ΠΟΛΥΤΕΧΝΕΙΟ
ΚΡΗΤΗΣ /
TECHNICAL
UNIVERSITY
OF CRETE

DOCTORAL THESIS

Internal Boundary Control of Lane-free Automated Vehicle Traffic

Author:

Milad Malekzadehkebria

Supervisors:

Prof. Markos Papageorgiou ***Prof. Ioannis Papamichail***

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

in the

Dynamic Systems & Simulations Laboratory
School of Production Engineering and Management

November 2023

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Τίτλος (ελληνικά/αγγλικά):

ΕΛΕΓΧΟΣ ΕΣΩΤΕΡΙΚΟΥ ΟΡΙΟΥ ΚΥΚΛΟΦΟΡΙΑΣ ΑΥΤΟΜΑΤΩΝ ΟΧΗΜΑΤΩΝ ΧΩΡΙΣ ΛΩΡΙΔΕΣ

INTERNAL BOUNDARY CONTROL OF LANE-FREE AUTOMATED VEHICLE TRAFFIC

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Milad Malekzadehkebria

ΤΡΙΜΕΛΗΣ ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΕΠΙΤΡΟΠΗ:

1. Καθ. Μάρκος Παπαγεωργίου (Επιβλέπων)
2. Καθ. Ιωάννης Παπαμιχαήλ (Συν-επιβλέπων)
3. Καθ. Ιάσων Καραφύλλης

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την: 10 / 11 / 2023

1. Καθ. Μάρκος Παπαγεωργίου
2. Καθ. Ιωάννης Παπαμιχαήλ
3. Καθ. Ιάσων Καραφύλλης (Εθνικό Μετσόβιο Πολυτεχνείο) Ι.Καραφύλλης
4. Καθ. Γεώργιος Χαλκιαδάκης
5. Επ. Καθ. Ελευθέριος Δοϊτσίδης
6. Επ. Καθ. Δημήτριος Ιψάκης
7. Prof. Klaus Bogenberger (Technical University of Munich)

TECHNICAL UNIVERSITY OF CRETE

Abstract

School of Production Engineering and Management

Doctor of Philosophy

Internal Boundary Control of Lane-free Automated Vehicle Traffic

By Milad Malekzadehkebria

Road transportation is undergoing a profound transformation, which is triggered by significant technological developments regarding the road infrastructure, the vehicles and the traffic management. In this context, vehicles communicate seamlessly with each other and with the infrastructure and are guided by real-time data and advanced algorithms, allowing them to move autonomously. Connected Automated Vehicles (CAVs) are equipped with an array of sensors, cameras, Lidar, radar, and V2X (Vehicle-to-Everything) communication systems. These technologies enable vehicles to "see" their surroundings, exchange information with nearby vehicles, traffic signals, and infrastructure, and make intelligent decisions based on this collective awareness.

One of the many advancements that have come about in this context is the TrafficFluid paradigm, which is a forward-thinking approach to road mobility that becomes increasingly important in the age of CAVs. TrafficFluid proposes lane-free movement of vehicles. Traditionally, roads have been demarcated by parallel lanes to ease human driving. However, the lane-free paradigm represents a departure from this rigid framework in the era of CAVs. Instead of confining vehicles to predefined lanes, it envisions a dynamic and adaptable traffic ecosystem.

Lane-free traffic implies that incremental road widening (narrowing) leads to corresponding incremental increase (decrease) of capacity; and this opens the way for consideration of real-time Internal Boundary Control (IBC) on highways and arterials, in order to flexibly share the total (both directions) road width and capacity among the two directions in dependence of the bi-directional demand and traffic conditions, so as to maximize the total (two directions) flow efficiency. The problem is initially presented in this thesis as a convex Quadratic Programming (QP) problem, which can be solved easily and effectively. Furthermore, the novel IBC control action is thoroughly investigated through various case studies, which highlight its features, capabilities and potential. This thesis efficiently addresses IBC also in a second step by developing suitable feedback-based Linear-Quadratic regulators, with or without integral action (LQI and LQ regulators). The proposed regulators have been put to the test through macroscopic simulation investigations that involve different demand scenarios on a realistic highway stretch. The findings indicate that the regulators can be just as efficient as an open-loop nonlinear constrained optimal control solution. Furthermore, the regulators do not require accurate

modeling or external demand prediction, thus being more robust. Nonetheless, centralized solutions that call for information from the entire highway stretch being considered can be burdensome for lengthy highways in relation to the scope of required real-time communication. Therefore, during the third step of this work, two different decentralized control schemes, specifically overlapping LQR method and Linear Matrix Inequality (LMI) method, are developed for IBC in lane-free CAV traffic.

As a final step, the present thesis completes and validates the IBC concept in a much more realistic way via microscopic simulation and active internal boundary moving, using the SUMO-based *TrafficFluid-Sim* simulation tool. To effectuate IBC, an LQR feedback control scheme is employed. In addition, to enhance the performance of the LQR controller, a feedforward term, accounting for external disturbances, i.e. mainstream entering flow and on-ramp flows, is designed, leading to an augmented LQR-FF control scheme. The LQR and LQR-FF controllers are tested and compared in the created realistic environment, demonstrating how IBC may operate in practice to combat traffic congestion on highways with unprecedented efficiency.

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

Περίληψη

Σχολή Μηχανικών Παραγωγής και Διοίκησης

Doctor of Philosophy

Έλεγχος Εσωτερικού Ορίου Κυκλοφορίας Αυτόματων Οχημάτων Χωρίς Λωρίδες

Του Milad Malekzadehkebria

Οι οδικές μεταφορές υφίστανται μια βαθιά μεταμόρφωση, που βασίζεται σε σημαντικές τεχνολογικές εξελίξεις αναφορικά με την οδική υποδομή, τα οχήματα και την διαχείριση της κυκλοφορίας. Σε αυτό το πλαίσιο, τα οχήματα επικοινωνούν αδιάλειπτα μεταξύ τους και με την υποδομή και καθοδηγούνται από δεδομένα πραγματικού χρόνου και προηγμένους αλγόριθμους, που τα επιτρέπουν να κινούνται αυτόνομα. Τα Συνδεδεμένα Αυτόματα Οχήματα (ΣΑΟ) είναι εξοπλισμένα με μια σειρά από αισθητήρες, κάμερες, Lidar, ραντάρ και συστήματα επικοινωνίας V2X (Vehicle-to-Everything). Αυτές οι τεχνολογίες επιτρέπουν στα οχήματα να «βλέπουν» το περιβάλλον τους, να ανταλλάσσουν πληροφορίες με κοντινά οχήματα, σήματα κυκλοφορίας και υποδομές και να λαμβάνουν έξυπνες αποφάσεις με βάση αυτή τη συλλογική ενημέρωση.

Μία από τις πολλές εξελίξεις σε αυτό το πλαίσιο είναι το πρότυπο TrafficFluid, μια εμπροσθομερης προσέγγιση στην οδική κινητικότητα που γίνεται ολοένα και πιο σημαντική στην εποχή των ΣΑΟ. Το TrafficFluid προτείνει την κυκλοφορία των οχημάτων χωρίς λωρίδες. Παραδοσιακά, οι δρόμοι οριοθετούνται με παράλληλες λωρίδες για να απλοποιηθεί η οδήγηση από ανθρώπους. Ωστόσο, το παράδειγμα χωρίς λωρίδα σηματοδοτεί μια απόκλιση από αυτό το άκαμπτο πλαίσιο. Αντί να περιορίζει τα οχήματα σε προκαθορισμένες λωρίδες, οραματίζεται ένα δυναμικό και προσαρμόσιμο κυκλοφοριακό οικοσύστημα.

Η κυκλοφορία χωρίς λωρίδες συνεπάγεται ότι η σταδιακή διαπλάτυνση (στένωση) του δρόμου οδηγεί σε αντίστοιχη σταδιακή αύξηση (μείωση) ικανότητας, και αυτό ανοίγει τον δρόμο για την εξέταση του Ελέγχου Εσωτερικού Ορίου (ΕΕΟ) σε πραγματικό χρόνο σε αυτοκινητόδρομους και αρτηρίες προκειμένου να μοιράζεται ευέλικτα το συνολικό πλάτος και ικανότητα του δρόμου (και στις δύο κατευθύνσεις) μεταξύ των δύο κατευθύνσεων σε εξάρτηση από την αμφίδρομη ζήτηση και συνθήκες κυκλοφορίας, έτσι ώστε να μεγιστοποιείται η συνολική απόδοση ροής (δύο κατευθύνσεων). Το πρόβλημα διατυπώνεται αρχικά στην παρούσα διατριβή ως ένα κυρτό πρόβλημα Τετραγωνικού Προγραμματισμού (ΤΠ), το οποίο μπορεί να λυθεί εύκολα και αποτελεσματικά. Επιπλέον, η νέα δράση ελέγχου διερευνάται διεξοδικά για διάφορες περιπτώσεις, οι οποίες αναδεικνύουν τα χαρακτηριστικά και τις δυνατότητες της. Επίσης, αυτή η διατριβή αντιμετωπίζει αποτελεσματικά τον ΕΕΟ, σε ένα δεύτερο βήμα, αναπτύσσοντας κατάλληλους Γραμμικούς-Τετραγωνικούς ρυθμιστές βασισμένους σε ανατροφοδότηση με ή χωρίς ολοκληρωμένη δράση (ρυθμιστές LQI και LQ). Οι προτεινόμενοι ρυθμιστές διερευνήθηκαν μέσω μακροσκοπικών προσομοιώσεων που περιλαμβάνουν διαφορετικά σενάρια ζήτησης σε ένα ρεαλιστικό τμήμα αυτοκινητόδρομου. Τα ευρήματα υποδεικνύουν ότι οι ρυθμιστές μπορούν να είναι εξίσου αποτελεσματικοί με μια λύση ανοιχτού βρόχου βέλτιστου ελέγχου με μη γραμμικούς περιορισμούς. Επιπλέον, οι προτεινόμενοι ρυθμιστές δεν απαιτούν

ακριβή μοντελοποίηση ή πρόβλεψη της εξωτερικής ζήτησης, επομένως είναι πιο εύρωστοι. Ωστόσο, οι κεντρικές λύσεις που απαιτούν πληροφορίες από ολόκληρο το τμήμα του αυτοκινητόδρομου που εξετάζεται μπορεί να είναι προβληματικές για μεγάλους αυτοκινητόδρομους σε σχέση με το εύρος της απαιτούμενης επικοινωνίας σε πραγματικό χρόνο. Σε ένα τρίτο βήμα αυτής της εργασίας, εξετάζονται δύο διαφορετικά σχήματα αποκεντρωμένου ελέγχου, συγκεκριμένα η επικαλυπτόμενης μέθοδος LQR και οι ανισότητες γραμμικού πίνακα (LMI) για τον ΕΕΟ σε κυκλοφορία ΣΑΟ χωρίς λωρίδες.

