate net-inflow and net-outflow respectively. Those vectors then have to satisfy $\sum_{v \in \mathcal{V}} \hat{\lambda}_v = \sum_{v \in \mathcal{V}} \hat{\mu}_v > 0$.

2.3 Dynamical Flow Networks

With the exogenous inflows and routing described, all the exogenous parameters to describe the dynamical flow network is in place. We let the traffic volume in each cell $i \in \mathcal{E}$ be the state-space of our dynamical model and denote the traffic volume of all cells by $x \in \mathcal{X}$ where $\mathcal{X} = \mathbb{R}_+^{\mathcal{E}}$.

The change of traffic volume, x_i, in each cell is given by the conservation of mass and is simply the sum of exogenous and internal inflows from upstream cells minus the outflow from the current cell. Hence

$$\dot{x}_i = \lambda_i + \sum_{j \in \mathcal{E}} R_{ji} z_j - z_i \,, \quad \forall i \in \mathcal{E} \,. \tag{2.1}$$

In vector form, the dynamics can be described as

$$\dot{x} = \lambda + (R^T - I)z \,. \tag{2.2}$$

Until now, the dynamical flow network model is just a linear system and the outflow from each cell is not yet prescribed. Depending on the application, constraints that are possibly non-linear may limit the flow between cells. Usually, the outflow from a cell $i \in \mathcal{E}$ is constrained by a demand function $z_i \leq d_i(x_i)$ that depends on the traffic volume x_i in the cell. The demand function is often assumed to be increasing and bounded from above, such that the cell's maximal outflow capacity is given by

$$c_i = \sup_{x_i \geq 0} d_i(x_i) \,.$$

In many models, the outflow from one cell is also limited by the storage capacity of the downstream cells. To each cell, there is an assigned supply function $s_i(x_i) \geq 0$ that limits how much flow from upstream cells that the cell can receive. This function is decreasing, and if a cell's maximum storage capacity is limited by $x_i^{\max} > 0$, it holds that

$$\lim_{x_i \to x_i^{\max}} s_i(x_i) = 0 \,.$$

Depending upon the situation that is modeled, the supply constraint can limit the outflow from a diverge junction in either a first-in-first-out (FIFO) way or in a non-FIFO way. For a narrow single-lane road, the traffic flow is forced to split up in a FIFO manner, while if there are several lanes so vehicles can overtake each other, a non-FIFO behavior can be expected. In the FIFO-case, if any of the downstream cells are full, the total outflow from cell i is given by

$$z_i \leq \gamma_i d_i(x_i) \,,$$

where

$$\gamma_i = \sup\left\{ \gamma \in [0,1] \mid \gamma \cdot \max_{k \in \mathcal{E} \mid (i,k) \in \mathcal{E}} \sum_{h \in \mathcal{E}} R_{hk} d_h(x_h) \leq s_k(x_k) \right\} .$$

For the non-FIFO-case we have to note that the actual turning ratios will be different from the desired ones. We therefore let R denote the prescribed turning ratios, i.e., the routing matrix in (2.2) and $\bar{R}$ the actual turning ratios that may be different from R due to congestion effects. Then the dynamics is described by

$$\dot{x}_i = \lambda_i + \sum_{j \in \mathcal{E}} \bar{R}_{ij} z_j - z_i \,, \quad z_i \leq d_i(x_i) \,,$$
$$\bar{R}_{ij} = \bar{\gamma}_{ij} R_{ij} \,, \quad \bar{\gamma}_{ij} = \sup\left\{ \gamma \in [0,1] \mid \gamma \cdot \sum_{h \in \mathcal{E}} R_{hj} d_h(x_h, \alpha_h) \leq s_j(x_j) \right\} .$$

Since the outflow z_i is only limited from above by inequality constraints, e.g., all out-flows equal to zero is still a valid choice, we will throughout the thesis make the following quite natural assumption about the outflows.

Assumption 2.1
When $x_i > 0$, the outflow z_i is the maximal possible. □

With the assumption above, it is only in the cases when

$$\lim_{x_i \to 0+} d_i(x_i) > 0$$

that the dynamics is not well-specified. We will see later on why this technicality is needed and when it can be guaranteed that the dynamical system has a unique solution.

We conclude this section by giving three examples of known models that fit into this framework. The two first are with traffic applications in mind, where the latter is a computer network example.

Example 2.1—Daganzo's Cell Transmission Model
Let each demand function be

$$d_i(x_i) = \min\left\{ \frac{v_i x_i}{L_i}, c_i \right\} ,$$

where $v_i > 0$ is the cell's free-flow speed (usually the speed-limit), $L_i > 0$ the cell's length and c_i the maximal outflow capacity. Moreover, let each supply function be

$$s_i(x_i) = \max\left\{ \frac{w_i}{L_i}(x_i^{\max} - x_i), 0 \right\} ,$$

where $w_i > 0$ is the cell's shock wave speed, i.e., the speed by which a congestion back-propagates. A plot of these demand and supply functions for a given cell i, is shown in Figure 2.4. This type of diagrams illustrating how the flow at one cell depends on the traffic volume, are referred to as fundamental diagrams for traffic flow

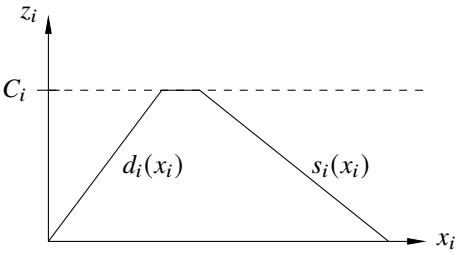


Figure 2.4 The demand and supply functions for a cell i, in Daganzo's cell transmission model. These type of diagrams showing the relationship between the flow and traffic volume, is commonly referred to as fundamental diagrams for traffic flow.

and have also been observed empirically. For examples of empirical observations, see e.g., [Dervisoglu et al., 2009].

A time-discretized version of (2.1) for a line graph where the cells are numbered in order $\{1, 2, \ldots, n\}$ and for a given initial state $x(0)$ then reads

$$\begin{aligned} x_1(k+1) &= x_1(k) + h\left(\lambda_1 - \min\{d_1(x_1), s_2(x_2)\}\right), \\ x_i(k+1) &= x_i(k) + h\left(\lambda_i + \min\{d_{i-1}(x_{i-1}), s_i(x_i)\} - \min\{d_i(x_i), s_{i+1}(x_{i+1})\}\right), \\ &\qquad i \in \{2, \ldots, n-1\} \\ x_n(k+1) &= x_n(k) + h\left(\lambda_n + \min\{d_{n-1}(x_{n-1}), s_n(x_n)\} - d_n(x_n)\right), \end{aligned}$$

where $h > 0$ is the discretization step. Under the assumption that the discretization step $h \leq \frac{L_i}{x_i}$ for all cells $i \in \mathcal{E}$, we have obtained the cell transmission model for traffic flow presented in [Daganzo, 1994; Daganzo, 1995]. Observe that in this example, the nodes between the cells are not real junctions and just introduced in order to make a suitable discretization. □

Example 2.2—Point Queue Model
Let

$$d_i(x_i) = c_i$$

with $c_i > 0$ and

$$s_i(x_i) = +\infty .$$

This model is a point queue model. If there are traffic present in the cell, they leave the cell with the maximum outflow capacity, since no downstream congestions are affecting the outflow. This point queue model has previously been used in [Muralidharan et al., 2015] and [Hosseini and Savla, 2017] to model signalized traffic networks. □

Observe that for the cell transmission model, the traffic volume x consists of traffic moving forward, while for the point queue model the traffic volume represents traffic standing still waiting for service.

While the two previous examples show how the dynamical network flow model connects to transportation networks, there are other applications where a similar model structure is used.

Example 2.3—Mean-Value Dynamics in Computer Networks
Let

$$d_i(x_i) = c_i \mu \frac{x_i}{1 + x_i},$$

where $1/\mu > 0$ is the average packet length, and

$$s_i(x_i) = +\infty .$$

This model was presented in [Tipper and Sundareshan, 1990] to have a fluid model of M/M/1-queues in a computer network, where the states correspond to the mean values of the queue-lengths. □

2.4 Some Control Challenges

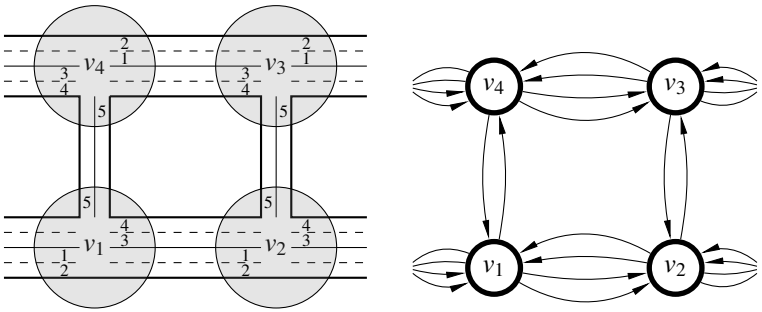
In this section, three different categories of control challenges in dynamical traffic flow networks will be presented, namely ramp-metering and variable speed limits, traffic signal control, and routing. While ramp-metering and variable speed limits are included for completeness of the presentation, the latter two control problems will be the focus of this thesis.

Ramp-metering and Variable Speed Limits

Ramp-metering and variable speed limits are used on highways to control the traffic flow. Ramp-metering is usually done by placing a traffic light on the on-ramps, enforcing the drivers to wait for a short time before entering the highway. In this way, a small decrease in the inflow to the highway can be achieved and may prevent the traffic volume to get above the free-flow region. Another way to avoid that a cell in highway gets congested is to limit the speed in the upstream cells by variable speed limit signs, as shown in Figure 1.2. Again, the idea is here to limit the inflow into one cell, to avoid that the traffic volume goes above the critical one.

For variable speed limits the control action is to reduce the free-flow speed. Since the free-flow speed is related to the slope of the demand function, the control action has to enter into a modified version of the demand function, $\tilde{d}_i(x_i, \alpha_i)$, such that the outflow $z_i \leq \tilde{d}_i(x_i, \alpha_i) \leq d_i(x_i)$ where $\alpha_i \in [0, 1]$.

For ramp-metering, the control action is rather limiting the fraction of time when outflow is allowed from the on-ramp cell. Hence the outflow from a ramp-metered cell is limited such that $z_i \leq \beta_i d_i(x_i)$, where $\beta_i \in [0, 1]$ if the control action is time-averaged and $\beta_i \in \{0, 1\}$ if the control action is not time-averaged.



(a) The network. (b) The graph representation.

Figure 2.5 A subset of a traffic network and its graph representation. In the graph, each link corresponds to one lane or cell wherein traffic is stored. The nodes in the graphs correspond to junctions.

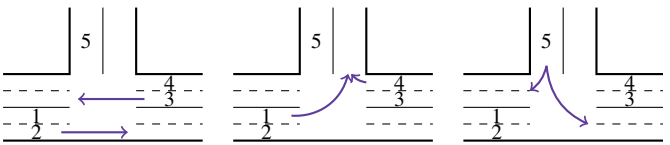


Figure 2.6 Example of a set of phases for a T-junction. The set of phases consist in this case of three different phases.

Traffic Signal Control

Traffic signals are introduced to avoid collisions and make the traffic flow smoother in signalized junctions. To each junction or set of junctions, there is a set of phases assigned where each phase contains a set of lanes that can receive green light simultaneously. The controller's task is then to decide which phase that should be activated and for how long.

To describe which cells that can receive service simultaneously in a junction, we introduce phases, where each phase contains a subset of edges. For a given junction $v \in \mathcal{V}$, with p phases and n incoming cells, the set of phases can be described as a phase matrix $P^{(v)} \in \mathbb{R}_+^{n \times p}$ where

$$P_{i,j}^{(v)} = \begin{cases} 1 & \text{if cell } i \text{ belongs to phase } j \\ 0 & \text{otherwise} \end{cases}.$$

Example 2.4
Consider the junction denoted v_1 in the network shown in Figure 2.5. One possible

set of phases is illustrated in Figure 2.6. The corresponding phase matrix is then

$$P^{(v_1)} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T .$$

We will in Chapter 4 show that the set of phases can span over more than one junction as well.

Just like the ramp-metering, the control action will limit the outflow from the cells such that $z_i \leq \zeta_i d_i(x_i)$, where $\zeta_i \in [0, 1]$ if the control action is time-averaged and $\zeta_i \in \{0, 1\}$ if the control action is not time-averaged.

Routing

The routing problem is about how traffic should be guided through a traffic network, i.e., which paths the traffic should take. While the traffic management often solves the previous two control problems, the routing problem is usually solved by the driver herself or some third-party service. Depending on which type of objective and where the routing decisions are taking place, the control actions can be very different. For instance, as shown earlier, the traffic can be guided towards a system optimal assignment either via direct guidance or via tolls. However, all control actions will affect the R matrix in (2.1).

3

On the Well-Posedness of a Feedback Controlled Point Queue Model

In this chapter, we introduce feedback to the point queue model presented in Chapter 2. As already observed in Chapter 2, the dynamics of the flow network may be specified through inequalities when the traffic volumes are zero in some cells. This means that classical existence and uniqueness results of a solution to a dynamical system do not apply. However, even if inequalities specify the outflows from some cells, we will in this chapter show that a unique solution exists given that the feedback controller is Lipschitz continuous.

This problem without feedback has been studied before, originally for Brownian motion in [Harrison and Reiman, 1981]. Later, the idea has been applied to open loop traffic signal control in [Muralidharan et al., 2015], where the authors assume that there is a strictly positive propagation delay between the junctions. In [Hosseini and Savla, 2017] the authors solve the existence and uniqueness problem with arbitrary delays and also provide an algorithm to compute the trajectories. While the previous work only considers an open-loop control, where the outflow control solely depends on time, we provide a result when the control signals are determined by feedback.

3.1 The Feedback Controlled Continuous Time Point Queue Model

The point queue model, sometimes referred to as a vertical queue model, is a simple fluid model for queuing networks. When studying the point queue model, we will assume that there are no supply constraints present in the network dynamical flow model presented in Chapter 2, i.e., $s_i(x_i) = +\infty$ for all cells $i \in \mathcal{E}$. Moreover, the outflow from each cell $i \in \mathcal{E}$ is limited by a feedback-controller, denoted $\zeta_i : \mathcal{X} \to \mathbb{R}_+$. Since we are considering a point queue dynamical system, the demand is equal to

the queue's outflow capacity whenever the traffic volume is strictly positive. Hence, $d_i(x_i) = c_i$ for all cells $i \in \mathcal{E}$. To simplify the notation, we let C denote the matrix with the outflow capacities on the diagonal, i.e., $C = \text{diag}(c)$. Within this setting, the dynamics for flow networks in (2.1) can now be described in vector form as

$$\dot{x} = \lambda + (R^T - I)z, \tag{3.1}$$

$$0 \leq z \leq C\zeta(x), \tag{3.2}$$

$$0 = x^T(z - C\zeta(x)). \tag{3.3}$$

Observe that the two last equations imply that when the traffic volume is non-zero in one cell, the outflow from the cell will be the maximum outflow that is allowed by the controller. In other words, if the traffic volume $x_i > 0$ for some cell $i \in \mathcal{E}$, then the outflow $z_i = c_i\zeta_i(x)$. Hence Assumption 2.1 is satisfied.

3.2 Existence and Uniqueness of Solution

In this point queue model, it may happen, just like in reality, that an empty cell receives a green-light. This is the reason why inequalities specify the outflow, since if the outflow was equal to the one prescribed by the controller, the traffic volume may have become negative. Moreover, the inflow to the downstream cells has to be correct in the model as well. That is why it is not sufficient to let the inflow to a downstream cell equal the control signal times the flow capacity and just set the cell's traffic volume to zero whenever it becomes negative.

In this section we will show that even if the outflow is just limited with inequalities when the traffic volume is zero, there exists a unique continuous trajectory as a solution to the dynamics (3.1) – (3.3). We start this section by giving a motivating example of such occurrence, and also show what a continuous solution can be. After that, a formal proof for the existence of solutions to the dynamics in (3.1) – (3.3) is provided.

A Motivating Example

To illustrate the issue of existence and uniqueness of solutions for a dynamical flow network, we provide a small motivating example.

Example 3.1
Consider the small network whose topology is depicted in Figure 3.1. Let the routing matrix be

$$R = \begin{bmatrix} 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0.3 & 0.7 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and the exogenous inflows be $\lambda_1 = 0.5$, $\lambda_2 = 0.4$, and $\lambda_3 = \lambda_4 = 0$. The outflow capacity is set to be $c_i = 1$ for all cells $i \in \{e_1, e_2, e_3, e_4\}$ and the controller for cell

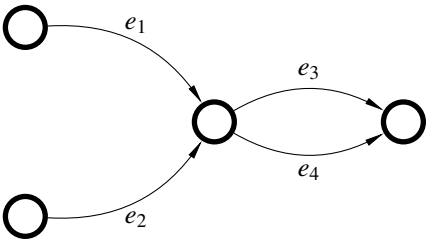


Figure 3.1 The network in Example 3.1.

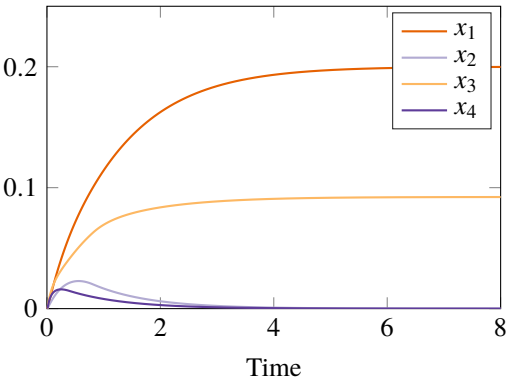


Figure 3.2 How the traffic volumes evolve in time in Example 3.1.

e_1 and e_2 is

$$\zeta_1(x) = \zeta_2(x) = x_1 + x_2 + 0.3\,,$$

while for cell e_3 and e_4, it is

$$\zeta_3(x) = \zeta_4(x) = \frac{x_3 + x_4}{x_3 + x_4 + 0.1}\,.$$

In this example, ζ_1 and ζ_2 will both be strictly positive, even when the traffic volumes in those cells are zero. Moreover, ζ_3 will be strictly positive when the traffic volume is strictly positive in cell e_4, even if cell e_3 is empty.

The unique trajectories for the traffic volumes, starting from $x(0) = 0$ are shown in Figure 3.2. Figure 3.3 shows the two control actions, and Figure 3.4 shows the actual outflows from each edge. For two of the edges, e_2 and e_4, the actual outflow is less than the one given by the controller, since the traffic volumes are zero on those edges.

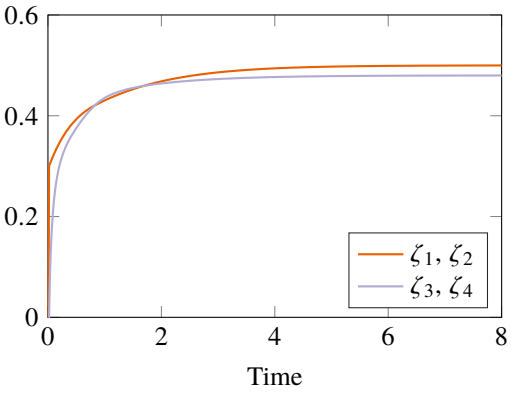


Figure 3.3 How the control actions evolve in time in Example 3.1.

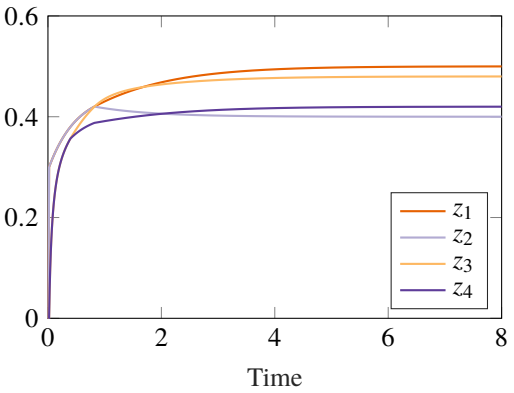


Figure 3.4 How the actual outflows evolve in time in Example 3.1.

A Well-Posedness Result

In this section, we present a proof of existence and uniqueness of a solution to the dynamical system (3.1) – (3.3). The proof is an extension of the reflection principle for Brownian motion, previously presented in [Harrison and Reiman, 1981]. In our case, the outflow from each cell is determined by a Lipschitz continuous feedback controller, while the proof in [Harrison and Reiman, 1981] considers the case when the outflow is limited by a predefined time dependent function, i.e., open-loop control. Apart from that the outflow is feedback-controlled, another key difference to [Harrison and Reiman, 1981] is that we allow for non-linear dynamics, which makes the contraction argument used in proof different.

Theorem 3.1
Let R be an out-connected routing matrix and $\lambda : \mathbb{R}_+ \to \mathbb{R}_+^{\mathcal{E}}$ a bounded measurable function describing the time-varying exogenous inflow vector. If $\zeta : \mathcal{X} \to \mathbb{R}_+^{\mathcal{E}}$ Lipschitz continuous, then for every initial condition $x(0) \in \mathcal{X}$ there exists a unique solution to the dynamics given by (3.1) – (3.3). □

The proof of Theorem 3.1 is divided into two main parts. First we will introduce an operator and show that this operator is a contraction and hence it has a fixed point. The operator depends on the routing matrix R, and the contraction can be shown due to the fact that the routing matrix is assumed to be out-connected. In the second part, we will show how the point queue dynamics in (3.1) – (3.3) relates to this operator and that a unique solution to the dynamics exists.

Throughout the proof the of Theorem 3.1, we will make use of the fact that from Proposition 2.1 that since the routing matrix is out-connected there exists a norm $\|\cdot\|_\dagger$ on $\mathbb{R}^{\mathcal{E}}$ such that the induced matrix norm of R satisfies $\|R\|_\dagger < 1$. For $T > 0$, we shall consider the space $\mathcal{C}_T$ of continuous vector-valued functions $f : [0, T] \to \mathbb{R}^{\mathcal{E}}$ equipped with the norm

$$\|f\|_{\mathcal{C}_T} = \left\| \sup_{0 \le t \le T} |f(t)| \right\|_\dagger .$$

As a first step towards the proof of Theorem 3.1, to a given continuous vector-valued function γ in $\mathcal{C}_T$, we associate the operator $\Pi_\gamma : \mathcal{C}_T \to \mathcal{C}_T$ defined as

$$\left[\Pi_\gamma(v)\right](t) = \sup_{0 \le s \le t} \left[R^T v(s) - \gamma(s)\right]_+ , \quad 0 \le t \le T , \tag{3.4}$$

where $[a]_+$ is the vector of positive parts of entries of $a \in \mathbb{R}^{\mathcal{E}}$. In other words, $[a]_+ = \max(a, 0)$ where the max is applied entry-wise. Some fundamental properties of the operator Π_γ are summarized in the following lemma:

Lemma 3.1
For every $T > 0$ and every continuous vector-valued function γ in $\mathcal{C}_T$, the operator Π_γ is a contraction on $\mathcal{C}_T$, hence it has a unique fixed point such that

$$\Psi(\gamma) = \Pi_\gamma(\Psi(\gamma)) \in \mathcal{C}_T . \tag{3.5}$$

Moreover, the operator $\Psi : \mathcal{C}_T \to \mathcal{C}_T$ that maps a continuous vector-valued function γ into the unique fixed point of the associated operator Π_γ is Lipschitz continuous.□

Proof. We will first prove that Π_γ is a contraction on $\mathcal{C}_T$. For any v, w in $\mathcal{C}_T$, $0 \le s \le t \le T$, and i in $\mathcal{E}$, put

$$\begin{aligned} f(s) &= [R^T v(s) - \gamma(s)]_i , \\ g(s) &= [R^T w(s) - \gamma(s)]_i , \\ h(s) &= f(s) - g(s) . \end{aligned} \tag{3.6}$$

Then choose some

$$s^* \in \arg\max_{0\le s\le t} [f(s)]_+, \qquad q^* \in \arg\max_{0\le s\le t} [g(s)]_+,$$

and observe that

$$[f(s^*)]_+ = [g(s^*) + h(s^*))]_+ \le [g(s^*)]_+ + [h(s^*))]_+, \tag{3.7}$$
$$[g(q^*)]_+ = [f(q^*) - h(q^*))]_+ \le [f(q^*)]_+ + [-h(q^*))]_+ . \tag{3.8}$$

Using (3.7) and $[f(s^*)]_+ = \sup_{0\le s\le t}[f(s)]_+$, we get

$$\sup_{0\le s\le t} [h(s)]_+ \ge [h(s^*)]_+ \ge [f(s^*)]_+ - [g(s^*)]_+ \ge \sup_{0\le s\le t} [f(s)]_+ - \sup_{0\le s\le t} [g(s)]_+ .$$

Analogously, (3.8) and $[g(q^*)]_+ = \sup_{0\le s\le t}[g(s)]_+$ give

$$\begin{aligned}\sup_{0\le s\le t} [h(s)]_- &= \sup_{0\le s\le t} [-h(s)]_+ \ge [-h(q^*)]_+ \\ &\ge [g(q^*)]_+ - [f(q^*)]_+ \ge -\left(\sup_{0\le s\le t} [f(s)]_+ - \sup_{0\le s\le t} [g(s)]_+\right) .\end{aligned}$$

Therefore,

$$\begin{aligned}\sup_{0\le s\le t} |h(s)| &= \max\left\{\sup_{0\le s\le t} [h(s)]_+, \sup_{0\le s\le t} [h(s)]_-\right\} \\ &\ge \left|\sup_{0\le s\le t} [f(s)]_+ - \sup_{0\le s\le t} [g(s)]_+\right| .\end{aligned}$$

Now, define the vector $\alpha \in \mathbb{R}^{\mathcal{E}}$ with entries

$$\alpha_i = \sup_{0\le t\le T} \left|\left[\Pi_\gamma(v)\right]_i(t) - \left[\Pi_\gamma(w)\right]_i(t)\right| ,$$

for all $i \in \mathcal{E}$. Using (3.4) and (3.6), we get

$$\begin{aligned}\alpha_i &= \sup_{0\le t\le T} \left|\sup_{0\le s\le t} [f(s)]_+ - \sup_{0\le s\le t} [g(s)]_+\right| \\ &\le \sup_{0\le t\le T} \sup_{0\le s\le t} |h(s)| \\ &= \sup_{0\le t\le T} \left|[R^T(v(t) - w(t))]_i\right| \\ &\le \sum_j R_{ji} \sup_{0\le t\le T} |v_j(t) - w_j(t)| .\end{aligned}$$

Hence,

$$\left\|\Pi_\gamma v - \Pi_\gamma w\right\|_{\mathcal{C}_T} = \|\alpha\|_\dagger \le \left\|R^T\right\|_\dagger \|v - w\|_{\mathcal{C}_T} .$$

Since $\left\|R^T\right\|_\dagger < 1$, the equation above proves that Π_γ is a contraction.

In order to prove the second part of the Lemma 3.1, the fact that the operator $\Psi : \mathcal{C}_T \to \mathcal{C}_T$ is a Lipschitz continues operator, for $k \geq 0$, let Π_γ^k be the composition of Π_γ with itself k times. Then for two functions $\gamma, \eta \in \mathcal{C}_T$ and $0 \leq t \leq T$, we have that

$$\begin{aligned}
&\left|\left[\Pi_\gamma^{k+1}(v)\right](t) - \left[\Pi_\eta^{k+1}(v)\right](t)\right| \\
&\quad = \left| \sup_{0\leq s\leq t} \left[R^T[\Pi_\gamma^k v](s) - \gamma(s)\right]_+ - \sup_{0\leq s\leq t} \left[R^T[\Pi_\eta^k v](s) - \eta(s)\right]_+ \right| \\
&\quad \leq \left| \sup_{0\leq s\leq t} \left[R^T\left([\Pi_\gamma^k v](s) - [\Pi_\eta^k v](s)\right) - (\gamma(s) - \eta(s))\right]_+ \right| \\
&\quad \leq \sup_{0\leq s\leq t} \left| R^T\left(\Pi_\gamma^k v(s) - \Pi_\eta^k v(s)\right)\right| + \sup_{0\leq s\leq t} |\gamma(s) - \eta(s)|,
\end{aligned}$$

so that

$$\left\|\Pi_\gamma^{k+1}(v) - \Pi_\eta^{k+1}(v)\right\|_{\mathcal{C}_T} \leq \left\|R^T\right\|_\dagger \left\|\Pi_\gamma^k v - \Pi_\eta^k v\right\|_{\mathcal{C}_T} + \|\gamma - \eta\|_{\mathcal{C}_T} .$$

It follows that, for all v in $\mathcal{C}_T$ and $k \geq 0$,

$$\left\|\Pi_\gamma^k(v) - \Pi_\eta^k(v)\right\|_{\mathcal{C}_T} \leq \sum_{l=0}^{k} \left\|R^T\right\|_\dagger^l \|\gamma - \eta\|_{\mathcal{C}_T} .$$

Since $\left\|R^T\right\|_\dagger < 1$ and Π_γ and Π_η are both contractions with fixed points $\Psi(\gamma)$ and $\Psi(\eta)$, respectively, taking the limit as k grows large in the above gives

$$\begin{aligned}
\|\Psi(\gamma) - \Psi(\eta)\|_{\mathcal{C}_T} &= \lim_{k\to\infty} \left\|\Pi_\gamma^k(v) - \Pi_\eta^k(v)\right\|_{\mathcal{C}_T} \\
&\leq \sum_{l=0}^{+\infty} \left\|R^T\right\|_\dagger^l \|\gamma - \eta\|_{\mathcal{C}_T} = \frac{\|\gamma - \eta\|_{\mathcal{C}_T}}{1 - \left\|R^T\right\|_\dagger} .
\end{aligned}$$

which concludes the proof of the Lemma 3.1. □

Our next step towards proving Theorem 3.1 consists of finding an equivalent formulation of the network flow dynamics described by (3.1) – (3.3). Towards this goal, we introduce two operators

$$\Phi, \Gamma : \mathcal{C}_T \to \mathcal{C}_T ,$$

where

$$\Phi(y) = y + (I - R^T)\Psi(y), \tag{3.9}$$

and

$$\Gamma(x)(t) = x(0) + \int_0^t \left(\lambda(s) - (I - R^T)C\zeta(x(s))\right) \mathrm{d}s . \tag{3.10}$$

The relationship between the solution to the controlled traffic network dynamics (3.1) – (3.3) and the operators (3.9) and (3.10) is given by the following lemma.

Lemma 3.2
Let R be an out-connected routing matrix, λ a possible, time varying exogenous inflow vector, $\zeta : \mathcal{X} \to \mathcal{Z}$ a Lipschitz continuous function, and $x(0) \in \mathcal{X}$. Then given any initial condition, $(x(t), z(t))$ is a solution of the controlled network flow dynamics (3.1) – (3.3) in a time interval $[0, T]$ if and only if there exist absolutely continuous $y, w \in \mathcal{C}_T$ such that

$$x = \Phi(y), \tag{3.11}$$
$$y = \Gamma(x), \tag{3.12}$$
$$w = \Psi(y), \tag{3.13}$$
$$z = C\zeta(x) - \dot{w}, \tag{3.14}$$

for almost all $t \in [0, T]$. □

Proof. We divide the proof into two parts: in the first part we prove the if part and in the second part the only if part.

If-part: Let $(x(t), z(t))$ be a solution of the controlled network flow dynamics (3.1) – (3.3) on $[0, T]$ with initial condition $x(0)$. For $0 \le t \le T$, let

$$w(t) = \int_0^t (C\zeta(x(s)) - z(s))\mathrm{d}s, \tag{3.15}$$
$$y(t) = x(t) - (I - R^T)w(t). \tag{3.16}$$

We will show that (3.11) – (3.14) are satisfied. Indeed, taking the time derivative of both sides of (3.15) gives (3.14). On the other hand, (3.16), (3.1), and (3.15) yield

$$\begin{aligned} y(t) &= x(t) - (I - R^T)w(t) \\ &= x(0) + \int_0^t (\lambda(s) - (I - R^T)z(s))\mathrm{d}s - (I - R^T)w(t) \\ &= x(0) + \int_0^t (\lambda(s) - (I - R^T)C\zeta(x(s)))\mathrm{d}s \\ &= \Gamma(x)(t), \end{aligned}$$

so that (3.12) is satisfied as well. Moreover, (3.16) and (3.13) clearly imply (3.11). Hence, it remains to prove (3.13). For that, first observe that (3.15) and (3.3) imply that

$$x \ge 0, \qquad x^T \dot{w} = 0, \qquad 0 \le \dot{w} \le C\zeta(x). \tag{3.17}$$

In turn, the above and (3.16) imply that, for $0 \le s \le t$,

$$w(t) \ge w(s) = R^T w(s) + x(s) - y(s) \ge R^T w(s) - y(s),$$

so that

$$w(t) \ge \sup_{0 \le s \le t} \left\{ R^T w(s) - y(s) \right\}.$$

Since $w(0)$ is non-increasing and $w(0) = 0$, we have $w(t) \geq 0$, which together with the above gives

$$w(t) \geq \sup_{0 \leq s \leq t} \left[R^T w(s) - y(s) \right]_+ = \Pi_y(w)(t) \,.$$

In fact, if the above were not an identity for some $0 \leq t \leq T$, there would exist some $0 \leq t^* \leq T$ and $i \in \mathcal{E}$ such that

$$w_i(t^*) > \sup_{0 \leq s \leq t^*} \left\{ \sum\nolimits_j R_{ji} w_j(s) - y_i(s) \right\}, \qquad \dot{w}_i(t^*) > 0 \,. \tag{3.18}$$

But the second inequality above and (3.17) imply that $x_i(t^*) = 0$ so that, by (3.16), $y_i(t^*) = \sum_j R_{ji} w_j(t^*) - y_i(t^*)$ which contradicts (3.18). Hence, we necessarily have

$$w(t) = \Pi_y(w)(t), \qquad 0 \leq t \leq T \,,$$

i.e., w is the fixed point Π_y on $\mathcal{C}_T$, so that (3.13) is satisfied.

Only-if-part: Let $w, x, y, z \in \mathcal{C}_T$ be such that y and w are absolutely continuous and (3.11) – (3.14) are satisfied. Then, for $0 \leq t \leq T$, an application of (3.11), (3.9), (3.12), (3.13), (3.10), and (3.14) give

$$\begin{aligned} x(t) &= \Phi(y)(t) = y(t) + (I - R^T)\Psi(y)(t) = \Gamma(x)(t) + (I - R^T)w(t) \\ &= x(0) + \int_0^t \Big(\lambda(s) - (I - R^T) C \zeta(x(s)) \Big) \, \mathrm{d}s + (I - R^T) \int_0^t (C\zeta(x(s)) - z(s)) \, \mathrm{d}s \\ &= x(0) + \int_0^t \Big(\lambda(s) - (I - R^T) z(s) \Big) \, \mathrm{d}s \,, \end{aligned}$$

hence (3.1) is satisfied. On the other hand, (3.11), (3.9), (3.5), and (3.4) give

$$\begin{aligned} x(t) &= \Phi(y)(t) \\ &= y(t) + (I - R^T)\Psi(y)(t) \\ &= y(t) - R^T \Psi(y)(t) + \sup_{0 \leq s \leq t} \left[R^T \Psi(y)(s) - y(s) \right]_+ \\ &\geq 0 \,. \end{aligned} \tag{3.19}$$

Moreover, (3.13), (3.5), and (3.4) yield

$$w(t) = \Psi(y)(t) = \sup_{0 \leq s \leq t} \left[R^T w(s) - y(s) \right]_+ , \tag{3.20}$$

so that $w_i(t)$ is non-decreasing for all $i \in \mathcal{E}$, hence $\dot{w} \geq 0$. Furthermore, let $\mathcal{I} := \{ i \in \mathcal{E} \mid \dot{w}_i(t) > 0 \}$ be the set of cells i such that $w_i(t)$ is strictly increasing at time t. It then follows from (3.20) that

$$w_i(t) = \sum_{j \in \mathcal{E}} R_{ji} w_j(t) - y_i(t), \qquad i \in \mathcal{I} \,. \tag{3.21}$$

Equation (3.21) implies that, for $i \in \mathcal{I}$,

$$\begin{aligned}\dot{w}_i(t) &= \sum_{j\in\mathcal{E}} R_{ji}\dot{w}_j(t) - \dot{y}_i(t)\\ &= \sum_{j\in\mathcal{I}} R_{ji}\dot{w}_j(t) - \lambda_i(t) + c_i\zeta_i(x(t)) - \sum_{j\in\mathcal{E}} R_{ji}c_j\zeta_j(x(t))\\ &\le \sum_{j\in\mathcal{I}} R_{ji}\dot{w}_j(t) - \lambda_i(t) + c_i\zeta_i(x(t)) - \sum_{j\in\mathcal{I}} R_{ji}c_j\zeta_j(x(t))\,.\end{aligned}$$

The above implies that

$$(I - R_{\mathcal{I}\mathcal{I}}^T)\dot{w}_{\mathcal{I}}(t) \le (I - R_{\mathcal{I}\mathcal{I}}^T)C\zeta_{\mathcal{I}}(x(t)) - \lambda_{\mathcal{I}}(t)\,, \tag{3.22}$$

where $R_{\mathcal{I}\mathcal{I}}$ is the $\mathcal{I}\times\mathcal{I}$ block of R and $\dot{w}_{\mathcal{I}}(t)$, $\zeta_{\mathcal{I}}(x(t))$, and $\lambda_{\mathcal{I}}(t)$ are the $\mathcal{I}$ blocks of the corresponding vectors $\dot{w}(t)$, $\zeta_{\mathcal{I}}(x(t))$, and $\lambda_{\mathcal{I}}(t)$. Since R is out-connected, it follows from Proposition 2.1 that each of its diagonal blocks such as $R_{\mathcal{I}\mathcal{I}}$ has spectral radius smaller than 1. Hence $(I - R_{\mathcal{I}\mathcal{I}}^T)$ invertible with nonnegative inverse $(I - R_{\mathcal{I}\mathcal{I}}^T)^{-1}$. Hence, (3.22) implies that

$$\dot{w}_{\mathcal{I}}(t) \le C\zeta_{\mathcal{I}}(x(t)) - (I - R_{\mathcal{I}\mathcal{I}}^T)^{-1}\lambda_{\mathcal{I}}(t) \le C\zeta_{\mathcal{I}}(x(t))\,.$$

Since $\dot{w}_{\mathcal{E}\setminus\mathcal{I}}(t) = 0$ by definition and we have already noticed that $\dot{w}(t) \ge 0$, we thus have that $z = C\zeta(x) - \dot{w}$ satisfies

$$0 \le z \le C\zeta(x)\,. \tag{3.23}$$

Finally, using again (3.11), (3.9), (3.13), and (3.21), one gets that

$$x_i(t) = y_i(t) + w_i(t) - \sum_j R_{ji}w_j(t) = 0$$

for every $i \in \mathcal{I}$. Along with (3.19) and (3.23), this implies that

$$x^T(C\zeta(x) - z) = x^T\dot{w} = 0\,. \tag{3.24}$$

From (3.19), (3.23), and (3.24) it follows that (3.3) is satisfied. Therefore (x, z) is a solution of (3.1) – (3.3). □

Remark 3.1
There is an interpretation of the quantities in Lemma 3.2. The quantity y can be seen as the traffic volumes in the cells *if* the volumes were allowed to be negative, and w is how much one must add to this quantity to make sure that the traffic volume x stays non-negative. In Figure 3.5 those trajectories are illustrated for a single cell, i.e., $R = 0$. Observe that $w(t)$ is non-decreasing and only increases when $x = 0$. □

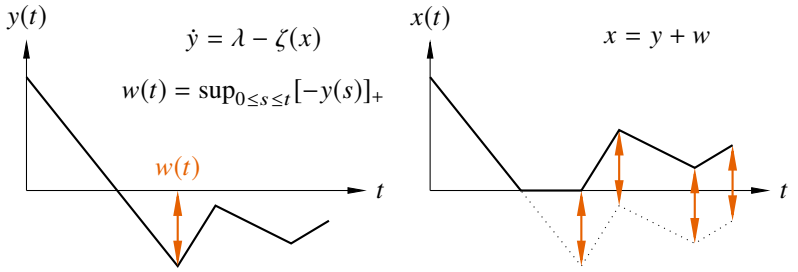


Figure 3.5 The connection between the quantities x, y and w in Lemma 3.2 for the case when the network consists of a single cell

Proof of Theorem 3.1. It follows from Lemma 3.1 that Ψ is a Lipschitz continuous operator on $\mathcal{C}_T$. Hence, the operator Φ is Lipschitz-continuous as well, and we shall denote by $\phi > 0$ its Lipschitz constant. Since $\zeta : \mathcal{X} \to \mathbb{R}^{\mathcal{E}}$ is a Lipschitz continuous function, the operator Γ is Lipschitz-continuous on $\mathcal{C}_T$ for all $T > 0$, with Lipschitz constant equal to ϖT for some constant $\varpi > 0$ that is independent of T. It then follows that, for $0 < T < (\varpi\phi)^{-1}$, the composition operator $\Phi \circ \Gamma : \mathcal{C}_T \to \mathcal{C}_T$ is Lipschitz continuous with Lipschitz constant

$$L = \varpi\phi T < 1 .$$

Therefore, $\Phi\circ\Gamma$ is a contraction on $\mathcal{C}_T$, hence it has a unique fixed point $x = \Phi(\Gamma(x))$. Let $y = \Gamma(x)$, $w = \Psi(y)$, and $z = \zeta(x) - \dot{w}$. By Lemma 3.2 we get that this (x, z) is the unique solution to (3.1) – (3.3) on $[0, T]$ with initial condition $x(0)$. Existence and uniqueness of the solution (x, z) of (3.1) – (3.3) can then be extended to the whole semi-infinite time interval $[0, +\infty)$ by standard arguments. □

3.3 Conclusions

In this chapter, we showed that the point queue dynamics introduced in Chapter 2 have a unique solution when the outflow is limited by a Lipschitz continuous feedback controller. Since it may can happen that controller allows for outflow even when a cell is empty, the actual outflow from one cell is only bounded by inequalities. Hence, classical results about well-posedness for dynamical systems can not be applied. Nevertheless, due to the fact that the routing matrix is out-connected combined with assumption of Lipschitz-continuity, existence and uniqueness of a solution to the dynamics can be ensured.

In the next chapter, we will show how a throughput optimal controller for this point queue dynamics can be designed, and utilize the results in this chapter to ensure well-posedness of the problem.

4

Decentralized Throughput Optimal Traffic Signal Control

In the chapter, we present a feedback control law for traffic signals named the Generalized Proportional Allocation (GPA). We start by introducing phase constraints to the dynamical point queue model presented in the previous chapter. Then we post a fundamental bound on how much exogenous inflow any controller for the point-queue dynamical network can handle, independent of the control strategy. In the following section, we present the GPA controller for the case when every cell belongs to only one phase, then show that the with the previously presented point-queue dynamics, the GPA controller can stabilize the network whenever the fundamental bound is satisfied. After illustrating the stability results with numerical simulations, two generalizations of the results are made. The first generalization is to allow the phases span over multiple junctions, and the second generalization is to the case when one cell may belong to several phases.

4.1 Point Queue Model with Phases

We recall from Chapter 2 that the control action in $\zeta(x)$ is limited by the set of phases assigned to each junction. To formalize the concept of phases introduced in Chapter 2, to each junction $v \in \mathcal{V}$ we assign a set of phases

$$\mathcal{P}^{(v)} \subseteq \{\mathcal{Q} \mid \mathcal{Q} \subset \mathcal{E}_v\}\,.$$

The elements in $\mathcal{P}^{(v)}$ are subsets of incoming lanes to the junction $\mathcal{E}_v$. The set $\mathcal{P}^{(v)}$ can then be described by the phase matrix, already introduced in Chapter 2, as

$$P^{(v)}_{i,j} = \begin{cases} 1 & \text{if cell } i \in \mathcal{E}_v \text{ belongs to phase } j \in \mathcal{P}^{(v)}\,, \\ 0 & \text{otherwise.} \end{cases}$$

To control which phase that should be activated in each junction in the network, we introduce for each junction $v \in \mathcal{V}$ the set of phase activations

$$\mathcal{U}^{(v)} = \left\{ u \in \mathbb{R}_+^{\mathcal{P}^{(v)}} \mid \mathbb{1}^T u \leq 1 \right\} .$$

For a given control signal $u \in \mathcal{U}^{(v)}$, the element u_i corresponds to the fraction of time that phase $i \in \mathcal{P}^{(v)}$ is activated.

While the previous sets only models the phases and phase activations for one isolated junction, we can construct corresponding sets for the whole network as well. Under the assumption that all junctions are signalized, the set of phases for the whole network can be constructed as

$$\mathcal{P} = \bigcup_{v \in \mathcal{V}} \mathcal{P}^{(v)},$$

where its corresponding phase matrix P is given by

$$P = \begin{bmatrix} P^{(1)} & & & \\ & P^{(2)} & & \\ & & \ddots & \\ & & & P^{(v)} \end{bmatrix} .$$

The set of all phase activations is the Cartesian product of the phase activations for each junction, hence

$$\mathcal{U} = \prod_{v \in \mathcal{V}} \mathcal{U}^{(v)} .$$

From a given phase activation vector u, the outflow controller ζ is given by

$$\zeta = Pu . \tag{4.1}$$

Equation (4.1) determines $\zeta(x)$ as a function of the phase activation. In effect, this imposes an additional constraint to the point-queue dynamics in (3.1) – (3.3). To summarize, the flow network dynamics for point queues constrained by phases is given by

$$\dot{x} = \lambda + (R^T - I)z, \tag{4.2}$$

$$0 \leq z \leq C\zeta, \quad x^T(z - C\zeta) = 0, \tag{4.3}$$

$$\zeta = Pu . \tag{4.4}$$

Our main goal of this chapter is to find a stabilizing decentralized outflow controller. We say that the dynamics in (4.2) – (4.4) is stable, when the traffic volume $x(t)$ remains bounded for all $t \geq 0$. Observe that, under the assumption of constant exogenous inflows, this definition of stability implies that all the traffic that exogenously flows into the network will also leave the network eventually.

4.2 Fundamental Bound on Exogenous Inflows

In this section, we present a fundamental bound on how large exogenous inflow a network can handle. Clearly, if the exogenous flows are too large, no control scheme can handle them. This bound states that any controller –independent of which information the controller has knowledge of– cannot manage a larger exogenous inflow than the one specified by this bound.

For a given exogenous inflow λ, it is obvious that if (λ, R) is inflow-connected, then routing matrix R has to be out-connected for the states to stay bounded. To describe which values the outflow vector z can take, we introduce the set of feasible outflow vectors

$$\mathcal{Z} = \left\{ z \in \mathbb{R}_+^{\mathcal{E}} \mid 0 \leq z \leq CPu \text{ for some } u \in \mathcal{U} \right\} . \tag{4.5}$$

The following proposition gives a necessary condition for any exogenous inflow vector λ to be stabilizable:

Proposition 4.1—Necessary condition for stability
Let R be an out-connected routing matrix and λ a possibly time-varying exogenous inflow vector. If the dynamics (4.2) – (4.4), with an initial state $x(0) \in \mathbb{R}_+^{\mathcal{E}}$ admit a stable solution, then the average inflow vector

$$\bar{\lambda}(t) = \frac{1}{t} \int_0^t \lambda(s) \mathrm{d}s ,$$

satisfies

$$\lim_{t \to +\infty} \operatorname{dist} \left((I - R^T)^{-1} \bar{\lambda}(t), \mathcal{Z} \right) = 0 , \tag{4.6}$$

where $\operatorname{dist}(x, \mathcal{A})$ denotes the shortest distance between an element $x \in \mathbb{R}_+^{\mathcal{E}}$ and a set $A \subset \mathbb{R}_+^{\mathcal{E}}$. In particular, if the exogenous inflow vector λ is constant, then

$$(I - R^T)^{-1} \lambda \in \mathcal{Z} . \tag{4.7}$$

□

Proof. For every $t > 0$ and initial state $x(0)$, integrating (4.2) gives

$$x(t) = x(0) + t\bar{\lambda}(t) - (I - R^T) \int_0^t z(s) \mathrm{d}s . \tag{4.8}$$

Since R is out-connected, it follows from Proposition 2.1 that $(I - R^T)$ is invertible. Multiplying both sides of (4.8) by $\frac{1}{t}(I - R^T)^{-1}$ and rearranging terms yields

$$(I - R^T)^{-1} \bar{\lambda}(t) = \bar{z}(t) + \varepsilon(t) , \tag{4.9}$$

where

$$\bar{z}(t) = \frac{1}{t} \int_0^t z(s) \mathrm{d}s , \qquad \varepsilon(t) = \frac{1}{t} (I - R^T)^{-1} \left(x(t) - x(0) \right) .$$

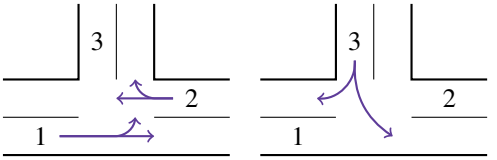


Figure 4.1 The T-junction in Example 4.1.

Since $z(s) \in \mathcal{Z}$ for $0 \leq s \leq t$ and $\mathcal{Z}$ is a convex set, it follows that $\bar{z}(t) \in \mathcal{Z}$. Hence (4.9) implies that

$$\operatorname{dist}\left((I - R^T)\bar{\lambda}(t), \mathcal{Z}\right) \leq \|\varepsilon(t)\|\,, \quad t \geq 0\,. \tag{4.10}$$

On the other hand, $x(t)$ is a stable solution of the dynamics (4.2) – (4.4), so $x(t)$ remains bounded in $t \geq 0$. This implies that $\|\varepsilon(t)\|$ converges to 0 as t grows large, so that (4.6) follows from (4.10). In the special case of constant inflow vector λ, we have $(I - R^T)^{-1}\bar{\lambda}(t) = \lambda$, so that (4.6) reduces to (4.7). □

Observe that if λ is a constant exogenous arrival vector, the quantities in (4.7) are the average inflows to each cell in the network. Henceforth, whenever they exist, we will denote these flows as

$$a = (I - R^T)^{-1}\lambda \in \mathbb{R}_+^{\mathcal{E}}\,.$$

Example 4.1
Consider the junction in Figure 4.1. Suppose that this junction has two phases, either lane 1 and lane 2 receive green light simultaneously or lane 3 receives green light. The phase matrix is then given by

$$P = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\,.$$

For this phase matrix, the set $\mathcal{Z}$ is depicted in Figure 4.2, where the axis shows the average inflow to each cell. □

Since by Proposition 4.1 the vector of average inflows, $a = (I - R^T)^{-1}\lambda$, has to belong to $\mathcal{Z}$, with full knowledge of the exogenous inflow rates and the routing matrix, it is possible to design an open-loop controller that keeps the queue lengths bounded. However, from a practical point of view both the inflow rates and the turning ratios are seldom known, and even if they were, controlling the service allocation based on those leads to poor robustness. Hence a feedback solution –that requires as little information about the network as possible– is strongly preferable. Such a controller will be presented in the next section.

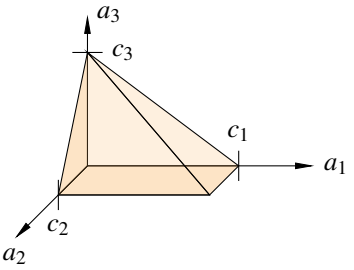


Figure 4.2 The set $\mathcal{Z}$ in Example 4.1.

4.3 Generalized Proportional Allocation (GPA) Control

In this section we will construct a decentralized feedback control policy for the point queue dynamics in (4.2) – (4.4) that is able to stabilize the network whenever it is possible for any control policy to do so. We will assume that the phases are *orthogonal*, i.e.,

$$P\mathbb{1} \leq \mathbb{1}\,.$$

This means that every cell belongs to at most one phase. On the other hand, every cell with some average inflow, i.e., cells $i \in \mathcal{E}$ with $a_i > 0$, must belong to one phase in order to be stabilizable, so under the assumption that (λ, R) is inflow-connected, the fact that the phases are orthogonal implies that every lane belongs to exactly one phase when the dynamics is stable. This means that if the network is inflow connected, for orthogonal phases it must hold that $P\mathbb{1} = \mathbb{1}$. Later in this chapter, we will show how this assumption can be relaxed.

Remark 4.1
The phases are orthogonal in both Example 2.4 and Example 4.1. □

The controller we propose relies only on local information on traffic volumes. With local information we mean that the only measurements the controller in a junction $v \in \mathcal{V}$ needs is the traffic volume in the in-coming cells $\mathcal{E}_v$. To stress this out, we let $x^{(v)}$ denote the projection of the traffic volumes $x \in \mathcal{X}$ onto $\mathcal{E}_v$.

Under the assumption that the phases are orthogonal, we introduce the *generalized proportional allocation* (GPA) that is for a junction $v \in \mathcal{V}$ given by

$$u_q^{(v)}(x^{(v)}) = \frac{(P^{(v)^T} x^{(v)})_q}{\kappa_v + \sum_{r \in \mathcal{P}^{(v)}} (P^{(v)^T} x^{(v)})_r} = \frac{\sum_i P_{iq}^{(v)} x_i^{(v)}}{\kappa_v + \sum_{j \in \mathcal{E}_v} x_j^{(v)}}, \quad q \in \mathcal{P}^{(v)}\,. \tag{4.11}$$

In the equation above, $\kappa_v > 0$ is a parameter. It is introduced to capture the fact that it is seldom in practice possible to switch between different phases without loosing some control action during the phase shift. However, the fraction of time when no

cell receives service is decreasing with the demand, something that well captures the fact that in applications such as traffic, one usually lets the traffic signal cycles be longer when the demand is higher [Roess et al., 2011].

The proposed control strategy has several benefits. First of all, it is distributed, the control action for each junction can be computed separately. Moreover, the controller does not require any information about the network topology or how the traffic propagate thought the network. These facts make the controller robust to perturbations, but it also makes it easy to deploy new controllers into the network, since one does not have to re-tune the already deployed ones.

The following example shows that any $\kappa > 0$, will make the cycle length depend on the demands λ.

Example 4.2
Consider a dynamical flow network consisting of one junction with two incoming lanes. The exogenous inflows to the lanes are $\lambda_1 > 0$, $\lambda_2 > 0$ and the capacities are $c_1 > 0$ and $c_2 > 0$. The junction is equipped with two phases, one for each lane. The dynamics is then described by

$$\begin{aligned}\dot{x}_1 &= \lambda_1 - c_1 \frac{x_1}{x_1 + x_2 + \kappa}, \\ \dot{x}_2 &= \lambda_2 - c_2 \frac{x_2}{x_1 + x_2 + \kappa}.\end{aligned}$$

The traffic volumes at equilibrium are

$$(x_1^*, x_2^*) = \left(\frac{\kappa\rho_1}{1 - \rho_1 - \rho_2}, \frac{\kappa\rho_2}{1 - \rho_1 - \rho_2} \right),$$

where $\rho_i = \lambda_i / c_i$. Observe that the necessary condition for stability is $\rho_1 + \rho_2 < 1$. The fraction of the cycle that will be allocated to phase shifts at the equilibrium is then given by

$$\frac{\kappa}{x_1^* + x_2^* + \kappa} = \frac{1}{1 + \frac{\rho_1}{1-\rho_1-\rho_2} + \frac{\rho_2}{1-\rho_1-\rho_2}} = 1 - \rho_1 - \rho_2 .$$

Since the total cycle length will be inverse proportional to the fraction allocated to phase shifts, we get that the cycle length at equilibrium $T(x^*)$ will be proportional to

$$T(x^*) \propto \frac{1}{1 - \rho_1 - \rho_2}.$$

One classical formula for computing the cycle length in a static traffic signal control setting is Webster's formula [Webster, 1958], which suggests that that the cycle length should be

$$T(x^*) = \frac{1.5L + 5}{1 - \frac{z_1^*}{c_1} - \frac{z_2^*}{c_2}}$$

where L is the total lost time, i.e., the total time where no phase is activated. Hence, for any $\kappa > 0$, the GPA will adjust the cycle length after the demand –without knowing the demand or the lanes outflow capacity– in the same way as Webster's formula suggests. □

The example above also shows that while the cycle length at equilibrium does not depend on κ, the traffic volumes at equilibrium do. To keep the traffic volumes low at equilibrium, one should choose a small κ. However, as we will show in the next chapter choosing a too small κ may decrease the performance of the controller when it is discretized.

As a corollary of Theorem 3.1, we state existence and uniqueness of the solution to the dynamics (4.2) – (4.4), with the GPA controller (4.11) for orthogonal phases when the routing matrix R is out-connected.

Corollary 4.1
Consider the network flow dynamics (4.2) – (4.4) with a given initial state $x(0) \in \mathcal{X}$. Then, for every out-connect routing matrix R the network flow dynamics (4.2) – (4.4) with GPA controller (4.11) has a unique solution. □

Proof. Since the GPA controller in (4.11) is Lipschitz continuous, the outflow controller ζ in (4.1) is Lipschitz continuous. The corollary is then a direct consequence of Theorem 3.1. □

4.4 Stability of the GPA Controller

In this section we will analyze the stability properties of the GPA control (4.11) for orthogonal phases. We will show that the network flow dynamics converge to a set of states $\mathcal{X}^*$ such that the phase activation is at least as much as average arrival rate. The remarkable feature of this result is that this happens even though the controller has no explicit information about the flows.

Theorem 4.1
Assume that the exogenous inflow vector λ and routing matrix R are such that (λ, R) is inflow-connected and R is out-connected. Introduce the set $\mathcal{X}^*$ as

$$\mathcal{X}^* = \left\{ x \in \mathcal{X} \mid \ C\zeta(x) \geq a, \ x^T(C\zeta(x) - a) \geq 0 \right\} . \tag{4.12}$$

If the vector of average arrival rates $a = (I - R^T)^{-1}$ lies in the interior of the set $\mathcal{Z}$, defined in (4.5), i.e.,

$$a = (I - R^T)^{-1}\lambda \in \text{int}(\mathcal{Z}), \tag{4.13}$$

the dynamical flow network with dynamics specified in (4.2) – (4.4) with the GPA controller (4.11) is stable and, for every initial state $x(0)$, the traffic volumes

$$x(t) \to \mathcal{X}^*,$$

as $t \to +\infty$. □

In order to prove Theorem 4.1 we shall use a LaSalle-Lyapunov argument. For every junction $v \in \mathcal{V}$, let $C^{(v)} = \mathrm{diag}((c_i)_{i \in \mathcal{E}_v})$ and

$$b_v = 1 - \min_{\substack{u \in \mathcal{U}^{(v)}: \\ C^{(v)} P^{(v)} u \geq a^{(v)}}} \mathbb{1}^T u \,. \tag{4.14}$$

Observe that the assumption $a \in \mathrm{int}(\mathcal{Z})$ implies that $b_v > 0$. Then, define the scalar fields

$$H : \mathcal{X} \times \mathcal{U} \to \mathbb{R}, \qquad V : \mathcal{X} \to \mathbb{R},$$

by

$$H(x, \upsilon) = \sum_{i \in \mathcal{E}} x_i \log \frac{(CP\upsilon)_i}{a_i} + \sum_{v \in \mathcal{V}} \kappa_v \log \frac{1 - \mathbb{1}^T \upsilon^{(v)}}{b_v}, \tag{4.15}$$

and

$$V(x) = \max_{\upsilon \in \mathcal{U}} H(x, \upsilon), \tag{4.16}$$

respectively.

As we shall see, the proof of Theorem 4.1 relies on showing that, when the generalized proportional allocation feedback controller (4.11) is employed, the quantity $V(x(t))$ is non-increasing in t along solutions of the network flow dynamics (4.2) – (4.4) and strictly decreasing outside the set $\mathcal{X}^*$ defined in (4.12). Let also $\omega : \mathcal{X} \to \mathbb{R}^{\mathcal{E}}$ be the vector field defined by

$$\omega_i(x) \coloneqq \log \frac{C_i \zeta_i(x)}{a_i}, \qquad i \in \mathcal{E} \,. \tag{4.17}$$

The following lemma gathers a few properties of the functions above.

Lemma 4.1
Let $u(x)$ be the generalized proportional allocation controller defined in (4.11), and let $H(x, \upsilon)$, $V(x)$, and $\omega(x)$ be defined as in (4.15), (4.16), and (4.17), respectively. Then, for every state vector $x \in \mathcal{X}$,

$$V(x) = H(x, u(x)) \geq 0 \,. \tag{4.18}$$

Moreover, $V(x)$ is absolutely continuous on $\mathcal{X}$ and

$$\frac{\partial V(x)}{\partial x_i} = \omega_i(x), \tag{4.19}$$

for all $i \in \mathcal{E}$ such that $x_i > 0$. □

Proof. To show the first equality in (4.18), we have to show that

$$u(x) = \operatorname*{argmax}_{\upsilon \in \mathcal{U}} H(x, \upsilon) \,. \tag{4.20}$$

Let $L : \mathcal{X} \times \mathcal{U} \times \mathbb{R}_+^{\mathcal{V}} \to \mathbb{R}$ defined as

$$\begin{aligned} L(x, \upsilon, \gamma) &= H(x, \upsilon) + \sum_{v \in \mathcal{V}} \gamma_v (1 - \mathbb{1}^T \upsilon^{(v)}) \\ &= \sum_{i \in \mathcal{E}} x_i \log \frac{(CP\upsilon)_i}{a_i} + \sum_{v \in \mathcal{V}} \kappa_v \log \frac{1 - \mathbb{1}^T \upsilon^{(v)}}{b_v} + \sum_{v \in \mathcal{V}} \gamma_v (1 - \mathbb{1}^T \upsilon^{(v)}) \\ &= \sum_{v \in \mathcal{V}} \left(\sum_{i \in \mathcal{E}_v} x_i \log \frac{(CP\upsilon)_i}{a_i} + \kappa_v \log \frac{1 - \mathbb{1}^T \upsilon^{(v)}}{b_v} + \gamma_v (1 - \mathbb{1}^T \upsilon^{(v)}) \right) \end{aligned}$$

be the Lagrangian associated with the optimization problem in (4.20). Then necessary conditions for optimum is that

$$\frac{\partial L}{\partial \upsilon_q^{(v)}} = \frac{1}{\upsilon_q^{(v)}} \sum_{i \in \mathcal{E}_v} P_{iq}^{(v)} x_i - \frac{1}{1 - \mathbb{1}^T \upsilon^{(v)}} \kappa_v - \gamma_v = 0 \,.$$

Moreover, since the problem in (4.20) is convex, using the complementary slackness principle [Boyd and Vandenberghe, 2004], we get that either $1 - \mathbb{1}^T \upsilon^{(v)}$ is zero, which clearly cannot be a maximum, or $\gamma_v = 0$. For the latter case, it holds that

$$\frac{1}{\kappa_v} \sum_{i \in \mathcal{E}_v} P_{iq}^{(v)} x_i = \frac{\upsilon_q^{(v)}}{1 - \mathbb{1}^T \upsilon^{(v)}} \,. \tag{4.21}$$

Summing up the expression above over all phases $q \in \mathcal{P}^{(v)}$ and using the fact that the phases are orthogonal yields

$$\frac{1}{\kappa_v} \sum_{i \in \mathcal{E}_v} x_i = \frac{\mathbb{1}^T \upsilon^{(v)}}{1 - \mathbb{1}^T \upsilon^{(v)}} \,,$$

and hence

$$\mathbb{1}^T \upsilon^{(v)} = \frac{\sum_{i \in \mathcal{E}_v} x_i}{\kappa_v + \sum_{i \in \mathcal{E}_v} x_i} \,. \tag{4.22}$$

By combining (4.21) and (4.22) we get

$$u_q^{(v)} = \frac{\sum_{i \in \mathcal{E}_v} P_{iq} x_i}{\kappa_v + \sum_{i \in \mathcal{E}_v} x_i} \,,$$

which, together with the concavity of (4.15), proves that (4.11) is a solution to (4.16).

The inequality in (4.18), that $V(x) \geq 0$ follows from the fact that

$$V(x) = \max_{\upsilon} H(x, \upsilon) \geq H(x, \tilde{\upsilon}) \geq 0 \,,$$

where $\tilde{\upsilon} \in \mathcal{U}$ is chosen such that $(CP\tilde{\upsilon})_i \geq a_i$ for all $i \in \mathcal{E}$ and $1 - \mathbb{1}^T \tilde{\upsilon}^{(v)} = b_v$ for all $v \in \mathcal{V}$. That such a choice exists follows from the definition of b_v in (4.14).

To show (4.19), we follow the idea presented in [Walton, 2014]. For a $x \in \mathcal{X}$, let $x^{(\epsilon)} \in \mathcal{X}$ be a vector such that $x_i^{(\epsilon)} = x_i + \epsilon$ for some $\epsilon > 0$ and $x_j^{(\epsilon)} = x_j$ for all $j \neq i \in \mathcal{E}$. Then

$$\begin{aligned}
V(x^{\epsilon}) - V(x) &= \sum_{j\in\mathcal{E}} x_j^{(\epsilon)} \log \frac{C_j\zeta_j(x^{(\epsilon)})}{a_j} + \sum_{v\in\mathcal{V}} \kappa_v \log \frac{1 - \mathbb{1}^T u^{(v)}(x^{(\epsilon)})}{b_v} \\
&\quad - \sum_{j\in\mathcal{E}} x_j \log \frac{C_j\zeta_j(x)}{a_j} + \sum_{v\in\mathcal{V}} \kappa_v \log \frac{1 - \mathbb{1}^T u^{(v)}(x)}{b_v} \\
&\geq \sum_{j\in\mathcal{E}} x_j^{(\epsilon)} \log \frac{C_j\zeta_j(x)}{a_j} + \sum_{v\in\mathcal{V}} \kappa_v \log \frac{1 - \mathbb{1}^T u^{(v)}(x)}{b_v} \\
&\quad - \sum_{j\in\mathcal{E}} x_j \log \frac{C_j\zeta_j(x)}{a_j} + \sum_{v\in\mathcal{V}} \kappa_v \log \frac{1 - \mathbb{1}^T u^{(v)}(x)}{b_v} \\
&= \epsilon \log \frac{C_i\zeta_i(x)}{a_i},
\end{aligned}$$

where the inequality follows from the fact that

$$H(x^{(\epsilon)}, u(x^{(\epsilon)})) = \max_{v\in\mathcal{U}} H(x^{(\epsilon)}, v) \geq H(x^{(\epsilon)}, u(x)).$$

In the same manner, we have that

$$\begin{aligned}
V(x^{(\epsilon)}) - V(x) &= \sum_{j\in\mathcal{E}} x_j^{(\epsilon)} \log \frac{C_j\zeta_j(x^{(\epsilon)})}{a_j} + \sum_{v\in\mathcal{V}} \kappa_v \log \frac{1 - \mathbb{1}^T u^{(v)}(x^{(\epsilon)})}{b_v} \\
&\quad - \sum_{j\in\mathcal{E}} x_j \log \frac{C_j\zeta_j(x)}{a_j} + \sum_{v\in\mathcal{V}} \kappa_v \log \frac{1 - \mathbb{1}^T u^{(v)}(x)}{b_v} \\
&\leq \sum_{j\in\mathcal{E}} x_j^{(\epsilon)} \log \frac{C_j\zeta_j(x^{(\epsilon)})}{a_j} + \sum_{v\in\mathcal{V}} \kappa_v \log \frac{1 - \mathbb{1}^T u^{(v)}(x^{(\epsilon)})}{b_v} \\
&\quad - \sum_{j\in\mathcal{E}} x_j \log \frac{C_j\zeta_j(x^{(\epsilon)})}{a_j} + \sum_{v\in\mathcal{V}} \kappa_v \log \frac{1 - \mathbb{1}^T v^{(v)}(x^{(\epsilon)})}{b_v} \\
&= \epsilon \log \frac{C_i\zeta_i(x^{(\epsilon)})}{a_i}.
\end{aligned}$$

The bounds combined together yields

$$\log \frac{C_i\zeta_i(x)}{a_i} \leq \frac{1}{\epsilon}(V(x^{(\epsilon)}) - V(x)) \leq \log \frac{C_i\zeta_i(x^{(\epsilon)})}{a_i}.$$

Since $\zeta(x)$ depends continuously on x, letting $\epsilon \to 0$ proves the statement. □

A key difficulty in proving that $V(x(t))$ is nondecreasing along solutions $x(t)$ of the network flow dynamics (4.2) – (4.4) consists in dealing with the time instants when some of the entries $x_i(t)$ are equal to 0. Towards this goal, it is convenient to introduce the following additional notation. Recall that $A_{\mathcal{I}\mathcal{J}}$ denotes the restriction of the matrix A, only containing the rows indexed by $\mathcal{I}$ and columns indexed $\mathcal{J}$. For a state vector $x \in \mathcal{X}$, define $\mathcal{I}(x) = \mathcal{I}$ and $\mathcal{J}(x) = \mathcal{J}$ as

$$\mathcal{I} = \{i \in \mathcal{E} \mid x_i = 0\}, \qquad \mathcal{J} = \{j \in \mathcal{E} \mid x_j > 0\}, \tag{4.23}$$

and the vector $\tilde{\lambda}(x) \in \mathbb{R}_+^{\mathcal{J}}$, the matrix $\tilde{R}(x) \in \mathbb{R}_+^{\mathcal{J}\times\mathcal{J}}$, and the scalar $W(x) \in \mathbb{R}$ as

$$\tilde{\lambda}(x) \coloneqq \lambda_{\mathcal{J}} + (R^T)_{\mathcal{J}\mathcal{I}}(I - R_{\mathcal{I}\mathcal{I}}^T)^{-1}\lambda_{\mathcal{I}}, \tag{4.24}$$

$$\tilde{R}^T(x) \coloneqq R_{\mathcal{J}\mathcal{J}}^T + (R^T)_{\mathcal{J}\mathcal{I}}(I - R_{\mathcal{I}\mathcal{I}}^T)^{-1}(R^T)_{\mathcal{I}\mathcal{J}}, \tag{4.25}$$

$$W(x) \coloneqq -\omega_{\mathcal{J}}^T(x)\left(\tilde{\lambda} - (I - \tilde{R}^T(x))C_{\mathcal{J}}\zeta_{\mathcal{J}}(x)\right), \tag{4.26}$$

where $C_{\mathcal{J}}$ is a matrix with the outflow capacities of the cells in $\mathcal{J}$ on the diagonal.

The following lemma states a fundamental property of $W(x)$.

Lemma 4.2
For every state vector $x \in \mathcal{X}$, it holds true that

$$W(x) \geq 0,$$

with equality if and only if

$$\zeta_j(x) = \frac{a_j}{C_j},$$

for all $j \in \mathcal{J}$. □

We prove Lemma 4.2 by combining two intermediate results. The first one is a lower bound on $W(x)$ as stated in the following.

Lemma 4.3
For every state $x \in \mathcal{X}$ we have

$$W(x) = \sum_{j\in\mathcal{J}} \tilde{\lambda}_j F_j(\omega_{\mathcal{J}}), \tag{4.27}$$

where

$$F(\omega_{\mathcal{J}}) = (I - \tilde{R})^{-1}\mathrm{diag}((I - \tilde{R})w_{\mathcal{J}})(e^{\omega_{\mathcal{J}}} - \mathbb{1}),$$

and $e^{\omega_{\mathcal{J}}}$ is the vector with entries $(e^{\omega_{\mathcal{J}}})_j = e^{\omega_j}$ for $j \in \mathcal{J}$. Moreover,

$$F_j(\omega_{\mathcal{J}}) \geq \chi_j, \qquad \forall j \in \mathcal{J}, \tag{4.28}$$

where

$$\chi_j = \sum_{i,k\in\mathcal{J}} N_{ik}^{(j)}\omega_i(e^{\omega_i} - 1) - \sum_{i,k\in\mathcal{J}} N_{ik}^{(j)}\omega_i(e^{\omega_k} - 1)$$

and, for every $i, j, k \in \mathcal{J}$,

$$N_{ik}^{(j)} = \sum_{h \geq 0} \tilde{R}_{ji}^{h} \left(\tilde{R}_{ik} + \delta_k^{(j)} \left(1 - \sum_{l \in \mathcal{J}} \tilde{R}_{il} \right) \right) . \tag{4.29}$$

□

Proof. It follows from $\lambda = (I - R^T)a$ that

$$\begin{aligned} \lambda_{\mathcal{I}} &= (I - R_{\mathcal{I}\mathcal{I}}^T)a_{\mathcal{I}} - (R^T)_{\mathcal{I}\mathcal{J}} a_{\mathcal{J}} , \\ \lambda_{\mathcal{J}} &= (I - R_{\mathcal{J}\mathcal{J}}^T)a_{\mathcal{J}} - (R^T)_{\mathcal{J}\mathcal{I}} a_{\mathcal{I}} . \end{aligned}$$

Using the above, as well as (4.24), we obtain that

$$\begin{aligned} (I - R_{\mathcal{J}\mathcal{J}}^T)a_{\mathcal{J}} &= \lambda_{\mathcal{J}} + (R^T)_{\mathcal{J}\mathcal{I}} a_{\mathcal{I}} \\ &= \tilde{\lambda} + (R^T)_{\mathcal{J}\mathcal{I}} (I - R_{\mathcal{I}\mathcal{I}}^T)^{-1} (R^T)_{\mathcal{I}\mathcal{J}} a_{\mathcal{J}} \end{aligned}$$

so that, by substituting (4.25), we get that

$$(I - \tilde{R}^T)a_{\mathcal{J}} = \tilde{\lambda} .$$

Let $A = \mathrm{diag}((a_j)_{j \in \mathcal{J}})$. Then, $C_{\mathcal{J}} \zeta_{\mathcal{J}}(x) = A e^{\omega_{\mathcal{J}}}$, so that

$$\begin{aligned} W(x) &= -\omega_{\mathcal{J}}^T \left(\tilde{\lambda} - (I - \tilde{R}^T) A e^{\omega_{\mathcal{J}}} \right) \\ &= -\omega_{\mathcal{J}}^T \left((I - \tilde{R}^T) A \mathbb{1} - (I - \tilde{R}^T) A e^{\omega_{\mathcal{J}}} \right) \\ &= -\omega_{\mathcal{J}}^T (I - \tilde{R}^T) A (\mathbb{1} - e^{\omega_{\mathcal{J}}}) \\ &= \tilde{\lambda}^T F(\omega_{\mathcal{J}}) , \end{aligned}$$

which proves the first part of the claim.