Ως τελευταίο βήμα, η παρούσα διατριβή ολοκληρώνει και επικυρώνει την λειτουργικότητα του ΕΕΟ με πολύ πιο ρεαλιστικό τρόπο μέσω μικροσκοπικής προσομοίωσης και ενεργού κίνησης εσωτερικών ορίων, χρησιμοποιώντας το εργαλείο προσομοίωσης TrafficFluid-Sim που βασίζεται στο SUMO. Για την υλοποίηση ΕΕΟ χρησιμοποιείται ένα σχήμα ελέγχου ανάδρασης LQR. Επιπλέον, για τη βελτίωση της απόδοσης του ελεγκτή LQR, σχεδιάστηκε ένας όρος εμπρόσθιας τροφοδοσίας, ο οποίος λαμβάνει υπόψη τις εξωτερικές διαταραχές, δηλαδή την ροή που εισέρχεται στο κυρίως ρεύμα του αυτοκινητόδρομου και τις ροές στις ράμπες εισόδου και οδηγεί σε ένα επαυξημένο σχήμα ελέγχου LQR-FF. Οι ελεγκτές LQR και LQR-FF διερευνώνται και συγκρίνονται σε ένα ρεαλιστικό περιβάλλον, αποδεικνύοντας πώς μπορεί να λειτουργήσει ο ΕΕΟ στην πράξη για την καταπολέμηση της κυκλοφοριακής συμφόρησης σε αυτοκινητόδρομους με πρωτόγνωρη αποτελεσματικότητα.

Acknowledgements

I am deeply indebted to my thesis advisors, Prof. Markos Papageorgiou and Prof. Ioannis Papamichail, for their unwavering support, expert guidance, and profound dedication throughout the entire research process. The shaping of this thesis was significantly influenced by their invaluable insights, profound knowledge, and constructive feedback. The mentorship and encouragement that they have provided me have been integral to my success and I am deeply grateful for their contribution to my life.

My heartfelt thanks go to my wife and my parents for their unwavering love, encouragement, and patience during this challenging academic journey. The support they have given me has been my bedrock, and I am immensely thankful for their understanding and trust in me.

I also extend a special note of appreciation to Dimitrios Troullinos for his generous support, which significantly contributed to the success of this thesis. His assistance has been invaluable.

To all those who have played a part in my academic journey, whether through direct collaboration, emotional support, or intellectual inspiration, I express my sincere gratitude. I would like to extend my heartfelt appreciation for the instrumental role you played in shaping this research and my development as a scholar.

This research received funding provided by the European Research Council under the European Union's Horizon 2020 Research and Innovation programme, ERC Grant Agreement n. [833915], for the project TrafficFluid. This funding played a pivotal role in facilitating this research and its outcomes.

Contents

List of Figures

List of Tables

1 Introduction and State of the Art

1.1 Introduction

1.1.1 Motivation

The honking of horns and the sight of bumper-to-bumper cars on urban freeways, highways, and arterials are a growing concern for most big cities around the world, leading to extended delays, higher fuel consumption, increased environmental pollution, and compromised traffic safety. The utilization of conventional traffic management techniques has proven to be highly effective in mitigating or preventing traffic congestion, making them a valuable tool. However, these approaches may not be able to cope with the demands leading to extremely congested traffic conditions.

As an example, that is most relevant with the scope of this work, elevating the throughput of the transportation infrastructure by reversing the direction of the lanes is one way to reduce traffic congestion without constructing additional roads in traditional traffic control schemes. Lane-based tidal flow control systems for conventional traffic may be very useful in certain situations, but they face a number of difficulties that limit their widespread use, such as:

- The resolution of infrastructure sharing among the two traffic directions cannot be higher than one lane, which may not be sufficiently fine-grained for many traffic situations.
- A reversible lane must extend over sufficient length (minimum of few kilometers) to mitigate counter-problems due to frequent merging or diverging of traffic streams.
- Whenever a reversible lane switch to the opposite direction is decided, a time-delay (corresponding to the travel time on the reversible lane) must be respected, before actually opening the lane to the opposite direction, so as to allow for the evacuation of the lane and avoid simultaneous opposite-direction movements. For the duration of this delay, the reversible lane is under-utilized, and this side-effect is an overhead that reduces the overall benefit of the measure. To mitigate the impact of this side-effect, the frequency of reversible lane switches must be limited (Frejo et al., 2015).

The fusion of forthcoming vehicle automation and communication systems with a smart road infrastructure presents a prospect for creating ingenious solutions that have the potential to transform the transportation industry . The past decade has witnessed an unprecedented effort by the automotive industry and research institutions to revolutionize vehicles using advanced automation and communication systems. These groundbreaking technologies are transforming the transportation industry, paving the way for a future where vehicles are safer, more efficient, and progressively autonomous. Modern vehicles are equipped with a range of features, from adaptive cruise control to self-driving capabilities, that improve safety, convenience, and driving experience. Through the utilization of artificial intelligence, sensor technologies, and advanced control algorithms, these automation systems empower vehicles to navigate roads, make informed decisions, and adapt to evolving conditions with unparalleled precision and efficiency.

The revolution in the transportation field has led to the recent launch of the TrafficFluid concept, which introduces a novel paradigm for vehicular traffic that is suitable for high penetration rates of vehicles equipped with advanced automation and communication systems (Papageorgiou et al., 2021), proposing the abandonment, in the age of connected automated vehicles (CAV), of a fundamental traditional constraint that stems from the physical perception limitation of human drivers. One of the main ideas behind the TrafficFluid concept is the elimination of fixed traffic lanes, allowing vehicles to move freely without being bound to specific lanes. A lane-free environment prompts vehicles to form ever-changing 2-D clusters through self-organization, which is influenced by several factors, including vehicle size, desired speed, movement strategies, and density, and which contributes to improved infrastructure utilization. Vehicles on a lane-free road adjust their movement to cover the expanded or narrowed 2-D road space, which can result in the formation of new clusters. As demonstrated by Papageorgiou et al. (2021), if the width of the road is increased by a certain amount, vehicles swiftly adjust their position to occupy the available space. This brings about an increase in the average inter-vehicle spacing, which in turn permits greater vehicle speeds, leading to higher flow and ultimately, higher capacity.

The concept of Internal Boundary Control (IBC) can be brought in here. It takes advantage of TrafficFluid's lane-free principle and the resulting road capacity property to show incremental changes when the road is widened or narrowed. While lane-based roads and traffic can only be modified by changing the road width by a lane or a lane-multiple, this is not the case in lane-free roads. Consider a road with two opposite traffic directions serving CAVs. The total available cross-road capacity (for both directions) may be shared between the two directions in a flexible way, according to the prevailing demand per direction, so as to maximize the infrastructure exploitation. Flexible capacity sharing may be achieved by virtually moving the internal boundary, which separates the two traffic directions, and communicating this decision to CAVs, so that they respect the changed road boundary. This way, the road width portion (and total capacity share) assigned to each direction can be changed in space and time (subject to constraints) according to an appropriate real-time control strategy.

In contrast to the lane reversal option, an environment for CAVs that is lane-free holds significant potential for successfully addressing the aforementioned difficulties. Specifically:

- The resolution of infrastructure sharing among the two directions can be high, still leading to corresponding intended capacity changes for the two opposite traffic streams.
- The potentially smooth (no lane changes) driving of CAV in a lane-free road surface would allow for the internal boundary to be a smooth space-function. This function may be softly changing in real time in response to the prevailing traffic conditions.
- Due to moderate changes of the internal boundary over time and space, CAVs may respect relatively promptly the changed boundary; hence the aforementioned safety-induced time-delay, required to avoid opposite movements on the same road surface, may be small.

Thanks to these characteristics, real-time IBC for lane-free CAV traffic has the potential to be broadly applicable to the high number of arterial or highway infrastructures that feature unbalanced demands during the day in the two traffic directions, so as to strongly mitigate or even avoid congestion.

1.1.2 Outline and Contribution

This section proceeds with an overview of the thesis. The literature review that is presented in Section 1.2 sheds light on the different lane-reversal methods and tidal flow policies that have been developed and implemented over time.

Chapter 2 presents a macroscopic model-based optimization scheme to elaborate on and demonstrate the characteristics of IBC. The well-known CTM (Daganzo, 1994) is employed to this end, leading to a convex Quadratic Programming (QP) problem. Carefully designed simulation scenarios highlight some very interesting and intriguing implications of this innovative control measure. Optimal internal boundary control scheme is rendered in section 2.2. The modified version of CTM, suited for lane-free characteristics, is outlined in this section. Subsequently, through the application of certain alterations, the nonlinear CTM model is converted into a linear system for the purpose of utilizing linear programming. The cost function is defined accordingly to incorporate QP in IBC. The performance of the proposed method is analyzed through two case studies, as introduced in Section 2.3. The initial scenario entails dissolvable congestion through IBC activation, whereas in the latter extreme scenario, IBC can only alleviate the congestion caused by excessive flow from both directions, exceeding highway capacity. Finally, the IBC's effectiveness is evaluated against the no-control case, and the simulation outcomes affirm the potential of the proposed approach.

Feedback-based Linear-Quadratic regulators with or without Integral action (LQI and LQ regulators) are appropriately developed in Chapter 3 to efficiently address the internal boundary control problem. Simulation investigations, involving a realistic highway stretch and different demand scenarios, demonstrate that the proposed simple regulators are robust and similarly efficient as an open-loop nonlinear constrained optimal control solution, while circumventing the need for accurate modelling and external demand prediction. Section 3.3 presents a comprehensive design for LQ regulators that incorporates relative density as a new variable in a lane-free paradigm, along with a detailed linearization process. The evidence presented in Section 3.4 indicates that the feedback controller outputs produced by LQR are closely correlated with the optimal QP solution.

Chapter 4 investigates two different overlapping decentralized control schemes for IBC in lane-free automated vehicle traffic. The first approach is based on a contractible controller, which is designed and employed in a decomposed way, for each subsystem of an extended highway system. In the second approach, the overall discrete-time LQ control problem is transformed into a linear matrix inequalities (LMI) problem. After selecting the overlapping feedback structure and solving the overall LMI problem, the correspondingly structured gain matrix is obtained, enabling a decentralized overlapping control deployment in real-time. Section 4.3 renders the details of two proposed overlapping control schemes. Initially Contractible controller design with

state and input inclusion is given and then overlapping control scheme using Linear Matrix Inequalities (LMIs) is rendered. Simulation investigations in section 4.4 first demonstrate the adequacy of previously developed centralized IBC approaches in sizable highway scenarios. Regarding the novel overlapping decentralized control schemes, it was found that, for the case of the uncongested demand scenario (where congestion is avoidable via IBC), the overlapping control schemes are as efficient as an open-loop optimal control solution (with perfect model knowledge and demand prediction) developed for the same problem by Malekzadeh et al. (2021a) using a convex QP problem formulation; and as efficient as the centralized LQR. On the other hand, for the more challenging case of the congested scenario (where congestion may be mitigated, but cannot be fully avoided via IBC), the LMI-based approach performs better than the contractible controller (, as it takes into account the whole system in the design phase.

Chapter 5 completes and validates the IBC concept in a much more realistic way via microscopic simulation and active internal boundary moving, using the SUMO-based *TrafficFluid-Sim* simulation tool. To effectuate IBC, a Linear Quadratic Regulator (LQR), which is a feedback control scheme, is employed. In addition, to enhance the performance of the LQR controller, a feedforward term, accounting for external disturbances, i.e. entering flow and on-ramp flows, is also designed, leading to an augmented LQR-FF control scheme. Section 5.2 renders details of LQR and LQR-FF design for IBC. Section 5.3 introduces lane-free microscopic implementation environment of IBC that is performed with TrafficFluid-Sim. Vehicle moving strategy based on double-double-integrator (DDI) model is given in section 5.4. The simulation results in section 5.5 confirm the efficiency and feasibility of IBC in a very realistic microscopic environment. Both controllers were found to act adequately to mitigate or avoid congestion creation in cases of unbalanced demands. LQR-FF was found to offer benefits versus LQR in better balancing the relative densities in both traffic directions, thus enlarging the margin to congestion creation in case of stronger demands.