In order to prove the second part, let

$$B(\omega_{\mathcal{J}}) = \mathrm{diag}((I - \tilde{R})\omega_{\mathcal{J}})(e^{\omega_{\mathcal{J}}} - \mathbb{1}) .$$

For $i \in \mathcal{J}$, rewrite $\omega_i = [\omega_i]_+ - [\omega_i]_-$ and observe that $e^{[\omega_i]_\pm} - 1 = [q_i]_\pm$, where $q_i = e^{\omega_i} - 1$. Then,

$$\begin{aligned} B_i(\omega_{\mathcal{J}}) &= q_i \left(\omega_i - \sum_{k \in \mathcal{J}} \tilde{R}_{ik} \omega_k \right) \\ &= [q_i]_+ \left([\omega_i]_+ - \sum_{k \in \mathcal{J}} \tilde{R}_{ik} [\omega_k]_+ \right) + [q_i]_- \left([\omega_i]_- - \sum_{k \in \mathcal{J}} \tilde{R}_{ik} [\omega_k]_+ \right) \\ &\quad + [q_i]_+ \sum_{k \in \mathcal{J}} R_{ik} [\omega_k]_- + [q_i]_- \sum_{k \in \mathcal{J}} R_{ik} [\omega_k]_- \\ &\geq B_i(\omega_{\mathcal{J}}^+) + B_i(\omega_{\mathcal{J}}^-) , \end{aligned}$$

where the fact that $[q_i]_\pm[\omega_i]_\mp = 0$ is used in the second equality. Since $(I - \tilde{R})^{-1}$ is a nonnegative matrix, the above implies that

$$\begin{aligned} F(\omega_{\mathcal{J}}) &= (I - \tilde{R})^{-1} B(\omega_{\mathcal{J}}) \\ &\geq (I - \tilde{R})^{-1} B([\omega_{\mathcal{J}}]_+) + (I - \tilde{R})^{-1} B([\omega_{\mathcal{J}}]_-) \\ &= F([\omega_{\mathcal{J}}]_+) + F([\omega_{\mathcal{J}}]_-) \,. \end{aligned}$$

Now, rewrite $F_j(w_{\mathcal{J}})$ as

$$\begin{aligned} F_j(\omega_{\mathcal{J}}) &= \sum_{i \in \mathcal{J}} \sum_{n \geq 0} \tilde{R}_{ji}^n (\omega_i - \sum_{k \in \mathcal{J}} \tilde{R}_{ik} \omega_k)(e^{\omega_i} - 1) \\ &= \sum_{i,k \in \mathcal{J}} N_{ik}^{(j)} (e^{\omega_i} - 1)\omega_i - \sum_{i,k \in \mathcal{J}} N_{ik}^{(j)} (e^{\omega_i} - 1)\omega_k \\ &\quad - \sum_{i \in \mathcal{J}} \sum_{n \geq 0} \tilde{R}_{ji}^n \left(1 - \sum_{l \in \mathcal{J}} \tilde{R}_{il}\right) (e^{\omega_i} - 1)\omega_j \,. \end{aligned}$$

It then follows that

$$\begin{aligned} F_j(\omega_{\mathcal{J}}) &\geq F_j([\omega_{\mathcal{J}}]_+) + F_j([\omega_{\mathcal{J}}]_-) \\ &\geq \sum_{i,k \in \mathcal{J}} N_{ik}^{(j)} (e^{[\omega_i]_+} - 1)[\omega_i]_+ \\ &\quad - \sum_{i,k \in \mathcal{J}} N_{ik}^{(j)} (e^{[\omega_i]_+} - 1)[\omega_k]_+ \\ &\quad + \sum_{i,k \in \mathcal{J}} N_{ik}^{(j)} (e^{[\omega_i]_-} - 1)[\omega_i]_- \\ &\quad - \sum_{i,k \in \mathcal{J}} N_{ik}^{(j)} (e^{[\omega_i]_-} - 1)[\omega_k]_- \\ &\geq \sum_{i,k \in \mathcal{J}} N_{ik}^{(j)} (e^{\omega_i} - 1)\omega_i - \sum_{i,k \in \mathcal{J}} N_{ik}^{(j)} (e^{\omega_i} - 1)\omega_k \\ &= \chi_j \,, \end{aligned}$$

thus completing the proof. □

The following lemma is generalization of the rearrangement inequality. It provides alternative deterministic proof without necessity to use the probabilistic interpretation used in [Massoulié, 2007] to show the same result.

Lemma 4.4
Let $\mu \in \mathbb{R}_{++}^n$ be a strictly positive vector and let

$$\mathcal{M} := \left\{ M \in \mathbb{R}_+^{n \times n} \mid M\mathbb{1} = M^T\mathbb{1} = \mu \right\}$$

be the set of nonnegative square matrices with both row and column sum vectors equal to μ. Let $f, g : \mathbb{R} \to \mathbb{R}$ be strictly increasing functions. Then, for every vector $v \in \mathbb{R}^n$, it holds true that

$$\sum_{i=1}^{n} \mu_i f(v_i) g(v_i) \geq \sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij} f(v_i) g(v_j),$$

for every $M \in \mathcal{M}$, with equality if and only if

$$M_{ij} = 0, \qquad \forall\, i, j : \; v_i \neq v_j \,. \tag{4.30}$$

□

Proof. Let us define the function $h : \mathcal{M} \to \mathbb{R}$ by

$$h(M) = \sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij} f(v_i) g(v_j) \,.$$

Observe that $h(M)$ is a continuous function and $\mathcal{M}$ is a compact set. Hence, $h(M)$ admits a maximum over $\mathcal{M}$. We shall prove the claim by showing that such a maximum value is

$$\max\{h(M) \mid M \in \mathcal{M}\} = \sum_{i=1}^{n} \mu_i f(v_i) g(v_i),$$

and that the set of maximum points

$$\operatorname{argmax}\{h(M) \mid M \in \mathcal{M}\} = \{M \in \mathcal{M} \mid (4.30)\},$$

coincides with the subset of matrices satisfying (4.30).

Without any loss of generality, we shall assume that

$$v_1 \leq v_2 \leq \cdots \leq v_{n-1} \leq v_n \,.$$

Now, let $m \leq n$ be the number of distinct entries of v and let $\mathcal{H}_1, \ldots, \mathcal{H}_m \subseteq \{1, \ldots, n\}$ be the subsets of indices such that $v_i = v_j$ if and only if $i, j \in \mathcal{H}_l$ for the same $1 \leq l \leq m$. Then, a matrix $M \in \mathcal{M}$ satisfies (4.30) if and only if is in the following block diagonal form

$$M = \begin{bmatrix} M^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & M^{(m)} \end{bmatrix},$$

with each block $M^{(l)} \in \mathbb{R}_+^{|\mathcal{H}_l|\times|\mathcal{H}_l|}$ for $1 \le l \le m$. Using the block diagonal form above, for an arbitrary selection of $k_l \in \mathcal{H}_l$, $1 \le l \le m$, one gets that

$$\begin{aligned} h(M) &= \sum_{l=1}^{m} \sum_{i,j\in\mathcal{H}_l} M_{ij} f(v_i) g(v_j) \\ &= \sum_{l=1}^{m} |\mathcal{H}_l| \mu_{k_l} f(v_{k_l}) g(v_{k_l}) \\ &= \sum_{i=1}^{n} \mu_i f(v_i) g(v_i), \end{aligned}$$

for every matrix $M \in \mathcal{M}$ satisfying (4.30).

We are then left with proving that no matrix $M \in \mathcal{M}$ not satisfying (4.30) can be a maximizer of $h(M)$ over $\mathcal{M}$. For any such M, let j be the unique value in $\{1, 2, \ldots, n-1\}$ such that $M_{ii} = \mu_i$ for all $1 \le i < j$ and $M_{jj} < \mu_j$ and let $1 \le q \le m$ be such that $j \in \mathcal{H}_q$. Then, since $M \in \mathcal{M}$ and it does not satisfy (4.30), there must exist indices $k \in \mathcal{H}_r$ and $l \in \mathcal{H}_s$, with $r, s \in \{q+1, \ldots, m\}$, such that

$$\epsilon = \min\{M_{jl}, M_{kj}\} > 0\,.$$

Define the matrix $\tilde{M} \in \mathbb{R}^{n\times n}$ with entries

$$\tilde{M}_{hi} = \begin{cases} M_{hi} + \epsilon & \text{if } i = j \text{ and } h = j\,, \\ M_{hi} + \epsilon & \text{if } i = l \text{ and } h = k\,, \\ M_{hi} - \epsilon & \text{if } i = l \text{ and } h = j\,, \\ M_{hi} - \epsilon & \text{if } i = j \text{ and } h = k\,, \\ M_{hi} & \text{otherwise}\,. \end{cases}$$

It is easily verified that $\tilde{M} \in \mathcal{M}$. Moreover, since $j \in \mathcal{H}_q$, $k \in \mathcal{H}_r$, and $l \in \mathcal{H}_s$, with $r, s \in \{q+1, \ldots, m\}$, we have that $v_k > v_j$ and $v_l > v_j$. Since the functions f and g are strictly increasing, this implies that

$$f(v_l) > f(v_j), \qquad g(v_k) > g(v_j)\,.$$

It follows that

$$\begin{aligned} 0 &< \epsilon(f(v_l) - f(v_j))(g(v_k) - g(v_j)) \\ &= \epsilon(f(v_j)g(v_j) + f(v_l)g(v_k) - f(v_l)g(v_j) - f(v_j)g(v_k)) \\ &= h(\tilde{M}) - h(M)\,. \end{aligned}$$

The above shows that no matrix $M \in \mathcal{M}$ that does not satisfy (4.30) can be a maximizer of $h(M)$ over $\mathcal{M}$, thus completing the proof. □

We are now ready to prove Lemma 4.2.

Proof of Lemma 4.2. For $i, j, k \in \mathcal{J}$, let $N_{ik}^{(j)}$ be defined as in (4.29) and let

$$\mu_i^{(j)} = \sum_{h \geq 0} \tilde{R}_{ji}^h \,.$$

Clearly, $\mu_j^{(j)} \geq 1 > 0$ and, more in general, $\mu_k^{(j)} > 0$ if and only if k is reachable from j through $\tilde{R}$. Let $\mathcal{K}_j$ be the the set reachable from j through $\tilde{R}$. Now observe that, for $i \in \mathcal{K}_j$,

$$\sum_{k \in \mathcal{K}_j} N_{ik}^{(j)} = \sum_{h \geq 0} \tilde{R}_{ji}^h = \mu_i^{(j)} ,$$

while, for $k \in \mathcal{K}_j$,

$$\sum_{i \in \mathcal{K}_j} N_{ik}^{(j)} = \sum_{h \geq 0} \tilde{R}_{jk}^{h+1} + \sum_{h \geq 0} (\tilde{R}_{jk}^h - \tilde{R}_{jk}^{h+1}) = \mu_k^{(j)} \,.$$

On the other hand, observe that, since $\mathcal{K}_j$ is the set reachable from j, the restriction of the matrix $N^{(j)}$ to $\mathcal{K}_j \times \mathcal{K}_j$ consists of a single diagonal block. Then, (4.28) and Lemma 4.4 imply that, for every $j \in \mathcal{J}$,

$$\begin{aligned} F_j(\omega_{\mathcal{J}}) &\geq \chi_j \\ &= \sum_{i,k \in \mathcal{K}_j} N_{ik}^{(j)} \omega_i (e^{\omega_i} - 1) - \sum_{i,k \in \mathcal{K}_j} N_{ik}^{(j)} \omega_i (e^{\omega_k} - 1) \\ &\geq 0 \,, \end{aligned} \tag{4.31}$$

where the last inequality holds true as an equality if and only if w is constant over $\mathcal{K}_j$. Observe that, in this case, there exists some $c \in \mathbb{R}$ such that

$$F_j(\omega_{\mathcal{J}}) = \sum_{i \in \mathcal{K}_j} \sum_{h \geq 0} \tilde{R}_{ji}^h \Big(1 - \sum_{l \in \mathcal{K}_j} \tilde{R}_{il}\Big)(e^c - 1)c \geq 0 \,. \tag{4.32}$$

However, since $\mathcal{K}_j$ is out-connected, then necessarily there must exist at least one $i \in \mathcal{K}_j$ such that $\sum_{l \in \mathcal{K}_j} \tilde{R}_{il} < 1$ and an $h \geq 0$ such that $\tilde{R}_{ji}^h > 0$. It then follows from (4.31) and (4.32) that

$$F_j(\omega_{\mathcal{J}}) \geq 0 \,, \qquad j \in \mathcal{J} \,,$$

with equality if and only if $\omega_i = 0$ for every $i \in \mathcal{K}_j$.

Finally, observe that

$$\bigcup_{j \in \mathcal{J} : \tilde{\lambda}_j > 0} \mathcal{K}_j = \mathcal{J} \,.$$

The above, (4.27), and (4.32) imply that

$$W(x) = \sum_{j \in \mathcal{J}} \tilde{\lambda}_j F_j(\omega_{\mathcal{J}}) \geq 0 \,,$$

with equality if and only if $\omega_i = 0$ for all $i \in \mathcal{J}$, i.e., if and only if

$$\zeta_i(x) = \frac{a_i}{C_i}, \qquad \forall i \in \mathcal{J}\,.$$

The proof of Lemma 4.2 is then complete. □

Proof of Theorem 4.1. For a state vector $x \in \mathcal{X}$, let the subsets of cells $\mathcal{I}(x) = \mathcal{I}$ and $\mathcal{J}(x) = \mathcal{J}$ be defined as in (4.23). Let $(x(t), z(t))$ be a solution of the dynamics (4.2) – (4.4) with the GPA controller (4.11). Observe that, within any open time interval (t_-, t_+) where no entry of $x(t)$ changes sign, so that the sets $\mathcal{I} = \mathcal{I}(x(t))$ and $\mathcal{J} = \mathcal{J}(x(t))$ remain constant, one has that $z_{\mathcal{J}} = \zeta_{\mathcal{J}}(x)$ and

$$0 = \dot{x}_{\mathcal{I}} = \lambda_{\mathcal{I}} + (R^T)_{\mathcal{I}\mathcal{J}} z_{\mathcal{J}} + R^T_{\mathcal{I}\mathcal{I}} z_{\mathcal{I}} - z_{\mathcal{I}}$$

so that the vector $z_{\mathcal{I}}$ of outflows from the cells in $\mathcal{I}$ satisfies

$$z_{\mathcal{I}} = (I - R^T_{\mathcal{I}\mathcal{I}})^{-1}(\lambda_{\mathcal{I}} + (R^T)_{\mathcal{I}\mathcal{J}}\zeta_{\mathcal{J}}(x))\,. \tag{4.33}$$

Moreover, the vector $x_{\mathcal{J}}$ of the states of the cells in $\mathcal{J}$ has time-derivative

$$\begin{aligned}\dot{x}_{\mathcal{J}} &= \lambda_{\mathcal{J}} + R^T_{\mathcal{J}\mathcal{J}}\zeta_{\mathcal{J}}(x) + (R^T)_{\mathcal{J}\mathcal{I}} z_{\mathcal{I}} \\ &= \tilde{\lambda}(x) - (I - \tilde{R}^T(x))\zeta_{\mathcal{J}}(x)\,.\end{aligned} \tag{4.34}$$

Now, let $\omega : \mathcal{X} \to \mathbb{R}^{\mathcal{E}}$ be the vector field defined by (4.17) and $V, W : \mathcal{X} \to \mathbb{R}$ be the scalar fields defined by (4.16) and (4.26), respectively. Then, for every solution $(x(t), z(t))$ of the dynamics (4.2) – (4.4) with the GPA controller (4.11) and for every time instant t belonging to an open interval where the sign of all entries of $x(t)$ are constant, Lemma 4.1 and (4.34) imply that

$$\begin{aligned}\dot{V}(x(t)) &= \sum_{j\in\mathcal{J}} \frac{\partial V}{\partial x_j}(x(t))\dot{x}_j(t) \\ &= \omega^T_{\mathcal{J}}(x(t))\big(\tilde{\lambda}(x(t)) - (I - \tilde{R}^T(x(t)))\zeta_{\mathcal{J}}(x(t))\big) \\ &= -W(x(t))\,.\end{aligned}$$

Since $V(x(t))$ is absolutely continuous as a function of t, it follows that

$$V(x(t)) = V(x(0)) - \int_0^t W(x(s))\mathrm{d}s\,.$$

By rearranging terms in the identity above and using Lemma 4.1 one gets that

$$\int_0^t W(x(s))\mathrm{d}s = V(x(0)) - V(x(t)) \le V(x(0))\,, \tag{4.35}$$

for all $t \geq 0$. For all $\mathcal{J} \subseteq \mathcal{E}$, let

$$\Omega_{\mathcal{J}} = \text{int}\{t \geq 0 \mid \mathcal{J}(x(t)) = \mathcal{J}\}\,.$$

Now, it follows from Lemma 4.2 that

$$W(x(t)) \geq 0\,, \qquad t \geq 0\,. \tag{4.36}$$

Hence $V(x(t)) \leq V(x(0))$ for all $t \geq 0$. We will now show that $x(t)$ will be bounded for all $t \geq 0$.

Due to the assumption in (4.13), there exists a $\tilde{u} \in \mathcal{U}$ such that $(CP\tilde{u})_i = a_i(1+\epsilon_i)$ for some $\epsilon_i > 0$ for all $i \in \mathcal{E}$.

$$\begin{aligned} V(x(0)) \geq V(x(t)) &= \max_{\upsilon \in \mathcal{U}} H(x, \upsilon) \geq H(x, \tilde{u}) \\ &= \sum_{i \in \mathcal{E}} x_i \log(1+\epsilon_i) + D = \sum_{i \in \mathcal{E}} |x_i| \log(1+\epsilon_i) + D\,, \end{aligned}$$

where

$$D = \sum_{v \in \mathcal{V}} \kappa_v \log \frac{1 - \mathbb{1}^T \tilde{u}^{(v)}}{b_v}\,.$$

Hence $x(t)$ will be bounded for all $t \geq 0$.

Now, inequality (4.36), combined with (4.35), implies that the integral

$$\int_{\Omega_{\mathcal{J}}} W(x(t))\mathrm{d}t \leq \lim_{t \to +\infty} \int_0^t W(x(t))\mathrm{d}t \leq V(x(0))$$

is finite for all $\mathcal{J} \subseteq \mathcal{E}$.

Since $x(t)$ is bounded and $W(x)$ is continuous, $W(x(t))$ is uniformly continuous on $\Omega_{\mathcal{J}}$. This implies that

$$\lim_{\substack{t \in \Omega_{\mathcal{J}} \\ t \to +\infty}} W(x(t)) = 0\,, \tag{4.37}$$

for all $\mathcal{J} \subseteq \mathcal{E}$ such that $\Omega_{\mathcal{J}}$ has infinite measure. Then, it follows from (4.37) and Lemma 4.2 that

$$\lim_{\substack{t \in \Omega_{\mathcal{J}} \\ t \to +\infty}} \zeta_j(x(t)) = \frac{a_j}{c_j}\,, \quad j \in \mathcal{J}\,. \tag{4.38}$$

On the other hand, one has that

$$\lambda_{\mathcal{I}} = ((I - R^T)a)_{\mathcal{I}} = (I - (R^T)_{\mathcal{I}\mathcal{I}})a_{\mathcal{I}} - (R^T)_{\mathcal{I}\mathcal{J}} a_{\mathcal{J}}\,. \tag{4.39}$$

Using (4.33), (4.38), and (4.39), one gets that

$$\begin{aligned} C_{\mathcal{I}}\zeta_{\mathcal{I}}(x(t)) &\geq z_{\mathcal{I}}(t) \\ &= (I - R^T_{\mathcal{I}\mathcal{I}})^{-1}(\lambda_{\mathcal{I}} + (R^T)_{\mathcal{I}\mathcal{J}}\zeta_{\mathcal{J}}(x)) \\ &\xrightarrow[t\to\infty]{t\in\Omega_{\mathcal{J}}} (I - R^T_{\mathcal{I}\mathcal{I}})^{-1}(\lambda_{\mathcal{I}} + (R^T)_{\mathcal{I}\mathcal{J}}a_{\mathcal{J}}) \\ &= a_{\mathcal{I}}\,. \end{aligned} \tag{4.40}$$

Together, (4.38) and (4.40) imply that

$$\liminf_{\substack{t\,\in\,\Omega_{\mathcal{J}} \\ t\,\to\,+\infty}} \zeta_i(x(t)) \geq \frac{a_i}{C_i}, \quad i \in \mathcal{E}\,,$$

so that, for every $\mathcal{J} \subseteq \mathcal{E}$ such that $\Omega_{\mathcal{J}}$ has infinite measure,

$$\lim_{\substack{t\,\in\,\Omega_{\mathcal{J}} \\ t\,\to\,+\infty}} \operatorname{dist}(x(t), \mathcal{X}^*) = 0\,.$$

The claim now follows from the fact that, on the one hand, since $x(t)$ is absolutely continuous,

$$\mathbb{R}_+ = \bigcup_{\mathcal{J}\subset\mathcal{E}} \Omega_{\mathcal{J}} \cup A$$

for some measure-0 subset of times $A \subseteq \mathbb{R}_+$, on the other hand,

$$\lim_{t\to+\infty} \mu(\Omega_{\mathcal{J}} \cap [t, +\infty)) = 0$$

for every $\mathcal{J} \subseteq \mathcal{E}$ such that $\Omega_{\mathcal{J}}$ has finite measure. □

Observe that the set $\mathcal{X}^*$ can consist of more than one point, as the following example shows.

Example 4.3
Consider a network with two cells and only one phase. Assume that the exogenous inflows to the cells are strictly positive, $\lambda_1 = \lambda_2 = \lambda > 0$ and that the outflow capacities for both cells are $c_1 = c_2 = 1$. Then, the dynamics is given by

$$\begin{aligned} \dot{x}_1 &= \lambda - z_1 \\ \dot{x}_2 &= \lambda - z_2 \end{aligned}$$

where

$$\begin{aligned} 0 \leq z_1 \leq u(x), &\quad x_1(z_1 - u(x)) = 0, \\ 0 \leq z_2 \leq u(x), &\quad x_2(z_2 - u(x)) = 0, \end{aligned}$$

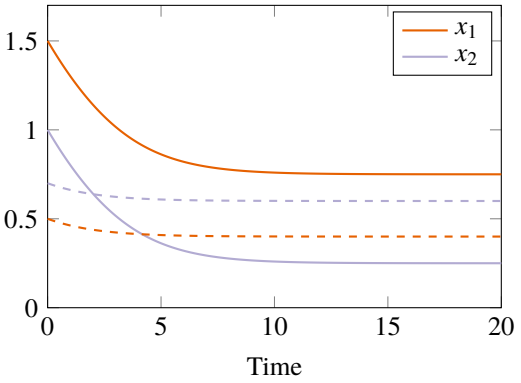


Figure 4.3 The trajectories in Example 4.3. The solid lines are for the initial state $(x_1(0), x_2(0)) = (1.5, 1)$, while the dashed lines are for $(x_1(0), x_2(0)) = (0.5, 0.7)$. For both simulations $\lambda = 0.5$ and $\kappa = 1$.

and the controller is given by

$$u(x) = \frac{x_1 + x_2}{x_1 + x_2 + \kappa}.$$

If $x_1(0) > x_2(0)$, then $\lim_{t\to+\infty} x_1(t) > \lim_{t\to+\infty} x_2(t)$. On the other hand, if $x_1(0) < x_2(0)$, $\lim_{t\to+\infty} x_1(t) < \lim_{t\to+\infty} x_2(t)$. The trajectories for the two different cases are shown in Figure 4.3. □

However, in the special case when every phase only consists of one cell, the following corollary states that $\mathcal{X}^*$ is a singleton.

Corollary 4.2
In the case when each phase only consists of one cell, the solution to the dynamics (4.2) – (4.4) with the GPA controller (4.11) converges to a point $x^* \in \mathcal{X}$, such that $x_i^* > 0$ for all $i \in \mathcal{E}$ and $\zeta_i(x) = a_i/c_i$ for all $i \in \mathcal{E}$. This equilibrium point is independent of the initial state $x(0)$. □

Proof. When every phase consists of one cell, it holds that when $x_i = 0$ for a cell i, $\zeta_i(x) = 0$. Since each cell is inflow-connected, this cannot be an equilibrium. Hence the equilibrium must be such that $x_i^* > 0$. From the definition of $\mathcal{X}^*$ in (4.12), it follows that

$$\zeta_i(x) = \frac{a_i}{c_i}, \quad \forall i \in \mathcal{E}.$$

Let $\{e_1, e_2, \ldots, e_l\}$ be an arbitrary block in $v \in \mathcal{V}$. Using the expression for the GPA-controller in (4.11), the equality above can be rewritten as

$$\begin{bmatrix} c_1 - a_1 & -a_1 & \cdots & -a_1 \\ -a_2 & c_2 - a_2 & \cdots & -a_2 \\ & & \ddots & \\ -a_l & -a_l & \cdots & c_l - a_l \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_l^* \end{bmatrix} = \kappa_v \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_l \end{bmatrix},$$

Let $a^{(v)} = (a_i)_{i \in \mathcal{E}_v}$ which can be written in compact form as

$$(C^{(v)} - a^{(v)} \mathbb{1}^T)(x^{(v)*}) = \kappa_v a^{(v)}$$

where the matrix $(C^{(v)} - a^{(v)} \mathbb{1}^T)$ is invertible if and only if $1 - \mathbb{1}^T \left(C^{(v)}\right)^{-1} a^{(v)} \neq 0$, which is clearly the case since $a_i < c_i$ for all $i \in \mathcal{E}$ and it follows that $\mathcal{X}^*$ only consists of one point. □

4.5 Simulation of Point Queue Dynamics with GPA

To illustrate the concepts presented in this paper, we will simulate the dynamical system using an Euler solver in MATLAB. Since the traffic volumes can not go below zero, we let the actual outflow from one cell be the minimum between the outflow in that discretization step, and the remaining traffic volume in the cell. The network topology we use is the one shown in Figure 2.5. We let the exogenous inflow rate be 0.2 on all incoming cells from the outside of the network, i.e., cell 1 and 2 for junction v_1 and v_3 and cells 3 and 4 for junction v_2 and v_4. For simplicity, we let the outflow capacity be 1 for every cell in the network.

For the vehicles propagating from junction v_1 to junction v_2, we let 20 percent go to the devoted turn cell and 80 percent to the cell that leaves the network (and denote this 20/80). For the vehicles propagating from junction v_2 to junction v_1, this ratio is 30/70 instead. For the vehicles propagation between junction v_3 and v_4, this ratio is set to be 40/60 and in the opposite direction it is 50/50. For the north-south cells, we assume that 65 percent of the vehicles will turn so they will leave the network, i.e., 65 percent do a right turn in cell v_1 and v_3 and 65 percent do a left turn in cell v_2 and v_4. To illustrate the controllers' ability to adopt a new traffic setting, when one-third of the simulation time has passed we change so that 60 percent of the vehicles are turning away from the network in all four junctions instead.

The trajectories for the dynamics (4.2) – (4.4) with GPA control (4.11) in the setting previously described are shown in Figure 4.4. For all four junctions, we let the initial traffic volume in the cells be $x(0) = \begin{bmatrix} 0.5 & 0.4 & 0.3 & 0.2 & 0.1 \end{bmatrix}^T$. As we can see, the controller manages to keep the traffic volume bounded, and adapt to a new setting when the routing is changed. We also see that the traffic volume in a few cells will be close to zero. This is expected since we have cells with different average inflow rate belonging to the same phase, so the cell with lower average inflow rate will stay at zero.

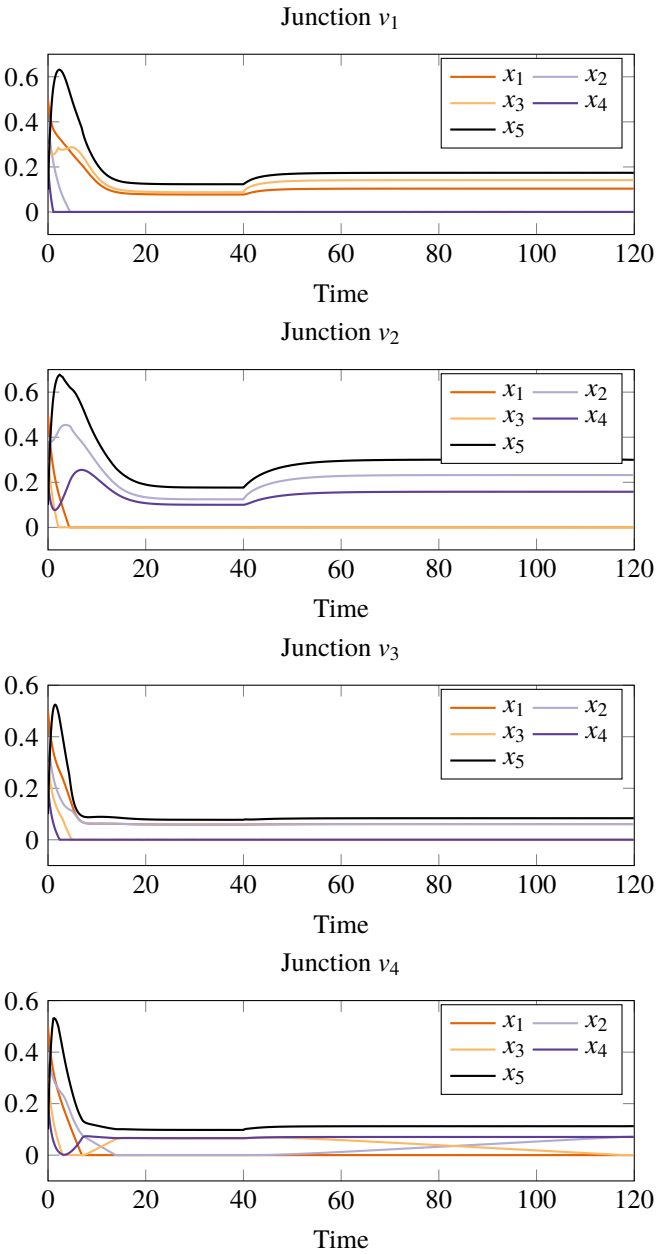


Figure 4.4 How the traffic volumes varies with time for all incoming lanes to the four nodes in the simulations described in Section 4.5.

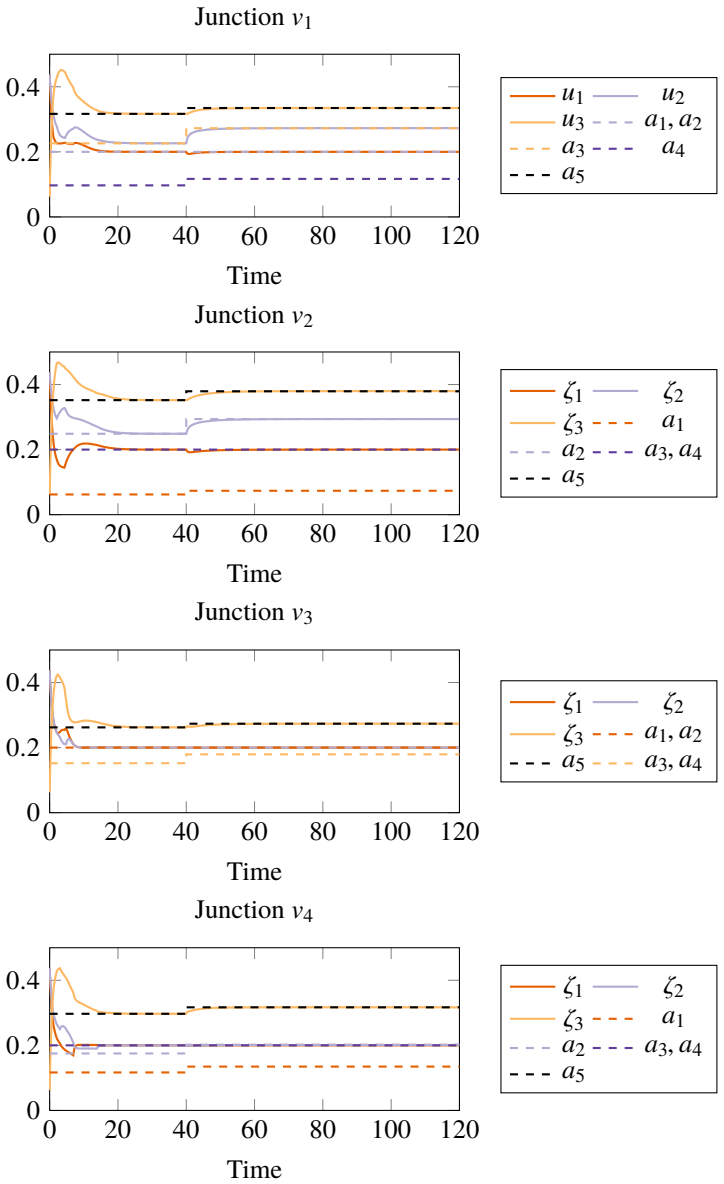


Figure 4.5 How the control signals ζ vary with time for all four junctions in the simulations described in Section 4.5. Observe that the control signal, and hence the controlled outflow is larger than or equal to the average inflow, i.e., $\zeta_1 \geq a_1, a_4$, $\zeta_2 \geq a_2, a_3$, and $\zeta_3 \geq a_5$.

In Figure 4.5 we show the control signals, together with the average inflow rates. We see that the control signals are always greater than or equal to the average inflow rates. In other words, for all cells at equilibrium, it holds that $a_1, a_4 \leq \zeta_1$, $a_2, a_3 \leq \zeta_2$, and $a_5 \leq \zeta_3$. This is, of course, necessary to keep the traffic volume bounded. For the lanes where the control signals are strictly greater than the average inflow rates, the traffic will stay zero, and the actual outflow from every such cell will equal its inflow.

4.6 Phases That Span over Multiple Junctions

In Chapter 4.1 we said that the set of phases was associated with each junction. In this section we will generalize the set of phases, and talk about phases assigned to partitions of the junctions instead. The reason for talking about partitions is that sometimes there is a need to activate green lights at nearby junctions simultaneously.

We let $\mathcal{W}$ be a partition of the cells, i.e.,

$$\mathcal{E} = \bigcup_{k \in \mathcal{W}} \mathcal{E}_k\,, \qquad \mathcal{E}_k \cap \mathcal{E}_h = \emptyset\,,\ \forall h \neq k \in \mathcal{W}\,.$$

Moreover, for every $k \in \mathcal{W}$, we let $\mathcal{P}^{(k)} \subseteq \{\mathcal{Q} : \mathcal{Q} \subseteq \mathcal{E}_k\}$ be a set of local phases of cardinality p_k and let n_k the number of cells in partition k.

For each set of local phases, i.e., the phases associated with a partition, we associate a local phase matrix. This phase matrix is constructed exactly in the same way as the phase matrix was constructed earlier in this chapter when the phases were associated with junctions. The phase matrix for a partition $k \in \mathcal{W}$ is a binary $n_k \times p_k$ matrix

$$P^{(k)} \in \{0,1\}^{\mathcal{E}_k \times \mathcal{P}^{(k)}}$$

that is defined as

$$P^{(k)}_{i,j} = \begin{cases} 1 & \text{if cell } i \in \mathcal{E}_k \text{ activated in phase } j \in \mathcal{P}^{(k)}, \\ 0 & \text{if cell } i \in \mathcal{E}_k \text{ is not activated in phase } j \in \mathcal{P}^{(k)}. \end{cases}$$

The global phase matrix is then defined as

$$P = \begin{bmatrix} P^{(1)} & & & \\ & P^{(2)} & & \\ & & \ddots & \\ & & & P^{(w)} \end{bmatrix}.$$

To control how the phases should be activated in each block of the partition, we introduce the set of control signals

$$\mathcal{U} = \prod_{k \in \mathcal{W}} \mathcal{U}^{(k)}\,,$$

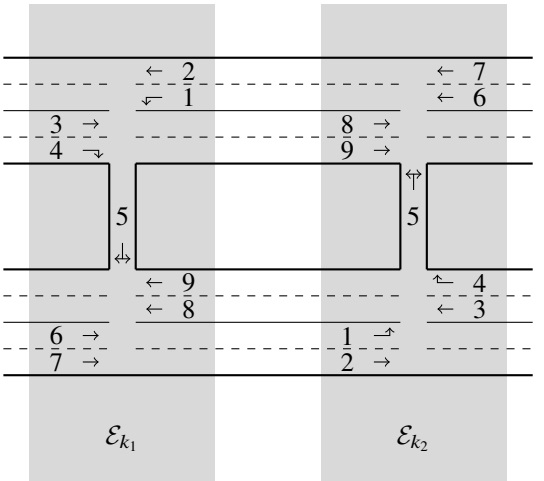


Figure 4.6 An example of when considering another partitioning than the standard partitioning can be useful. In this case, the short one-directional cross-roads may should receive green light simultaneously with the turning lanes on the main roads, in order to avoid spill-back from the cross-roads into the main roads.

where

$$\mathcal{U}^{(k)} = \left\{ u \in \mathbb{R}_+^{\mathcal{P}^{(k)}} : \mathbb{1}^T u \leq 1 \right\}$$

is the set of local controls.

The following example shows when partitions different from the standard one may be useful.

Example 4.4
Consider the small traffic network shown in Figure 4.6. If vehicles are allowed to turn into the short cross-roads, we want to make sure that the cross-roads have green light as well. This to avoid back-spill into the main-roads. Hence, instead of constructing the phases over the four junctions in the network, we construct the phases over two partitions $\mathcal{E}_{k_1}$ and $\mathcal{E}_{k_2}$. A set of phases that makes sure that the cross-roads receives green light whenever there is a inflow to the cross-roads and prevents collisions is

$$P^{(k)} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}^T,$$

where $k \in \{k_1, k_2\}$. □

The previously stated stability results can be generalized to when the phases are constructed over partitions as well

Proposition 4.2
If the set of phases and control signals are constructed over a partition of the network instead of over the junctions, the stability results in Theorem 4.1 still hold true. □

The proof of the proposition above goes in the same way as the proof of Theorem 4.1, but with summations over partitions instead of junctions.

4.7 Non-Orthogonal Phases

In the case when the phases are not orthogonal, it is still possible to compute a control action that will stabilize the system.

In the general case, the controller $u(x)$ is computed by solving the following convex optimization problem

$$u(x) \in \underset{\upsilon \in \mathcal{U}}{\operatorname{argmax}} H(x, \upsilon) \tag{4.41}$$

This is a generalization of the results presented in Chapters 4.3 – 4.4, since the controller in (4.11) is the solution to (4.41) when the phases are orthogonal.