Here below is a list of the publications that are encompassed within this thesis:

Journals:

- **Malekzadeh**, **M**., Papamichail, I., Papageorgiou, M. and Bogenberger, K., 2021. Optimal internal boundary control of lane-free automated vehicle traffic. *Transportation Research Part C: Emerging Technologies*, 126, p.103060.
- **Malekzadeh, M**., Papamichail, I. and Papageorgiou, M., 2021. Linear–quadratic regulators for internal boundary control of lane-free automated vehicle traffic. *Control Engineering Practice*, *115*, p.104912.
- **Malekzadeh**, **M**., Yanumula, V.K., Papamichail, I. and Papageorgiou, M., 2023. Overlapping internal boundary control of lane-free automated vehicle traffic. *Control Engineering Practice*, *133*, p.105435.
- **Malekzadeh**, **M**., Troullinos, D., Papamichail, I., Papageorgiou, M. and Bogenberger, K. Internal Boundary Control in Lane-free Automated Vehicle Traffic: Comparison of

Approaches via Microscopic Simulation. Transportation Research Part C: Emerging Technologies (revised).

- Yanumula, V.K., Typaldos, P., Troullinos, D., **Malekzadeh, M**., Papamichail, I. and Papageorgiou, M., 2023. Optimal trajectory planning for connected and automated vehicles in lane-free traffic with vehicle nudging. *IEEE Transactions on Intelligent Vehicles*, *8*(3), pp.2385-2399.

Conferences:

- **Malekzadeh, M.**, Papamichail, I. and Papageorgiou, M., 2021, June. Internal Boundary Control of Lane-free Automated Vehicle Traffic using a Linear Quadratic Integral Regulator. *In 2021 European Control Conference (ECC)* (pp. 35-41). IEEE.
- **Malekzadeh, M.,** Papamichail, I. and Papageorgiou, M., 2021. Internal Boundary Control of Lane-Free Automated Vehicle Traffic using a Model-Free Adaptive Controller. *IFAC-PapersOnLine*, *54*(2), pp.99-106.
- **Malekzadeh, M.**, Papamichail, I., Papageorgiou, M. and Bogenberger, K., 2021, June. Optimal Control of Internal Road Boundary for Lane-free Automated Vehicle Traffic. In *2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)* (pp. 1-6). IEEE.
- **Malekzadeh, M.**, Papamichail, I. and Papageorgiou, M., 2022, June. Overlapping Internal Boundary Control of Lane-free Automated Vehicle Traffic with State and Input Inclusion. In *2022 30th Mediterranean Conference on Control and Automation (MED)* (pp. 1066-1073). IEEE.
- **Malekzadeh, M.**, Manolis, D., Papamichail, I. and Papageorgiou, M., 2022, October. Empirical investigation of properties of lane-free automated vehicle traffic. In *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)* (pp. 2393-2400). IEEE.
- **Malekzadeh, M.**, Yanumula, V.K., Papamichail, I., Papageorgiou, M.,2023. June. Overlapping internal boundary control of lane-free automated vehicle traffic – An LMI approach. *8th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS 2023)*.
- **Malekzadeh, M.**, Troullinos, D., Papamichail, I., Papageorgiou, M.,2023, September. Microscopic simulation-based testing of internal boundary control of lane-free automated vehicle traffic. *11th Symposium of the European Association for Research in Transportation (hEART 2023)*.
- Yanumula, V.K., Typaldos, P., Troullinos, D., **Malekzadeh, M**., Papamichail, I. and Papageorgiou, M., 2021, September. Optimal path planning for connected and automated vehicles in lane-free traffic. *In 2021 IEEE International Intelligent Transportation Systems Conference (ITSC)* (pp. 3545-3552). IEEE.

- Jin, X., Yu, X., Hu, Y., Wang, Y., Papageorgiou, M., Papamichail, I., **Malekzadeh**, **M**. and Guo, J., 2022, October. Integrated Control of Internal Boundary and Ramp Inflows for Lane-free Traffic of Automated Vehicles on Freeways. In *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)* (pp. 1234-1239). IEEE.
- Papamichail, I., Schoenn-Anchling, N., **Malekzadeh, M.,** Markantonakis, V., Papageorgiou, M.: Macroscopic traffic flow model calibration for lane-free automated vehicle traffic. *In 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC).* IEEE.

1.2 State-of-the-Art

1.2.1 Lane reversal control

The idea of sharing the total road capacity among the two traffic directions is not new and has been occasionally employed for conventional lane-based traffic in many countries, typically offline or manually (Wolshon and Lambert, 2004). The measure is known as tidal flow or reversible lane control, and its main principle is to adapt the total available cross-road supply to the bi-directional demand. Its most basic form is the steady allocation of one (or more) lanes of one direction to the opposite direction for a period of time (ranging from few hours to many days) in the aim of addressing abnormal traffic supply or demand. This happens indeed often at work zones, in order to compensate for the capacity loss in one direction due to road works; or in cases of big events (sport events, concerts, holiday departure or return, evacuation etc.) due to excessive demand in one traffic direction, while demand in the opposite direction is low. More advanced reversible lane systems may operate in real time, see (Frejo et al., 2015; Ma, et al., 2018; Ampountolas et al., 2020), to balance delays on both sides of a known bottleneck (e.g. bridge or tunnel) by assigning a lane to the each of the two directions in alternation in response to the prevailing traffic conditions. To this end, optimal control or feedback control algorithms of various types are proposed by Frejo et al. (2015), Ma et al. (2018) and Ampountolas et al. (2020).

Reversible lanes have also been considered in connection with lane-based CAV driving. To start with, Hausknecht et al. (2011) formulate an integer linear problem and find the lane configuration that is able to serve the maximum amount of traffic a network can handle satisfying constant demand for each origin-destination pair. They also formulate a bi-level programming problem, whereby the upper level includes the allocation of capacity to each of the links, while the lower-level addresses User Equilibrium (UE) conditions; and employ genetic algorithms for the solution. Duell et al. (2015; 2016) use the system optimal dynamic traffic assignment models formulated by Ziliaskopoulos (2000) for a single destination and by Li et al. (2003) for more general networks, using the Cell Transmission Model (CTM) by Daganzo (1994). Lanes are reflected with integer variables, and the problem is formulated as a mixed integer linear programming (MILP) problem that has, however, high (exponential) complexity due to the many integer variables involved. Levin and Boyles (2016) use this model for a single

link and utilize stochastic demand as a Markov decision process. The MILP problem is solved using a heuristic and is incorporated within a UE routing problem. Finally, Chu et al. (2019) address the reservation-based routing and scheduling of CAV, considering dynamically reversible lanes, as an integer linear program.

1.2.2 Overlapping control

Generally speaking, control decentralization in large-scale and geographically extended systems is a desirable and reasonable control structure, as it is simpler to monitor and maintain and more reliable in cases of device failures. Such features appear, for example, in the control of interconnected power systems (Patil et al. 2019), water systems (Hafeez et al. 2021), as well as in space or traffic systems (Pasquale et al. 2020), where analysis, synthesis and implementation is distributed to interacting subsystems (Siljak., 2011) under the motto "think global, act local". Such systems are typically composed of various physically or functionally local, possibly interacting control sub-stations, where each sub-station is in charge only of the operation of a part of the overall system, and the overall control system objective is reached by the actions of the local control subsystems.

Various methodologies were proposed in the past decades for the design of such decentralized control systems, the structure of which may be dictated by the structure of the process under control; or may have to be defined based on entailed performance, maintenance and communication requirements, reliability and other properties. In the latter case, there are four well-known system decomposition structures: disjoint decomposition, overlapping decomposition, border block diagonal decomposition, and epsilon decomposition (Bakule, 2014). The disjoint decomposition of the overall system is not reasonable if the subsystems are strongly interconnected, in which case an overlapping control structure may be employed, where the overlapping parts create connections between subsystems to enhance the local and overall control performance. Several overlapping control methods and applications have been proposed, see for example (Trudnowski and Pierre, 1992), (Bakule et al., 2006), (Hug-Glanzmann and Andersson, 2009), (Ahmadi and Aldeen, 2017).

2 Optimal IBC of lane-free automated vehicle traffic

2.1 Abstract

A recently proposed paradigm for vehicular traffic in the era of CAV (connected and automated vehicles), called TrafficFluid, involves lane-free vehicle movement. Lane-free traffic implies that incremental road widening (narrowing) leads to corresponding incremental increase (decrease) of capacity; and this opens the way for consideration of real-time internal boundary control on highways and arterials, in order to flexibly share the total (both directions) road width and capacity among the two directions in dependence of the bi-directional demand and traffic conditions, so as to maximize the total (two directions) flow efficiency. The problem is formulated as a convex QP (Quadratic Programming) problem that may be solved efficiently, and representative case studies shed light on and demonstrate the features, capabilities and potential of the novel control action.

2.2 Introduction

Vehicular traffic is crucial for the transport of persons and goods, but traffic congestion, which appears on a daily basis, particularly in and around metropolitan areas, around the world has been an increasingly serious problem that calls for drastic solutions. Traffic congestion causes substantial delays, excessive environmental pollution and reduced traffic safety. Conventional traffic management measures are valuable (Papageorgiou et al., 2003; Papageorgiou et al., 2007; Kurzhanskiy and Varaiya, 2010), but not always sufficient to tackle the heavily congested traffic conditions, which must be addressed in a more comprehensive way that exploits gradually emerging and future ground-breaking new capabilities of vehicles and the infrastructure.

During the last decade, there has been an enormous effort by the industry and by numerous research institutions to develop and deploy a variety of vehicle automation and communication systems that are revolutionizing the vehicle capabilities. Vehicle automation ranges from different kinds of driver support to highly or fully automated driving; and vehicle communication enables V2V (vehicle-to-vehicle) and V2I (vehicle-to-infrastructure) communication that may support various potential applications. Many automotive and information-technology companies, as well as research institutions, have been developing and testing in real traffic conditions high-automation or virtually driverless autonomous vehicles that monitor their environment and make sensible driving decisions based on appropriate decision and control methods (Ardelt et al., 2012; Aeberhard et al., 2015; Kamal et al., 2016; Makantasis and Papageorgiou, 2018).

A recent paper (Papageorgiou et al., 2021) launched the TrafficFluid concept, which is a novel paradigm for vehicular traffic, applicable at high levels of vehicle automation and communication (SAE levels 4 or 5) and high penetration rates that are expected to prevail in the not-too-far future. The TrafficFluid concept is based on the following two combined principles: (1) *Lane-free traffic*, whereby vehicles are not bound to fixed traffic lanes, as in conventional traffic, but may drive anywhere on the 2-D surface of the road; (2) *Vehicle nudging*, whereby vehicles communicate their presence to other vehicles in front of them (or are sensed by them), and this may exert a “nudging” effect on the vehicles in front, i.e. vehicles in front may, under

appropriate circumstances, experience (apply) a pushing influence. Several advantages and challenges related to this novel traffic paradigm are discussed by Papageorgiou et al. (2021) (see also (Papageorgiou et al. 2019) for more modelling details).

This paper exploits the lane-free property of TrafficFluid, i.e. the possibility for vehicles to drive on the 2-D road surface without being bound to lanes. As demonstrated in a small experiment by Papageorgiou et al. (2021) and is also intuitively sensible, lane-free traffic implies that the traffic capacity may exhibit incremental (increasing or decreasing) changes in response to corresponding incremental (widening or narrowing) changes of the road width. This is in contrast to lane-based roads and traffic, where capacity changes may only occur if the road width is changed by lane "quanta".

To illustrate this feature of lane-free connected and automated vehicle (CAV) traffic, the snapshot of a lane-free microscopic simulation video using an ad-hoc vehicle movement strategy (see (Papageorgiou et al., 2019) for modelling details) in Fig. 2.1 illustrates that, on a highway with constant road width, vehicles, driven by their smooth (no lane changes) 2-D movement strategy, spread on the available 2-D road surface without necessarily forming lanes, except, at times, along the road boundaries, so as to maximize the infrastructure utilization. If the road would be widened by some amount, vehicles would immediately move laterally to cover the free space. This would increase the average inter-vehicle spacing, thus allowing for higher vehicle speeds and hence higher flow and capacity.

Figure 2.1: Snapshot of microscopic simulation with lane-free CAV traffic

Consider a road or highway with two opposite traffic directions, where CAV are driving. The total carriageway width, and hence the total capacity (for both directions), could be shared among the two directions in a flexible way, according to the prevailing demand, so as to maximize the infrastructure exploitation. Flexible width and capacity sharing may be achieved via virtual moving of the internal boundary, which separates the two traffic directions, and corresponding communication to the CAV to respect the changed internal boundary. Such internal boundary shift includes shifting of possible safety margins among the two directions. This way, the carriageway's width portion (and corresponding total capacity share) assigned to each traffic direction can be changed in space and time (subject to constraints) according to an appropriate control strategy, as illustrated in Fig. 2.2, so as to maximize the total traffic efficiency in both directions.

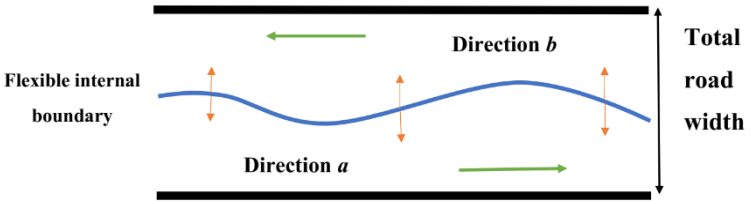


Figure 2.2: Space-time flexible internal road boundary..

The idea of sharing the total road capacity among the two traffic directions is not new and has been occasionally employed for conventional lane-based traffic in many countries, typically offline or manually (Wolshon and Lambert, 2004). The measure is known as tidal flow or reversible lane control, and its main principle is to adapt the total available cross-road supply to the bi-directional demand. Its most basic form is the steady allocation of one (or more) lanes of one direction to the opposite direction for a period of time (ranging from few hours to many days) in the aim of addressing abnormal traffic supply or demand. This happens indeed often at work zones, in order to compensate for the capacity loss in one direction due to road works; or in cases of big events (sport events, concerts, holiday departure or return, evacuation etc.) due to excessive demand in one traffic direction, while demand in the opposite direction is low. More advanced reversible lane systems may operate in real time, see (Frejo et al., 2015; Ma, et al., 2018; Ampountolas et al., 2020), to balance delays on both sides of a known bottleneck (e.g. bridge or tunnel) by assigning a lane to the each of the two directions in alternation in response to the prevailing traffic conditions. To this end, optimal control or feedback control algorithms of various types are proposed by Frejo et al. (2015), Ma et al. (2018) and Ampountolas et al. (2020).

Reversible lanes have also been considered in connection with lane-based CAV driving. To start with, Hausknecht et al. (2011) formulate an integer linear problem and find the lane configuration that is able to serve the maximum amount of traffic a network can handle satisfying constant demand for each origin-destination pair. They also formulate a bi-level programming problem, whereby the upper level includes the allocation of capacity to each of the links, while the lower level addresses User Equilibrium (UE) conditions; and employ genetic algorithms for the solution. Duell et al. (2015; 2016) use the system optimal dynamic traffic assignment models formulated by Ziliaskopoulos (2000) for a single destination and by Li et al. (2003) for more general networks, using the Cell Transmission Model (CTM) by Daganzo (1994). Lanes are reflected with integer variables, and the problem is formulated as a mixed integer linear programming (MILP) problem that has, however, high (exponential) complexity due to the many integers variables involved. Levin and Boyles (2016) use this model for a single link and utilize stochastic demand as a Markov decision process. The MILP problem is solved using a heuristic and is incorporated within a UE routing problem. Finally, Chu et al. (2019) address the reservation-based routing and scheduling of CAV, considering dynamically reversible lanes, as an integer linear program.

Lane-based tidal flow control systems for conventional (manually driven vehicle) traffic may be very useful in certain situations, but they face a number of difficulties that limit their widespread use:

- The resolution of infrastructure sharing among the two traffic directions cannot be higher than one lane, which may not be sufficiently fine-grained for many traffic situations.
- A reversible lane must extend over sufficient length (minimum of few kilometers) to mitigate counter-problems due to frequent merging or diverging of traffic streams.
- Whenever a reversible lane switch to the opposite direction is decided, a time-delay (corresponding to the travel time on the reversible lane) must be respected, before actually opening the lane to the opposite direction, so as to allow for the evacuation of the lane and avoid simultaneous opposite-direction movements. For the duration of this delay, the reversible lane is under-utilized, and this side-effect is an overhead that reduces the overall benefit of the measure. To mitigate the impact of this side-effect, the frequency of reversible lane switches must be limited (Frejo et al., 2015).

Due to these limitations, reversible lane control has not evolved as a major traffic management measure, even less so in real time. Even for future CAV traffic, however, some of the mentioned difficulties would persist in lane-based conditions, notably the low capacity sharing resolution, the merging nuisance and, last but not least, the complex (integer-based) nature of the corresponding optimization problems that hinder real-time feasibility.

In contrast, in a lane-free CAV traffic environment, there is a clear prospect for the mentioned difficulties to be largely mitigated. Specifically:

- The resolution of infrastructure sharing among the two directions can be high, still leading to corresponding intended capacity changes for the two opposite traffic streams.
- The potentially smooth (no lane changes) driving of CAV in a lane-free road surface would allow for the internal boundary to be a smooth space-function, as illustrated in Fig. 2.2. This function may be softly changing in real time in response to the prevailing traffic conditions.
- Due to moderate changes of the internal boundary over time and space and the lack of physical boundary, CAV may respect relatively promptly the changed boundary; hence the aforementioned safety-induced time-delay, required to avoid opposite movements on the same road surface, may be small.
- As practiced in this paper, the resulting optimization problems include only real-valued variables (no integers are necessary) and may therefore be solved very efficiently, so as to be readily real-time feasible.

Thanks to these characteristics, real-time internal boundary control for lane-free CAV traffic has the potential to be broadly applicable to the high number of arterial or highway infrastructures that feature unbalanced demands during the day in the two traffic directions, so as to strongly mitigate or even avoid congestion. Even for infrastructures experiencing strong demand in both directions during the peak periods, real-time internal boundary control may intensify the road utilization and lead to sensible improvements. In fact, the possibility to control the internal boundary in real time may prove to be one of the major advantages of the TrafficFluid concept.

This paper proposes a macroscopic model-based optimization scheme to elaborate on and demonstrate the characteristics of internal boundary control. The well-known CTM (Daganzo, 1994) is employed to this end, leading to a convex Quadratic Programming (QP) problem. Carefully designed simulation scenarios highlight some very interesting and intriguing implications of this innovative control measure. Section 2.2 presents the general problem

statement, followed by the CTM presentation, the transformation of CTM equations to linear equality and inequality constraints, and the proposed quadratic cost function. Two considered case studies are introduced in Section 2.3, along with the obtained results. Conclusions work are summarized in Section 2.4.

2.3 Optimal internal boundary control scheme

2.3.1 General modelling and problem statement

The low lane discipline and vehicle size diversity encountered in several developing countries motivated, in the last few years, some microscopic modelling works, which proposed, using various approaches, models for heterogeneous and lane-less traffic; similar developments were triggered by the recent interest in shared spaces, which are used by pedestrians and vehicles of various types; see e.g. (Manjunatha et al., 2013; Rudloff et al., 2013); see also (Kanagaraj and Treiber, 2018; Mulla et al., 2018), where such models are proposed, validated with real traffic data and analyzed with respect to stability and other properties. Clearly, these modelling works attempt to describe the driving behavior of real vehicles and drivers. In the case of lane-free CAV traffic, as proposed in the TrafficFluid concept, we need to design (rather than model) opportune movement strategies for safe and efficient traffic flow. Such an ad-hoc CAV movement strategy was reported in (Papageorgiou et al., 2019), but more systematic developments in this direction are in progress.

Lane-free traffic is not expected to give rise to structural changes of existing macroscopic models. It is reasonable to assume, as also supported by results in (Bhavathrathan and Mallikarjuna, 2012; Asaithambi et al., 2016; Munigety et al., 2016; Papageorgiou et al., 2019), that notions and concepts like the conservation equation, the fundamental diagram, as well as moving traffic waves will continue to characterize macroscopic traffic flow modelling in the case of lane-free CAV traffic. By the same token, specific physical traffic parameters, such as free speed, critical density, flow capacity, jam density, are also relevant for lane-free traffic, but may of course take different values than in lane-based traffic. However, the exact values that these parameters will take in lane-free CAV traffic are of minor importance for the presented demonstration of a novel control measure.

In the present context, it is crucial to elaborate on the impact of internal boundary control on the respective Fundamental Diagrams (FDs) of the two opposite traffic directions, which we will call directions a and b, respectively (Fig. 2.2). Let us assume that directions a and b are assigned respective road widths (in m) $w^a = \varepsilon \cdot w$ and $w^b = (1-\varepsilon) \cdot w$, where $0 \le \varepsilon \le 1$ is the sharing factor and w is the total road width (both directions). Let $Q(\rho)$, where ρ is the traffic density in veh/km, be the total FD (both directions) of a highway section, which would prevail if the whole carriageway would be assigned to only one of the two opposite traffic directions (i.e. for ε equal to 0 or 1), with total critical density ρ_{cr}, total capacity q_{cap} and total jam density $\rho_{\max}$. Let us consider the case of partial road sharing, i.e. $\varepsilon_{\min} \le \varepsilon \le \varepsilon_{\max}$, where $\varepsilon_{\min}, \varepsilon_{\max} \in (0,1)$ are appropriate bounds to be specified later. We want to derive the corresponding FDs and parameter values for the two directions a and b. For easier understanding, let us assume for a moment that w is the total number of lanes and $\varepsilon \cdot w$, $(1-\varepsilon) \cdot w$ are the respective integer numbers of lanes

assigned to the two directions. Then, $Q(\rho^1 \cdot w)/w$ is the FD per lane, where ρ^1 is the density per lane, and, analogously, the FDs for the two directions, which are functions of ε, are given by

$$Q^a(\rho^a, \varepsilon) = \varepsilon \cdot Q(\rho^a / \varepsilon),\ Q^b(\rho^b, \varepsilon) = (1-\varepsilon) \cdot Q(\rho^b / (1-\varepsilon)), \tag{2.1}$$

where ρ^a and ρ^b (in veh/km) are the respective densities of the two directions. Clearly, (2.1) applies in the exact same way if we return to the original meaning of ε as the real-valued sharing factor.