If $x_i = 0$ for a subset of cells, the optimization problem in (4.41) is not necessary strictly convex anymore, and the control action may not be a unique one, as the following example illustrates

Example 4.5
Consider a junction v with three cells (indexed $\{1, 2, 3\}$), all with unit outflow capacity. Let the phase matrix be

$$P^{(v)} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} .$$

The maximization problem in (4.41) can then be equivalently written as

$$\begin{aligned} u^{(v)}(x) \in \underset{\upsilon \in \mathcal{U}^{(v)}}{\operatorname{argmax}} \quad & \{x_1 \log(\upsilon_1) + x_2 \log(\upsilon_1 + \upsilon_2) \\ & + x_3 \log(\upsilon_2) + \kappa_v \log(1 - \upsilon_1 - \upsilon_2)\} . \end{aligned}$$

The solution to the maximization problem is:

- If $x_1 = 0, x_2 > 0, x_3 = 0$, then

$$0 \leq u_1 \leq \frac{x_2}{x_2 + \kappa_v}, \quad u_2 = \frac{x_2}{x_2 + \kappa_v} - u_1 .$$

- For all other cases,

$$u_1 = \frac{x_1(x_1 + x_2 + x_3)}{(x_1 + x_3)(x_1 + x_2 + x_3 + \kappa_v)}, \quad u_2 = \frac{x_3}{x_1} u_1 .$$

We will now show for this example that there exits a control signal in u such that the trajectory is absolutely continuous. Assume that the cells have exogenous inflows, λ_1, λ_2 and λ_3, respectively, and no inflows from other cells. Let $x_1 = 0$ and $x_3 = 0$, then

$$u_1 + u_2 = \frac{x_2}{x_2 + \kappa_v} .$$

Now suppose that $\lambda_1 + \lambda_3 < u_1 + u_2$. To keep $x(t) \geq 0$, we have to choose z_1, z_2 such that $z_1 \leq \lambda_1$ and $z_3 \leq \lambda_3$. However, choosing $z_1 < \lambda_1$ or $z_3 < \lambda_3$, will make $\dot{x}_1 > 0$ or $\dot{x}_3 > 0$, and the traffic volumes will become positive. Let us for simplicity assume that $z_1 = 0$ and $z_3 = \lambda_3$, then after a sufficiently small time, $x_1 > 0$ and

$$u_1 = \frac{x_1 + x_2}{x_1 + x_2 + \kappa_v} > \lambda_1 ,$$

and x_1 will immediately go back to zero again. Therefore this solution cannot be absolutely continuous. To get an absolutely continuous solution in this case one has to choose $z_1 = \lambda_1$ and $z_3 = \lambda_3$. □

Although, the existence and possible uniqueness is topic for further research, the previously developed stability theory applies to all absolutely continuous trajectories.

Proposition 4.3
Consider the network flow dynamics (4.2) – (4.4) with the controller in (4.41). If an absolutely continuous trajectory $x(t)$ exists for the dynamics, the trajectory will converge to the set $\mathcal{X}^*$ as given in (4.12). In other words,

$$x(t) \to \mathcal{X}^* ,$$

when $t \to +\infty$. □

The proof of this proposition follows in the same way as the proof of Theorem 4.1, since there it is has already been shown that all absolutely continuous trajectories for the dynamics (4.2) – (4.4) with the controller in (4.41), we converge to the set $\mathcal{X}^*$.

4.8 Conclusions

In this chapter, we have presented a decentralized feedback controller for control of traffic signals, the Generalized Proportional Allocation controller. We have shown that the controller is both stabilizing, i.e., the traffic volumes will stay bounded over time, and throughput optimal, i.e., no other controller is able to stabilize the network for a larger amount of exogenous inflows than this one. While the dynamics we have analyzed in this chapter should be seen as an averaged dynamics for the control signal, we will in the next chapter discuss how to discretize the controller and evaluate the controller's performance in a microscopic traffic simulator.

5

Evaluation of GPA Control in a Microscopic Traffic Simulator

In this chapter, we perform simulations in a micro-simulator of the Generalized Proportional Allocation policy presented in the previous chapter. We start by showing a discretized version of the controller suitable for implementation. The micro-simulator we will be using is SUMO [Krajzewicz et al., 2012], which is an open-source micro-simulator for traffic, mainly developed by the Institute of Transportation Systems at the German Aerospace Center. The GPA controller for two different topologies will be tested, one artificial Manhattan-like grid and one real scenario covering the city of Luxembourg. For the artificial scenario, we will compare the controller's performance with the MaxPressure-controller, but also with standard fixed-time control and with a pure proportional fairness controller that does not adapt the cycle length. For the Luxembourg scenario, we compare the GPA with the standard fixed-time controller that comes with the scenario.

5.1 Signal Control with Discrete Control Actions

In the previous chapters a time-averaged control signal for the outflows from the point queue dynamics was studied. However, in reality traffic signals are either allowing outflow or not allowing out flow. This means to implement the controller in a traffic simulator, a discretized version of the controller is needed. In other words, the outflow controller, $\zeta(x)$, now either gives the control signal 0 or 1, and hence $\zeta \in \{0,1\}^{\mathcal{E}}$.

When discretizing the control action, we also explicitly need to define the clearance phases, i.e., the phases that need to be activated between two real phases to give the vehicles that already are in the junction time to pass through. For each phase $p \in \mathcal{P}$, we will denote its corresponding clearance phase as p'. We then introduce

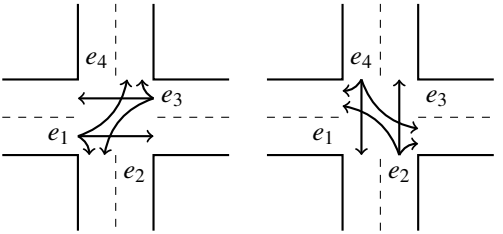


Figure 5.1 The phases for the junction in Example 5.1. This junction has four incoming lanes and two phases, $p_1 = \{e_1, e_3\}$ and $p_2 = \{e_2, e_4\}$. Hence there is no specific lane for the drivers who are turning left.

the extended set of phases, which we will denote $\bar{\mathcal{P}}$, that is constructed such that for every phase $p \in \mathcal{P}$, both the phase and its corresponding clearance phase is in $\bar{\mathcal{P}}$. Hence $\bar{\mathcal{P}}$ has twice the cardinality compared to $\mathcal{P}$ and $p, p' \in \bar{\mathcal{P}}$.

With the clearance phases defined, we can now define a signal program for every signalized junction. A signal program tells at which time each phase should be deactivated and which phase that should be activated next. For each signalized junction $v \in \mathcal{V}$ a signal program is assigned. The signal program will be denoted $\mathcal{T}^{(v)}$, and is such that

$$\mathcal{T}^{(v)} = \{(p, t_{\text{end}}) \mid (p, t_{\text{end}}) \in \bar{\mathcal{P}}^{(v)} \times \mathbb{R}_+\},$$

where p is the phase and t_{end} is the time when the phase should be deactivated. Specifying the signal program with only the deactivation time is enough, since the phase that will be active at time t can then be determined by the phase with the smallest deactivation time that is larger than t. Formally, we can define the function $\rho^{(v)}(t)$ that gives the phase that is activated at time t as follows

$$\rho^{(v)}(t) = \{p : (p, t_{\text{end}}) \in \bar{\mathcal{T}}^{(v)} \mid t_{\text{end}} > t \text{ and } t_{\text{end}} \leq t'_{\text{end}} \text{ for all } (p', t'_{\text{end}}) \in \bar{\mathcal{T}}^{(v)}\}.$$

To illustrate how a signal program can look like, we provide an example for a junction with two phases:

Example 5.1
Consider the junction in Figure 5.1 with the incoming lanes numbered as in the figure. For this junction the phase matrix is

$$P = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}^T.$$

An example of signal program is shown in Figure 5.2. The program is $\mathcal{T} = \{(p_1, 25), (p'_1, 30), (p_2, 55), (p'_2, 60)\}$, which means that both the phases are activated for 25 seconds each, and the clearance phases are activated for 5 seconds each. □

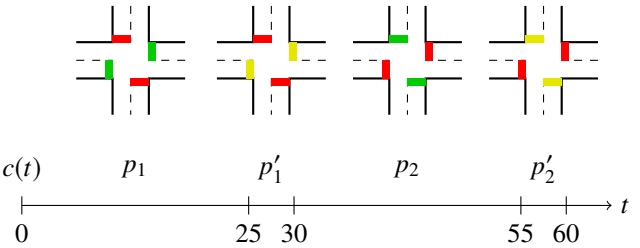


Figure 5.2 Example of a signal program for the junction in Example 5.1. In this example the signal program is $\mathcal{T} = \{(p_1, 25), (p'_1, 30), (p_2, 55), (p'_2, 60)\}$.

Moreover, we let

$$T^{(v)} = \max\{t_{\text{end}} \mid (p, t_{\text{end}}) \in \bar{\mathcal{T}}^{(v)}\}$$

denote the time when the signal program ends for junction v. At latest at time $T^{(v)}$ a new signal timing program has to be determined.

5.2 Discretization of the GPA Controller

In this section, we present two different discretization schemes of the GPA. The first one makes sure that all the clearance phase are activated during one cycle. The second one only activates the clearance phases if their corresponding phases have been activated.

As mentioned earlier, the GPA controllers are fully distributed, in the sense that to determine the signal program in one junction, the controller only needs information about the queue lengths on the incoming lanes for that junction.

For all of the controllers presented in this section, we assume for simplicity of the presentation that after a phase has been activated, a clearance phase has to be activated for a fixed amount of time, which we will denote $T_w > 0$. This time is independent of which phase that has just been activated and which junction the program is computed for.

GPA with Full Clearance Cycles

For this discretization of the GPA, we assume that all the clearance phases have to be activated for each cycle. When $t = T^{(v)}$, a new signal program is computed for

junction $v \in \mathcal{V}$ by solving the following convex optimization problem:

$$\begin{aligned} \underset{\substack{u \in \mathcal{U}^{(v)} \\ w \in \mathbb{R}_+}}{\text{maximize}} \quad & \sum_{i \in \mathcal{E}_v} x_i(t) \log\left((Pu)_i\right) + \kappa_v \log(w), \\ \text{subject to} \quad & \sum_{p \in \mathcal{P}^{(v)}} u_p + w = 1, \\ & w \geq \bar{w}. \end{aligned} \tag{5.1}$$

Recall from Chaper 4 that κ_v is a tuning parameter for the controller. The optimization problem above is a restatement of the optimization problem in (4.41). In (5.1) we have introduced the variable $w = 1 - \mathbb{1}^T u$ to explicitly have a quantity for how large fraction of a cycle that should be used for the clearance phases. Moreover, we introduce the possibility to explicitly impose a lower bound $\bar{w}$ on this quantity. The parameter $\bar{w}$ is a design-parameter for the controller and we will later see why this lower bound is needed.

The solution (u, w) to the optimization problem in (5.1) determines the fraction of a cycle that each phase should be activated, where each element in u contains this fraction, but also how large fraction of the cycle that should be allocated to clearance phases, w. Observe that as long as the measured queue lengths are finite, which they will always be due to limited sensor coverage area, w will be strictly greater than zero. Since we assume that each clearance phase has to be activated for a fixed amount of time, $T_w > 0$, the total cycle length T_{cyc} for the upcoming cycle can be computed by

$$T_{\text{cyc}} = \frac{|\mathcal{P}^{(v)}| \cdot T_w}{w}. \tag{5.2}$$

From the expression above we see that imposing a lower bound on w will impose an upper bound on the cycle length.

With the knowledge of the full-cycle length, the signal program for the upcoming cycle can be computed according to Algorithm 1.

For the case when $\bar{w} = 0$, we recall from Chapter 4, that when the phases are orthogonal, an explicit solution to the optimization problem (5.1) can be found. Hence an explicit expression for the fraction of the cycle that should be used for clearance phases can be obtained as well, and the full explicit solution to (5.1) for every junction $v \in \mathcal{V}$ is

$$\begin{aligned} u_p(x(t)) &= \frac{\sum_{i \in \mathcal{E}} P_{ip} x_i(t)}{\kappa_v + \sum_{i \in \mathcal{E}_v} x_i(t)}, \quad \forall p \in \mathcal{P}^{(v)}, \\ w(x(t)) &= \frac{\kappa_v}{\kappa_v + \sum_{i \in \mathcal{E}_v} x_i(t)}. \end{aligned} \tag{5.3}$$

Algorithm 1: GPA with Full Clearance Cycles

Data: Current time t, local queue lengths $x^{(v)}(t)$, phase matrix $P^{(v)}$, clearance time T_w, tuning parameters $\kappa, \bar{w}$
Result: Signal program $\mathcal{T}^{(v)}$
$\mathcal{T}^{(v)} \leftarrow \emptyset$
$n_{p_v} \leftarrow$ Number of columns in $P^{(v)}$
$(u, w) \leftarrow$ Solution to (5.1) given $x^{(v)}(t), P^{(v)}, \kappa, \bar{w}$
$T_{\text{cyc}} \leftarrow n_{p_v} \cdot T_w / w$
$t_{\text{end}} \leftarrow t$
for $i \leftarrow 1$ **to** n_{p_v} **do**
  $t_{\text{end}} \leftarrow t_{\text{end}} + u_i \cdot T_{\text{cyc}}$
  $\mathcal{T}^{(v)} \leftarrow \mathcal{T}^{(v)} + (p_i, t_{\text{end}})$ ▷ Add phase p_i
  $t_{\text{end}} \leftarrow t_{\text{end}} + T_w$
  $\mathcal{T}^{(v)} \leftarrow \mathcal{T}^{(v)} + (p_i', t_{\text{end}})$ ▷ Add clearance phase p_i'
end

From (5.2) a direct expression for the total cycle length can be obtained

$$T_{\text{cyc}} = T_w |\mathcal{P}^{(v)}| + \frac{T_w |\mathcal{P}^{(v)}|}{\kappa_v} \sum_{i \in \mathcal{E}_v} x_i(t) \,.$$

From the expressions above we can observe a few things. First, we see that the fraction of the cycle that each phase is activated is proportional to the queue lengths in that phase, and this explains why we denote this control strategy Generalized Proportional Allocation. Moreover, we get an interpretation of the tuning parameter κ: it tells how the cycle length T_{cyc} should scale with the current queue lengths. If κ_v is small, even small demands will cause longer cycles, while if κ is large, the cycles will be short even for high demands. Hence, a too small κ_v may give too long cycles, which can result in that lanes get more green-light than needed and the controller ends up giving green light to empty lanes. This while vehicles in other lanes are waiting for service. On the other hand, a too large κ may make the cycle lengths so short that the fraction of the cycle that each phase gets activated is too short for the drivers to react on.

The following example shows that one needs to explicitly impose an upper bound on the cycle length when the controller is discretized:

Example 5.2
Consider a junction with two incoming lanes with unit flow capacity, both having their own phase, and let the exogenous inflows $\lambda_1 = \lambda_2 = \lambda$, $T_w = 1$, $\bar{w} = 0$, $x_1(0) = A > 0$, and $x_2(0) = 0$. The control signals and the cycle time for the first

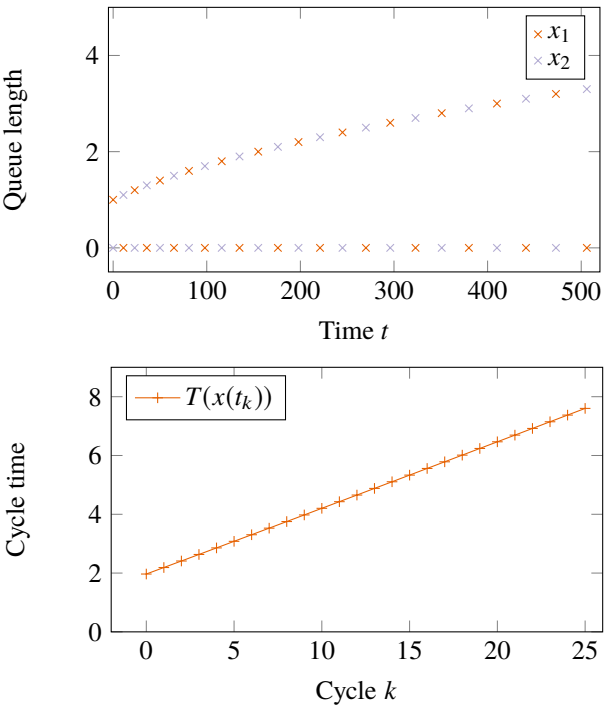


Figure 5.3 How the traffic volumes evolve in time together with the cycle times for the system in Example 5.2. We can observe that the cycle length increases for each cycle.

iteration is then given by

$$\begin{aligned} u_1(x(0)) &= \frac{A}{A+\kappa}, \\ u_2(x(0)) &= 0, \\ T(x(0)) &= \frac{A+\kappa}{\kappa}. \end{aligned}$$

Observe that the cycle time $T(x(0))$ is strictly increasing with A. After one full service cycle, i.e., at $t_1 = T(x(0))$ the queue lengths are

$$\begin{aligned} x_1(t_1) &= A + T(x(0))\left(\lambda - \frac{A}{A+\kappa}\right) = \overbrace{A + \lambda\frac{A+\kappa}{\kappa} - \frac{A}{\kappa}}^{f(A)}, \\ x_2(t_1) &= T(x(0))\lambda = \lambda\left(\frac{A+\kappa}{\kappa}\right). \end{aligned}$$

If $x_1(t_1) = 0$, then due to symmetry, the analysis of the system can be repeated in the same way with a new initial condition. To make sure that one queue always get empty during the service cycles, it must hold that $f(A) \leq 0$. Moreover, to make sure that the other queue grows, it must also hold that $x_2(t_1) > A$ which can be equivalently expressed as

$$\begin{aligned} A\kappa + \lambda(A + \kappa) - A &\leq 0\,, \\ A\kappa - \lambda(A + \kappa) &< 0\,. \end{aligned}$$

The choice of $\lambda = \kappa = 0.1$ and $A = 1$ is one set of parameters satisfying the constraints above, and will hence make the queue lengths and cycle times grow unboundedly. How the traffic volume in the lanes and cycle times evolve in this case is shown in Figure 5.3. □

GPA with Shortened Cycles

One possible drawback of the GPA controller with full clearance cycles is that it has to activate all the clearance phases in one cycle. This property implies that if the junction is empty when the signal program is computed, it will take $|\mathcal{P}^{(v)}|T_w$ seconds until a new signal program is computed. Therefore we also present a version of the GPA where the clearance phases only get activated if their corresponding phases have been activated. If we let n'_{p_v} denote the number of phases that will be activated during the upcoming cycle, the total cycle time is given by

$$T_{\text{cyc}} = \frac{n'_{p_v} T_w}{w}\,.$$

How to compute the signal program in this case with shortened cycles, is shown in Algorithm 2.

5.3 Comparison with MaxPressure Control in a Manhattan Grid

In this section we will compare the two discretized versions of the GPA with the MaxPressure controller in an artificial Manhattan-like grid. As stated in Chapter 1.2, the MaxPressure controller and variants of it has been well-studied in the literature, which motivates a comparison.

MaxPressure Controller

While the GPA controller only needs information about the incoming queues to the junction it is controlling, the MaxPressure controller needs information about the queues at the down-stream junctions as well. The idea behind the MaxPressure controller is that for each lane compute the difference between its queue length and the queue lengths immediately downstream, i.e., the queues the vehicles will proceed to when they receive a green light in the current junction. Then a phase's pressure is

Algorithm 2: GPA with Shortened Cycles

Data: Current time t, local queue lengths $x^{(v)}(t)$, phase matrix $P^{(v)}$, clearance time T_w, tuning parameters $\kappa, \bar{w}$
Result: Signal program $\mathcal{T}^{(v)}$
$\mathcal{T}^{(v)} \leftarrow \emptyset$
$n_{p_v} \leftarrow$ Number of columns in $P^{(v)}$
$(u, w) \leftarrow$ Solution to (5.1) given $x^{(v)}(t), P^{(v)}, \kappa, \bar{w}$
▷ Compute the number of phases to be activated
$n'_{p_v} \leftarrow 0$
for $i \leftarrow 1$ **to** n_{p_v} **do**
  if $u_i > 0$ **then**
    $n'_{p_v} \leftarrow n'_{p_v} + 1$
  end
end
if $n'_{p_v} > 0$ **then**
  ▷ If vehicles are present on some phases, activate those
  $T_{\text{cyc}} \leftarrow n'_{p_v} \cdot T_w / w$
  $t_{\text{end}} \leftarrow t$
  for $i \leftarrow 1$ **to** n_{p_v} **do**
    if $u_i > 0$ **then**
      $t_{\text{end}} \leftarrow t_{\text{end}} + u_i \cdot T_{\text{cyc}}$
      ▷ Add phase p_i
      $\mathcal{T}^{(v)} \leftarrow \mathcal{T}^{(v)} + (p_i, t_{\text{end}})$
      $t_{\text{end}} \leftarrow t_{\text{end}} + T_w$
      ▷ Add clearance phase p'_i
      $\mathcal{T}^{(v)} \leftarrow \mathcal{T}^{(v)} + (p'_i, t_{\text{end}})$
    end
  end
else
  ▷ If no vehicles are present, hold a clearance phase for one time unit
  $\mathcal{T}^{(v)} \leftarrow (p'_1, t + 1)$
end

Algorithm 3: MaxPressure

Data: Current time t, local queue lengths $x(t)$, phase matrix $P^{(v)}$, routing matrix R, phase duration d
Result: Signal program $\mathcal{T}^{(v)}$
$\mathcal{T}^{(v)} \leftarrow \emptyset$
$n_{p_v} \leftarrow$ Number of columns in $P^{(v)}$
for $i \leftarrow 1$ **to** n_{p_v} **do**
  for $l \in \mathcal{E}_v$ **do**
    if $l \in p_i^{(v)}$ **then**
      $w_i \leftarrow w_i + x_l(t) - \sum_k R_{lk} x_k(t)$
    end
  end
end
$i \leftarrow \text{argmax}_i \, w_i$
▷ Add phase p_i
$\mathcal{T}^{(v)} \leftarrow \mathcal{T}^{(v)} + (p_i, t + d)$
▷ Add clearance phase p_i'
$\mathcal{T}^{(v)} \leftarrow \mathcal{T}^{(v)} + (p_i', t + d + T_w)$

defined as the sum of these differences for all lanes in the phase. Since vehicles in one lane may proceed to several different lanes, the MaxPressure controller needs to know the routing matrix R, to compute the right pressures. This means that, compared to the GPA strategies, the MaxPressure requires both more state information and more exogenous information.

While the MaxPressure allows for taking different flow rates for different phases into account when determining the phases' pressures, we will assume that the flow rates are the same for all the phases. Under this assumption, the pressure, w_p, for each phase $p \in \mathcal{P}^{(v)}$ can be computed as

$$w_p = \sum_{i \in p} \left(x_i(t) - \sum_k R_{ik} x_k(t) \right) .$$

The phase that should be activated is then any phase in the set $\text{argmax}_i \, w_i$.

When implementing the MaxPressure in a discretized signal control setting, one has to determine how long a phase should be activated. We will denote this activation time parameter d, and will assume that after a phase has been activated for d seconds, its corresponding clearance phase will be activated for T_w seconds. After that, all the pressures associated with the junction will be recomputed, and a new phase will be activated for d seconds. The algorithm for computing a signal program with the MaxPressure controller is given in Algorithm 3.

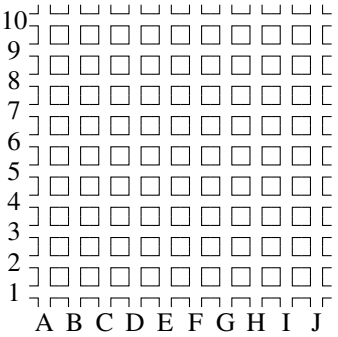


Figure 5.4 The Manhattan-like network used in the comparison between GPA and MaxPressure.

Simulation Setting

To compare the two different discretized versions of the GPA and the MaxPressure controller, we simulate both controllers on an artificial Manhattan-like grid with artificial demand. A schematic drawing of the network is shown in Figure 5.4. In a setting like this, we can experiment with the turning ratios, and provide the MaxPressure controller both correct and incorrect turning ratios, to investigate the robustness properties of both the MaxPressure and the variants of the GPA controller.

The Manhattan grid we will be using has ten bidirectional north to south streets (indexed A to J) and ten bidirectional east to west streets (indexed 1 to 10). All streets with an odd number or indexed by letter A, C, E, G or I consist of one lane in each direction, while the others consist of two lanes in each direction. The speed limit on each lane is 50 km/h. The distance between each junction is three hundred meters. Fifty meters before each junction, every street has an additional lane, reserved for vehicles that want to turn left. Due to the varying number of lanes, four different junction topologies exist, all shown in Figure 5.5, together with the set of possible phases. Each junction is equipped with sensors on the incoming lanes that can measure the number of vehicles queuing up to fifty meters from the junction. The sensors measure the queue lengths by the number of stopped vehicles.

Since the scenario is artificial, we can generate demand with prescribed turning ratios and hence let the MaxPressure controller run in an ideal setting. For the demand generation, we assume that at each junction a vehicle will with probability 0.2 turn left, with probability 0.6 go straight and with probability 0.2 turn right. We do assume that all vehicles depart from lanes connected to the boundary of the network, and all vehicles will also end their trips when they have reached the boundary of the network. In other words, no vehicles will depart or arrive inside the grid. We will study the controllers' performance for three different demands. The demands are determined by the rate vehicles are inserted on the boundary of the network. Every second and for every lane directly connected to the boundary

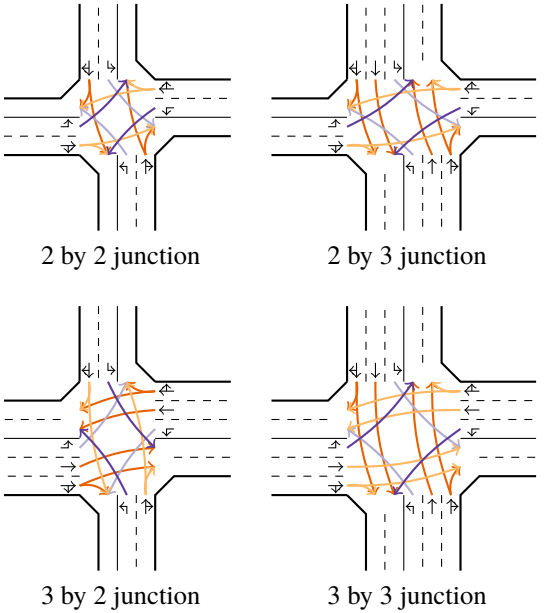


Figure 5.5 The four different types of junctions present in the Manhattan grid, together with their phases.

of the network, a new vehicle will depart with probably δ. The probabilities for the three different demands are $\delta = 0.05$, $\delta = 0.1$ and $\delta = 0.15$. We generate vehicles for 3600 seconds and then let the simulations run until all the vehicles have left the network.

Heuristic Tuning of the GPA Controller

To find a suitable choice of the parameter κ for the GPA controller with shorted cycles, we test the controller with five different values of $\kappa = 1, 5, 10, 15, 20$. In Table 5.1 we show how the total travel time varies for the GPA controller with shortened cycles for different values of κ, which is set to be the same for all junctions in the network. Since the phases in this scenario are all orthogonal, the expressions in (5.3) can be used to solve the optimization problem in (5.1). The tuning parameter $\bar{w}$ is set to $\bar{w} = 0$ for all simulations. How the total queue length varies with time is shown in Figure B.1–B.3. From the figures, we can conclude that $\kappa = 5$ and $\kappa = 10$ give the best performance in this setting for the cases when $\delta = 0.10$ and $\delta = 0.15$. In Figure 5.6 we have plotted the total queue lengths in those cases, in order to get a side by side comparison. In Table 5.1 the total travel time (TTT) for all vehicles loaded into the network is shown.

Table 5.1 GPA with Shorted Cycles - Manhattan Scenario

κ	δ	Total Travel Time [h]
1	0.05	1398
5	0.05	715
10	0.05	699
15	0.05	696
20	0.05	690
1	0.10	7636
5	0.10	1898
10	0.10	1992
15	0.10	2263
20	0.10	2495
1	0.15	$+\infty$
5	0.15	5134
10	0.15	4498
15	0.15	5140
20	0.15	6050

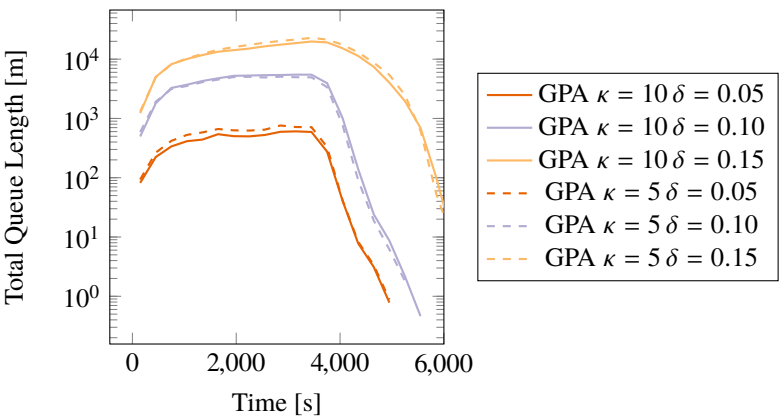


Figure 5.6 How the queue length varies with time when the GPA with shortened cycles are used in the Manhattan grid. The GPA is tested with two different values of $\kappa = 5, 10$ for the three demand scenarios $\delta = 0.05, 0.10, 0.15$. To improve the readability of the results, the queue lengths are averaged over 300 seconds interval.

Heuristic Tuning of the MaxPressure Controller

Since the MaxPressure controller will decide its control action based on queue lengths on the downstream lanes, it is not always clear which downstream lane a vehicle can end up in. Because a vehicle can choose between several lanes that are all valid for its path, the vehicle's lane choice will be determined during the simulation, and depend upon how many other vehicles that are occupying the possible lanes. Because of this, we assume that if a vehicle can choose between several lanes, it will try to join the shortest one. This implies that if for example, a vehicle going straight can choose between lane l_1 and l_2, but l_2 is also used by vehicles turning right (the probability that a vehicle is turning right is 0.2, and going straight is 0.6), the probability that the vehicle will queue up in lane l_1 is assumed to be 0.4 and the probability that the vehicle will queue up in lane l_2 is 0.2.

To also investigate the MaxPressure controller's robustness with respect to the routing information, we perform simulations both when the controller has the correct information about the turning probabilities, i.e., that a vehicle will turn left with probability 0.2, continue straight with probability 0.6 and turn right with probability 0.2, and when it has incorrect information. For the simulations when the MaxPressure has the wrong turning information, the controller instead has the information that with probability 0.6 the vehicle will turn right, with probability 0.3 the vehicle will proceed straight and with probability 0.1 the vehicle will turn left. In the simulations, we consider four different phase durations: $d = 5$ seconds, $d = 10$ seconds, $d = 20$ seconds, and $d = 30$ seconds.

How the total queue lengths vary over time for the different demands is shown in Figure B.4, Figure B.5 and Figure B.6. The total travel time, both when the MaxPressure controller is operating with the right and the wrong turning ratios, are shown in Table 5.2. From these results, we conclude that a phase duration of $d = 10$, is the most efficient for the demands $\delta = 0.10$ and $\delta = 0.15$. For $\delta = 0.05$, a phase duration of $d = 5$ is most efficient. That the phase durations should be low has probably has to do with that with a longer phase duration the activation time is becoming larger than the time it takes to empty the measurable part of the queue. Another interesting observation is that if the MaxPressure controller has wrong information about the turning ratios, it seems like, for at least in this case, that its performance does not decrease significantly.

Results of the Comparison

To better observe the difference between the GPA and MaxPressure, we have taken the best GPA configuration with $\kappa = 20$ for $\delta = 0.05$ and $\kappa = 10$ for $\delta = 0.10$ and $\delta = 0.15$, and the best MaxPressure configuration with $d = 5$ for $\delta = 0.05$ and $d = 10$ for $\delta = 0.10$ and $\delta = 0.15$. The results are shown in Figure 5.7. From the simulations, we can conclude that, for this scenario and during high demands, the MaxPressure controller performs better than the GPA controller. On the other hand, during low demands the GPA performs better. One explanation for this could

Table 5.2 MaxPressure - Manhattan Scenario

d	δ	TTT correct TR [h]	TTT incorrect TR [h]
5	0.05	770	772
10	0.05	858	856
20	0.05	1079	1102
30	0.05	1172	1193
5	0.10	1910	1906
10	0.10	1865	1864
20	0.10	2254	2312
30	0.10	2690	2718
5	0.15	4873	4861
10	0.15	3511	3488
20	0.15	3992	4102
30	0.15	5579	5590

be that during low demands, adapting the cycle length is critical, while during high demands when almost all the sensors are covered, it is more important to keep the queue balanced between the current and downstream lanes.

For reference, we also compare the results for GPA and MaxPressure with a standard fixed-time (FT) controller and a proportional fair (PF) controller, i.e., the GPA controller with $\kappa = 0$ and a prescribed fixed cycle length of 110 seconds. For the fixed-time controller, we use the standard phase activation times given by SUMO. The phases which contain a straight movement are activated for 30 seconds and phases only containing left or right turn movements are activated for 15 seconds. The clearance time for each phase is still set to 5 seconds. This means that the cycle lengths for each of the four types of junctions will be 110 seconds.

The total travel times for the fixed-time controller and proportional fairness controller are shown in Table 5.3 and how the queue lengths compared to the other controllers are shown in Figure B.7–B.9. Observe that when the demand $d = 0.15$, the proportional fair controller creates a gridlock situation. Since the controller is known to be throughput optimal in an ideal setting [Walton, 2014], the reason for this gridlock is probably due to back-spill of vehicles. This occurs when a queue for one junction is getting so long so that it blocks another junction.

A interesting observation from Figure B.7 and B.8 is that the fixed-time controller performs better than the proportional fair controller, but worse than the GPA and MaxPressure controller. This means that the idea of imposing dynamic cycle lengths into the proportional fair controller –as the GPA controller does– is quite significant to achieve a good performing controller.

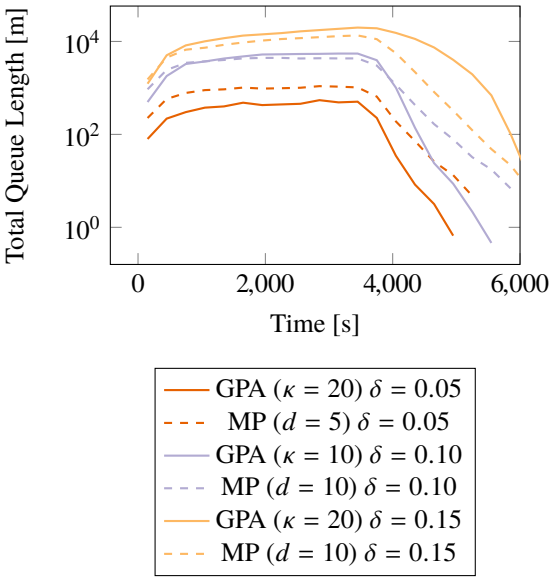


Figure 5.7 Comparison between the GPA controller and the MaxPressure controller for the three different demand levels in the Manhattan scenario. In order to improve the readability of the results, the queue lengths are averaged over 300 seconds interval.

Table 5.3 Fixed-Time and Proportional Fair Control - Manhattan Scenario

Controller	δ	Total Travel Time [h]
FT	0.05	1201
FT	0.10	2555
FT	0.15	4642
PF	0.05	1694
PF	0.10	4165
PF	0.15	$+\infty$

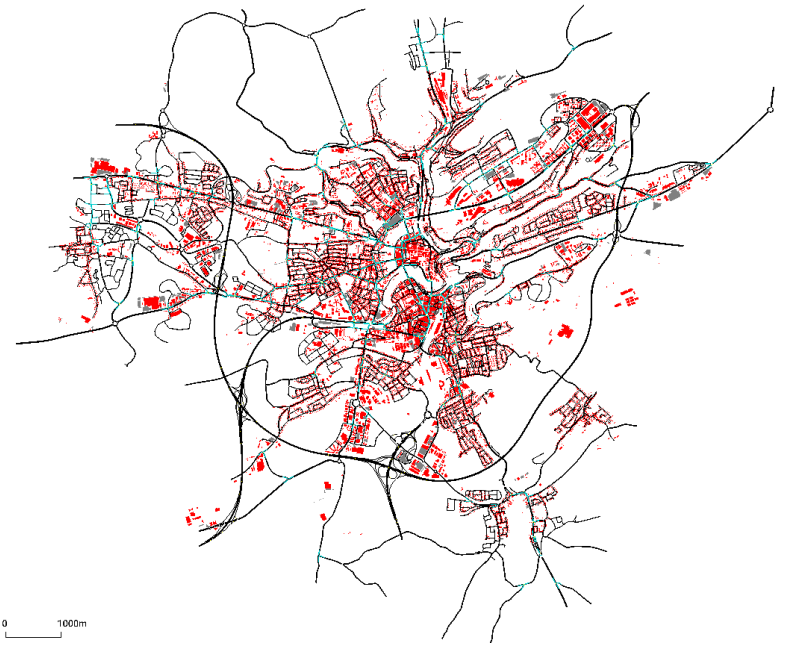


Figure 5.8 The roads in the city of Luxembourg in the LuST scenario, imported as a screenshot from the SUMO micro-simulator.

5.4 Simulations of a Luxembourg Scenario

To test the proposed controller in a realistic scenario, we make use of the Luxembourg SUMO Traffic (LuST) scenario presented in [Codecá et al., 2017][1]. The scenario models the city center of Luxembourg during a full day, and the authors of [Codecá et al., 2017] have made several adjustments from some given population data when creating the scenario, to make it as realistic as possible.

To each of the 199 signalized junctions, we have added a lane area detector to each incoming lane. The length of the detectors are 100 meters, or as long as the lane is if it is shorter than 100 meters. Those sensors are added to give the controller real-time information about the queue lengths at each junction.

As input to the system, we are using the Dynamic User Assignment demand data. For this data-set, the drivers try to take their shortest path (with respect to time) between their current position and destination. It is assumed that 70 percent of the vehicles can recompute their shortest path while driving, and will do so every fifth minute. This rerouting possibility is introduced in order to model the fact that more and more drivers are using online navigation with real-time traffic state information,

[1] The scenario files are obtained from https://github.com/lcodeca/LuSTScenario/tree/v2.0

and will hence get updates about what the optimal route choice is.

In the LuST scenario, the phases are constructed in a bit more complex way and are not always orthogonal. For non-orthogonal phases, it is not always the case that all lanes receive yellow light when a clearance phase is activated. If the lane receives a green light in the next phase as well, it will receive green light during the clearance phase too. This property makes it more difficult to shorten the cycle, and for that reason, we choose to implement the controller which activates all the clearance phases in the cycle, i.e., the controller is given in Algorithm 1.