We now wish to derive the shared critical density $\rho^a_{cr}(\varepsilon)$, capacity $q^a_{cap}(\varepsilon)$, and jam density $\rho^a_{\max}(\varepsilon)$ for direction a as functions of the sharing factor ε. To this end, we need the derivative $\partial Q^a(\rho^a, \varepsilon)/\partial\rho = Q^{a\prime}(\rho^a, \varepsilon)$, for which, using (2.1), we obtain $Q^{a\prime}(\rho^a, \varepsilon) = Q'(\rho^a / \varepsilon)$. For the critical density, this derivative equals zero, i.e. we have $Q'(\rho^a_{cr} / \varepsilon) = 0$, but we also have $Q'(\rho_{cr}) = 0$, hence we deduce that $\rho^a_{cr}(\varepsilon) = \varepsilon \cdot \rho_{cr}$. For the capacity, we have $q^a_{cap}(\varepsilon) = Q^a(\rho^a_{cr}, \varepsilon) = \varepsilon \cdot Q(\rho_{cr}) = \varepsilon \cdot q_{cap}$. Finally, we have $Q(\rho_{\max}) = Q^a(\varepsilon.\rho_{\max}, \varepsilon) = 0$, and hence $\rho^a_{\max}(\varepsilon) = \varepsilon \cdot \rho_{\max}$. Similarly, we have $\rho^b_{cr}(\varepsilon) = (1-\varepsilon) \cdot \rho_{cr}$, $q^b_{cap}(\varepsilon) = (1-\varepsilon) \cdot q_{cap}$ and $\rho^b_{\max}(\varepsilon) = (1-\varepsilon) \cdot \rho_{\max}$ for direction b.

In short, the sharing factor ε scales density (veh/km) and flow (veh/h), leaving unaffected the speed (km/h). The derived equations for the FD parameters can be easily shown to hold also for FDs that are not differentiable at ρ_{cr}, such as the triangular FD, as we will see later, see also (Ampountolas et al., 2020).

The above derivations rely on two main assumptions. The first assumption, which is the most crucial one, is that any incremental widening (narrowing) of the road width entails a corresponding incremental increase (decrease) of capacity; or, in other words, that the capacity is a monotonic continuous function of the sharing factor, i.e. we have $q^a_{cap}(\varepsilon)$ and $q^b_{cap}(\varepsilon)$ as continuous functions. Indeed, the highway may hold vehicles of different dimensions and speeds. These vehicles occupy, in a lane-free structure, the road surface according to their movement strategies, which generate a "spread" of lateral vehicle positions, as illustrated earlier (Fig. 2.1). Thus, every incremental widening of the road increases the average inter-vehicle spacing and offers possibilities for higher speed, and hence higher flow and capacity. This is evidenced in FDs collected from microscopic simulation experiments and reported in (Papageorgiou et al., 2019). It should be noted that the capacity change, entailed by road width change, may be stochastic in practice and even subject to occasional random discontinuities due to the virtually unpredictable exact outcome of 2-D vehicle movement on the 2-D road surface. However, using continuous functions $q^a_{cap}(\varepsilon)$ and $q^b_{cap}(\varepsilon)$ is deemed a reasonable modelling idealization, like many others in traffic flow and other kinds of modelling.

The second, less crucial assumption is that the functions $q^a_{cap}(\varepsilon)$ and $q^b_{cap}(\varepsilon)$ are linear and symmetric, i.e. that have $q^a_{cap}(\varepsilon) = q^b_{cap}(1-\varepsilon)$. Linearity seems reasonable, and FDs reported in (Papageorgiou et al., 2019) verify it, as capacity was found to increase roughly proportionally to the sharing factor. Note, however, that the relations have $q^a_{cap}(\varepsilon)$ and $q^b_{cap}(\varepsilon)$ depend on the 2-D CAV movement strategy employed, and functional forms other than linear cannot be excluded. On the other hand, symmetry is present if the total FD $Q(\rho)$ is the same in either direction. Due to road grade or other reasons, we may sometimes have asymmetric cases with the total FD

being different, i.e. $Q^A(\rho)$ in direction a and $Q^B(\rho)$ in direction b. Cases involving nonlinear or non-symmetric functions $q_{cap}^a(\varepsilon)$ and $q_{cap}^b(\varepsilon)$ call for a slightly generalized treatment along the same lines, but are not considered here for simplicity.

A general macroscopic dynamic traffic model for internal boundary control may be expressed in discrete-time state-space form as follows:

$$\mathbf{x}^a(k+1) = \mathrm{f}^a[\mathbf{x}^a(k), \boldsymbol{\varepsilon}(k_c), \mathbf{d}^a(k)], \tag{2.2}$$

$$\mathbf{x}^b(k+1) = \mathrm{f}^b[\mathbf{x}^b(k), \boldsymbol{\varepsilon}(k_c), \mathbf{d}^b(k)], \tag{2.3}$$

where $\mathbf{x}^a$ and $\mathbf{x}^b$ are the state vectors for traffic directions a and b, respectively, comprising section-based traffic densities and, in case of second-order models, also mean speeds; the model time step is typically 5 – 10 s for section lengths of some 500 m, and $k = 0,1,\ldots$ is the corresponding discrete time index; $\mathbf{d}^a$ and $\mathbf{d}^b$ are vectors of external variables in the respective traffic directions a and b (upstream mainstream demand, on-ramp flows etc.); and $\boldsymbol{\varepsilon}$ is the vector of the sharing factors (control variables), one per section. The control time step T_c does not need to be equal to the model time step T, but is assumed to be a multiple of T, in which case, the control time index is given by $k_c = \lfloor kT/T_c \rfloor$, where $\lfloor . \rfloor$ is the integer part notation. It is noted that the notation $\varepsilon(k_c)$ indicates that the specific sharing factor is applied for the duration of the control time interval $[k_c \cdot T_c, (k_c+1) \cdot T_c)$. Some constraints regarding the control input will be detailed in the next section. The control inputs ε_i, one per section i, will need to be spatially smoothed prior to their actual application to the road width. This may be done via appropriate spline interpolation, the details of which will be elaborated in future work, which will test the presented concept via microscopic simulation, where safety as well as convenience implications will be carefully addressed.

Once a dynamic model is selected, the general state-space equations (2.2), (2.3) may be used to evaluate different control strategies for specific scenarios. Beyond the road's traffic characteristics, a scenario over a time horizon K is defined by the initial state $\mathbf{x}^a(0)$, $\mathbf{x}^b(0)$ and the trajectories of the external variables $\mathbf{d}^a(k)$, $\mathbf{d}^b(k)$, $k = 0,1,\ldots,K-1$. The control variables determine the share of the overall road width (and capacity) among the two opposite directions as a function of space (sections) and time (control time steps). Evaluation of the quality of a control strategy calls for the specification of an objective criterion to be minimized, as will be detailed in Section 2.2.3.

2.3.2 CTM-based optimal control problem

2.3.2.1 CTM relevance

While other control design approaches are possible and indeed in preparation, we have opted, as a first approach, to formulate this novel traffic control action as a model-based optimal control problem. Optimization methods, when properly applied, are known to deliver solutions that may reveal novel control aspects (Papageorgiou, 1997), something that is particularly interesting in the case of still unexplored control measures, like internal boundary control.

Various dynamic traffic flow models have been employed for optimal control problems, among which a simple but realistic possibility is CTM (Daganzo, 1994); see (Ziliaskopoulos, 2000; Gomes and Horowitz, 2006; Roncoli et al., 2015) for CTM-based optimal control formulations (among many others). CTM is a first-order model deriving from the LWR model (Lighthill and Whitham, 1955; Richards, 1955), which attains a space-time discretized form by application of the Godunov numerical scheme. CTM employs a triangular FD, and its main advantage, when used within an optimal control setting, is that it may lead to a convex, hence globally optimizable, linear or quadratic programming problem, which can be solved numerically using very efficient available codes. The reason behind this property is that the nonlinearities that every traffic flow model must necessarily feature to realistically reflect the traffic flow dynamics, have, in CTM, a piecewise linear form that is amenable to linear constraints for the optimization problem; and, hence, to a convex admissible region. Thus, even large-scale traffic control problems may be solved very efficiently computationally, something that allows for real-time model predictive control (MPC). MPC is widely applied in many control application areas; and is employed, in particular, in hundreds of published works on traffic control (Burger et al., 2013). MPC's feedback structure allows for correction of inaccuracies of the employed model and demand predictions over time. MPC is not directly considered in this work, as our primary focus is to develop its kernel optimal control problem for internal boundary control.

The LWR model, and hence also CTM, are strictly based on the following three principles:

1. The conservation-of-vehicles, which is a universal physical law applying to any kind of fluid and certainly also to lane-free CAV traffic.

2. The FD with a characteristic inverse-U shape (triangular for CTM). As explicitly evidenced in (Papageorgiou et al., 2019), such a flow-density relation applies indeed also to lane-free CAV traffic.

3. The transport equation $q = \rho v,$ which is also a universal physical law applying to any kind of fluid and certainly also to lane-free CAV traffic.

Given these facts, CTM appears to offer a good basis for our purposes.

In the following, we will use CTM, appropriately adjusted to incorporate the internal boundary control action, so as to cast the control problem in the form of a convex Quadratic Programming (QP) problem.

2.3.2.2 CTM equations

Consider a highway stretch holding two opposite traffic directions a (from left to right) and b (from right to left). The stretch is subdivided in n road sections, each some 500 m in length. As explained in the previous section, the total road width, which is assumed constant over all sections for simplicity, can be flexibly shared among the two directions in real time. As the sharing may be different for every section, we have corresponding sharing factors $\varepsilon_i,\ i = 1, 2, ..., n$; and (2.1) applies to each section. As a consequence, the total section capacity, as well as the critical density and jam density, are shared among traffic directions a and b according to

$$q_{i,cap}^{a}(\varepsilon_i)=\varepsilon_i\cdot q_{cap},\ q_{i,cap}^{b}(\varepsilon_i)=(1-\varepsilon_i)\cdot q_{cap},$$
$$\rho_{i,cr}^{a}(\varepsilon_i)=\varepsilon_i\cdot \rho_{cr},\ \rho_{i,cr}^{b}(\varepsilon_i)=(1-\varepsilon_i)\cdot \rho_{cr}, \tag{2.4}$$
$$\rho_{i,\max}^{a}(\varepsilon_i)=\varepsilon_i\cdot \rho_{\max},\ \rho_{i,\max}^{b}(\varepsilon_i)=(1-\varepsilon_i)\cdot \rho_{\max}.$$

The corresponding changes of the triangular FD that may occur at each section and traffic direction are illustrated in Fig. 2.3. More specifically, when the value of control input is 0.5, i.e., the flow capacities of the two directions are equal, their FDs are "nominal" (blue line with $(.)^N$ parameters); when the control input is different than 0.5, we have two FDs: the extended one (green line with $(.)^E$ parameters) applies to the direction that is assigned higher width, and the reduced, complementary FD (orange line with $(.)^R$ parameters) applies to the other direction that is assigned less flow capacity. Based on (4), all FD parameters of a section change, whenever it is decided to change the corresponding sharing factor in real time, i.e. at discrete times $k_c=1,2,\ldots..$

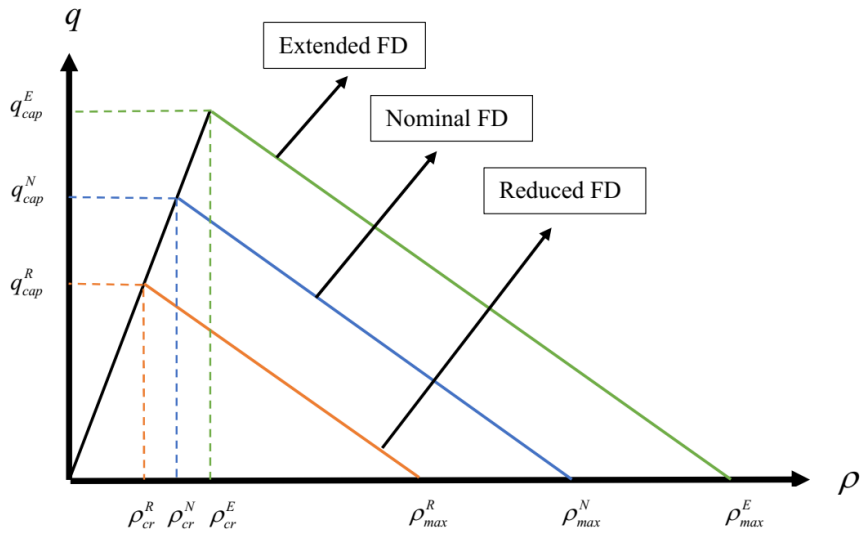


Figure 2.3: The triangular fundamental diagram with flexible internal boundary.