As mentioned, the phases in the LuST scenario are not orthogonal in each junction. Hence we have to solve the convex optimization problem in (5.1) to compute the phase activation. The computation is done by using the solver CVXPY[2] in Python. Although the controller can be implemented in a distributed manner, the simulations are in this thesis performed on a single computer. Despite the size of the network, and that the communication via TraCI between the controller written in Python and SUMO slows down the simulations significantly, the simulations are still running about 2.5 times faster than real-time. Hence there is no problem with running this controller in a real-time setting.

Since the demand is high during the peak-hours in the scenario, gridlock situations occur. Those kinds of situations are unavoidable since there will be conflicts in the car following model. To make the simulation continue to run, SUMO has a teleporting option that is utilized in the original LuST scenario. The original LuST scenario is configured such that if a vehicle has been locked for more than 10 minutes, it will teleport along its route until there is free space. Without the possibility for the vehicles to teleport, gridlocks occur with the standard open-loop controller that comes with the scenario, so it is not just for this comparison it is needed to have the teleportation activated. It is therefore important when we evaluate the control strategies that we keep track of the number of teleports, to make sure that the control strategy will not create a significantly larger amount of gridlocks, compared to the original fixed-time controller. In Table 5.4 the number of teleports are reported for each controller. It is also reported how many of those teleports that are caused directly due to traffic jam, but one should have in mind that, e.g., a gridlock caused by that two vehicles want to swap lanes, is often a consequence of congestion.

The total travel time and the number of teleports for different choices of tuning parameters are shown in Table 5.4. For the fixed-time controller, we keep the standard fixed-time plan provided with the LuST scenario. Based on our conclusions from the Manhattan grid simulations, we choose to run the GPA controller for two different values of κ: 5 and 10. The values of $\bar{w}$ are chosen from 0 up to 0.50 with increments of 0.05. How the queue lengths vary with time is shown in Figure 5.9 for $\kappa = 5$ and in Figure 5.10 for $\kappa = 10$.

From the results, we can see that any controller with $\kappa = 10$ and $\bar{w}$ within the range of investigation will improve the traffic situation. However, the controller that

[2] https://cvxpy.org

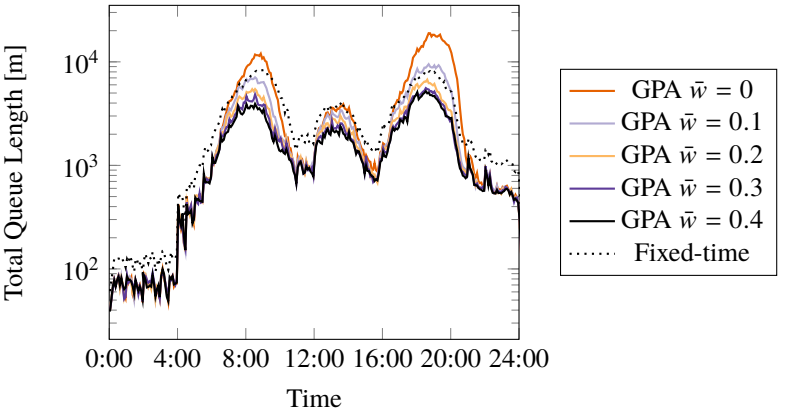


Figure 5.9 How the queue lengths vary with time when the traffic lights in the LuST scenario are controlled with the GPA controller and the standard fixed-time controller. For the GPA controller the parameters $\kappa = 5$ and different values of $\bar{w}$ are tested. In order to improve the readability of the results, the queue lengths are averaged over 300 seconds interval.

yields the overall shortest total travel time is the one with $\kappa = 5$ and $\bar{w} = 0.40$, which means that when we bound the cycle length, a lower κ may yield better performance. This result suggests that tuning the GPA only with respect to κ, and keeping $\bar{w} = 0$, may not lead to the best performance with respect to total travel time, although it gives higher throughput.

5.5 Conclusions

In this chapter, we performed evaluations of the Generalized Proportional Allocation controller in a microscopic traffic simulator. First, we showed two different discretization schemes of the controller: one scheme that ensures that all clearance phases are activated during one cycle, and one scheme that just activates the necessary clearance phases. We then performed simulations for two different scenarios. In the first scenario, the network topology was a Manhattan-like grid with artificial demands. The reason behind designing and testing in a scenario like this one is that it is easy to implement another well known decentralized traffic signal controller, the MaxPressure controller. For the investigated scenario, it can be concluded that the GPA controller performs better than the MaxPressure controller when the demands are low, while the MaxPressure performs better during high demands. Although the MaxPressure controller needs information about the turning ratios, it seems to be quite robust to deviant traffic flow behavior.

To evaluate the GPA controller in a more realistic scenario, simulations were

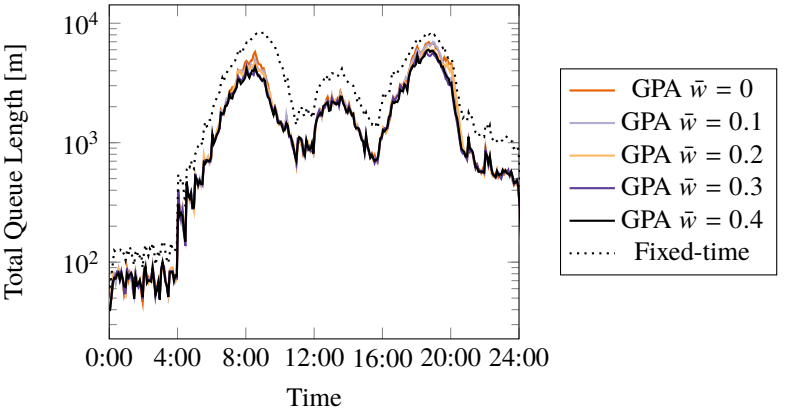


Figure 5.10 How the queue lengths vary with time when the traffic lights in the LuST scenario are controlled with the GPA controller and the standard fixed-time controller. For the GPA controller the parameters $\kappa = 10$ and different values of $\bar{w}$ are tested. In order to improve the readability of the results, the queue lengths are averaged over 300 seconds interval.

also done for a scenario covering the city of Luxembourg. In this case, it could be seen that the GPA controller outperforms the standard fixed-time controller included in the scenario.

Table 5.4 Comparison of the different control strategies - LuST scenario

	κ	$\bar{w}$	Teleports (jam)	Total Travel Time [h]
GPA	10	0	76 (6)	49 791
GPA	10	0.05	65 (1)	49 708
GPA	10	0.10	37 (0)	49 519
GPA	10	0.15	57 (19)	49 408
GPA	10	0.20	50 (10)	49 380
GPA	10	0.25	35 (0)	49 265
GPA	10	0.30	30 (0)	48 930
GPA	10	0.35	25 (1)	48 922
GPA	10	0.40	51 (0)	48 932
GPA	10	0.45	49 (5)	49 076
GPA	10	0.50	42 (15)	49 383
GPA	5	0	668 (76)	57 249
GPA	5	0.05	234 (62)	54 870
GPA	5	0.10	68 (10)	52 038
GPA	5	0.15	47 (9)	50 696
GPA	5	0.20	50 (6)	49 904
GPA	5	0.25	41 (3)	49 454
GPA	5	0.30	23 (0)	48 964
GPA	5	0.35	30 (1)	48 643
GPA	5	0.40	35 (5)	48 445
GPA	5	0.45	39 (1)	48 503
GPA	5	0.50	42 (10)	48 772
Fixed-time	–	–	122 (80)	54 103

6

Two-Tier Traffic Assignment

While we in the previous three chapters have focused on the control of traffic signals, we will in this and next chapter cover two different aspects of the routing problem. In this chapter, we will study the routing problem mentioned in Chapter 2, but for two different classes of vehicles.

This work in this chapter is motivated by the eventual introduction of autonomous vehicles onto roads, which will result in new possibilities for technological impact in traffic route planning. We will study the assignment of traffic where the network includes the operation of a large fleet of autonomous vehicles among many ordinary drivers. This situation may fast become a reality, as the tenth principle of the Shared Mobility Principles for Livable Cities [Chase, 2018] states:

> 10. We support that autonomous vehicles (AVS) in dense urban areas should be operated in only shared fleets.

The restriction to autonomous traffic to be operated solely by fleets would incentivize the need for novel methods of traffic control for use in networks with fleets.

In this chapter, we study a two-tier assignment problem, where one class of users seeks user optimality, i.e., minimizes their own traveling time in the network, while the other class of users seeks fleet optimality, i.e., minimizes the total travel time for the entire fleet. We provide simpler conditions for uniqueness of equilibrium than [Yang et al., 2007], where the problem is previously mentioned, and introduce two different algorithms to obtain the equilibrium, one centralized with proven convergence properties and the other decentralized. We also show the interconnection between the two-tier assignment problem and a multicommodity dynamical network flow model, proving stability properties in the case of acyclic networks. By having this linkage between the static assignment problem and the dynamical network flow problem, we allow for further research of feedback controllers in order to improve the robustness of the assignment.

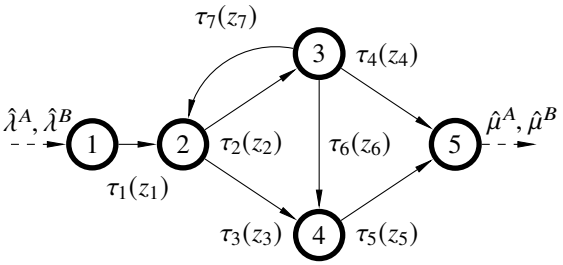


Figure 6.1 Example of a traffic network where delay functions are assigned to the cells.

6.1 The Two-Tier Assignment Problem

In this chapter, we consider a traffic network facilitating two types of vehicular traffic. The first type, which we denote Class A, consists of vehicles optimizing a user-optimal policy, i.e., vehicles minimizing their marginal delay. The second type, which we denote Class B, consists of vehicles optimizing a fleet-optimal policy, i.e., vehicles minimizing the total delay of the fleet. The flows of Class A and Class B vehicles along each cell are denoted by $z^A, z^B \in \mathbb{R}_+^{\mathcal{E}}$, and the exogenous inflows and outflows for Class A and Class B vehicles are given as $\hat{\lambda}^A, \hat{\mu}^A \in \mathbb{R}_+^{\mathcal{V}}$ and $\hat{\lambda}^B, \hat{\mu}^B \in \mathbb{R}_+^{\mathcal{V}}$, respectively, where the entries of $\hat{\lambda}^A$, $\hat{\mu}^A$, $\hat{\lambda}^B$, and $\hat{\mu}^B$ are nonnegative and are associated with the junctions of the network. We further assume that the exogenous flows are feasible, i.e., there exist $z^A, z^B \in \mathbb{R}^{\mathcal{E}}$ such that $Bz^A = \hat{\lambda}^A - \hat{\mu}^A$ and $Bz^B = \hat{\lambda}^B - \hat{\mu}^B$.

Definition 6.1
To each cell $i \in \mathcal{E}$ we associate a *delay function* $\tau_i : \mathbb{R}_+ \to \mathbb{R}_+$, mapping the flows z_i^A and z_i^B to a delay value. Delay functions are assumed to be twice continuously differentiable, strictly increasing, and nonnegative at 0. □

Class A and Class B vehicles optimize different policies. Class A vehicles minimize the amount of time taken for each vehicle to arrive at its destination given the current state of traffic. In this case, the assignment of traffic flow along each cell is given as the solution to the following optimization problem,

$$\underset{z^A \in \mathbb{R}_+^{\mathcal{E}}}{\text{minimize}} \quad g^A(z^A, z^B) \coloneqq \sum_{i \in \mathcal{E}} \int_0^{z_i^A} \tau_i(s + z_i^B) \mathrm{d}s\,, \tag{6.1a}$$

$$\text{subject to} \quad Bz^A = \hat{\lambda}^A - \hat{\mu}^A\,. \tag{6.1b}$$

Class B vehicles minimize the average amount of time taken for each vehicle to arrive at its destination given the current state of traffic. In this case, the assignment of traffic flow along each cell is given as the solution to the following optimization

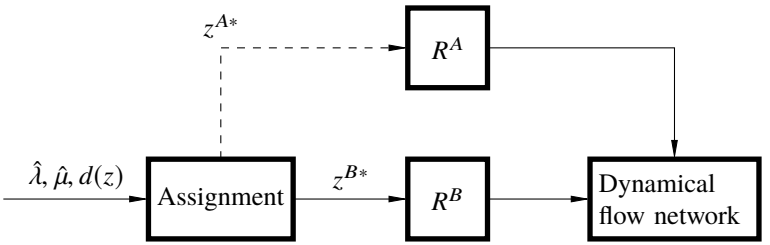


Figure 6.2 Conceptual schematic of the assignment and control. Observe that Class A vehicles are assumed to find the optimal assignment by their own.

problem,

$$\underset{z^B \in \mathbb{R}_+^{\mathcal{E}}}{\text{minimize}} \quad g^B(z^A, z^B) := \sum_{i \in \mathcal{E}} z_i^B \tau_i(z_i^A + z_i^B), \tag{6.2a}$$

$$\text{subject to} \quad Bz^B = \hat{\lambda}^B - \hat{\mu}^B . \tag{6.2b}$$

We study the assignment of traffic flow based on the solution to (6.1) and (6.2), the optimal behavior of vehicles when the dynamical traffic network is at equilibrium. Based on this assignment, it becomes necessary to route the traffic to achieve the desired flows. The assignment and routing scheme is presented in Figure 6.2. In the figure, the assignment block computes the optimal traffic flows $z^{A*}, z^{B*} \in \mathbb{R}^{\mathcal{E}}$ based on delay functions τ_i and exogenous inflows and outflows. Once the assignment is computed, the desired flows are achieved by employing a routing policy which routes the fleet of Class B vehicles, denoted R^B, in the presence of a routing policy of the opportunistic behavior of Class A drivers, denoted R^A.

6.2 Existence and Uniqueness of an Assignment

The solutions to the assignment problems in (6.1) and (6.2) depend on each other. Therefore it is of interest to investigate the properties of equilibrium flows, z^{A*} and z^{B*}, which solve both equations at the same time.

In this section, we provide sufficient conditions for the existence and uniqueness of an equilibrium which depend solely on the properties of the delay functions τ_i. The proof relies on results of [Rosen, 1965], which provides conditions for existence and uniqueness of equilibria in convex games. The optimization problems (6.1) and (6.2) are both convex under certain choices of the delay functions τ_i. Here, we present the requirements imposed on τ_i which imply existence and uniqueness. We begin with a sufficient condition for existence.

Proposition 6.1
Assume that the delay functions $\tau_i(z_i)$, as given in Definition 6.1, satisfies

$$2\tau_i'(z_i^A + z_i^B) + z_i^B \tau_i''(z_i^A + z_i^B) \geq 0, \tag{6.3}$$

for all $i \in \mathcal{E}$ and all feasible $z^A, z^B \in \mathbb{R}^{\mathcal{E}}$. Then there exist z^{A*} and z^{B*} such that z^{A*} is a solution to (6.1) given $z^B = z^{B*}$, and z^{B*} is a solution to (6.2) given $z^A = z^{A*}$.□

Proof. Theorem 1 of [Rosen, 1965] states that there exists a solution to any concave (resp. convex) game. A game is concave (resp. convex) if the cost functions are concave (resp. convex) and individual constraints of every strategy are convex, i.e., the constraints can be written in the form $h(z) \geq 0$ where every h is a convex function. Since the constraints are convex in both (6.1) and (6.2), all that is left to show is that the cost functions are both convex.

In the case of (6.1), the cost function is convex because τ_i is monotonically increasing for all $i \in \mathcal{E}$, implying that its integral is monotonically increasing and therefore is convex. In the case of (6.2), convexity follows from (6.3). □

Results for uniqueness are less straightforward to obtain than results for existence. We proceed by providing a sufficient condition for uniqueness, which is a generalization of Proposition 4 of [Yang et al., 2007], in which the authors require $\tau_i(z_i)$ to be affine and strictly monotone.

Proposition 6.2
The pair z^{A*} and z^{B*} of Proposition 1 is unique if τ_i satisfies the following relationship,

$$2\tau_i'(z_i^A + z_i^B) > z_i^B \tau_i''(z_i^A + z_i^B), \tag{6.4}$$

for all $i \in \mathcal{E}$ and feasible $z^A > 0$ and $z^B > 0$. □

Proof. Let $S(z^A, z^B, y^A, y^B) = g^A(z^A, y^B) + g^B(y^A, z^B)$ and let,

$$h(z^A, z^B) = \begin{bmatrix} \nabla_{z^A} g^A(z^A, z^B) \\ \nabla_{z^B} g^B(z^A, z^B) \end{bmatrix},$$

denote the pseudogradient of S. We let the cells in $\mathcal{E}$ be indexed $\{1, 2, \ldots, n\}$. Since τ_i is differentiable for any $i \in \mathcal{E}$, the pseudogradient of S is given by,

$$h(z^A, z^B) = \begin{bmatrix} \tau_1(z_1^A + z_1^B) \\ \vdots \\ \tau_n(z_n^A + z_n^B) \\ \tau_1(z_1^A + z_1^B) + z_1^B \tau_1'(z_1^A + z_1^B) \\ \vdots \\ \tau_n(z_n^A + x_n^B) + z_n^B \tau_n'(z_n^A + z_n^B) \end{bmatrix}.$$

Let H denote the Jacobian of S, which is equal to

$$H = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \tag{6.5}$$

where,

$$A = B = \operatorname{diag}\left(\tau_i'(z_i^A + z_i^B)\right)_{i\in\mathcal{E}},$$
$$C = \operatorname{diag}\left(\tau_i'(z_i^A + z_i^B) + z_i^B\tau_i''(z_i^A + z_i^B)\right)_{i\in\mathcal{E}},$$
$$D = \operatorname{diag}\left(2\tau_i'(z_i^A + z_i^B) + z_i^B\tau_i''(z_i^A + z_i^B)\right)_{i\in\mathcal{E}}.$$

Now let,

$$F = H + H^T = \begin{bmatrix} 2A & A+C \\ A+C & 2D \end{bmatrix} = \begin{bmatrix} 2A & D \\ D & 2D \end{bmatrix}.$$

According to [Rosen, 1965, Theorem 7], S is strictly diagonally convex if F is positive definite for all feasible $z^A > 0$ and $z^B > 0$. Furthermore, according to [Rosen, 1965, Theorem 3], if S is strictly diagonally convex then the flow equilibrium z^{A*} and z^{B*} is unique. We now show that F is positive definite by considering its Schur complement $\bar{F} = 2D + D^{\mathrm{T}}(2A)^{-1}D = 2D + (2A)^{-1}D^2$. The matrix F is then positive definite if and only if $\bar{F}$ and A are positive definite. The matrix A is positive definite because τ_i is strictly increasing. The matrix $\bar{F}$ is given by,

$$\begin{aligned}\bar{F} &= \operatorname{diag}\left(4\tau_i'(z_i^A + z_i^B) + 2z_i^B\tau_i''(z_i^A + z_i^B)\right. \\ &\qquad \left. - \frac{(2\tau_i'(z_i^A + z_i^B) + z_i^B\tau_i''(z_i^A + z_i^B))^2}{2\tau_i'(z_i^A + z_i^B)}\right)_{i\in\mathcal{E}} \\ &= \operatorname{diag}\left(2\tau_i'(z_i^A + z_i^B) - \frac{(z_i^B\tau_i''(z_i^A + z_i^B))^2}{2\tau_i'(z_i^A + z_i^B)}\right)_{i\in\mathcal{E}}.\end{aligned}$$

Clearly, the elements in $\bar{F}$ are positive when (6.4) is satisfied. □

We now consider the special case where the delay functions are given by an affine relation to the power of a positive exponent,

$$\tau_i(z_i) = (\alpha_i z_i + \beta_i)^{c_i}. \tag{6.6}$$

We introduce this form of delay function because it is useful in modeling the behavior of traffic flow in the free-flow regime, where the flows do not saturate. In the sequel, we will show how these delay functions can recover the free-flow regime of the fundamental diagram.

Corollary 6.1
Suppose the delay functions take the form in (6.6) with $\alpha_i, c_i > 0$ and $\beta_i \geq 0$ for all $i \in \mathcal{E}$. The solution to the assignment problem (6.1) – (6.2) exists and is unique if $c_i \leq 3$. □

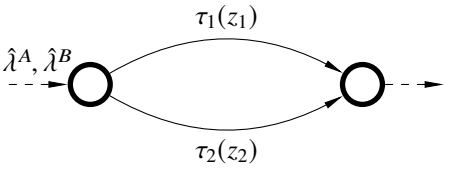


Figure 6.3 The network of Example 6.1.

Proof. Existence follows from the fact that τ_i satisfies the requirements imposed on the delay function and Proposition 1. To show uniqueness, we compute the first two derivatives,

$$\tau_i'(z_i) = \alpha_i c_i(\alpha_i z_i + \beta_i)^{c_i - 1},$$
$$\tau_i''(z_i) = \alpha_i^2 c_i(c_i - 1)(\alpha_i z_i + \beta_i)^{c_i - 2}.$$

Then the expression

$$\begin{aligned} &2\tau_i'(z_i^A + z_i^B) - z_i^B \tau_i''(z_i^A + z_i^B) \\ &\quad = \alpha_i c_i(\alpha_i(z_i^A + z_i^B) + \beta_i)^{c_i - 2}(2(\alpha_i(z_i^A + z_i^B) + \beta_i) - z_i^B \alpha_i(c_i - 1)) \\ &\quad = \alpha_i c_i(\alpha_i(z_i^A + z_i^B) + \beta_i)^{c_i - 2}(2\alpha_i z_i^A + (3 - c_i)\alpha_i z_i^B + 2\alpha_i \beta_i), \end{aligned}$$

is greater than 0 for all positive flows if $c_i \leq 3$. □

We have thus far shown that an equilibrium solution to the optimization problem (6.1) and (6.2) always exists and given conditions that guarantee this equilibrium to be unique. We now present a case in which a slight modification, in terms of a maximum flow capacity constraint, no longer guarantees uniqueness of the equilibrium solution. Specifically, if we introduce a capacity constraint on the flow of each cell, then uniqueness may not be guaranteed since capacity constraints are concave constraints.

Example 6.1
Consider the network consisting of two cells in Figure 6.3 with exogenous inflows $\hat{\lambda}^A = 1$, $\hat{\lambda}^B = 1$, and delay functions $\tau_1(z_1) = z_1$, $\tau_2(z_2) = 2z_2$. Impose a constraint on the flow capacity of cell 1 so that $z_1^A + z_1^B \leq 1$. Then $z^A = (0, 1)$, $z^B = (1, 0)$ is an equilibrium because the minimizer of

$$z_1^B \tau_1(z_1^B) + z_2^B \tau_2(1 + z_2^B) = \left(z_1^B\right)^2 + 2z_2^B(1 - z_2^B)$$

is $(1, 0)$ and the minimizer of the other optimization must satisfy

$$0 \leq z_1^A + z_1^B = z_1^A + 1 \leq 1,$$

implying that $z_1^A = 0$ and therefore $z^A = (0, 1)$. Furthermore, $z^A = (1, 0)$, $z^B = (0, 1)$ is also an equilibrium because the minimizer of

$$\int_0^{z_1^A} \tau_1(s)\mathrm{d}s + \int_0^{z_2^A} \tau_2(s+1)\mathrm{d}s = \frac{1}{2}\left(z_1^A\right)^2 + \left(z_2^A\right)^2 + 2z_2^A$$

is $(1, 0)$ and the minimizer of the other optimization must satisfy

$$0 \leq z_1^A + z_1^B = 1 + z_1^B \leq 1\,,$$

implying that $z_1^B = 0$ and therefore $z^B = (0, 1)$. □

6.3 Algorithms for Determining the Traffic Assignment

We propose two algorithms for determining the equilibrium flows z^{A*} and z^{B*}. The first is a centralized algorithm based on the idea that the optimal flow assignments for each class of vehicles are updated simultaneously in the constraint-admissible opposite direction of the gradient of the objective functions.

This is done by keeping the total flow in the system the same as an initial assignment, while making sure that the flow on each cell remains non-negative.

Centralized algorithm

We begin by defining the Lagrangian corresponding to the optimization problem (6.1),

$$L^A(z^A, z^B, \gamma^A, \xi^A) = -g^A(z^A, z^B) + (\gamma^A)^T(Bz^A - \hat{\lambda}^A + \hat{\mu}^A) + (\xi^A)^T z^A\,,$$

where $\gamma^A \in \mathbb{R}^{\mathcal{V}}$ and $\xi^A \in \mathbb{R}^{\mathcal{E}}$. We define L^B analogously,

$$L^B(z^A, z^B, \gamma^B, \xi^B) = -g^B(z^A, z^B) + (\gamma^B)^T(Bz^B - \hat{\lambda}^B + \hat{\mu}^B) + (\xi^B)^T z^B\,.$$

The algorithm we propose updates the flow assignments by following the gradients of the Lagrangians,

$$\dot{z}^A = \nabla_{z^A} L^A = -\nabla_{z^A} g^A(z^A, z^B) + B^{\mathrm{T}}\gamma^A + \xi^A =: f^A(z^A, z^B, \gamma^A, \xi^A), \tag{6.7}$$

$$\dot{z}^B = \nabla_{x^B} L^B = -\nabla_{z^B} g^B(z^A, z^B) + B^{\mathrm{T}}\gamma^B + \xi^B =: f^B(z^A, z^B, \gamma^B, \xi^B)\,. \tag{6.8}$$

The values of γ^A, γ^B, ξ^A, and ξ^B are chosen according to the optimization,

$$(\gamma^A, \xi^A) = \underset{\xi^A \geq 0}{\operatorname{argmin}} \|f^A(z^A, z^B, \gamma^A, \xi^A)\|^2\,, \tag{6.9}$$

$$(\gamma^B, \xi^B) = \underset{\xi^B \geq 0}{\operatorname{argmin}} \|f^B(z^A, z^B, \gamma^B, \xi^B)\|^2\,, \tag{6.10}$$

subject to the constraint that $\xi_i^A = 0$ and $\xi_i^B = 0$ whenever $z_i^A > 0$ and $z_i^B > 0$, respectively.

Proposition 6.3
Given feasible initial states $z^A(0)$, $z^B(0)$, i.e, $Bz^k(0) = \hat{\lambda}^k - \hat{\mu}^k$ and $z^k(0) \geq 0$ are satisfied for $k \in \{A, B\}$, and assume that the delay functions $\tau_i(x_i)$ stratifies the conditions in Proposition 6.1 and Proposition 6.2. Then algorithm (6.7)–(6.10) converges to the unique equilibrium of the two tier assignment problem (z^{A*}, z^{B*}), where z^{A*} is a solution to (6.1) given $z^B = z^{B*}$, and z^{B*} is a solution to (6.2) given $z^A = z^{A*}$ □

Proof. We begin by showing that as dynamics proceeds, the constraint will always be satisfied. This means that if we start the algorithm with a feasible solution, $z^k(t)$ will be feasible for all $t > 0$. Suppose that for $k \in \{A, B\}$, there exists $\tilde{z}^k$ such that $B\tilde{z}^k \neq \hat{\lambda}^k - \hat{\mu}^k$ or $\tilde{z}^k \ngeq 0$. By continuity of the solution and the assumption of feasible initialization, there must exist a point $\bar{z}^k$ where $B\bar{z}^k = \hat{\lambda}^k - \hat{\mu}^k$ and $\bar{z}^k \geq 0$, with corresponding Lagrange multipliers $\bar{\gamma}^k$ and $\bar{\xi}^k$, which are solutions to (6.9) or (6.10), and for which either $B\tilde{z}^k \neq \hat{\lambda}^k - \hat{\mu}^k$ or $\tilde{z}^k \ngeq 0$.

In the first case, it must hold that $B\dot{\bar{z}}^k \neq 0$. Observe that,

$$\left\|f^k(z^A, z^B)\right\|^2 = (\nabla_{z^k} g^k(z^A, z^B))^T \nabla_{z^k} g^k(z^A, z^B) + 2(\nabla_{z^k} g^k(z^A, z^B))^T (B^T \bar{\gamma}^k + \bar{\xi}^k) + 2(\bar{\xi}^k)^T B^T \bar{\gamma}^k + (\bar{\gamma}^k)^T BB^T \bar{\gamma}^k + (\bar{\xi}^k)^T \bar{\xi}^k .$$

Taking the partial derivative of $\left\|f^k(z^A, z^B)\right\|^2$ with respect to $\bar{\gamma}^k$ gives,

$$\frac{\partial}{\partial \bar{\gamma}^k} \left\|f^k(z^A, z^B)\right\|^2 = 2B(\nabla_{z^k} g^k(z^A, z^B) + \bar{\xi}^k + B^T \bar{\gamma}^k) = 2B\dot{\bar{z}}^k .$$

If $(2B\dot{\bar{z}}^k)_i > 0$ (respectively < 0) for some i, the norm $\left\|f^k(z^A, z^B)\right\|^2$ can be decreased by decreasing (respectively increasing) $\bar{\gamma}_i^k$. Therefore, $\bar{\gamma}^k$ cannot be a solution to (6.9) or (6.10) implying a contradiction. In the second case, it must hold that $\dot{\bar{z}}_i^k < 0$ for some i. Taking the partial derivative of $\left\|f^k(z^A, z^B)\right\|^2$ with respect to $\bar{\xi}^k$ gives,

$$\frac{\partial}{\partial \bar{\xi}^k} \left\|f^k(z^A, z^B)\right\|^2 = 2(\nabla_{z^k} g^k(z^A, z^B) + \bar{\xi}^k + B^T \bar{\gamma}^k) = 2\dot{\bar{z}}^k .$$

Since there exists i such that $\dot{\bar{z}}_i^k < 0$, the norm can be decreased by increasing $\bar{\xi}_i^k$, which contradicts that $\bar{\xi}^k$ is a solution to (6.9) or (6.10). From the above contradictions, it follows that the algorithm will always stay inside the feasible region.

To prove convergence, we will show that the time derivative of $\left\|f(z^A, z^B)\right\|^2$ is always negative, apart from when $f(z^A, z^B) = 0$. As a first step, we find an explicit solution to (6.9) and (6.10). From the necessary KKT conditions it follows that, for $k \in \{A, B\}$,

$$B(\nabla_{z^k} g^k(z^A, z^B) + \bar{\xi}^k + B^T \bar{\gamma}^k) = 0 . \tag{6.11}$$

Moreover, for any cell $i \in \mathcal{E}$ such that $\bar{\xi}_i^k > 0$, it must hold that

$$(\nabla_{z^k} g^k(z^A, z^B) + \bar{\xi}^k + B^T \bar{\gamma}^k)_i = 0 . \tag{6.12}$$

From the assumption that the graph $\mathcal{G}$ is connected, it follows that $\text{rank}(B) = |\mathcal{V}| - 1$. In order to solve the equations (6.11) and (6.12) we have to fix one element in γ^k for each k. For simplicity, let $\gamma_1^k = 0$. Moreover, let $\bar{B}$ be the matrix B with the first column removed. Then it follows that,

$$\gamma^k = \begin{bmatrix} 0 \\ -(\bar{B}\bar{B}^T)^{-1}\bar{B}(\nabla_{z^k} g^A(z^A, z^B) + \bar{\xi}^k) \end{bmatrix}.$$

Observe that,

$$f^k(z^A, z^B) = -\nabla_{z^k} g^k(z^A, z^B) + B\bar{\gamma}^k + \bar{\xi}^k,$$

and that,

$$\begin{bmatrix} \dot{f}^A(z^A, z^B) \\ \dot{f}^B(z^A, z^B) \end{bmatrix} = -H \begin{bmatrix} z^A \\ z^B \end{bmatrix} + \begin{bmatrix} B\bar{\gamma}^A + \bar{\xi}^A \\ B\bar{\gamma}^B + \bar{\xi}^B \end{bmatrix},$$

with H given as in (6.5). Taking the time-derivative of the norm yields,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|f(z^A, z^B)\right\|^2 = (f(z^A, z^B))^T \dot{f}(z^A, z^B) = -f^T H f + f^T \begin{bmatrix} B\bar{\gamma}^A + \bar{\xi}^A \\ B\bar{\gamma}^B + \bar{\xi}^B \end{bmatrix}.$$

The second term in the sum is zero due to the fact that $\bar{\gamma}$ satisfies (6.11) and $\bar{\xi}_i$ satisfies (6.12) for all i where $\bar{\xi}_i > 0$. Hence it holds that,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|f(z^A, z^B)\right\|^2 = -\frac{1}{2}(f(z^A, z^B))^T [H + H^T] f(z^A, z^B) \leq -\epsilon \left\|f(z^A, z^B)\right\|^2,$$

where the last inequality with $\epsilon > 0$ follows from the fact that $H + H^T$ is positive definite. Therefore, $\left\|f(z^A, z^B)\right\|^2 \to 0$ as $t \to \infty$ for all $\left\|f(z^A, z^B)\right\|^2 > 0$. □

Decentralized algorithm

We will now present a decentralized algorithm for obtaining traffic assignment equilibrium. The benefits of decentralized algorithms are several. Since traffic networks often are large-scale networks, we are able to decrease the amount of computations by solving part of the optimization locally in the network and restricting communication to the local neighborhood.

Furthermore, it imposes desired scalability problems to the algorithm. If a part of the topology of the network changes, only the algorithms associated with the neighboring areas need to be updated. The algorithm presented in the previous section is in general not decentralized due to the fact that γ in (6.11) is computed using $(\bar{B}^T\bar{B})^{-1}\bar{B}^T$, which is in general not limited to having entries between neighboring junctions. Furthermore, when one of the cells has zero flow, the ξ value associated with it may affect the computations of γ all over the network. For our proposed decentralized algorithm, we propose a dual descent scheme. We utilize the idea that, in order to compute the optimal flow in one cell $i \in \mathcal{E}$, we only need information about the Lagrange multipliers $\gamma^{A,B}$ associated with the tail junction i^- and head junction

i^+. The dynamics to update the Lagrange multiplier for each junction $v \in \mathcal{V}$ are then only dependent on the incoming flows $z_i^{A,B}$ for the cells $i \in \mathcal{E}_v$ and the outgoing flows $z_i^{A,B}$ for the cells in $\mathcal{E}_v^+$. The algorithm consists of solving the following,

$$\begin{aligned} z^A &= \operatorname*{argmin}_{z^A \geq 0} L^A(z^A, z^B, \gamma^A, 0), \\ z^B &= \operatorname*{argmin}_{z^B \geq 0} L^B(z^A, z^B, \gamma^B, 0), \\ \dot{\gamma}^A &= Bz^A - \hat{\lambda}^A + \hat{\mu}^A, \\ \dot{\gamma}^B &= Bz^B - \hat{\lambda}^B + \hat{\mu}^B. \end{aligned}$$

Since z^A and z^B depend only on the differences between different components in γ^A and γ^B, the system above will have multiple equilibria in γ^A and γ^B. These will be related to each other by a constant offset term so that if $(\tilde{\gamma}^A, \tilde{\gamma}^B)$ is an equilibrium, then $(\tilde{\gamma}^A + c\mathbb{1}_{n_e}, \tilde{\gamma}^B + d\mathbb{1}_{n_e})$ will be an equilibrium for any scalars c and d. This is not a limitation, however, due to the fact that the difference in Lagrange multipliers determines the flow.

The decentralized aspect of the algorithm can be seen by deriving the necessary conditions,

$$\begin{aligned} 0 &= \tau_i(z_i^A + z_i^B) + \gamma_{i^+}^A - \gamma_{i^-}^A, \\ 0 &= z_i^B \tau_i'(z_i^A + z_i^B) + \tau_i(z_i^A + z_i^B) + \gamma_{i^+}^B - \gamma_{i^-}^B, \end{aligned}$$

for all $i \in \mathcal{E}$. Defining $\tilde{\tau}^{-1}$ as,

$$\tilde{\tau}^{-1}(\gamma_i) = \begin{cases} 0 & \text{if } \gamma_i < \tau_i(0), \\ \tau_i^{-1}(\gamma_i) & \text{otherwise.} \end{cases}$$

Then,

$$\begin{aligned} z_i^A &= \max\left\{\tilde{\tau}_i^{-1}(\Gamma_i^A) - z_i^B, 0\right\}, \\ z_i^B &= \max\left\{\frac{\Gamma_i^B - \Gamma_i^A}{\tau_i'(\tilde{\tau}_i^{-1}(\Gamma_i^A))}, 0\right\}, \end{aligned}$$

where $\Gamma_i^k = \gamma_{i^+}^k - \gamma_{i^-}^k$ for $k \in \{A, B\}$, thus showing that updates of the flow are independent of nonadjacent cells.