By their nature, the sharing factors take values $\varepsilon_i\in(0,1)$. However, for the internal boundary control problem, we would like to disallow the utter closure of either direction; hence, the assigned road width in either direction should never be smaller than the widest vehicles driving on the road. This requirement gives rise to stricter constraints for the sharing factors as follows

$$0<\varepsilon_{i,\min}\le\varepsilon_i\le\varepsilon_{i,\max}<1, \tag{2.5}$$

where $\varepsilon_{i,\min}\cdot w$ and $(1-\varepsilon_{i,\max})\cdot w$ are the minimum admissible widths to be assigned to directions a and b, respectively. If the two minimum widths are equal, then we have $\varepsilon_{i,\min}+\varepsilon_{i,\max}=1$. The minimum admissible width does not necessarily define a lane. The highway holds vehicles of

different dimensions (trucks, small and big cars, vans, buses, motorcycles), the widest of which should be able to drive on a road with width equal to the minimum admissible width.

Another restriction to be applied to the sharing factors concerns the time-delay needed to evacuate traffic on the direction that receives a restricted width, compared with the previous control time step. As discussed earlier, this time-delay may be considerable in the case of long reversible lanes in lane-based traffic with physical reversible lane separation (Frejo et al., 2015). In contrast, this time-delay is much smaller in lane-free CAV traffic with moderate changes of the sharing factors that are applied to short sections; but needs nevertheless to be considered. Clearly, the time-delay should apply only to the traffic direction that is being widened, compared to the previous control interval; while the direction that is restricted should promptly apply the smaller width, so that CAVs therein may move out of the reduced-width zone. Assume that the required time-delay is smaller than or equal to the control time interval T_c; then, the time-delay requirement is automatically fulfilled for each section i, if the sharing factors that are actually applied to the two directions, i.e. ε_i^a and ε_i^b, respectively, are calculated as follows

$$\varepsilon_i^a(k_c) = \min\left\{\varepsilon_i(k_c), \varepsilon_i(k_c - 1)\right\}, \tag{2.6}$$

$$\varepsilon_i^b(k_c) = \min\left\{1 - \varepsilon_i(k_c), 1 - \varepsilon_i(k_c - 1)\right\}. \tag{2.7}$$

These equations may be readily extended if the required time-delay is a multiple of the control time interval T_c.

We are now ready to present the CTM equations, considering the changing sharing factors and their constraints. We recall that we consider a highway stretch with n sections, with respective lengths L_i. Traffic flows from section 1 to section n for direction a; and from section n to section 1 for direction b (see Fig. 2.4 for an example). We denote ρ_i^a, $i = 1, 2, ..., n$, the traffic density of section i, direction a; and ρ_i^b, $i = 1, 2, ..., n$, the traffic density of section i, direction b. Similarly, we have the mainstream exit flows of section i being denoted q_i^a for direction a and q_i^b for direction b. Thus, q_0^a is the feeding upstream mainstream inflow for direction a; and q_{n+1}^b is the feeding upstream mainstream inflow for direction b. In addition, every section may have an on-ramp or an off-ramp at its upstream boundary. The on-ramp flow (if any) for section i, direction a, is denoted r_i^a; and the on-ramp flow (if any) for section i, direction b, is denoted r_i^b. The off-ramp flow (if any) of section i, direction a, is calculated based on known exit rates β_i^a multiplied with the upstream-section flow, i.e. $\beta_i^a q_{i-1}^a$; and the off-ramp flow (if any) of section i, direction b, is calculated based on known exit rates β_i^b multiplied with the upstream-section flow, i.e. $\beta_i^b q_{i+1}^b$.

We now deliver the CTM equations for each direction of the highway stretch.

Direction a:

The conservation equations for the sections of direction a read:

$$\rho_1^a(k+1) = \rho_1^a(k) + \frac{T}{L_i}(q_0^a(k) - q_1^a(k))$$
$$\rho_i^a(k+1) = \rho_i^a(k) + \frac{T}{L_i}((1-\beta_i^a)q_{i-1}^a(k) - q_i^a(k) + r_i^a(k)),\ i = 2,3,\ldots,n \quad . \tag{2.8}$$

According to CTM, the traffic flows are obtained as the minimum of demand and supply functions, except for the last section, where we consider only the demand function, as we assume that the downstream traffic conditions are uncongested. Clearly, when writing the demand and supply functions for the case of internal boundary control, we need to consider the impact of the respective sharing factors $\varepsilon_i^a(k_c)$ on the FDs. Thus we have

$$q_i^a(k) = \min\left\{Q_D(\rho_i^a(k), \varepsilon_i^a(k_c)), \frac{Q_S(\rho_{i+1}^a(k), \varepsilon_{i+1}^a(k_c))}{(1-\beta_{i+1}^a)} - \lambda_r r_{i+1}^a(k)\right\},\ i = 1,2,\ldots,n-1,$$
$$q_n^a(k) = Q_D(\rho_n^a(k), \varepsilon_n^a(k_c)), \tag{2.9}$$

where the demand and supply functions are given by the following respective equations

$$Q_D(\rho_i^a(k), \varepsilon_i^a(k_c)) = \min\left\{\varepsilon_i^a(k_c)q_{cap} + \lambda_d q_{cap}\frac{\rho_i^a(k) - \varepsilon_i^a(k_c)\rho_{cr}}{\rho_{cr} - \rho_{\max}}, v_f\rho_i^a(k)\right\},$$
$$Q_S(\rho_i^a(k), \varepsilon_i^a(k_c)) = \min\left\{\varepsilon_i^a(k_c)q_{cap}, w_s(\varepsilon_i^a(k_c)\rho_{\max} - \rho_i^a(k))\right\}, \tag{2.10}$$

while v_f is the free speed (which is assumed equal for all sections for simplicity) and w_s is the back-wave speed.

It is well-known that CTM does not reproduce the capacity drop, i.e. the empirical finding that, at the head of congestion, the observed flow in real traffic is reduced compared to the carriageway capacity. Capacity drop is deemed to occur in conventional traffic due to bounded and differing accelerations of different vehicles (Yuan et al., 2015; Yuan, 2016). Recently, CTM has been extended in a number of possible ways to enable the reproduction of capacity drop, see (Kontorinaki et al., 2017) for an overview and comparison. The presence of capacity drop in conventional motorway traffic is a major reason for infrastructure degradation and for the need of introducing traffic control measures to restore capacity (Papageorgiou et al., 2003; Papageorgiou et al., 2008). In contrast, in the present context of internal boundary control, the presence of capacity drop is a secondary source of amelioration of the traffic conditions, because the potential benefits achievable via opportune capacity sharing are expected to be much higher. In fact, it is unknown at the moment, if and to what extent capacity drop may occur in lane-free CAV traffic. To be able to investigate the impact of possible capacity drop, we have incorporated in the above equations the option of introducing capacity drop according to (Kontorinaki et al., 2017). More specifically, this option is enabled via the parameters λ_d and λ_r in the above equations. If these parameters are set $\lambda_r = 1$ and $\lambda_d = 0$, no capacity drop is reproduced, as typical for CTM; if these values are set between 0 and 1, a corresponding level of capacity drop is produced by the model.

Direction b :

The equations for direction b are analogous to those of direction a, with few necessary index modifications. Section numbers in direction b are descending, hence we have

$$\begin{aligned}\rho_i^b(k+1) &= \rho_i^b(k) + \frac{T}{L_i}((1-\beta_i^b)q_{i+1}^b(k) - q_i^b(k) + r_i^b(k)),\ i = 1,2,\ldots,n-1 \\ \rho_n^b(k+1) &= \rho_n^b(k) + \frac{T}{L_i}(q_{n+1}^b(k) - q_n^b(k))\end{aligned} \tag{2.11}$$

and the flows are given by

$$\begin{aligned}q_i^b(k) &= \min\left\{Q_D(\rho_i^b(k), \varepsilon_i^b(k_c)), \frac{Q_S(\rho_{i-1}^b(k), \varepsilon_{i-1}^b(k_c))}{(1-\beta_{i-1}^b)} - \lambda_r r_{i-1}^b(k)\right\},\ i = 2,3,\ldots,n, \\ q_1^b(k) &= Q_D(\rho_1^b(k), \varepsilon_1^b(k_c)),\end{aligned} \tag{2.12}$$

where

$$\begin{aligned}Q_D(\rho_i^b(k), \varepsilon_i^b(k_c)) &= \min\left\{\varepsilon_i^b(k_c)q_{cap} + \lambda_d q_{cap}\frac{\rho_i^b(k) - \varepsilon_i^b(k_c)\rho_{cr}}{\rho_{cr} - \rho_{\max}}, v_f\rho_i^b(k)\right\}, \\ Q_S(\rho_i^b(k), \varepsilon_i^b(k_c)) &= \min\left\{\varepsilon_i^b(k_c)q_{cap}, w_s(\varepsilon_i^b(k_c)\rho_{\max} - \rho_i^b(k))\right\}.\end{aligned} \tag{2.13}$$

2.3.2.3 CTM-based and further linear inequality constraints

The conservation equations (2.8) and (2.11) are linear, but, due to the presence of the min-operator in (2.9), (2.10), (2.12) and (2.13), the CTM flow equations presented in the previous section are nonlinear. As proposed by Papageorgiou (1995) and practiced in most previous utilizations of CTM for optimal control (Ziliaskopoulos, 2000; Gomes and Horowitz, 2006; Roncoli et al., 2015), such nonlinearities may be transformed to linear inequalities by requesting the left-hand side of the equation, where the min-operator appears, to be smaller than or equal to each of the terms included in the min-operator. For equations (2.9), (2.10) of direction a, this yields the following four inequalities

$$q_i^a(k) \le v_f\rho_i^a(k), \quad i = 1,2,\ldots,n, \tag{2.14}$$

$$q_i^a(k) \le \varepsilon_i^a(k_c)q_{cap} + \lambda_d q_{cap}\frac{\rho_i^a(k) - \varepsilon_i^a(k_c)\rho_{cr}(k_c)}{\rho_{cr} - \rho_{\max}}, \quad i = 1,2,\ldots,n, \tag{2.15}$$

$$q_i^a(k) \le \frac{w_s}{(1-\beta_{i+1}^a)}(\varepsilon_{i+1}^a(k_c)\rho_{\max} - \rho_{i+1}^a(k)) - \lambda_r r_{i+1}^a(k), \quad i = 1,2,\ldots,n-1, \tag{2.16}$$

$$q_i^a(k) \le \frac{\varepsilon_{i+1}^a(k_c)q_{cap}^{total}}{(1-\beta_{i+1}^a)} - \lambda_r r_{i+1}^a(k), \quad i = 1,2,\ldots,n-1. \tag{2.17}$$

Similarly, we obtain from (2.6) the following two linear inequalities

$$\varepsilon_i^a(k_c) \le \varepsilon_i(k_c), \quad i = 1, 2, ..., n, \tag{2.18}$$

$$\varepsilon_i^a(k_c) \le \varepsilon_i(k_c - 1), \quad i = 1, 2, ..., n. \tag{2.19}$$

For direction b, we obtain

$$q_i^b(k) \le v_f \rho_i^b(k), \quad i = 1, 2, ..., n, \tag{2.20}$$

$$q_i^b(k) \le \varepsilon_i^b(k_c) q_{cap} + \lambda_d q_{cap} \frac{\rho_i^b(k) - \varepsilon_i^b(k_c)\rho_{cr}(k_c)}{\rho_{cr} - \rho_{\max}}, \quad i = 1, 2, ..., n, \tag{2.21}$$

$$q_i^b(k) \le \frac{w_s}{(1 - \beta_{i-1}^b)} (\varepsilon_{i-1}^b(k_c)\rho_{\max} - \rho_{i-1}^b(k)) - \lambda r_{i-1}^b(k), \quad i = 2, 3, ..., n, \tag{2.22}$$

$$q_i^b(k) \le \frac{\varepsilon_{i-1}^b q_{cap}}{(1 - \beta_{i-1}^b)} - \lambda r_{i-1}^b(k), \quad i = 2, 3, ..., n, \tag{2.23}$$

$$\varepsilon_i^b(k_c) \le 1 - \varepsilon_i(k_c), \quad i = 1, 2, ..., n, \tag{2.24}$$

$$\varepsilon_i^b(k_c) \le 1 - \varepsilon_i(k_c - 1), \quad i = 1, 2, ..., n. \tag{2.25}$$

In addition, the constraints (2.5) must be considered to appropriately limit the sharing factors.

In summary, inequalities (2.14) and (2.15) for direction a, and (2.20) and (2.21) for direction b represent the demand part of the FD; while inequalities (2.16) and (2.17) for direction a, and (2.22) and (2.23) for direction b represent the supply part of the FD.

It must be emphasized, however, that the min-operator is equivalent to the corresponding set of linear inequalities only if one of the inequality constraints is actually activated. If none of the inequalities is activated, then the corresponding flow takes a lower value than the one prescribed by CTM, and this occurrence, known as the flow holding-back effect, has been addressed in different ways in the literature (Doan and Ukkusuri, 2012). Since, in presence of holding-back, CTM is accordingly distorted, we consider, in the optimal control problem formulation, non-negativity and upper-bound constraints for all flow and density variables to ensure that non-physical values are excluded from the solution.

Holding-back may occur in the solution of the QP-problem if it is beneficial for the cost function, in particular for minimization of the Total Time Spent (TTS). Holding-back may indeed be beneficial in the present context if there is congestion in the QP-problem solution and if capacity drop is activated in the model. In fact, if no congestion is present, holding back traffic can only produce unnecessary delays and TTS increase. In presence of congestion, but without capacity drop, the incentive to hold back traffic is limited to special circumstances. In contrast, in presence of congestion and capacity drop, holding back traffic may be beneficial because it may

mitigate the capacity drop and restore capacity flow at the bottleneck location; something that is indeed a major motivation for traffic control in conventional traffic by use of various active holding-back control measures, such as VSL (variable speed limits) and ramp metering (Papageorgiou et al., 2008).

The presence of holding-back in the QP-problem solution could be interpreted as an opportunity to apply, in addition to internal boundary control, also speed control, something that is not difficult in CAV traffic, see also (Han et al. 2017). As the primary focus of this article is to elaborate on the features of internal boundary control, we will adopt the following policy in the investigations of Section 3:

- On one hand, we will report on the direct outcome of the QP-problem solution. If this solution contains no holding-back, then it is equivalent to the CTM simulation outcome;
- On the other hand, if flow holding back is observed in the QP-problem solution, then we will use the delivered sharing factor trajectories to feed the CTM equations and obtain the corresponding traffic states without holding-back. Thus, we will have two TTS values: one stemming from the QP-problem solution (with possible holding-back that could be interpreted as the result of additionally applying speed control); and a second obtained from CTM equations fed with the QP-optimal ε-trajectories. The deviation between these two values will indicate the extent of holding-back; or, in other words, the additional benefit that could be obtained by applying speed control on the top of internal boundary control.

Finally, the following inequalities are introduced to guarantee that a reduction of the width for any section and direction will not result in a density value that exceeds the jam density:

$$\rho_i^a(k+1) \le \varepsilon_i^a(k_c)\,\rho_{\max}, \quad i=1,2,...,n, \tag{2.26}$$

$$\rho_i^b(k+1) \le \varepsilon_i^b(k_c)\,\rho_{\max}, \quad i=1,2,...,n. \tag{2.27}$$

2.3.2.4 Objective function and QP problem formulation

The objective function to be minimized is defined as follows:

$$\begin{aligned} J = {} & T\sum_{k=1}^{K}\sum_{i=1}^{n}\left(L_i\rho_i^a(k)+L_i\rho_i^b(k)\right) - w_1\sum_{k_c=0}^{K_c-1}\sum_{i=1}^{n}\left(\varepsilon_i^a(k_c)+\varepsilon_i^b(k_c)\right) \\ & + w_2\sum_{k_c=1}^{K_c-1}\sum_{i=1}^{n}\left(\varepsilon_i(k_c)-\varepsilon_i(k_c-1)\right)^2 + w_3\sum_{k_c=0}^{K_c-1}\sum_{i=2}^{n}\left(\varepsilon_i(k_c)-\varepsilon_{i-1}(k_c)\right)^2 \\ & + w_4\sum_{k_c=0}^{K_c-1}\sum_{i=1}^{n}\left(\frac{\varepsilon_i(k_c)^2}{d_i^a(k_c)}+\frac{(1-\varepsilon_i(k_c))^2}{d_i^b(k_c)}\right). \end{aligned} \tag{2.28}$$

The proposed cost function extends over a time horizon of K model time steps or K_c control time steps, where $K_c = K\cdot T/T_c$; it includes five terms, the first two being linear and the rest of them quadratic. We will now comment on each of these terms.

The first term represents the TTS. This is the most important term in the cost function as it determines the traffic efficiency resulting from the proposed control action. TTS attains its minimum value if the road is shared among the two directions in such a way that no congestion, which would give rise to according vehicle delays, is present in any section and direction during the time horizon. If, despite optimal internal boundary control, the creation of congestion is inevitable due to strong external demands, then TTS minimization leads to minimization of the incurred delays.

As mentioned earlier, equations (2.6) and (2.7) were replaced by the linear inequalities (2.18), (2.19) and (2.24), (2.25), respectively. This transformation bears the risk that the resulting ε_i^a, ε_i^b might not activate either of the two values of the original min-operator, in which case a part of the road width would remain unexploited. To suppress this possibility, the second term of the objective function is introduced (with a sufficiently high weight w_1) to ensure that at least one of the inequalities (2.18), (2.19) and at least one of the inequalities (2.24), (2.25) will be activated, i.e. that one of the two terms included in each min-operator (2.6) and (2.7), respectively, will actually materialize.

The next three terms are quadratic and reflect secondary (as contrasted to the TTS minimization) operational and policy objectives. Therefore, the weights of these terms must be selected small enough, so that their impact on the TTS outcome is small (if any). The first quadratic term penalizes the variation of the control input in consecutive time steps, so that changes of the internal boundary of each section from one control time step to the next remain small. The second quadratic term penalizes the space variation of the control input from section to section, so as to suppress strong changes in the road width assigned to each direction within a short distance. The need for smooth space-time changes of the internal boundary is motivated by safety and convenience concerns. However, related implications will be addressed in more detail in the forthcoming microscopic testing of the concept.

The last quadratic term is policy related. Consider cases where appropriate sharing factors at the highway sections provide sufficient respective capacities for the two directions to utterly avoid congestion forming. As it will become clearer in Section 3, in such cases there is a range of possible ε_i values that lead to minimum TTS, i.e. the TTS-minimum is not unique. Under these conditions, a policy question arising is: Which one out of those TTS-optimal ε_i-values is preferable? One possible answer to this question is to express a preference for ε_i-values that are closer to 0.5, i.e. closer to the middle of the road. For such a policy, this quadratic term should be chosen as $(\varepsilon_i - 0.5)^2$.

In contrast, the policy pursued with the last quadratic term in (2.28) is a different one, as it attempts to assign to the two directions respective capacity shares that balance the respective capacity reserves for each section. To this end, we use $d_i^a(k_c)$ and $d_i^b(k_c)$, $i = 1,..,n$, which are the projected demand trajectories for each section of the two respective directions, and are calculated as follows. For a given highway stretch and demand scenario, we consider the respective upstream mainstream entry flow, as well as all on-ramp and off-ramp flows over the whole time horizon in each direction; then, the projected demand trajectories are obtained by running the CTM equations, with these entering and exiting flows, assuming that the capacity is

sufficiently high. In other words, the projected demands in each section and direction are obtained by propagating the external demands at free speed. Then, for given projected demands, the minimization of the last quadratic term w.r.t. ε_i is readily seen to lead to

$$\frac{\varepsilon_i(k_c)}{d_i^a(k_c)} = \frac{1-\varepsilon_i(k_c)}{d_i^b(k_c)},$$

hence

$$\frac{\varepsilon_i(k_c)}{d_i^a(k_c)} q_{cap} = \frac{1-\varepsilon_i(k_c)}{d_i^b(k_c)} q_{cap},$$

which corresponds indeed to the balancing of the capacity reserves in the two traffic directions in every section. On the other hand, if congestion is unavoidable due to strong demands, the utilized weight w_4 is sufficiently small, so that the impact of this term on TTS-minimization is marginal.

We are now ready to formulate the general convex QP problem for optimal internal boundary control for any highway stretch with known external (mainstream and on-ramp) demand trajectories over the considered time horizon as follows:

$$\min J = \mathbf{C}^T\mathbf{x} + \frac{1}{2}\mathbf{x}^T\mathbf{H}\mathbf{x} \tag{2.29}$$

subject to

$$\mathbf{A}_i\mathbf{x} \le \mathbf{b}_i \tag{2.30}$$

$$\mathbf{A}_e\mathbf{x} = \mathbf{b}_e \tag{2.31}$$

$$\mathbf{b}_{lb} \le \mathbf{x} \le \mathbf{b}_{ub} \tag{2.32}$$

where $\mathbf{x}$ is the decision vector including all the states (densities), flows and control variables (sharing factors) at all times. The two terms in the cost function (2.29) represent the linear and quadratic terms of equation (2.28), respectively. The linear inequality (2.30) derives from all inequalities mentioned in Section 2.2.2; whereas the linear equality (2.31) represents the conservation equations (2.8) and (2.11). Finally, (2.32) provides upper and lower bounds for the decision variables, including (2.5) for the sharing factors.

2.4 Case studies

2.4.1 Introduction of case studies

Two scenarios are considered in order to investigate the performance of the proposed method:

- *Uncongested scenario*, where congestion is created in one or both directions without internal boundary control (no-control case); but the congestion can be utterly avoided

with activation of internal boundary control. Such situations are likely to constitute the majority of real congestion cases on highways and arterials.

- *Congested scenario*, where no-control congestion can be mitigated with internal boundary control, but cannot be utterly suppressed due to strong bi-directional external demands.