Example 6.2
In the case when the delay functions are affine, i.e., $d_i(x_i) = \alpha_i x_i + \beta_i$, with $\alpha_i > 0$

and $\beta_i \geq 0$, the expressions introduced above become,

$$z_i^A = \max\left\{\frac{\Gamma_i^A - \beta_i}{\alpha_i} - x_i^B, 0\right\},$$
$$z_i^B = \max\left\{\frac{\Gamma_i^B - \Gamma_i^A}{\alpha_i}, 0\right\}. \qquad \square$$

6.4 Routing of Two-Tier Traffic

In this section, we study the stability of a routing scheme that statically allocates traffic to each node dependent on the solution to the assignment problem. To model the flow dynamics in the network, we introduce traffic density vectors $x^A, x^B \in \mathcal{X}$. A static relationship between the densities and the flows on each cell can be derived from delay functions, in the same way as proposed for one class of vehicles in [Como et al., 2013b]. Specifically, since the outflow on each cell is given by the traffic volume over the delay, it must hold that

$$x_i = \frac{z_i \tau_i(z_i)}{\ell_i},$$

where $\ell_i > 0$ is the length of the cell. Under the assumption that the supply function is not limiting the outflow in our model for dynamical flow networks, we have that $z_i = d_i(x_i)$, where $d_i(x_i)$ is the cell's demand function as described in Chapter 2. Using the above relationship, we get that,

$$\tau_i(z_i) = \begin{cases} \dfrac{d_i^{-1}(z_i)}{z_i} & \text{if } z_i > 0, \\ \dfrac{1}{d_i'(0)} & \text{if } z_i = 0. \end{cases}$$

We observe that the demand functions $d_i(x_i)$ are always strictly increasing, due to the fact that

$$d_i'(d_i^{-1}(z_i)) = \frac{\ell_i}{\tau_i(z_i) + z_i \tau_i'(z_i)} > 0.$$

Example 6.3
Let the delay function be given by $\tau_i(z_i) = \alpha_i z_i + \beta_i$. Then,

$$x_i = \frac{z_i(\alpha_i z_i + \beta_i)}{\ell_i},$$

and hence,

$$z_i = d_i(x_i) = \sqrt{\frac{1}{4}\left(\frac{\beta_i}{\alpha_i}\right)^2 + \frac{x_i \ell_i}{\alpha_i}} - \frac{1}{2}\frac{\beta_i}{\alpha_i}.$$

For all junctions with an exogenous inflow, we add a cell that acts as an on-ramp with the exogenous inflow instead. Then, combined with the static relationship between flow and density, the flow dynamics of the network can be described as

$$\dot{x}_i^A = \lambda_i^A + \sum_{j \in \mathcal{E}_v} \frac{x_j^A}{x_j} d_j(x_j) R_{j,i}^A - \frac{x_i^A}{x_i} d_i(x_i), \tag{6.13a}$$

$$\dot{x}_i^B = \lambda_i^B + \sum_{j \in \mathcal{E}_v} \frac{x_j^B}{x_j} d_j(x_j) R_{j,i}^B - \frac{x_i^B}{x_i} d_i(x_i). \tag{6.13b}$$

For a given flow assignment z^{k*}, the routing policies are given by,

$$R_{ji}^k := \begin{cases} \dfrac{z_i^{k*}}{\sum_{\ell \in \mathcal{E}_{i^-}^+} z_\ell^{k*}} & \text{if } \sum_{\ell \in \mathcal{E}_{i^-}^+} z_\ell^{k*} > 0 \text{ and } j^+ = i^-, \\ 0 & \text{otherwise.} \end{cases} \tag{6.14}$$

In the following proposition, we show that the dynamics (6.13) – (6.14) is globally asymptotically stable under the assumption that the network $\mathcal{G}$ is acyclic.

Proposition 6.4
Suppose the network $\mathcal{G}$ is acyclic. Then the dynamics given by (6.13) – (6.14) converges to the assigned equilibrium. □

Proof. Since the graph is acyclic, according to [Leiserson et al., 2001], a topological ordering exists among the junctions. Starting from the first junction $v_1 \in \mathcal{V}$, we first show that the aggregate traffic volumes of each cell x_i where $i \in \mathcal{E}_{v_1}^+$ converge if the inflows converge. Then we show that convergence in the aggregate traffic volume implies convergence of the traffic volume of each class.

Let

$$f_i(x_i, \hat{\lambda}^A(t), \hat{\lambda}^B(t)) = \lambda_i^A + \sum_{j \in \mathcal{E}_v} \frac{x_j^A}{x_j} d_j(x_j) R_{j,i}^A + \lambda_i^B + \sum_{j \in \mathcal{E}_v} \frac{x_j^B}{x_j} d_j(x_j) R_{j,i}^B - d_i(x_i),$$

for all cells $i \in \mathcal{E}_{v_1}^+$ where all $x_j^k(t)$ with $j \in \mathcal{E}_{v_1}$ and $k \in \{A, B\}$ are converging. For all $i \in \mathcal{E}_{v_1}^+$, we have that

$$\frac{\partial f_i}{\partial x_j} = 0, \quad \frac{\partial f_i}{\partial x_i} = -d_i'(x_i) < 0.$$

Hence the system is a monotone system with an equilibrium, so the assignment and convergence of the aggregate traffic volume on each cell can be assured [Lovisari et al., 2014, Lemma 6].

Since the aggregate density is converging, i.e., $x_i \to x_i^*$ for all $i \in \mathcal{E}_v^+$, there exists $\epsilon_i > 0$ such that for all $t \geq t^*$ it holds that

$$|x_i(t) - x_i^*| < \epsilon_i \,.$$

This implies that there also exists an ϵ_i' such that

$$|d_i(x_i(t)) - d_i(x_i^*)| < \epsilon_i',$$

and since the inflows are converging, there also exists $\epsilon_i'' > 0$ such that

$$\left| \sum_{j \in \mathcal{E}_v} \frac{x_j^k}{x_j} d_j(x_j) R_{j,i}^k - \sum_{j \in \mathcal{E}_v} \frac{x_j^{k*}}{x_j^*} d_j(x_j^*) R_{j,i}^k \right| \leq \epsilon_i'',$$

for all $t \geq t^*$. Hence it holds that,

$$\dot{x}_i^k \leq \lambda_i^k + \sum_{j \in \mathcal{E}_v} \frac{x_j^{k*}}{x_j^*} d_j(x_j^*) R_{j,i}^A + \epsilon_i'' - \frac{x_i^k}{x_i^* + \epsilon_i'} (d_i(x_i^*) - \epsilon_i),$$

and

$$\dot{x}_i^k \geq \lambda_i^k + \sum_{j \in \mathcal{E}_v} \frac{x_j^{k*}}{x_j^*} d_j(x_j^*) R_{j,i}^A - \epsilon_i'' - \frac{x_i^k}{x_i^* - \epsilon_i'} (d_i(x_i^*) + \epsilon_i) \,.$$

Direct use of Gronwall's Lemma yields that $x_i^k(t)$ converges.

We have thus shown convergence of the outflows from the node v_1, which means that node v_2 must have converging inflows. Similar reasoning to the above can be applied to show that the outflows from v_2 converge and the rest of the proof follows by the application of inductive reasoning. □

6.5 Numerical Example

In this section, we perform numerical simulations in order to test the schemes proposed in this chapter against a network with dynamics given in (6.13). The network, shown in Figure 6.1, is similar to the one used to illustrate Braess paradox, but with one additional cell. With the right choice of parameters, the user- and fleet-optimal assignments will be different. Moreover, the optimal assignment in the first node will depend on the delay functions in other parts on the network, rather than just the outgoing cells from the first node. In order to test if our algorithm can handle cycles, we have added an extra cell to the Braess network.

Let the delay functions be affine, given by $\tau_i(z_i) = \alpha_i z_i + \beta_i$ for all cells with the values of α_i and β_i specified in Table 6.1. We let $\hat{\lambda}^A = \hat{\mu}^A = 1$ and $\hat{\lambda}^B = \hat{\mu}^B = 4$ and, for simplicity, we let all the cells be of unit length, i.e., $\ell_i = 1$ for all cells. The results of simulations using the centralized and decentralized algorithm are shown

Table 6.1 Parameters and assigned flows for simulations

e_i	α_i	β_i	z_i^{A*}	z_i^{B*}
e_1	1	1	1	4
e_2	1	1	0.25	2.5
e_3	2	1	0.75	1.5
e_4	1	3	0.125	2.25
e_5	1	1	0.875	1.75
e_6	2	1	0.125	0.25
e_7	1	2	0	0

in Figure 6.4. In both simulations, the dynamics of the algorithms are simulated by using an Euler solver with a step length of 0.1. While the decentralized algorithm does not require a start from a feasible solution, i.e., the initial flows z^A and z^B can be any non-negative value, the centralized algorithm needs to start from a feasible solution because the algorithm has no information about exogenous inflows.

The routing polices R^A and R^B of (6.14) are determined according to the desired assignment. Figure 6.5 presents the outflows of each class of vehicles from each cell. Two simulations are performed. The first simulation corresponds to setting all initial densities to zero, i.e., $x_i^A(0) = x_i^B(0) = 0$ for all $i \in \mathcal{E}$. The second simulation corresponds to setting all initial densities to 5, i.e., $x_i^A(0) = x_i^B(0) = 5$ for all $i \in \mathcal{E}$. In the first simulation, cell e_7 can be removed from the network without causing any effect on the dynamics. This is because $R_{2,7} = 0$ for $k \in \{A, B\}$, thus making the network equivalent to an acyclic network with proven convergence properties according to Proposition 6.4. In the second simulation, cell e_7 contributes with a converging inflow to node v_1 but has no inflow to itself. Therefore it can be seen as a converging inflow to node v_1 and the proof of Proposition 6.4 can be slightly modified to show convergence in this case as well.

6.6 Conclusions

In this chapter, we have analyzed the routing assignment problem for the two classes of vehicles, where vehicles in each class either follow user-optimal or fleet-optimal paths. For the assignment problem, we have provided sufficient conditions for the existence and uniqueness of an equilibrium of the assignment, as well as two algorithms to compute it. We have also showed how the assignment may be achieved using static routing.

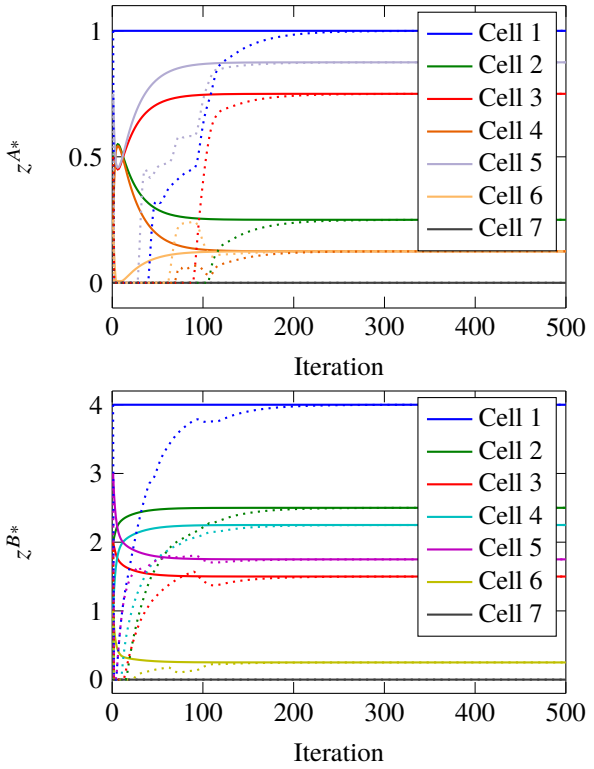


Figure 6.4 The assigned equilibria z^{A*} and z^{B*} computed by centralized (solid) and decentralized (dotted) algorithms, plotted as a function of algorithm iteration.

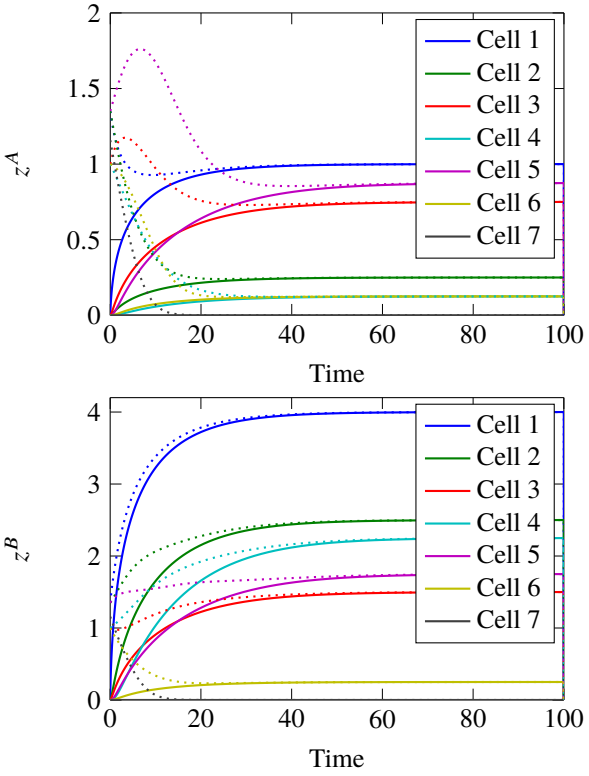


Figure 6.5 The evolution of the outflows $z^A(t)$ and $z^B(t)$ for first (solid) and second (dotted) simulations.

7

Resilience of Dynamically Routed Multicommodity Flows

In this chapter, we extend the model in Chapter 2 in two ways. First we allow the model to contain several classes of vehicles, a so called multicommodity network, where each class has its own routing preferences. This extension was already partly done in the end of the previous chapter, where we modeled two different classes of flows. Moreover, we will enhance the routing to be dynamic, so that each class of vehicles will adjust their path in the network, depending upon the current congestion levels in the cells. Situations like this can for example occur when commuters, not using a route guidance application, have to avoid a congested route due to an accident or roadworks.

Models for multicommodity flow networks based on PDEs and the celebrated LWR model have been studied [Lebacque and Khoshyaran, 2002; Herty et al., 2008], but solutions are usually difficult to obtain even in simple settings. In this chapter, we propose and analyze a model for dynamical flow networks with heterogeneous routing. Differently from single-commodity scenario, in which all vehicles belong to the same class and hence there is no competition among different classes, multi-commodity networks show a complex behavior even in the static setting [Leighton and Rao, 1999], in which it has been shown that the maximum flow that can flow in a multicommodity network is bounded away from the value predicted by the celebrated max-flow min-cut theorem. The main results of this chapter are the following: 1) Under certain assumptions on the constant inflows in the network, the network admits a globally asymptotically stable equilibrium for each commodity, and 2) When the network is not single-commodity, it can be extremely fragile with respect to perturbations. In particular, if a network is at equilibrium is perturbed, e.g,, the flow capacities of some cells decrease, a cascade effect can be triggered that destroys the stability of the network. This behavior arises in a multicommodity setting with heterogeneous routing only and has no counterpart in single-commodity networks.

7.1 A Motivating Example

Let us start with the case when only one commodity is present in the network, and introduce the dynamical routing.

Let us consider the network displayed in Figure 7.1. First, we focus on single-commodity dynamical flows, where the traffic volume in each cell $i \in \mathcal{E}$ is described by the dynamics in (2.2). We will in this chapter assume that the demand functions are always limiting the outflow from the cells, i.e., $s_i(x_i) = +\infty$ for all $i \in \mathcal{E}$. We let the demand functions be of the form

$$d_i(x_i) = c_i(1 - e^{-x_i}),$$

where $c_i > 0$ is the cell's maximal outflow capacity and x_i the cell's traffic volume.

While the routing matrix R previously has been assumed to be static, we will now let it be state-dependent. This means that, under the assumption that $d_i(0) = 0$, the dynamics in (2.2) can be rewritten as

$$\dot{x} = \lambda + (R^T(x) - I)d(x),$$

where the notation $R(x)$ stresses out that the routing now will be dependent on the state. The routing polices are constructed as

$$R_{1,2}(x_1, x_2) = \frac{e^{-x_1}}{e^{-x_1} + e^{-x_2}}, \qquad R_{3,4}(x_4, x_5) = \frac{e^{-x_4}}{e^{-x_4} + e^{-x_5}},$$
$$R_{1,3}(x_1, x_2) = \frac{e^{-x_2}}{e^{-x_1} + e^{-x_2}}, \qquad R_{3,5}(x_4, x_5) = \frac{e^{-x_5}}{e^{-x_4} + e^{-x_5}},$$

and $R_{4,5} = 1$. We let the exogenous inflow to each cell be zero, apart from $\lambda_1 = 2$. Moreover, we let the cells' flow capacity, $c_i = 2$, for all cells in the network apart from cell e_5, which has capacity $c_5 = 0.7$. In this setting, it can be verified that the dynamical flow network admits an equilibrium with corresponding flow vector z^* whose entries are specified in Figure 7.1. It has already been shown in [Como et al., 2013a] that the equilibrium is globally asymptotically stable.

We want to study how the equilibrium flows change when the network is perturbed, namely, when the flow capacity is reduced from c_i to $\tilde{c}_i < c_i$ for some cells $i \in \mathcal{E}$, and denote the corresponding perturbed demand functions $\tilde{d}$, such that

$$\tilde{c}_i = \sup_{x_i \geq 0} \tilde{d}_i(x_i),$$

for all cells $i \in \mathcal{E}$. We will define the network's margin of resilience to be the infimum aggregate flow capacity reduction $\sum_{i \in \mathcal{E}}(c_i - \tilde{c}_i)$, or *perturbation magnitude*, such that the perturbed system

$$\dot{x} = \lambda + (R^T(x) - I)\tilde{d}(x),$$

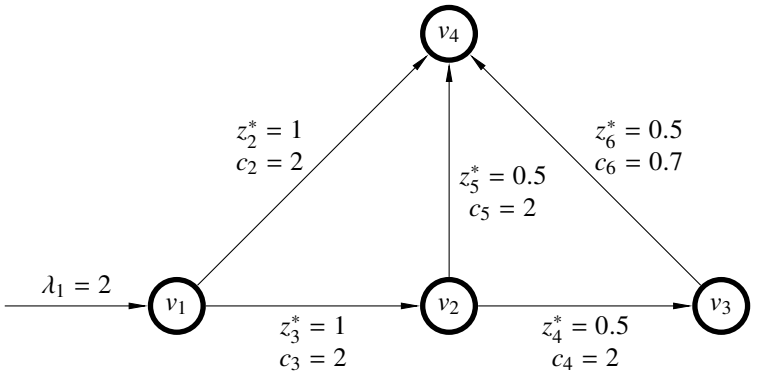


Figure 7.1 A single-commodity network. The minimum residual capacity 0.2 is achieved at junction v_3. Hence, under any perturbation of magnitude smaller than 0.2 the network is still able to transfer the external inflow λ_1 to the junction v_4.

is unstable, i.e., the traffic volume goes to infinity for at least one cell $i \in \mathcal{E}$ when time goes to infinity. For a single-commodity network, it was shown that the margin of resilience equals the minimum junction residual capacity [Como et al., 2013a; Como et al., 2013b], where the junction residual capacity is the difference between the equilibrium flow and the maximum capacity for each outgoing cell from a junction. This implies that the network in Figure 7.1, with the given routing policies, can handle any capacity perturbation of a magnitude smaller than 0.2.

Now, let us assume that there are two different classes, denoted A and B, of flows present in the network. Our state-space will then consist of traffic volumes for both the classes, and we will denote those volumes $x^A, x^B \in \mathcal{X}$. The aggregate traffic volume x_i in each cell $i \in \mathcal{E}$ is then given by

$$x_i = x_i^A + x_i^B \,.$$

We assume that the vehicles are fully mixed, so that the dynamics for vehicles of class $k \in \{A, B\}$ is

$$\dot{x}_i^k = \lambda_i^k + \sum_j R_{ji}^k(x) \frac{x_j^k}{x_j} d_j(x_j) - \frac{x_i^k}{x_i} d_i(x_i) \,.$$

We let the classes of vehicles have different routing policies. In particular, we

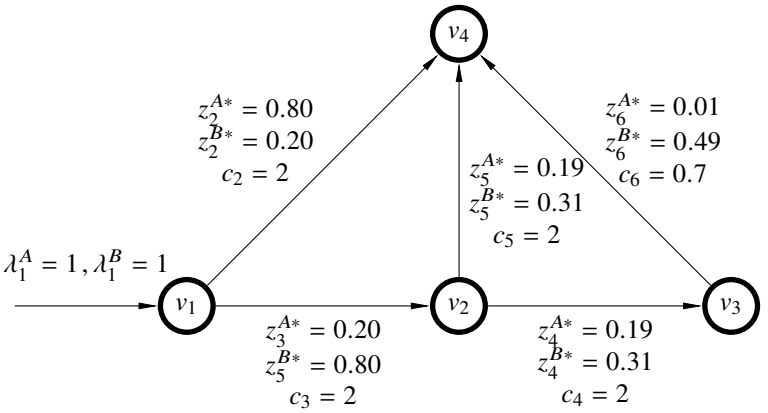


Figure 7.2 Equilibrium flows for a simple network. All flows are less than the cells' capacities, so that the network is able to fully transfer the external inflows λ_A, λ_B to their destination, junction v_4.

consider routing polices of the form

$$R^k_{1,2}(x_2, x_3) = 1 - R^k_{1,3}(x_2, x_3) = \frac{\bar{z}_2^k \cdot e^{-\alpha_2^k x_2}}{\bar{z}_2^k \cdot e^{-\alpha_2^k x_2} + \bar{z}_3^k \cdot e^{-\alpha_3^k x_3}},$$

$$R^k_{3,4}(x_4, x_5) = 1 - R^k_{3,5}(x_4, x_5) = \frac{\bar{z}_4^k \cdot e^{-\alpha_4^k x_4}}{\bar{z}_4^k \cdot e^{-\alpha_4^k x_4} + \bar{z}_5^k \cdot e^{-\alpha_5^k x_5}},$$

and $R^A_{4,6}(x) = R^B_{4,6}(x) \equiv 1$. Here, $\bar{z}_i^k$ is the desired limit flow for class k in cell i, and are specified in Figure 7.2. Observe that the aggregate limit flows coincide with those in the single-commodity case. On the other hand, $\alpha_i^k > 0$ are parameters which do not effect the limit flows. However, these parameters do affect how the dynamical flow network responds to perturbations.

In order to illustrate the fragility of the multicommodity setting, we let $\alpha_2^A = \alpha_3^B = 1000$ and $\alpha_2^B = \alpha_3^B = 1$ and $\alpha_4^k = \alpha_5^k = 0.01$, and consider now a perturbation of magnitude 0.01 which reduces $c_2 = 2$ to $\tilde{c}_2 = 1.99$. The limit flows for the perturbed dynamics are shown in Figure 7.3. The perturbation causes the limit flow in cell 3 to increase and exceed the capacity of the subsequent cell 5. Consequently, the traffic volume in cell 5 grows unbounded. This implies that the margin of resilience in the multicommodity case is not larger than 0.01. This example then indicates that a dynamical multicommodity network can be much more fragile than a single-commodity network with the same topology and same aggregate equilibrium flows.

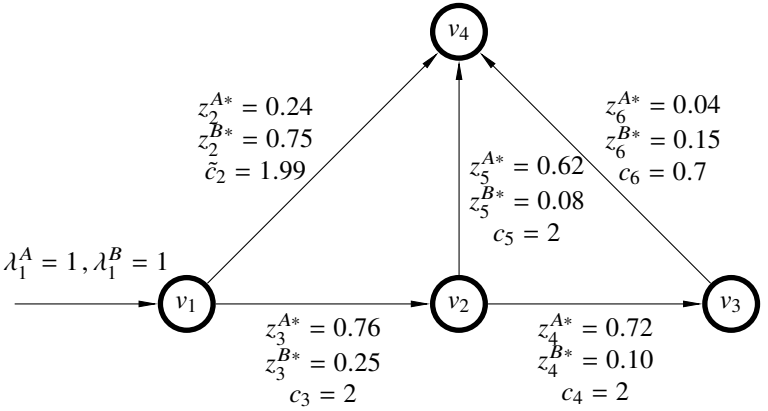


Figure 7.3 The same network when cell e_2's capacity is slightly decreased. Now the inflow to cell e_6 is larger than its capacity, and hence the network is not able to handle the flow demands.

7.2 Flow Networks with Dynamical Heterogeneous Routing

We will in this chapter study different classes of vehicles and we will denote the set of classes $\mathcal{K}$. It is assumed that each class of vehicles $k \in \mathcal{K}$ has a given destination junction[1], which is denoted $v^k \in \mathcal{V}$. For each class of vehicles, we will denote the subset of cells that the class has access to $\mathcal{E}^k \subseteq \mathcal{E}$. In the same manner, we let $\mathcal{E}_v^{k+}$ denote the subset of outgoing cells from junction v that class k has access to. This restriction is introduced for modeling purposes since we do not want the vehicles to end up in cells from where it is possible for the vehicles to reach their destinations. In practice, this situation can occur when heavier vehicles such as trucks may not be allowed to enter all roads in a city. For a subset of classes $\mathcal{J} \subset \mathcal{K}$, we denote the cells that the classes in $\mathcal{J}$ have access to as $\mathcal{E}_v^{\mathcal{J}+} = \cup_{k\in\mathcal{J}} \mathcal{E}_v^{k+}$. The traffic volume of each vehicle-class in each cell will now be the state-space, so the overall state-space from Chapter 2 is now extended to $\mathcal{X}^{\mathcal{K}}$. We will denote the traffic volume for a given class $k \in \mathcal{K}$ as $x^k \in \mathcal{X}$. In the same way, the exogenous static inflows and routing matrices for each class are denoted as $\lambda^k \in \mathbb{R}_+^{\mathcal{E}}$ and $R^k \in \mathbb{R}_+^{\mathcal{E}\times\mathcal{E}}$, where $\lambda_i^k = 0$ for all $i \in \mathcal{E} \setminus \mathcal{E}^k$ and $R_{ij}^k = 0$ for all $i \in \mathcal{E}$ and all $j \in \mathcal{E} \setminus \mathcal{E}^k$.

Throughout the chapter, we will make the following assumption about the demand functions:

Assumption 7.1—Demand functions
For each cell $i \in \mathcal{E}$ the demand function $d_i : \mathbb{R}_+ \to \mathbb{R}_+$ is assumed to

[1] This assumption is not restricting the model, since if a class should have the possibility to end up at several different junction, all those junctions can be connected to a fictive destination junction

be strictly increasing, continuously differentiable, with bounded derivative, and $c_i = \sup_{x_i \geq 0} d_i(x_i) < +\infty$. □

A class of vehicles k passing through a junction v splits among its accessible out-links $\mathcal{E}_v^{k+}$. Following [Como et al., 2013a; Como et al., 2013b], we allow for routing through distributed routing policies that are functions of the local aggregate traffic volumes $x^{(v+)} := \{x_i\}_{i \in \mathcal{E}_v^+} \in \mathbb{R}_+^{\mathcal{E}_v^+}$ on the outgoing cells from the junction v. The key novelty with respect to [Como et al., 2013a; Como et al., 2013b] is that we allow the different classes of vehicles to have different routing policies. We define distributed routing policies formally as follows:

Definition 7.1—Distributed routing policy
A distributed routing policy is a family of differentiable functions

$$\mathcal{R} := \{R_{ij}^k(x) : \mathbb{R}_+^{\mathcal{E}} \to \mathbb{R}_+\}_{k \in \mathcal{K}, i, j \in \mathcal{E}},$$

satisfying the following properties for all classes $k \in \mathcal{K}$ and all cells $i \in \mathcal{E}^k$:

a) The routing policy adheres the network topology,

$$R_{ij}^k(x) \equiv 0, \quad \forall j \in \mathcal{E} \setminus \mathcal{E}_{i^+}^{k+}.$$

b) The routing policy only depends on local information,

$$\frac{\partial}{\partial x_\ell} R_{ij}^k(x) \equiv 0, \quad \forall j \in \mathcal{E}_{i^+}^{k+}, \forall \ell \in \mathcal{E} \setminus \mathcal{E}_{i^+}^{k+}.$$

c) The routing policy is mass-conserving:

$$\sum_{j \in \mathcal{E}_{i^+}^{k+}} R_{ij}^k(x) \equiv \begin{cases} 1 & \text{if } i^+ \neq v^k, \\ 0 & \text{otherwise.} \end{cases}$$

d) The routing policy is congestion avoiding:

$$\frac{\partial}{\partial x_\ell} R_{ij}^k(x) \geq 0, \quad \forall \ell \neq j \in \mathcal{E}_{i^+}^{k+}.$$

e) The routing policy does not route traffic into congested cells. This means that for every nonempty proper subset $\mathcal{I} \subsetneq \mathcal{E}_{i+}^{k+}$ there exists a continuously differentiable family of functions $\bar{\mathcal{R}} := \{\bar{R}_{ij}^k(x) : \mathbb{R}_+^{\mathcal{E}} \to \mathbb{R}_+\}_{k \in \mathcal{K}, i, j \in \mathcal{E}}$ such that

$$\sum_{j \in \mathcal{E}_{i^+}^{k+}} \bar{R}_{ij}^k(x) \equiv 1, \quad \text{if } i^+ \neq v^k,$$

and such that if

$$x_j \to +\infty, \quad \forall j \in \mathcal{E}_{i^+}^{k+} \setminus \mathcal{I}, \quad x_\ell \to x_\ell^{\mathcal{I}}, \quad \forall \ell \in \mathcal{I},$$

then

$$R_{ij}^k(x) \to 0, \quad \forall j \in \mathcal{E}_{i^+}^{k+} \setminus \mathcal{I},$$
$$R_{i\ell}^k(x) \to \bar{R}_{i\ell}^k(x), \quad \forall \ell \in \mathcal{I}.$$

Having defined the distributed routing policies, we can now formally define dynamical multicommodity networks with distributed routing.

Definition 7.2—Dynamical multicommodity flow network with distributed routing
A dynamical multicommodity flow network is a dynamical flow network with distributed routing, with classes of vehicles $\mathcal{K}$ and a family of distributed routing policies satisfying Definition 7.1. The dynamics for each cell $i \in \mathcal{E}$ and each class $k \in \mathcal{K}$ is given by

$$\dot{x}_i^k = \lambda_i^k + \sum_{j \in \mathcal{E}} z_j^k R_{ji}^k(x) - \frac{x_i^k}{x_i} d_i(x_i).$$

The most important aim of a dynamical flow network is that all vehicles will reach their destinations. Therefore, we define a fully transferring network as:

Definition 7.3—Fully transferring
A dynamical flow network is said to be fully transferring if

$$\liminf_{t \to \infty} \sum_{i \in \mathcal{E}_v, k} z_i^k(t) = \sum_{i \in \mathcal{E}} \lambda_i^k, \qquad \forall\, k \in \mathcal{K}. \qquad \square$$

The definition above states that the outflow to the destination junctions (which is only one junction for each class $k \in \mathcal{K}$) should equal the exogenous inflows in the limit.

7.3 Stability Analysis

In this section we will state a sufficient condition for an acyclic dynamical multicommodity network to have finite cell volumes in equilibrium and a unique limit flow. First of all, we analyze a *local network* consisting of only the outgoing cells from one junction, as shown Figure 7.4, namely a network with a single junction. For in-coming cells $i \in \mathcal{E}_v$, we assume that they all have converging outflows, which we denote $y \in \mathbb{R}_+^{\mathcal{E}_v \times \mathcal{K}}$, such that

$$\lim_{t \to +\infty} y_j^k(t) = y_j^{k*}, \forall j \in \mathcal{E}_v, \forall k \in \mathcal{K}.$$

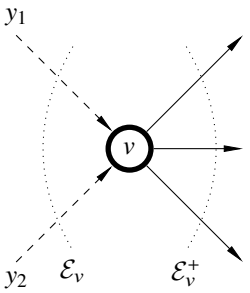


Figure 7.4 An example of a local network. In a local network, only the dynamics on the outgoing cells from a junction is studied. The set of those cells are denoted $\mathcal{E}_v^+$, while the incoming cells in $\mathcal{E}_v$ are assumed to only have converging outflows, which we denote y.

For a local network, i.e., for a fixed junction $v \in \mathcal{V}$, the multicommodity dynamics reads

$$\dot{x}_i^k = \lambda_i^k + \sum_{j \in \mathcal{E}} y_j^k R_{ji}^k(x^{(v+)}) - \frac{x_i^k}{x_i} d_i(x_i), \quad \forall i \in \mathcal{E}_v^+, k \in \mathcal{K}. \tag{7.1}$$

The next result, which is the building block of the whole theory, offers a necessary and sufficient condition for the local network to admit a globally asymptotically equilibrium.

Theorem 7.1
Consider a local dynamical multicommodity network with a junction $v \in \mathcal{V}$ as in (7.1). Assume moreover that the inflows are converging, namely $\lim_{t\to+\infty} y_j^k(t) = y_j^{k*}$ for all $k \in \mathcal{K}$ and all $j \in \mathcal{E}_v$. Then it holds that

a) if

$$\sum_{k\in\mathcal{J}} \sum_{j\in\mathcal{E}} y_j^{k*} < \sum_{i\in\mathcal{E}_v^{\mathcal{J}+}} \left(c_i - \sum_{k\in\mathcal{J}} \lambda_i^k \right),$$

for every nonempty $\mathcal{J} \subseteq \mathcal{K}$, then there exists a finite traffic volume x^* such that $\lim_{t\to\infty} x_i^k(t) = x_i^{k*}$ for every $i \in \mathcal{E}_v^+$ and $k \in \mathcal{K}$.

b) if there exists a nonempty $\mathcal{J} \subseteq \mathcal{K}$ such that

$$\sum_{k\in\mathcal{J}} \sum_{j\in\mathcal{E}} y_j^{k*} \geq \sum_{i\in\mathcal{E}_v^{\mathcal{J}+}} \left(c_i - \sum_{k\in\mathcal{K}} \lambda_i^k \right),$$

then there exists at least one $k \in \mathcal{J}$ such that $\lim_{t\to+\infty} x_i(t) = +\infty$ for all $i \in \mathcal{E}_v^{k+}$. □

Proof of Theorem 7.1

The proof of Theorem 7.1 is divided into several steps. First, we prove convergence of the aggregate flow on each link in the local network. While the aggregate flow will always converge, the aggregate traffic volume in each cell may not converge if the outflow converges to the flow capacity. If the converging outflow is less than the maximum flow capacity, we show that the aggregate traffic volumes converge too.

Lemma 7.1
Suppose that the inflows to the local network are constant, i.e., $y_j^k(t) \equiv y_j^k$ for all $k \in \mathcal{K}$ and all $j \in \mathcal{E}_v$, the dynamics is given by (7.1) and the demand-functions satisfies Assumption 7.1. Then there exists a unique aggregate limit outflow $z \in \mathbb{R}_+^{\mathcal{E}_v^+}$, such that for every initial traffic volumes $x(0) \in \mathbb{R}_+^{\mathcal{E}_v^+}$,

$$\lim_{t\to\infty} z_i(t) = z_i^*, \quad \forall i \in \mathcal{E}_v^+ .$$

Moreover, if $z_i^* < c_i$ or a cell $i \in \mathcal{E}_v^+$, then also the limit traffic volume x_i^* is unique.□

Proof. Consider the aggregate dynamics for a junction $v \in \mathcal{V}$, given by

$$\dot{x}_i = \sum_{k\in\mathcal{K}} \dot{x}_i^k = \sum_{k\in\mathcal{K}} \lambda_i^k + \sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} y_j^k R_{ji}^k(x^{(v+)}) - d_i(x_i) .$$

Let

$$H_i(x^{(v+)}) = \frac{\sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} y_j^k R_{ji}^k(x^{(v+)})}{\sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} y_j^k}.$$

Then $H_i(x^{(v+)})$ is a distributed routing policy. The aggregate dynamics for the local network can the be written as

$$\dot{x}_i = \sum_{k\in\mathcal{K}} \lambda_i^k + \left(\sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} y_j^k\right) H_i(x^{(v+)}) - d_i(x_i), \quad \forall i \in \mathcal{E}_v^+. \tag{7.2}$$

The aggregate dynamics then have the same properties as in the single commodity case, and therefore [Como et al., 2013a, Lemma 2] can be used to show that there exists a unique aggregate limit flow, z_i^*, for all $i \in \mathcal{E}_v^+$. Due to Assumption 7.1, there is a bijection between the aggregate flow and the aggregate traffic volume on each link when the limit flow is less than the capacity, and therefore the aggregate traffic volumes will converge as well. □

Next, a condition for having finite aggregate traffic volumes will be stated.

Lemma 7.2
A sufficient and necessary condition for a local network to have aggregate finite limit traffic volumes when the inflows $y^* \in \mathbb{R}_+^{\mathcal{E}_v\times\mathcal{K}}$ is constant, is that for every subset

$\mathcal{J} \subseteq \mathcal{K}$ it holds that

$$\sum_{k\in\mathcal{J}}\sum_{j\in\mathcal{E}_v} y_j^{k*} < \sum_{i\in\mathcal{E}_v^{\mathcal{J}+}}\left(c_i - \sum_{k\in\mathcal{J}}\lambda_i^k\right). \tag{7.3}$$

□

Proof. *Sufficiency*: Let $\mathcal{I} := \{i \in \mathcal{E} \mid x_i^* = +\infty\}$. Observe then that

$$\limsup_{t\to\infty} \dot{x}_i(t) \geq 0, \quad \forall i \in \mathcal{I}.$$

For a commodity $k \in \mathcal{K}$ two scenarios are possible. Either $\mathcal{E}_v^{k+} \not\subseteq \mathcal{I}$, but then the property e) in Definition 7.1 gives that $R_{ji}(x^{(v*)}) = 0$ for all cells $i \in \mathcal{I}$ and those classes of vehicles are not contributing to the infinite limit traffic volume. In the other case, we have that $\mathcal{E}_v^{k+} \subseteq \mathcal{I}$. Introduce the set $\mathcal{J} := \{k \in \mathcal{K} \mid \mathcal{E}_v^{k+} \subseteq \mathcal{I}\}$ and sum up equation (7.1) over the classes in $\mathcal{J}$,

$$\begin{aligned}\limsup_{t\to+\infty}\sum_{k\in\mathcal{J}}\sum_{i\in\mathcal{E}_v^{\mathcal{J}+}}\dot{x}_i(t) &= \limsup_{t\to+\infty}\sum_{k\in\mathcal{J}}\sum_{i\in\mathcal{E}_v^{\mathcal{J}+}}\left(\lambda_i^k + \sum_{j\in\mathcal{E}_v} y_j^{k*}R_{ji}(x^{(v)}) - \frac{x_i^k}{x_i}d_i(x_i)\right)\\ &= \sum_{k\in\mathcal{J}}\sum_{j\in\mathcal{E}_v} y_j^{k*} - \sum_{i\in\mathcal{E}_v^{\mathcal{J}+}}\left(c_i - \sum_{k\in\mathcal{J}}\lambda_i^k\right) \geq 0.\end{aligned}$$

In the equality we use the fact that the image of the routing policy is a simplex, together with the fact that the aggregate flow is bounded. Then inequality (7.3) is violated. Hence $\mathcal{I} = \emptyset$.

Necessity: Let $\mathcal{J} \subseteq \mathcal{K}$ be a nonempty subset such that

$$\sum_{k\in\mathcal{J}}\sum_{j\in\mathcal{E}_v} y_j^{k*} \geq \sum_{i\in\mathcal{E}_v^{\mathcal{J}+}}\left(c_i - \sum_{k\in\mathcal{J}}\lambda_i^k\right).$$

Then

$$\begin{aligned}\sum_{k\in\mathcal{J}}\sum_{i\in\mathcal{E}_v^{\mathcal{J}+}}\dot{x}_i^k &= \sum_{k\in\mathcal{J}}\left(\sum_{j\in\mathcal{E}_v} y_j^{k*} + \sum_{i\in\mathcal{E}_v^{\mathcal{J}+}}\lambda_i^k\right) - \sum_{i\in\mathcal{E}_v^{\mathcal{J}+}}\left(\frac{\sum_{k\in\mathcal{J}} x_i^k}{x_i}\right)d_i(x_i)\\ &\geq \sum_{k\in\mathcal{J}}\left(\sum_{j\in\mathcal{E}_v} y_j^{k*} + \sum_{i\in\mathcal{E}_v^{\mathcal{J}+}}\lambda_i^k\right) - \sum_{i\in\mathcal{E}_v^{\mathcal{J}+}} d_i(x_i).\end{aligned}$$

Taking the limit of both sides gives

$$\begin{aligned}\liminf_{t\to+\infty}\sum_{k\in\mathcal{J}}\sum_{i\in\mathcal{E}_v^{\mathcal{J}+}}\dot{x}_i^k &\geq \sum_{k\in\mathcal{J}}\left(\sum_{j\in\mathcal{E}_v} y_j^{k*} + \sum_{i\in\mathcal{E}_v^{\mathcal{J}+}}\lambda_i^k\right) - \sum_{i\in\mathcal{E}_v^{\mathcal{J}+}} d_i(x_i^*)\\ &\geq \sum_{k\in\mathcal{J}}\left(\sum_{j\in\mathcal{E}_v} y_j^{k*} + \sum_{i\in\mathcal{E}_v^{\mathcal{J}+}}\lambda_i^k\right) - \sum_{i\in\mathcal{E}_v^{\mathcal{J}+}} c_i \geq 0,\end{aligned}$$

where the central inequality is strict for any $i \in \mathcal{E}_v^{\mathcal{J}+}$ such that $x_i^* < +\infty$. This implies that there exists at least one cell $i \in \mathcal{E}_v^{\mathcal{J}+}$ such that $x_i \to +\infty$ as $t \to +\infty$ and the network is not fully transferring. Moreover, if $x_i \to +\infty$, there must exist at least one $k \in \mathcal{K}$ such that $x_i^k \to +\infty$. Then

$$\limsup_{t \to +\infty} \dot{x}_i^k \geq 0\,,$$

and

$$\limsup_{t \to +\infty} \lambda_i^k + \sum_{j \in \mathcal{E}_v} y_j^{k*} R_{ji}^k(x^{(v+)}) - \frac{x_i^k}{x_i} d_i(x_i) \geq 0\,,$$

which implies that either $\lambda_i^k \geq c_k$ and then the inequality in (7.3) is clearly violated or

$$\limsup_{t \to +\infty} R_{ji}^k(x^{(v+)}) > 0\,.$$

Then Definition 7.1 e), gives us that none of the traffic volumes x_j, $\forall j \in \mathcal{E}_v^k$ can be finite. □

Next, we establish convergence in each class of vehicles under the assumption that the aggregate traffic volumes on the outgoing cells converge to finite values.

Lemma 7.3
If the local network (7.1) with static inflows satisfies the condition given in Lemma 7.2, then the traffic volume for each class of vehicles on the outgoing cells will also converge to a unique finite value, i.e.,

$$\lim_{t \to +\infty} x_i(t) = x_i^* < +\infty \implies \lim_{t \to +\infty} x_i^k(t) = x_i^{k*}, \quad \forall i \in \mathcal{E}_v^+, \forall k \in \mathcal{K}\,.$$

Proof. First observe that if $x_i^* = 0$, then $x_i^{k*} = 0$ as well for all $k \in \mathcal{K}$, since the traffic volume for each class of vehicles cannot be below zero. Let us therefore for the rest of the proof assume that $x_i^* > 0$.

Consider an arbitrary cell $i \in \mathcal{E}_v^+$. Since the aggregate traffic volumes converge, for every $\epsilon > 0$ there exists a $t_0 > 0$ such that

$$|x_i(t) - x_i^*| < \epsilon, \quad \forall t \geq t_0\,.$$

Due to the continuity assumption of the routing policy and since the aggregate densities are finite, there also exists a $\xi > 0$ such that

$$\left| \sum_{j \in \mathcal{E}_v} y_j^{k*} R_{ji}^k(x(t)) - \sum_{j \in \mathcal{E}_v} y_j^{k*} R_{ji}^k(x^*) \right| < \xi\,,$$

and an $\eta > 0$ such that

$$\left|d_i(x_i(t)) - d_i(x_i^*)\right| < \eta\,.$$

Using the local system dynamics from (7.1) together with the inequalities, gives for $t > t_0$

$$\begin{aligned}\dot{x}_i^k &= \lambda_i^k + \sum_{j\in\mathcal{E}} y_j^{k*} R_{ji}^k(x^{(\nu+)}) - \frac{x_i^k}{x_i} d_i(x_i) \\ &\leq \lambda_i^k + \sum_{j\in\mathcal{E}} y_j^{k*} R_{ji}^k(x^{(\nu+)*}) + \xi_k - \frac{x_i^k}{x_i^* + \epsilon}(d_i(x_i^*) - \eta), \quad \forall k \in \mathcal{K}\,. \end{aligned} \tag{7.4}$$

Let $\alpha = \lambda_i^k + \sum_{j\in\mathcal{E}} y_j^{k*} R_{ji}^k(x^{(\nu+)*})$ and $\beta = d_i(x_i^*)$, then the equation above reads

$$\dot{x}_i^k \leq \alpha + \xi - x_i \cdot \frac{\beta - \eta}{x_i^* + \epsilon}\,.$$

By applying the affine transformation,

$$\hat{x}_i^k(t) = x_i^k(t) - \frac{(\alpha + \xi)(x_i^* + \epsilon)}{\beta - \eta},$$

Equation (7.4) can be written as

$$\dot{\hat{x}}_i^k \leq -x_i^k(t) \cdot \frac{\beta - \eta}{x_i^* + \epsilon}\,.$$

Direct use of Gronwall's inequality [Gronwall, 1919] yields

$$\hat{x}_i^k(t) \leq \hat{x}_i^k(0) \cdot \exp\left\{-\tfrac{\beta-\eta}{x_i^*+\epsilon} t\right\} \to 0 \quad \text{when } t \to +\infty,$$

if we choose η such that $\eta < \beta$. With $x_i^{k*} = \lim_{t\to+\infty} x_i^k(t)$, the inequality above can be written as

$$x_i^{k*} \leq \frac{(\alpha + \xi)(x_i^* + \epsilon)}{\beta - \eta}\,.$$

In the same way it holds

$$\begin{aligned}\dot{x}_i^k &= \lambda_i^k + \sum_{j\in\mathcal{E}} y_j^{k*} R_{ji}^k(x^{(\nu+)}) - \frac{x_i^k}{x_i} d_i(x_i) \\ &\geq \lambda_i^k + \sum_{j\in\mathcal{E}} z_j^{k*} R_{ji}^k(x^{(\nu+)*}) - \xi - \frac{x_i^k}{x_i^* - \epsilon}(d_i(x_i^*) + \eta), \quad \forall k \in \mathcal{K}.\end{aligned}$$

Using the same technique again gives a lower bound on x_i^{k*}, and the limit traffic volume can be bounded as follows

$$\frac{(\alpha-\xi)(x_i^*-\epsilon)}{\beta+\eta} \leq x_i^{k*} \leq \frac{(\alpha+\xi)(x_i^*+\epsilon)}{\beta-\eta}\,.$$

Since we have the freedom to choose ξ, ϵ, and η arbitrarily small, this implies that

$$\lim_{t\to+\infty} x_i^k(t) = x_i^{k*}, \quad \forall i \in \mathcal{E}_v^+\,,\ \forall k \in \mathcal{K}\,. \qquad \square$$

The next step is to show that if the aggregate traffic volumes converge when the inflows are static, they will also converge for converging inflows.

Lemma 7.4
Consider a local network given by (7.1) with converging inflows, such that $\lim_{t\to+\infty} y_j^k(t) = y_j^k$ for all $j \in \mathcal{E}_v$ and all $k \in \mathcal{K}$. Then the aggregate densities on each outgoing link also converge, such that

$$\lim_{t\to+\infty} x_i(t) = x_i^*, \quad \forall i \in \mathcal{E}_v^+\,. \qquad \square$$

Proof. Denote the right hand side of (7.2) $F_i(x^{(v+)}, z)$. Then

$$\frac{\partial}{\partial x_j} F_i(x^{(v+)}, y) \geq 0\,, \quad \forall j \neq i \in \mathcal{E}_v^+\,,$$

and

$$\frac{\partial}{\partial y_j^k} F_i(x^{(v+)}, y) \geq 0 \quad \forall j \in \mathcal{E}_v\,, \forall k \in \mathcal{K}.$$

Hence the system is a controlled monotone system in the sense of Angeli and Sontag [Angeli and Sontag, 2003]. If $y \in \mathbb{R}_+^{\mathcal{E}_v\times\mathcal{K}}$ is converging to y^*, for each $\epsilon > 0$ there exists a $t_0 > 0$ such that $|y_i^k(t) - y_i^{k*}| < \epsilon$ for $t > t_0$. Due to the monotonicity of the system it holds that

$$\Phi_t(y^* - \epsilon, x(0)) \leq \Phi_t(y(t), x(0)) \leq \Phi_t(y^* + \epsilon, x(0)),\ t \geq t_0,$$

where $\Phi_t : \mathbb{R}_+ \times (\mathbb{R}_+^{\mathcal{E}_v\times\mathcal{K}}, \mathbb{R}_+^{\mathcal{E}_v{}^+}) \to \mathbb{R}^{\mathcal{E}_v^+}$ is the semiflow. For a given inflow y, denote the corresponding limit traffic volume $x(y)$. As $t \to +\infty$ it holds that

$$x^*(y-\epsilon) \leq \lim_{t\to+\infty} x(y(t)) \leq x^*(y+\epsilon),$$

where Lemma 7.1 guarantees that $x^*(y-\epsilon)$ and $x^*(y+\epsilon)$ converge. Since x^* depends continuously on y, see [Como et al., 2013a, Lemma 3], by letting $\epsilon \to 0$ the aggregate on each link will converge to a unique limit traffic volume. $\square$

Proof of Theorem 7.1. Lemmas 7.1, 7.2, 7.3 and 7.4 together prove the first part of the theorem. The necessity condition in Lemma 7.2 proves the second part. □

Theorem 7.1 deals with stability of a local system. In the rest of this section we shall address the stability of an acyclic network with a single origin junction, denoted $o \in \mathcal{V}$, where all the exogenous inflow enter. The stability is then shown by interpreting it as a cascade of local networks. To this end, let $\mathcal{J} \subset \mathcal{K}$ and

$$\mathcal{V}^{\mathcal{J}} := \{v \in \mathcal{V} \mid \mathcal{E}_v^{\mathcal{J}+} \neq \emptyset\}.$$

In other words, $\mathcal{V}^{\mathcal{J}}$ contains all the junctions that allow flow of class $k \in \mathcal{J}$ on their outgoing cells. Moreover, let $\mathcal{U}^{\mathcal{J}} \subset \mathcal{V}^{\mathcal{J}}$ and $\partial\mathcal{U}^{\mathcal{J}} := \{e = (a, b) \in \mathcal{E}^{\mathcal{J}} \mid a \in \mathcal{U}, b \notin \mathcal{U}\}$. Define the minimum cut capacity between two junctions $o, s \in \mathcal{V}$, $\mathcal{C}^k_{o\to s}$ as

$$\mathcal{C}^{\mathcal{J}}_{o\to s} := \min_{\mathcal{U}^{\mathcal{J}} \subset \mathcal{V}^{\mathcal{J}} \text{ s.t } o\in\mathcal{U}^{\mathcal{J}}, s\notin\mathcal{U}^{\mathcal{J}}} \sum_{e\in\mathcal{E}^+_{\partial\mathcal{U}}} C_e.$$

For sake of simplicity, consider now an acyclic network with the same origin $o \in \mathcal{V}$ for all the classes of vehicles, i.e., $\lambda_i^k = 0$ for all $i \notin \mathcal{E}_o$ and all $k \in \mathcal{K}$. The following proposition offers a sufficient condition for such a network.

Proposition 7.1
Consider an acyclic dynamical multicommodity network with single origin, i.e., $\lambda_i^k = 0$ for all $i \notin \mathcal{E}_o$. Then a sufficient condition for it to admit a unique limit traffic volume and a unique limit flow is that for every $k \in \mathcal{K}$ and for every $v \in \mathcal{V}^k$ it holds that

$$\min\left(\mathcal{C}^{\mathcal{J}}_{o\to v}, \sum_{j\in\mathcal{E}_o}\sum_{k\in\mathcal{J}} \lambda_j^k\right) < \sum_{e\in\mathcal{E}_v^{\mathcal{J}}} c_e\,. \qquad \square$$

Proof. Since the graph is acyclic, a topological ordering exists [Leiserson et al., 2001, Theorem 22.12], so we can proceed by induction over the junctions of the graph. Number the junctions $v = 0, 1, .., n-2$. For $v = 0$, we only have converging from cells with exogenous inflows, and then Theorem 7.1 guarantees that the network is fully transferring. Now consider an arbitrary node w such that $0 < w < n-1$ and assume that the traffic volumes on the outgoing cells from junction up to index w are converging, and hence the outflows are converging as well. Also observe that the inflow to junction w cannot be greater than

$$\min\left(\mathcal{C}^{\mathcal{J}}_{o\to w}, \sum_{j\in\mathcal{E}_o}\sum_{k\in\mathcal{J}}\right).$$

Then it follows from Theorem 7.1 that the densities on the links in $\mathcal{E}_w^+$ admit a unique asymptotically stable equilibrium if and only if

$$\min\left(\mathcal{C}^{\mathcal{J}}_{o\to w}, \sum_{j\in\mathcal{E}_o}\sum_{k\in\mathcal{J}}\lambda_j^k\right) < \sum_{i\in\mathcal{E}_w^{\mathcal{J}}} c_i\,.$$

This is indeed satisfied by assumption, so the induction step is valid. Therefore the network admits a unique limit traffic volume. Uniqueness of the limit flow follows immediately from Assumption 7.1. □

7.4 Resilience

In this section we investigate how the dynamic multicommodity network responds to perturbations. Following [Como et al., 2013a; Como et al., 2013b], a perturbation is modeled as a family of perturbed demand functions, $\{\tilde{d}_i(x_i)\}_{\forall i\in\mathcal{E}}$ such that $\tilde{d}_i(x_i) \le d_i(x_i)$, for all $i \in \mathcal{E}$ and $\tilde{d}_i$ satisfies Assumption 7.1. The *magnitude of the perturbation* on one cell $i \in \mathcal{E}$ is then defined as $\delta_i := \sup_{x_i\ge 0}(\tilde{d}_i(x_i) - d_i(x_i))$ and the *total magnitude* of the perturbation is then given by $\delta := \sum_{e\in\mathcal{E}}\delta_e$. The *resilience* of a dynamical flow network is then defined as the infimum total magnitude of perturbations making the resulting dynamical flow network not fully transferring.

It was proven in [Como et al., 2013a; Como et al., 2013b] that, in the single commodity case, the resilience of an acyclic dynamical flow network coincides with the *minimum residual capacity*, defined as

$$\min_{v\neq v^d}\left\{\sum_{i\in\mathcal{E}_v^+} c_i - z_i^*\right\},$$

where z^* is the limit flow of the unperturbed dynamical flow network and v^d the destination junction of the flow. At the core of the proof is a *diffusivity property* of single-commodity local dynamical flow networks (cf. [Como et al., 2013b, Lemma 1]) guaranteeing that a perturbation of total magnitude δ in either some of the outgoing cells, or the inflow, does not increase the limit flow of any out-link by more than δ. In other words, the network does not overreact to perturbations.

The goal of this section is to show that, when more than one class of vehicles are present, dynamical flow networks can be instead arbitrarily fragile. In particular, we will construct a family of simple examples of multicommodity dynamical flow networks (with topology illustrated in Figure 7.2) that, irrespective of their minimal residual capacity, can lose their fully transferring property even by means of arbitrarily small perturbations. This will show that their resilience equals 0.

We will proceed by first stating some properties of local multicommodity dynamical flow networks that have the fully accessible properties. The first one can be considered as a weaker version of the aforementioned diffusivity property for multicommodity dynamical networks.

Lemma 7.5
Consider a fully accessible local dynamical multicommodity network for a junction $v \in \mathcal{V}$, with a fixed inflow y_j^k for all $j \in \mathcal{E}_v$ and all $k \in \mathcal{K}$. Assume that $\lambda_i^k = 0$ for all $i \in \mathcal{E}_v^+$ and all $k \in \mathcal{K}$, and

$$\sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} y_j^k < \sum_{i\in\mathcal{E}_v^+} c_i \,.$$

Let z^* denote the vector of limit outflows for this junction. Moreover, let $\tilde{y}$ be perturbed inflows and $\tilde{c}$ perturbed capacities such that

$$\sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} \tilde{y}_j^k < \sum_{i\in\mathcal{E}_v^+} \tilde{c}_i \,.$$

Let $\tilde{z}^*$ denote the limit outflows from the perturbed junction, with the inflows $\tilde{y}$. Then for every $\mathcal{I} \subseteq \mathcal{E}_v^+$ it holds that

$$\sum_{i\in\mathcal{I}} \left(\tilde{z}_i^* - z_i^*\right) \le \sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} \left[\tilde{y}_j^k - y_j^k\right]_+ + \sum_{i\in\mathcal{E}_v^+} \delta_i \,. \qquad \square$$

Proof. First, let $\hat{z} := \max(z, \tilde{z})$ and $\hat{y} := \max(y, \tilde{y})$ where max() applies componentwise. Moreover denote the solution of the local aggregate system (7.2) after perturbation $\hat{x}$ with the inflow $\hat{z}$ and initial condition $\hat{x}(0) = x^*$, where x^* is the limit traffic volume before perturbation. As a first step, we will prove that

$$\hat{z}_i(t) \ge \tilde{d}_i(x_i^*), \quad \forall t \ge 0, \quad \forall i \in \mathcal{E}_v^+ \,. \tag{7.5}$$

Consider a point in the space for the aggregate traffic volumes $\hat{x} \in \mathbb{R}_+^{\mathcal{E}_v^+}$, such that $\hat{x} > x^*$ and there exits a cell $i \in \mathcal{E}_v^+$ such that $\hat{x}_i = x_i^*$. Then [Como et al., 2013a, Lemma 1] implies that $R_{ji}^k(\hat{x}^{(v+)}) \ge R_{ji}^k(x^{(v+*)})$ for all $k \in \mathcal{K}$. Since we also know that $\hat{y} \ge y$ and

$$\tilde{d}_i(\hat{x}_i) \le d_i(\hat{x}_i) = d_i(x_i^*),$$

it holds that

$$\sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} \hat{y}_j^k R_{ji}^k(\hat{x}^{(v+)}) - d_i(\hat{x}_i) \ge \sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} y_j^k R_{ji}^k(x^{(v*+)}) - d_i(x_i^*) = 0 \,.$$

Let $\Omega := \{\hat{x} \in \mathbb{R}_+^{\mathcal{E}_v^+} \mid \hat{x}_i \ge x_i^*, \forall i \in \mathcal{E}_v^+\}$ and $\omega \in \mathbb{R}^{\mathcal{E}_v^+}$ the unit outpointing normal vector to the boundary of the set Ω. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}(\hat{x} \cdot w) = \left[\sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} \hat{y}_j^k R_{ji}^k(\hat{x}^{(v+)}) - d_i(\hat{x}_i)\right]_{i\in\mathcal{E}_v^+} \cdot \omega \le 0,$$

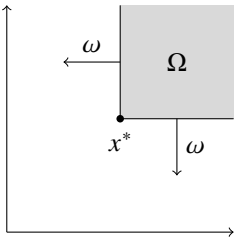


Figure 7.5 Schematic sketch of the invariant set used in the proof of the diffusivity lemma.

for all $\hat{x} \in \partial\Omega$ and for all $t \geq 0$. Hence Ω is an invariant set, see Figure 7.5, so that the aggregate traffic volumes will always be larger or equal to the limit volumes for the unperturbed system. This proves the inequality in (7.5).

Introduce now $\mathcal{I} \subseteq \mathcal{E}_v^+$ and $\mathcal{L} = \mathcal{E}_v^+ \setminus \mathcal{I}$. Since there exists an equilibrium for the perturbed system it follows that

$$\begin{aligned}
\sum_{i\in\mathcal{I}} \hat{z}_i^* &= \sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} \hat{y}_j^k - \sum_{l\in\mathcal{L}} \hat{z}_l^* \leq \sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} \hat{y}_j^k - \sum_{\ell\in\mathcal{L}} \tilde{d}_\ell(x_\ell^*) \\
&= \sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} \left(\hat{y}_j^k - y_j^k\right) + \sum_{i\in\mathcal{I}} z_i^* + \sum_{\ell\in\mathcal{L}} d_\ell(x_\ell) - \sum_{\ell\in\mathcal{L}} \tilde{d}_\ell(x_\ell^*) \\
&\leq \left[\sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} \hat{y}_j^k - y_j^k\right]_+ + \sum_{i\in\mathcal{I}} z_i^* + \sum_{\ell\in\mathcal{L}} \delta_\ell \\
&\leq \left[\sum_{k\in\mathcal{K}} \sum_{j\in\mathcal{E}_v} \hat{y}_j^k - y_j^k\right]_+ + \sum_{i\in\mathcal{I}} z_i^* + \sum_{\ell\in\mathcal{E}_v^+} \delta_\ell \,.
\end{aligned}$$

Since the aggregate system is monotone and $\lambda \leq \hat{\lambda}$ it also follows that

$$\tilde{z}_i^*(y) \leq \tilde{z}_i^*(\hat{y}) = \hat{z}_i^*, \quad \forall i \in \mathcal{E}_v^+ ,$$

which implies

$$\sum_{i\in\mathcal{I}} \tilde{z}_i^*(y) \leq \sum_{i\in\mathcal{I}} \hat{z}_i^* . \qquad \square$$

Lemma 7.5 provides a bound on the difference between aggregate limit flows before and after the perturbation in terms of its magnitude and of the difference between the inflows. Observe that when there is only one class of vehicles, i.e., $|\mathcal{K}| = 1$, Lemma 7.5 reduces to Lemma 1 in [Como et al., 2013b]. On the other hand, the following two results show that when more than one class of vehicle is present, each class can change in an arbitrary way as long as the bound on the aggregate flow provided by Lemma 7.5 is satisfied.

Lemma 7.6
Consider a local dynamical network with two outgoing cells e_2, e_3 and two classes of inflows y_1^A, y_1^B enter through one incoming link in the network . Let z^{k*} be a feasible equilibrium flow. Then, for $\epsilon > 0$ small enough, there exists distributed routing policies, R^A, R^B, such that

a) z^{k*} is the equilibrium flow of the dynamical local network,

b) there exists a perturbation of magnitude ϵ such that the perturbed limit flow, for one class $k \in \mathcal{E}$ and for one of the cells $i \in \{e_2, e_3\}$, satisfies

$$\tilde{z}_i^{k*} > \min(y_1^k, z_i^*) - \delta\,,$$

where $\delta > 0$ can be chosen arbitrary small. □

Proof. From a given limit flow z^*, the monotonicity of the flows in Assumption 7.1 gives unique corresponding limit traffic volumes $\tilde{x}$. By constructing the distributed routing policies as

$$R_{1,i}^k = \frac{z_i^{k*} e^{-\alpha_i^k (x_i - x_i^*)}}{\sum_{j \in \mathcal{E}_v^+} z_j^{k*} e^{-\alpha_j^k (x_j - x_j^*)}}\,, \quad \forall k \in \mathcal{K}, i \in \mathcal{E}_v^+\,,$$

where $\alpha_i^k > 0$, the equilibrium point to the dynamics given by (7.1) is x^*.

After a small capacity perturbation of $\epsilon > 0$ on one cell, assume without loss of generality e_2, such that $\tilde{d}_2(x_2) < d_2(x_2)$ it holds that

$$\begin{aligned}\tilde{z}_2^* &= y_1^A R_{1,2}^A(x_2, x_3) + y_1^B R_{1,2}^B(x_2, x_3)\,,\\ \tilde{z}_3^* &= y_1^A R_{1,3}^A(x_2, x_3) + y_1^B R_{1,3}^B(x_2, x_3)\,.\end{aligned}$$

The equations above are linearly dependent since $R_{1,2}^k = R_{1,3}^k = 1$ for all $k \in \mathcal{K}$. Hence we have one degree of freedom in the construction of the routing policies, which determines the ratio between the classes. Choose $\tilde{z}_2^* = z_3^*$ and $\tilde{z}_3^* = z_3^*$. Then, due to the perturbation $\tilde{x}_2 > x_3$ (also at limit) and $\tilde{x}_3 = x_2$. Moreover, either $\lambda_1^A > \tilde{z}_2^*$ or $\lambda_1^B > \tilde{z}_2^*$ or both. In the first case, let $\alpha_2^B = \beta$ and keep the other coefficients fixed. Then as $\beta \to +\infty$, $R_{1,2}^B(x_2, x_3) \to 0$ and

$$\begin{aligned}\tilde{z}_2^* &= y_1^A R_{1,2}^A(x_2, x_3)\,,\\ \tilde{z}_3^* &= y_1^A R_{1,3}^A(x_2, x_3) + y_1^B\,.\end{aligned}$$

Hence all flow of class B is routed to cell e_3. If instead $y_1^B > \tilde{z}_2^*$, choosing $\alpha_2^A = \beta$ will route all flow of class A to link e_3. □

Remark 7.1
In the single commodity case, it holds that

$$\tilde{z}_2 = yR_{1,2}(x_2, x_3),$$
$$\tilde{z}_3 = yR_{1,3}(x_2, x_3),$$

and since x is uniquely determined by the limit flows, there is no room for a parameter. This motivates why this fragile behavior can occur in a multicommodity setting but not in a single commodity setting. □

Notice in particular that the perturbation considered in Lemma 7.6 does not change the inflows y^A and y^B, and hence by Lemma 7.5 $\tilde{z}_i^* \leq z_i^* + \epsilon$ for $i \in \{e_2, e_3\}$. Also notice that trivially $\tilde{z}_i k* \leq \min\{y_1^k, \tilde{z}_i^*\} \leq \min\{y_1^k, z_i^* + \epsilon\}$. Lemma 7.6 ensures then that after perturbation we get

$$\min(y_1^k, z_i^*) - \delta \leq \tilde{z}_i^{k*} \leq \min(y_1^k, z_i^* + \epsilon).$$

Since ϵ and δ are arbitrary, we can steer $\tilde{z}_i^{k*}$ arbitrarily close to $\min(y_1^k, z_i^*)$.

Lemma 7.7
Consider a local dynamical network, with two outgoing links e_1, e_2 and two classes of inflows y_1^A, y_1^B. Let z^* be a feasible limit flow. Then, if the inflows change to $\tilde{y}$ and the new limit flows satisfy $c_2 > \tilde{z}_2^* > z_2^*$ and $\tilde{z}_3^* < z_3^*$, there exist routing policies R^A, R^B, such that for a given $\delta > 0$

$$\tilde{z}_2^* > \frac{z_2^{A*}}{y_1^A}\tilde{y}_1^A + \frac{z_2^{B*}}{y_1^B}\tilde{y}_1^B - \delta. \qquad \square$$

Proof. Construct the routing policies as in Lemma 7.6 with $\alpha_e^k = \frac{1}{\beta}$. Then, since $\tilde{z}_2^* < c_2$, $\tilde{x}_i(\beta) - x_i^*$ is bounded and

$$\tilde{z}_1^* = \tilde{y}_1^A R_{1,2}^A(\tilde{x}^*(\beta)) + \tilde{y}_1^B R_{1,2}^B(\tilde{x}^*(\beta)) \to \frac{z_2^{B*}}{y_1^B}\tilde{y}_1^B + \frac{z_2^{A*}}{y_1^A}\tilde{y}_1^A,$$

when $\beta \to +\infty$. □

We are now ready to construct an example showing that resilience can be arbitrarily low. To this aim, consider the network in Figure 7.1. Start from a given feasible limit flow z^* such that

$$\gamma_1 = \frac{z_4^{A*}}{z_3^{A*}} > \gamma_2 = \frac{z_4^{B*}}{z_3^{B*}},$$

and assume that

$$\min(\lambda_1^A, z_3^*) > \frac{c_6 - z_3^* \gamma_2}{\gamma_1 - \gamma_2} .$$

We claim that we can construct routing policies such that the network will not be fully transferring after an arbitrarily small perturbation.

Consider first the local network around junction v_1. Using Lemma 7.6, we know that we can construct routing policies such that after a small perturbation on link e_2 the flow of class A on link e_3 is steered close to the value

$$\tilde{z}_3^{A*} \approx \min(\lambda_1^A, z_3^*) > \frac{c_6 - z_3^* \gamma_2}{\gamma_1 - \gamma_2} .$$

In junction v_3, we construct then the routing policies according to Lemma 7.7. In this way, when, after perturbation, f_3^A approaches $\tilde{z}_3^A$ the perturbed limit flow on link e_4 converges to

$$\tilde{z}_4^* = \frac{z_4^{A*}}{z_4^{A*}} \tilde{z}_3^{A*} + \frac{z_4^{B*}}{z_3^{B*}} \tilde{z}_3^{B*} > c_6 .$$

Since the perturbed limit flow on e_4 is greater than the capacity of e_6, the network loses the fully transferring property, and the claim is proved.

To illustrate this behavior numerically, let the demand functions be

$$d_i(x_i) = c_i(1 - e^{-x_i}) \quad \forall i \in \mathcal{E} .$$

Moreover, let $\lambda_1^A = \lambda_1^B = 1.5$, $c_2 = c_3 = 1.55$, $c_4 = c_5 = 2$ and $c_6 = 1.45$. Since $\mathcal{C}_{2\to 4} = 1.51 > 1.45$ the sufficient condition stated in Proposition 7.1 is violated. However, by letting $z_3^{B*} = 1$, $z_3^{A*} = 0.5$, $z_4^{A*} = 0.49$ and $z_4^{B*} = 0.1$, distributed routing policies can be constructed such that the network is still fully transferring. Let $\alpha_2^A = \alpha_3^B = 50$ and $\alpha_3^A = \alpha_2^B = 0.01$ and $\alpha_4^A = \alpha_4^B = 0.01$. Then one can perturb the network in such a way that $c_2 = 1.55$ decreases to $c_2 = 1.54$ and z_4^{A*} to $\tilde{z}_4^{A*} = 1.454 > 1.45$. With these values after perturbation, the network is not fully transferring. In Figure 7.6 we show how the flows on link e_3 and e_4 evolve, starting from zero initial state. The perturbation occurs at $t = t_p$.

7.5 Conclusions

In this chapter, we have introduced dynamic routing to a multi-commodity version of the dynamical flow model presented in Chapter 2. We showed that if an equilibrium exists, the system will converge to this equilibrium. For the single-commodity case, it has previously been shown that the class of dynamical routing policies we are using, makes the network resilient to perturbations. However, we have in this chapter shown that when more than one class of vehicles are present in the network, those resilience results do not necessarily hold anymore.

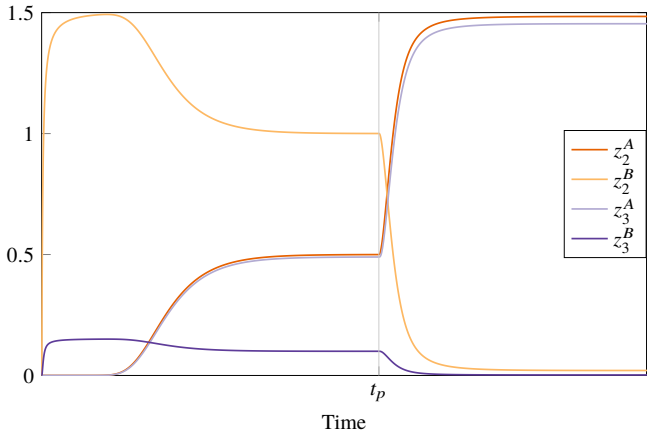


Figure 7.6 The propagation of flows on cells e_2 and e_3. At time $t = t_p$ a perturbation occurs such that z_3 goes above the capacity of cell e_5.

8

Conclusions and Future Work

8.1 Conclusions

In the thesis, we studied two control problems related to transportation networks, namely signal control and routing. For the first problem, we presented a fluid model for point queues, with the possibility for empty lanes to receive green-light. Because of this, the outflow from one lane cannot be equal to the amount of green light received. Instead, the actual outflow has to be upper bounded by the amount of green light. Although inequities in the model bound the actual outflow, we showed that there exists a unique solution, when the traffic signal control is feedback-based and Lipschitz continuous.

We also proposed a fully decentralized controller for traffic signals, the generalized proportional allocation (GPA) controller. We showed that the GPA controller could stabilize the queue lengths in the traffic network. Moreover, we showed that the GPA controller is throughput optimal, i.e., no other controller can handle a larger amount of exogenous inflows than the GPA controller. While there exist different throughput optimal controllers for signal control, the GPA has the benefit that it does not require any information about the network topology, average arrival rates or how the drivers propagate through the network. Those properties make the GPA controller very robust since it can allow for a new setting while maintaining its stability properties. Also, the GPA controller takes into account that a fraction of a cycle in signalized junction has to be devoted to phase shifts. The GPA controller adjusts the cycle length with the demand, such that a higher demand yields longer cycles, and a less fraction of the time is wasted for phase shifts.

Since the theoretical analysis of the GPA is done for an averaged model, we also evaluated the GPA controller's performance in a microscopic traffic simulator. The purpose of this was to investigate how well the GPA controller performed in a setting with discretized control action and a close-to-real traffic dynamics. We showed that the GPA controller outperformed standard fixed-time control for a traffic scenario

covering all the traffic in Luxembourg during a full day. We also compared the GPA controller with the MaxPressure controller in the simulator on an artificial Manhattan-like grid. In this case, we observed that the GPA controller performs well when the demands are low, but during high demands, the MaxPressure controller performs better.

A few routing problems were also studied in the thesis. First, we considered a static assignment problem between different classes of vehicles. One class of vehicles wants a route assignment that is fleet-optimal, i.e., the average delay for the whole fleet is optimized. The other class of vehicles is individual drivers that all try to take the route that is the fastest for just themselves. We present a sufficient condition for when a static routing assignment for both classes of vehicles exists. Also, we showed two algorithms to compute this assignment.

At last, we studied the robustness of local congestion-avoiding routing policies in a multicommodity setting. First, we presented a model for dynamically routed multicommodity flows and showed the stability of the model. We also showed that, if there are several classes of vehicles present, and they avoid congestions in a heterogeneous manner, cascade failures may occur. This is different to already existing results for single-commodity flows, where congestion-avoiding routing policies make the network robust to perturbations.

8.2 Directions for Future Work

Distributed Traffic Light Control

The fluid point queue model we are using has no propagation delay in it. While this simplifies the analysis, it may be too much of simplification if one considers the timing of signals between different junctions within a partition. It would therefore be nice to extend the analysis when the fluid model for queuing networks also has propagation delays, just like the way it has already been done in [Muralidharan et al., 2015] for open-loop control.

Another topic for further investigation is the existence and uniqueness of solutions for non-orthogonal phases. Although we conjecture that a solution always exists, it would be preferable to have a formal proof. Also, either proof of uniqueness or a counterexample for non-uniqueness is a point for further research.

For the discretized version of the GPA, a formal stability analysis is a topic for future work. In our discretization, two tuning parameters come into play, κ and $\bar{w}$. In the thesis, to illustrate that the controller conceptually works, we only did some experiments by trying out different values for the parameters. It would, therefore, be interesting to develop more sophisticated tuning rules for the parameters so that they can be tuned in an optimal way for each junction. While the controller already can achieve good throughput, the tuning rules should aim to achieve some optimality in e.g., the number of cycles vehicles that are queuing have to observe before passing

through the junction, or make sure that green-time is not over-allocated such that empty lanes receive green-light while vehicles in other lanes are waiting.

Also, in practice, the measurement of the queue lengths will saturate. Therefore a theoretical investigation on how saturation in the measurements of queue lengths affects the stability region is needed. If the arrival rates are high, the current controller is not able to "see behind" the area covered by the sensor, which means that it can not achieve throughput optimality when the measurement is saturated.

Another topic of future investigation is a weighting of the lanes in each phase. As it is now, if two lanes belong to the same phase, the fraction of the cycle allocated to that phase will be the same independent of if it is two vehicles queuing up in one of the lanes, or one vehicle in each lane. In the first case, it will, however, take a longer time to empty the lanes in the phase compared to the second case. By introducing weights, one can then get a more fair comparison between the demand on the different phases, while hopefully still keeping the stability properties of the controller.

In the presentation of the controller, we have not said anything about how to order the phases within a cycle. While keeping the same order for each makes the traffic signals behave as the drivers expect, one can think of either a few occasions doing recording or extend the clearance phases a bit to accommodate green-waves. With this, all the traffic signals will not be fully decentralized anymore. However, since accommodating green waves is usually only feasible when one has a traffic corridor, it may be worth implementing a centralized solution for this corridor.

One challenging problem within the field of distributed traffic signal control is when the controllers should take into account that each lane can only accommodate a limited number of vehicles. While attempts to enhance the MaxPressure controller to handle finite storage capacities, e.g., in [Gregoire et al., 2015], formal guarantees when the network is close to its maximal capacity seem still to be missing.

Two Tiered Traffic Assignment

For the moment, the routing assignment for the fleet is done in an open-loop setting. It would, therefore, be of interest to develop local routing policies that can respond to sudden changes in the network state, and still stay close to their global objective. If the fleet is not operated in a centralized way, there will be a need of design tolling system such that there are economic incentives to follow the fleet optimal assignment. Ideally, the tolling system should be combined with a contract on how to respond to disturbances along the way, to avoid that the price of the trip becomes more expensive during the journey due to raised tolls when congestion occurs. The contract on how to respond to perturbations should be such that cascade failures in the network are avoided as well.

Resilience of Multicommodity Flows

As we showed by example in Chapter 7, the routing policies are not robust to perturbations when more than one commodity is present in the network. It would, therefore, be desirable to find ways to construct routing policies that are robust for the multicommodity setting as well. Also, a stability proof for the case when the network contains cycles would be desirable.

Bibliography

Angeli, D. and E. D. Sontag (2003). "Monotone control systems". *IEEE Transactions on Automatic Control* **48**:10, pp. 1684–1698. DOI: 10.1109/TAC.2003.817920.

Ba, Q. and K. Savla (2016). "On distributed computation of optimal control of traffic flow over networks". In: *54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1102–1109. DOI: 10.1109/ALLERTON.2016.7852358.

Beckman, M., C. B. McGuire, and C. B. Winsten (1956). *Studies in the economics of transportation*. Yale University Press, New Heaven.

Bianchin, G. and F. Pasqualetti (2018). "A network optimization framework for the analysis and control of traffic dynamics and intersection signaling". In: *57th IEEE Conference on Decision and Control*, pp. 1017–1022.

Boyd, S. and L. Vandenberghe (2004). *Convex optimization*. Cambridge university press, Cambridge.

Carlson, R. C., I. Papamichail, and M. Papageorgiou (2011). "Local feedback-based mainstream traffic flow control on motorways using variable speed limits". *IEEE Transactions on Intelligent Transportation Systems* **12**:4, pp. 1261–1276. DOI: 10.1109/TITS.2011.2156792.

Chase, R. (2018). *Shared Mobility Principles for Livable Cities*. URL: http://www.sharedmobilityprinciples.org/ (visited on 2018-11-14).

Citilog (2018). *Xcam-ng*. URL: http://www.citilog.com/product/en/xcam-ng (visited on 2018-11-29).

Codecá, L., R. Frank, S. Faye, and T. Engel (2017). "Luxembourg SUMO Traffic (LuST) Scenario: Traffic Demand Evaluation". *IEEE Intelligent Transportation Systems Magazine* **9**:2, pp. 52–63.

Como, G., K. Savla, D. Acemoglu, M. A. Dahleh, and E. Frazzoli (2013a). "Robust distributed routing in dynamical networks–Part I: Locally responsive policies and weak resilience". *IEEE Transactions on Automatic Control* **58**:2, pp. 317–332. DOI: 10.1109/TAC.2012.2209951.

Como, G., K. Savla, D. Acemoglu, M. A. Dahleh, and E. Frazzoli (2013b). "Robust distributed routing in dynamical networks–Part II: Strong resilience, equilibrium selection and cascaded failures". *IEEE Transactions on Automatic Control* **58**:2, pp. 333–348. DOI: 10.1109/TAC.2012.2209975.

Como, G., K. Savla, D. Acemoglu, M. Dahleh, and E. Frazzoli (2013c). "Stability analysis of transportation networks with multiscale driver decisions". *SIAM Journal on Control and Optimization* **51**:1, pp. 230–252. DOI: 10.1137/110820804. eprint: https://doi.org/10.1137/110820804.

Como, G. (2017). "On resilient control of dynamical flow networks". *Annual Reviews in Control* **43**, pp. 80–90. DOI: 10.1016/j.arcontrol.2017.01.001.

Como, G. and F. Fagnani (2016). "From local averaging to emergent global behaviors: The fundamental role of network interconnections". *Systems & Control Letters* **95**, pp. 70–76. DOI: 10.1016/j.sysconle.2016.02.003.

Como, G., E. Lovisari, and K. Savla (2016). "Convexity and robustness of dynamic traffic assignment and freeway network control". *Transportation Research Part B: Methodological* **91**, pp. 446–465. DOI: 10.1016/j.trb.2016.06.007.

Coogan, S. and M. Arcak (2015). "A compartmental model for traffic networks and its dynamical behavior". *IEEE Transactions on Automatic Control* **60**:10, pp. 2698–2703. DOI: 10.1109/TAC.2015.2411916.

Coogan, S. and M. Arcak (2016). "Stability of traffic flow networks with a polytree topology". *Automatica* **66**, pp. 246–253. DOI: 10.1016/j.automatica.2015.12.015.

Coogan, S., C. Flores, and P. Varaiya (2017a). "Traffic predictive control from low-rank structure". *Transportation Research Part B: Methodological* **97**, pp. 1–22. DOI: 10.1016/j.trb.2016.11.013.

Coogan, S., E. Kim, G. Gomes, M. Arcak, and P. Varaiya (2017b). "Offset optimization in signalized traffic networks via semidefinite relaxation". *Transportation Research Part B: Methodological* **100**, pp. 82–92. DOI: 10.1016/j.trb.2017.01.016.

Cookson, G. (2018). *INRIX Global Traffic Scorecard*. INRIX Research.

Daganzo, C. F. (1994). "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory". *Transportation Research Part B: Methodological* **28**:4, pp. 269–287.

Daganzo, C. F. (1995). "The cell transmission model, Part II: Network traffic". *Transportation Research Part B: Methodological* **29**:2, pp. 79–93.

Dervisoglu, G., G. Gomes, J. Kwon, R. Horowitz, and P. Varaiya (2009). "Automatic calibration of the fundamental diagram and empirical observations on capacity". In: *Transportation Research Board 88th Annual Meeting*. Vol. 15.

Ferrara, A., A. N. Oleari, S. Sacone, and S. Siri (2015a). "Freeways as systems of systems: A distributed model predictive control scheme". *IEEE Systems Journal* **9**:1, pp. 312–323. DOI: 10.1109/JSYST.2014.2317931.

Ferrara, A., S. Sacone, and S. Siri (2015b). "Event-triggered model predictive schemes for freeway traffic control". *Transportation Research Part C: Emerging Technologies* **58**. Special Issue: Advanced Road Traffic Control, pp. 554–567. DOI: 10.1016/j.trc.2015.01.020.

Geroliminis, N., J. Haddad, and M. Ramezani (2013). "Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach". *IEEE Transactions on Intelligent Transportation Systems* **14**:1, pp. 348–359. DOI: 10.1109/TITS.2012.2216877.

Gomes, G. (2015). "Bandwidth maximization using vehicle arrival functions". *IEEE Transactions on Intelligent Transportation Systems* **16**:4, pp. 1977–1988. DOI: 10.1109/TITS.2014.2387731.

Gomes, G. and R. Horowitz (2006). "Optimal freeway ramp metering using the asymmetric cell transmission model". *Transportation Research Part C: Emerging Technologies* **14**:4, pp. 244–262. DOI: 10.1016/j.trc.2006.08.001.

Gomes, G., R. Horowitz, A. A. Kurzhanskiy, P. Varaiya, and J. Kwon (2008). "Behavior of the cell transmission model and effectiveness of ramp metering". *Transportation Research Part C: Emerging Technologies* **16**:4, pp. 485–513. DOI: 10.1016/j.trc.2007.10.005.

Grandinetti, P., C. Canudas-de-Wit, and F. Garin (2018). "Distributed optimal traffic lights design for large-scale urban networks". *IEEE Transactions on Control Systems Technology*, pp. 1–14. DOI: 10.1109/TCST.2018.2807792.

Gregoire, J., X. Qian, E. Frazzoli, A. de La Fortelle, and T. Wongpiromsarn (2015). "Capacity-aware backpressure traffic signal control". *IEEE Transactions on Control of Network Systems* **2**:2, pp. 164–173. DOI: 10.1109/TCNS.2014.2378871.

Gregoire, J., S. Samaranayake, and E. Frazzoli (2016). "Back-pressure traffic signal control with partial routing control". In: *2016 IEEE 55th Conference on Decision and Control (CDC)*, pp. 6753–6758. DOI: 10.1109/CDC.2016.7799309.

Gregoire, J., E. Frazzoli, A. de La Fortelle, and T. Wongpiromsarn (2014). "Back-pressure traffic signal control with unknown routing rates". *IFAC Proceedings Volumes* **47**:3. 19th IFAC World Congress, pp. 11332–11337. DOI: 10.3182/20140824-6-ZA-1003.01585.

Gronwall, T. H. (1919). "Note on the derivatives with respect to a parameter of the solutions of a system of differential equations". *Annals of Mathematics* **20**:4, pp. 292–296. DOI: 10.2307/1967124.

Hao, Z., R. Boel, and Z. Li (2018a). "Model based urban traffic control, Part I: Local model and local model predictive controllers". *Transportation Research Part C: Emerging Technologies* **97**, pp. 61–81. DOI: 10.1016/j.trc.2018.09.026.

Hao, Z., R. Boel, and Z. Li (2018b). "Model based urban traffic control, Part II: Coordinated model predictive controllers". *Transportation Research Part C: Emerging Technologies* **97**, pp. 23–44. DOI: 10.1016/j.trc.2018.09.025.

Harrison, J. M. and M. I. Reiman (1981). "Reflected Brownian motion on an orthant". *The Annals of Probability* **9**:2, pp. 302–308. ISSN: 00911798. URL: http://www.jstor.org/stable/2243462.

Hegyi, A., B. D. Schutter, and H. Hellendoorn (2005). "Model predictive control for optimal coordination of ramp metering and variable speed limits". *Transportation Research Part C: Emerging Technologies* **13**:3, pp. 185–209. DOI: 10.1016/j.trc.2004.08.001.

Herty, M., C. Kirchner, S. Moutari, and M. Rascle (2008). "Multicommodity flows on road networks". *Commun. Math. Sci.* **6**:1, pp. 171–187. URL: https://projecteuclid.org:443/euclid.cms/1204905782.

Horn, R. A. and C. R. Johnson (2013). *Matrix Analysis*. 2nd. Cambridge University Press, New York, NY, USA. ISBN: 978-0-521-83940-2.

Hosseini, P. and K. Savla (2017). "Queue length simulation for signalized arterial networks and steady state computation under fixed time control". *ArXiv e-prints*. arXiv: 1705.07493 [math.DS].

Kelly, F. (1997). "Charging and rate control for elastic traffic". *European Transactions on Telecommunications* **8**:1, pp. 33–37. DOI: 10.1002/ett.4460080106. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ett.4460080106.

Kotsialos, A., M. Papageorgiou, C. Diakaki, Y. Pavlis, and F. Middelham (2002). "Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool METANET". *IEEE Transactions on Intelligent Transportation Systems* **3**:4, pp. 282–292. DOI: 10.1109/TITS.2002.806804.

Kovacs, P., T. Le, R. Nunez Queija, H. L. Vu, and N. Walton (2016). "Proportional green time scheduling for traffic lights". Preprint. URL: https://ir.cwi.nl/pub/25248.

Krajzewicz, D., J. Erdmann, M. Behrisch, and L. Bieker (2012). "Recent development and applications of SUMO - Simulation of Urban MObility". *International Journal On Advances in Systems and Measurements* **5**:3&4, pp. 128–138.

Le, T., H. L. Vu, N. Walton, S. P. Hoogendoorn, P. Kovács, and R. N. Queija (2017). "Utility optimization framework for a distributed traffic control of urban road networks". *Transportation Research Part B: Methodological* **105**, pp. 539–558. DOI: 10.1016/j.trb.2017.10.004.

Lebacque, J. P. and M. M. Khoshyaran (2002). "First order macroscopic traffic flow models for networks in the context of dynamic assignment". In: Patriksson, M. et al. (Eds.). *Transportation Planning: State of the Art*. Springer US, Boston, MA, pp. 119–140. ISBN: 978-0-306-48220-5. DOI: 10.1007/0-306-48220-7_8.

Leighton, T. and S. Rao (1999). "Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms". *J. ACM* **46**:6, pp. 787–832. DOI: 10.1145/331524.331526.

Leiserson, C. E., R. L. Rivest, C. Stein, and T. H. Cormen (2001). *Introduction to algorithms*. The MIT press, Cambridge, MA.

Lighthill, M. J. and G. B. Whitham (1955). "On kinematic waves II. a theory of traffic flow on long crowded roads". *Proc. R. Soc. Lond. A* **229**:1178, pp. 317–345.

Lovisari, E., G. Como, and K. Savla (2014). "Stability of monotone dynamical flow networks". In: *53rd IEEE Conference on Decision and Control*, pp. 2384–2389. DOI: 10.1109/CDC.2014.7039752.

Massoulié, L. (2007). "Structural properties of proportional fairness: stability and insensitivity". *Ann. Appl. Probab.* **17**:3, pp. 809–839. DOI: 10.1214/105051606000000907.

Mauro, V. and C. Di Taranto (1990). "Utopia". In: *Control, computers, communications in transportation*. Elsevier, pp. 245–252.

McShane, C. (1999). "The origins and globalization of traffic control signals". *Journal of Urban History* **25**:3, pp. 379–404.

Mehr, N., J. Lioris, R. Horowitz, and R. Pedarsani (2017). "Joint perimeter and signal control of urban traffic via network utility maximization". In: *20th IEEE International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–6. DOI: 10.1109/ITSC.2017.8317837.

Mehr, N., M. Sanselme, N. Orr, R. Horowitz, and G. Gomes (2018). "Offset selection for bandwidth maximization on multiple routes". In: *2018 Annual American Control Conference (ACC)*, pp. 6366–6371. DOI: 10.23919/ACC.2018.8431660.

Merchant, D. K. and G. L. Nemhauser (1978a). "A model and an algorithm for the dynamic traffic assignment problems". *Transportation Science* **12**:3, pp. 183–199. DOI: 10.1287/trsc.12.3.183.

Merchant, D. K. and G. L. Nemhauser (1978b). "Optimality conditions for a dynamic traffic assignment model". *Transportation Science* **12**:3, pp. 200–207. DOI: 10.1287/trsc.12.3.200. eprint: https://doi.org/10.1287/trsc.12.3.200.

Messmer, A. and M. Papageorgiou (1990). "METANET: a macroscopic simulation program for motorway networks". *Traffic Engineering & Control* **31**, pp. 466–470.

Munoz, L., X. Sun, R. Horowitz, and L. Alvarez (2003). "Traffic density estimation with the cell transmission model". In: *Proceedings of the 2003 American Control Conference, 2003.* Vol. 5, 3750–3755 vol.5. DOI: 10.1109/ACC.2003.1240418.

Muralidharan, A. and R. Horowitz (2012). "Optimal control of freeway networks based on the link node cell transmission model". In: *2012 American Control Conference (ACC)*, pp. 5769–5774. DOI: 10.1109/ACC.2012.6315236.

Muralidharan, A., R. Pedarsani, and P. Varaiya (2015). "Analysis of fixed-time control". *Transportation Research Part B: Methodological* **73**, pp. 81–90. DOI: 10.1016/j.trb.2014.12.002.

Neely, M. J. (2003). *Dynamic power allocation and routing for satellite and wireless networks with time varying channels*. PhD thesis. Massachusetts Institute of Technology.

Nilsson, G., G. Como, and E. Lovisari (2014). "On resilience of multicommodity dynamical flow networks". In: *53rd IEEE Conference on Decision and Control*, pp. 5125–5130. DOI: 10.1109/CDC.2014.7040190.

Nilsson, G., P. Hosseini, G. Como, and K. Savla (2015). "Entropy-like Lyapunov functions for the stability analysis of adaptive traffic signal controls". In: *2015 54th IEEE Conference on Decision and Control (CDC)*, pp. 2193–2198. DOI: 10.1109/CDC.2015.7402532.

Nilsson, G. and G. Como (2017). "On generalized proportional allocation policies for traffic signal control". *IFAC-PapersOnLine* **50**:1. 20th IFAC World Congress, pp. 9643–9648. DOI: 10.1016/j.ifacol.2017.08.1728.

Nilsson, G. and G. Como (2018a). "A micro-simulation study of the generalized proportional allocation traffic signal control". Submitted.

Nilsson, G. and G. Como (2018b). "Evaluation of decentralized feedback traffic light control with dynamic cycle length". *IFAC-PapersOnLine* **51**:9. 15th IFAC Symposium on Control in Transportation Systems CTS 2018, pp. 464–469. DOI: 10.1016/j.ifacol.2018.07.076.

Nilsson, G. and G. Como (2019a). "Generalized proportional allocation policies for robust control of dynamical flow networks". In preparation.

Nilsson, G. and G. Como (2019b). "On well-posedness of feedback-controlled outflows in dynamical flow networks". In preparation.

Nilsson, G., P. Grover, and U. Kalabic (2018). "Assignment and control of two-tiered vehicle traffic". In: *57th IEEE Conference on Decision and Control*, pp. 1023–1028.

Nisan, N., T. Roughgarden, E. Tardos, and V. Vazirani (2007). *Algorithmic Game Theory*. Cambridge University Press. ISBN: 9781139466547.

Organisation for Economic Co-operation and Development (2018). *OECD Data - Gross domestic product (GDP)*. URL: https://data.oecd.org/gdp/gross-domestic-product-gdp.htm (visited on 2018-11-29).

Ozdaglar, A. and I. Menache (2011). *Network Games: Theory, Models, and Dynamics*. Synthesis Lectures on Communication Networks. Morgan & Claypool Publishers. ISBN: 9781608454099.

Papageorgiou, M., H. Hadj-Salem, J.-M. Blosseville, et al. (1991). "ALINEA: a local feedback control law for on-ramp metering". *Transportation Research Record* 1320, pp. 58–64. ISSN: 0361-1981.

Pasquale, C., I. Papamichail, C. Roncoli, S. Sacone, S. Siri, and M. Papageorgiou (2015). "Two-class freeway traffic regulation to reduce congestion and emissions via nonlinear optimal control". *Transportation Research Part C: Emerging Technologies* **55**. Engineering and Applied Sciences Optimization (OPT-i) - Professor Matthew G. Karlaftis Memorial Issue, pp. 85–99. DOI: 10.1016/j.trc.2015.01.013.

Patriksson, M. (2015). *The Traffic Assignment Problem: Models and Methods*. Dover Publications. ISBN: 9780486802275.

Pisarski, D. and C. Canudas-de-Wit (2012). "Analysis and design of equilibrium points for the cell-transmission traffic model". In: *2012 American Control Conference (ACC)*, pp. 5763–5768. DOI: 10.1109/ACC.2012.6315050.

Pisarski, D. and C. Canudas-de-Wit (2016). "Nash game-based distributed control design for balancing traffic density over freeway networks". *IEEE Transactions on Control of Network Systems* **3**:2, pp. 149–161. DOI: 10.1109/TCNS.2015.2428332.

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

Robertson, D. I. (1969). *TRANSYT: a traffic network study tool*. Transport and Road Research Laboratory, Crowthorne.

Robertson, D. I. and R. D. Bretherton (1991). "Optimizing networks of traffic signals in real time-the SCOOT method". *IEEE Transactions on vehicular technology* **40**:1, pp. 11–15.

Roess, R. P., E. S. Prassas, and W. R. McShane (2011). *Traffic engineering*. Prentice Hall, Upper Saddle River, NJ.

Rosdahl, C., G. Nilsson, and G. Como (2018). "On distributed optimal control of traffic flows in transportation networks". In: *2018 IEEE Conference on Control Technology and Applications (CCTA)*, pp. 903–908. DOI: 10.1109/CCTA.2018.8511601.

Rosen, J. B. (1965). "Existence and uniqueness of equilibrium points for concave N-person games". *Econometrica* **33**:3, pp. 520–534. DOI: 10.2307/1911749.

Roughgarden, T. and É. Tardos (2002). "How bad is selfish routing?" *J. ACM* **49**:2, pp. 236–259. DOI: 10.1145/506147.506153.

Samaranayake, S., W. Krichene, J. Reilly, M. L. D. Monache, P. Goatin, and A. Bayen (2018). "Discrete-time system optimal dynamic traffic assignment (SO-DTA) with partial control for physical queuing networks". *Transportation Science* **52**:4, pp. 982–1001. DOI: 10.1287/trsc.2017.0800.

Schmitt, M., C. Ramesh, and J. Lygeros (2017). "Sufficient optimality conditions for distributed, non-predictive ramp metering in the monotonic cell transmission model". *Transportation Research Part B: Methodological* **105**, pp. 401–422. DOI: 10.1016/j.trb.2017.10.001.

Shah, D. and J. Shin (2010). "Dynamics in congestion games". In: *Proceedings of the ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*. SIGMETRICS '10. ACM, New York, New York, USA, pp. 107–118. ISBN: 978-1-4503-0038-4. DOI: 10.1145/1811039.1811052.

Sims, A. G. and K. W. Dobinson (1980). "The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits". *IEEE Transactions on vehicular technology* **29**:2, pp. 130–137.

Spieser, K., K. Treleaven, R. Zhang, E. Frazzoli, D. Morton, and M. Pavone (2014). "Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: A case study in Singapore". In: Meyer, G. et al. (Eds.). *Road Vehicle Automation*. Springer International Publishing, Cham, pp. 229–245. ISBN: 978-3-319-05990-7. DOI: 10.1007/978-3-319-05990-7_20.

Tassiulas, L. and A. Ephremides (1992). "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks". *IEEE Transactions on Automatic Control* **37**:12, pp. 1936–1948. DOI: 10.1109/9.182479.

Tipper, D. and M. K. Sundareshan (1990). "Numerical methods for modeling computer networks under nonstationary conditions". *IEEE Journal on Selected Areas in Communications* **8**:9, pp. 1682–1695. DOI: 10.1109/49.62855.

Varaiya, P. (2013a). "Max pressure control of a network of signalized intersections". *Transportation Research Part C: Emerging Technologies* **36**, pp. 177–195. DOI: 10.1016/j.trc.2013.08.014.

Varaiya, P. (2013b). "The max-pressure controller for arbitrary networks of signalized intersections". In: Ukkusuri, S. V. et al. (Eds.). *Advances in Dynamic Network Modeling in Complex Transportation Systems*. Springer New York, New York, NY, pp. 27–66. ISBN: 978-1-4614-6243-9. DOI: 10.1007/978-1-4614-6243-9_2.

Walton, N. S. (2014). "Concave switching in single and multihop networks". In: *The 2014 ACM International Conference on Measurement and Modeling of Computer Systems*. SIGMETRICS '14. ACM, Austin, Texas, USA, pp. 139–151. ISBN: 978-1-4503-2789-3. DOI: 10.1145/2591971.2591987.

Wang, Y., M. Papageorgiou, J. Gaffney, I. Papamichail, and J. Guo (2010). "Local ramp metering in the presence of random-location bottlenecks downstream of a metered on-ramp". In: *13th International IEEE Conference on Intelligent Transportation Systems*, pp. 1462–1467. DOI: 10.1109/ITSC.2010.5625170.

Wardrop, J. G. (1952). "Road paper. Some theoretical aspects of road traffic research." *Proceedings of the Institution of Civil Engineers* **1**:3, pp. 325–362. DOI: 10.1680/ipeds.1952.11259.

Webster, F. V. (1958). *Traffic signal settings*. H.M.S.O., London.

Wongpiromsarn, T., T. Uthaicharoenpong, Y. Wang, E. Frazzoli, and D. Wang (2012). "Distributed traffic signal control for maximum network throughput". In: *2012 15th International IEEE Conference on Intelligent Transportation Systems*, pp. 588–595. DOI: 10.1109/ITSC.2012.6338817.

Yang, H., X. Zhang, and Q. Meng (2007). "Stackelberg games and multiple equilibrium behaviors on networks". *Transportation Research Part B: Methodological* **41**:8, pp. 841–861. DOI: 10.1016/j.trb.2007.03.002.

Yazicioglu, A. Y., M. Roozbehani, and M. A. Dahleh (2018). "Resilient control of transportation networks by using variable speed limits". *IEEE Transactions on Control of Network Systems (Early Access)*, pp. 1–1. DOI: 10.1109/TCNS.2017.2782364.

Zaidi, A. A., B. Kulcsár, and H. Wymeersch (2016). "Back-pressure traffic signal control with fixed and adaptive routing for urban vehicular networks". *IEEE Transactions on Intelligent Transportation Systems* **17**:8, pp. 2134–2143. DOI: 10.1109/TITS.2016.2521424.

Ziliaskopoulos, A. K. (2000). "A linear programming model for the single destination system optimum dynamic traffic assignment problem". *Transportation Science* **34**:1, pp. 37–49. DOI: 10.1287/trsc.34.1.37.12281.

A

Additional Proofs

A.1 Proposition 2.1

Proof.

i) This fact has been proven before, for example in [Como and Fagnani, 2016], but we give a proof here to make the presentation complete. Since R is a non-negative matrix, it follows from the Perron-Frobenious theory [Horn and Johnson, 2013][Theorem 8.3.1] that there exists an eigenvector $w \geq 0$ with all non-negative elements with corresponding eigenvalue $\rho \geq 0$ such that all other eigenvalues of R^T are not larger than ρ in magnitude. Hence it holds that

$$\rho w = R^T w \,. \tag{A.1}$$

Let $\mathcal{S}$ be the support of w. Since every cell is connected to a sink cell it must hold that $\min_{j\in\mathcal{S}} \sum_{i\in\mathcal{S}} R^T_{ij} < 1$, since if this was not the case, then all cells in $\mathcal{S}$ would not be connected to any sink cell, and the routing matrix would then not be out-connected. Multiplying equation (A.1) with $\mathbb{1}^T$ from the left yields

$$\sum_i \rho w_i = \sum_i \sum_j R^T_{ij} w_j < \sum_i w_i \,.$$

Hence $\rho < 1$, and it implies that the spectral radius of R^T and hence also R is strictly less than 1.

ii) Since R has spectral radius strictly less than one, this is a consequence of [Horn and Johnson, 2013][Lemma 5.6.10].

iii) A cell $i \in \mathcal{I}$ is either connected to a sink-cell in the set $\mathcal{I}$ or it is connected to a sink-cell in $\mathcal{E} \setminus \mathcal{I}$. In the latter case, there must exist a cell $j \in \mathcal{I}$ such that $\sum_{k\in\mathcal{I}} R_{jk} + \sum_{k\in\mathcal{E}\setminus\mathcal{I}} R_{jk} = 1$, where $\sum_{k\in\mathcal{E}\setminus\mathcal{I}} R_{jk} > 0$.

iv) Follows from the fact that R has spectral radius strictly less than one. □

B

Additional Simulation Results

B.1 The Manhattan Grid

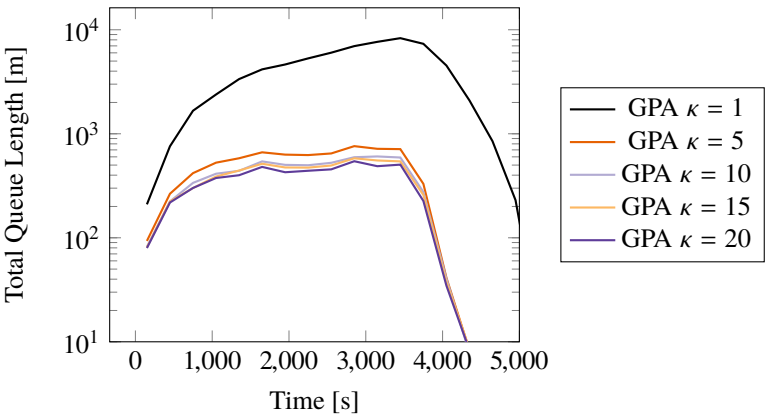


Figure B.1 The total queue length over time in the Manhattan grid with the GPA controller for different values of κ. The demand is $\delta = 0.05$. To improve the readability of the results, the queue lengths are averaged over 300 seconds intervals.

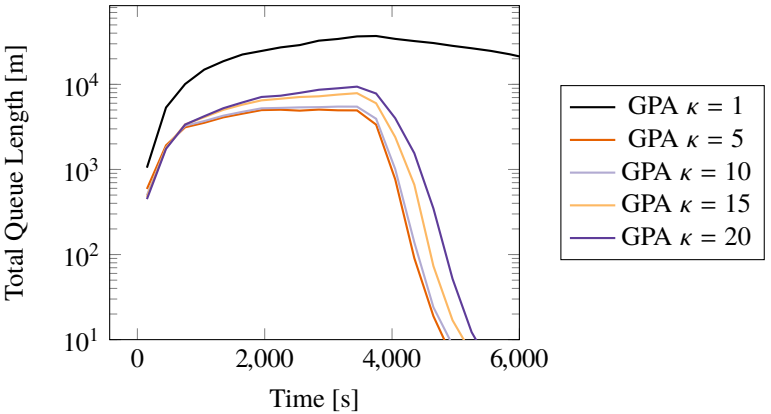


Figure B.2 The total queue length over time in the Manhattan grid with the GPA controller for different values of κ. The demand is $\delta = 0.10$. To improve the readability of the results, the queue lengths are averaged over 300 seconds intervals.

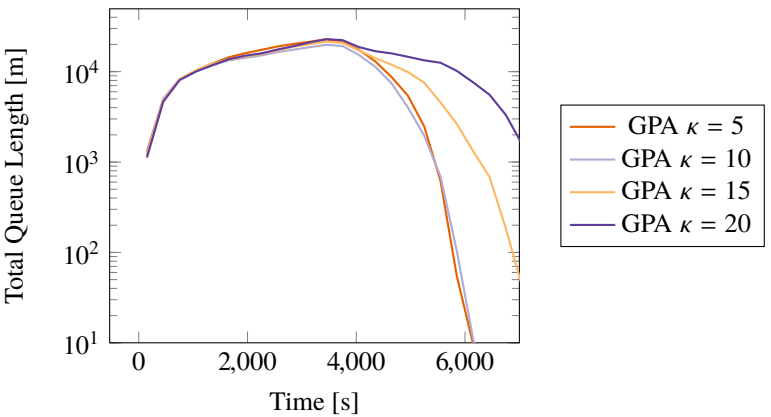


Figure B.3 The total queue length over time in the Manhattan grid with the GPA controller for different values of κ. For the value $\kappa = 1$ a grid lock situation occurs, and because of that the trajectory is not included in the plot. The demand is $\delta = 0.15$. To improve the readability of the results, the queue lengths are averaged over 300 seconds intervals.

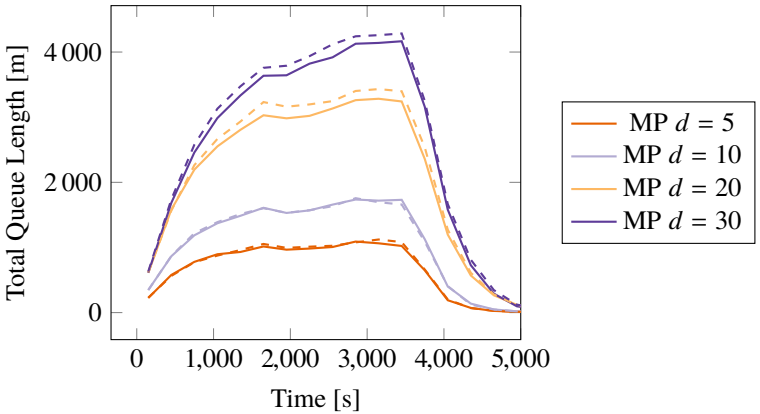


Figure B.4 The total queue length over time in the Manhattan grid with the Max-Pressure controller with right turning ratios (solid) and wrong turning ratios (dashed). The demand is $\delta = 0.05$. To improve the readability of the results, the queue lengths are averaged over 300 seconds intervals.

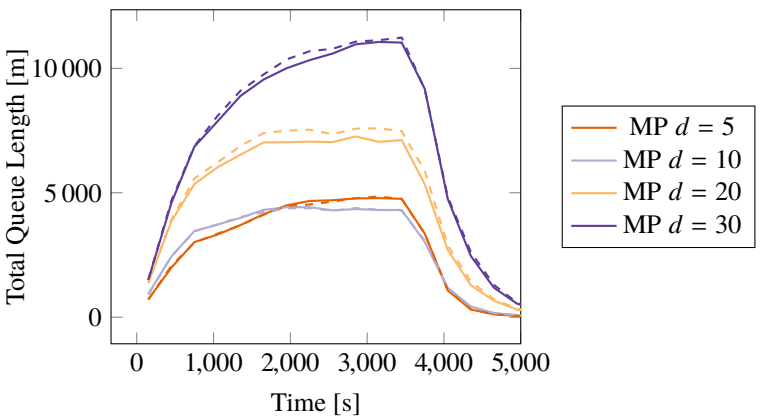


Figure B.5 The total queue length over time in the Manhattan grid with the Max-Pressure controller with right turning ratios (solid) and wrong turning ratios (dashed). The demand is $\delta = 0.10$. To improve the readability of the results, the queue lengths are averaged over 300 seconds intervals.

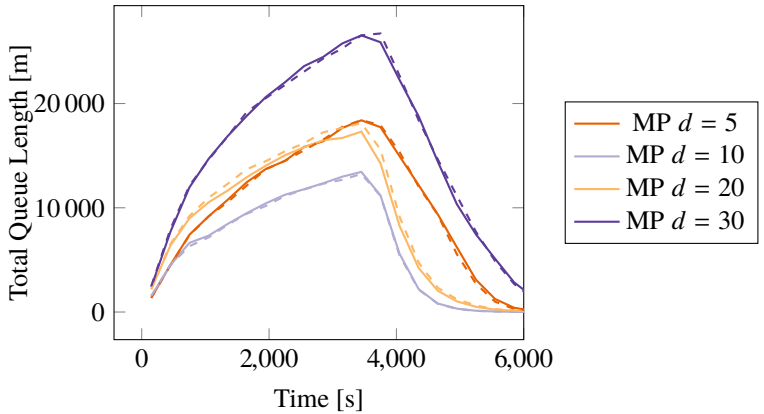


Figure B.6 The total queue length over time in the Manhattan grid with the Max-Pressure controller with right turning ratios (solid) and wrong turning ratios (dashed). The demand is $\delta = 0.15$. To improve the readability of the results, the queue lengths are averaged over 300 seconds intervals.

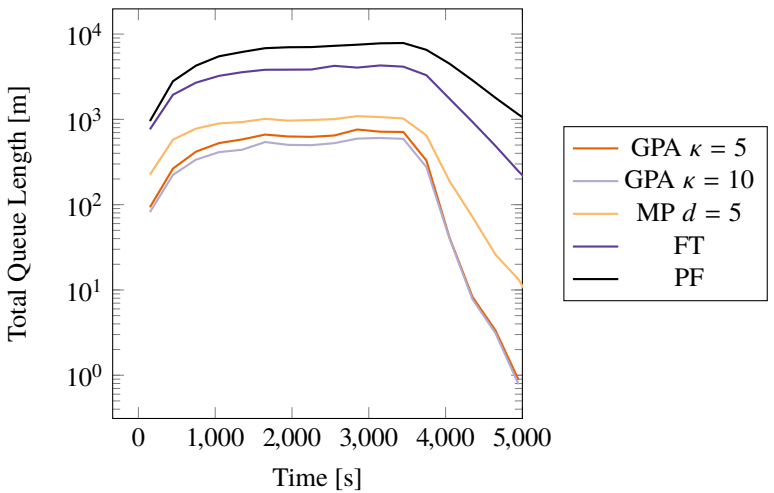


Figure B.7 A comparison between different control strategies for the Manhattan grid with the demand $\delta = 0.05$. To improve the readability of the results, the queue lengths are averaged over 300 seconds intervals.

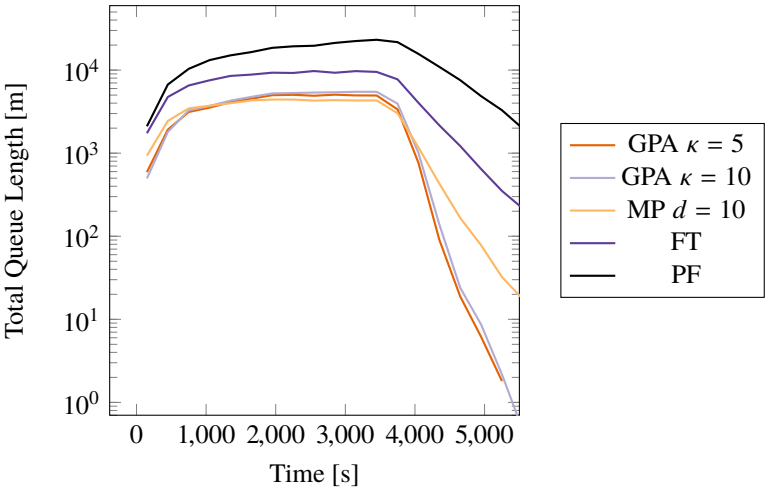


Figure B.8 A comparison between different control strategies for the Manhattan grid with the demand $\delta = 0.10$. To improve the readability of the results, the queue lengths are averaged over 300 seconds intervals.

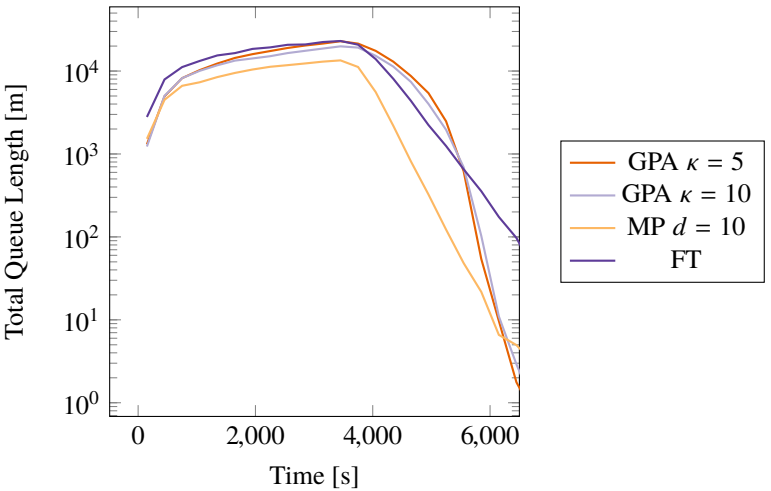


Figure B.9 A comparison between different control strategies for the Manhattan grid with the demand $\delta = 0.15$. Since the proportional fair controller (PF) creates a gridlock, it is not included in the comparison. To improve the readability of the results, the queue lengths are averaged over 300 seconds intervals.