The considered highway stretch is depicted in Fig. 2.4. It has a length of 3 km and is subdivided in 6 sections of 0.5 km each. In direction a, there is one off-ramp in section 2 and one on-ramp in section 5. In direction b, there is one off-ramp in section 4 and one on-ramp in section 3. The exit rates of the off-ramps are both equal to 0.1. The modelling time step is $T = 10\,s$, and the control time step $T_c = 60\,s$. The parameter values used to enable capacity drop, when this possibility is mentioned to be activated in the scenarios, are $\lambda_d = 0.4$ and $\lambda_r = 0.7$; when no capacity drop is activated in the model, these parameters are set equal to 1. The considered time horizon is 1 h, hence $K = 360$ and $K_c = 60$. The CTM parameters are $v_f = 100\,\text{km/h}$ and $w_s = 12\,\text{km/h}$; while the total cross-road capacity to be shared among the two directions is $q_{cap} = 12{,}000\,\text{veh/h}$. The upper and lower bounds for the sharing factors, so as to avoid utter blocking of any of the two directions, are equal for all sections and are given the values $\varepsilon_{\min} = 0.16$ and $\varepsilon_{\max} = 0.84$. For both scenarios, the initial density values are $\rho_i^a(0) = [5,5,5,5,18.5,29.4]\text{veh/km}$, $\rho_i^b(0) = [14.4,14.4,14,5,5,5]\text{veh/km}$.

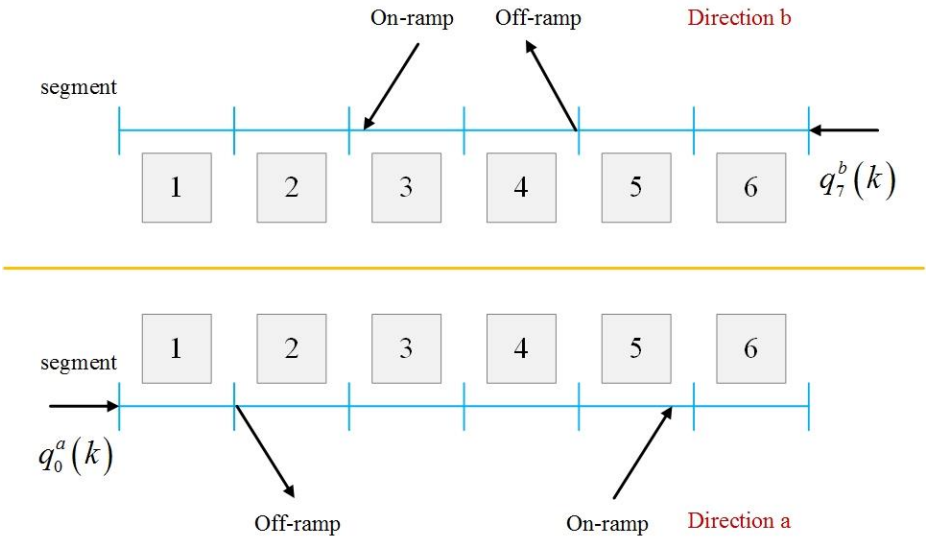


Figure 2.4: The considered highway stretch

For each scenario, the simulation results for the no-control case will be presented first, followed by the results obtained with optimal internal boundary control. The same weight parameters were used in the cost function of the QP-problem for the respective control cases of the two scenarios, namely $w_1 = 10^{-1}, w_2 = 10^{-4}, w_3 = 10^{-5}, w_4 = 10^{-3}$. These values were obtained with manual fine-tuning on the grounds outlined in Section 2.2.3, taking into account the magnitude of each related term. More specifically:

- The weight w_1 was selected sufficiently high to ensure that at least one of the inequalities (2.18), (2.19) and at least one of the inequalities (2.24), (2.25) is always activated, i.e. that one of the two terms included in each respective min-operator (2.6) and (2.7) actually applies.
- The weights of the three quadratic terms were initially set to zero, in order to obtain the minimum achievable TTS value. Eventually, the three weights w_2, w_3, w_4 were gradually increased to the above-mentioned values, such that the corresponding smoothness and balancing sub-objectives are achieved at a sufficient level; while, at the same time, TTS increases only marginally (less than 1%) compared to its minimum achievable value.

The results were found to be little sensitive around this choice; indeed, differing control behavior, if desired, may call for corresponding weight changes by one or more orders of magnitude.

The QP algorithm employed was interior point convex in MATLAB 2018 and was executed on a personal computer with processor Intel(R) Core (TM) i5-7500 CPU @ 3.4GHz and RAM of 8 GB. For an optimization horizon of 360 model time-steps (1 hour) the computational times were around 25 s.

2.4.2 Uncongested scenario

2.4.2.1 Scenario description

The demand flows for this scenario are displayed in Fig. 2.5a for both directions. It may be seen that the two directions feature respective peaks in their upstream mainstream demands that are slightly overlapping. In addition, the on-ramp demands are constant, with the on-ramp demand in direction a being higher than in direction b.

The demand and supply situation for this scenario can be analyzed based on Fig. 2.6. The figure displays, for each highway section, a flow-versus-time window, where time extends over the time-horizon of the scenario. The height of each window equals the total road capacity q_{cap} that must be shared among the two directions at each section. The lower displayed curve (blue) is the projected demand $d_i^a(k_c)$ in direction a for each section; and the upper curve (red), displayed upside-down, starting from the upper window edge, is the projected demand $d_i^b(k_c)$ in direction b. The fact that these two curves do not intersect at any section, indicates that flexible sharing may be applied so as to avoid any congestion forming in either direction. In fact, the finite distance among the two curves at any time for all sections indicates that there is a whole range of possible sharing factor trajectories that lead to congestion avoidance, and one such trajectory is indeed displayed (green) in each diagram, specifically one that balances approximately the capacity reserves among the two directions. In addition, each diagram displays a horizontal line (black) at 6000 veh/h, which is the capacity of each direction if no internal boundary control is applied. It may be seen that this line intersects with the projected demand curves at some sections during some time periods. Specifically, the projected demand in direction a exceeds the fixed capacity first at section 5 at around $k = 60$; while in direction b, the projected demand

exceeds capacity first in section 3 at around $k = 200$. In both cases, the exceeding of capacity is due to the presence of on-ramps in the respective sections. Obviously, these sections and time periods are candidates for congestion forming in the no-control case.

2.4.2.2 No control case

Using the entering flows of the uncongested scenario in the CTM equations of Section 2.2.1 with constant internal boundary at $\varepsilon_i = 0.5$ for all sections, we obtain the simulation results of the no-control case. Figure 2.7 displays the corresponding spatio-temporal density evolution. More precisely, the variable displayed in Fig. 2.7 for each direction is the relative density, which is defined as $\tilde{\rho}^a(k) = \rho^a(k) / \rho^a_{cr}(k) = \rho^a(k) / (\varepsilon^a(k)\rho_{cr})$ for direction a and $\tilde{\rho}^b(k) = \rho^b(k) / \rho^b_{cr}(k) = \rho^b(k) / (\varepsilon^b(k)\rho_{cr})$ for direction b. Note that density (in veh/h) by itself is not sufficient, in the internal boundary control environment, to distinguish between under-critical and congested conditions, because the critical density is also changing according to the applied control, see (2.4). Of course, the critical density is not changing in the no-control case, but we use already here relative densities for consistency with the control case. According to the definition, relative density values lower than 1 refer to uncongested traffic; while values higher than 1 refer to congested traffic; clearly, when the relative density equals 1, and the downstream section is uncongested, we have capacity flow at the corresponding section.

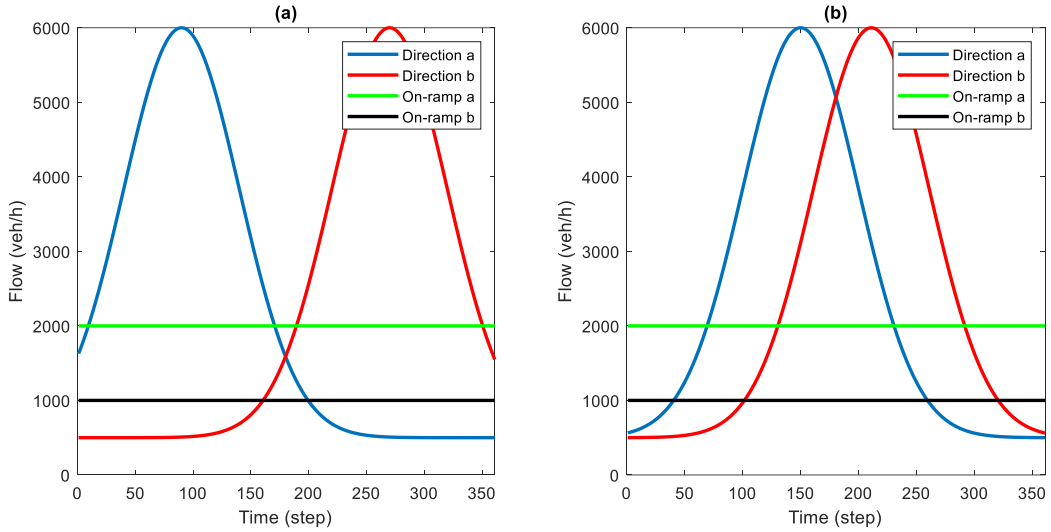


Figure 2.5: Demand flows of each direction in: (a) uncongested scenario; (b) congested scenario

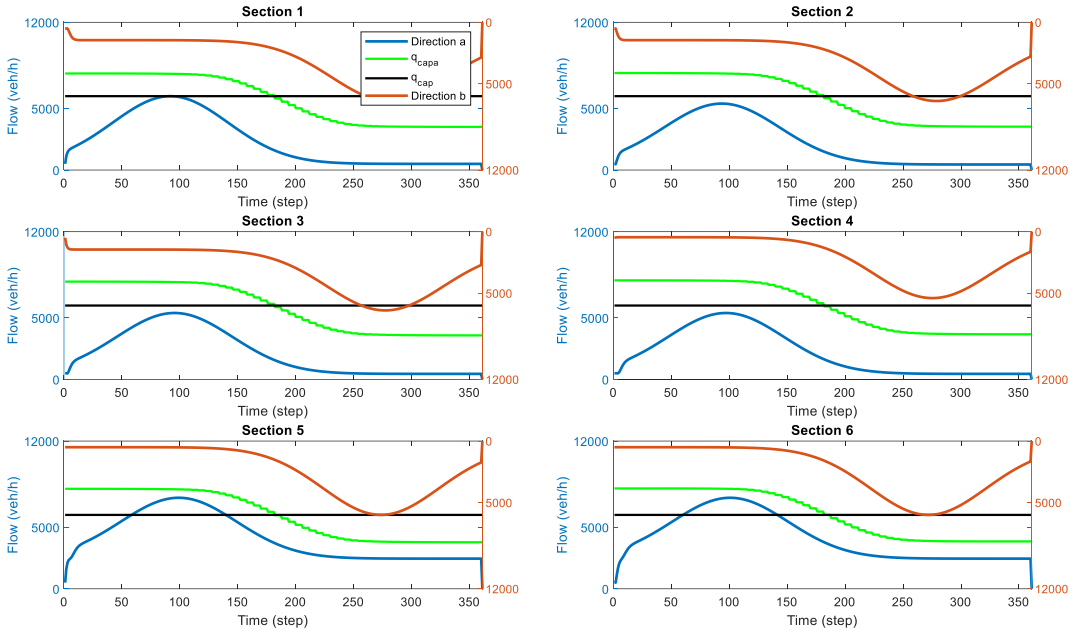


Figure 2.6: Demand-supply analysis for the uncongested scenario

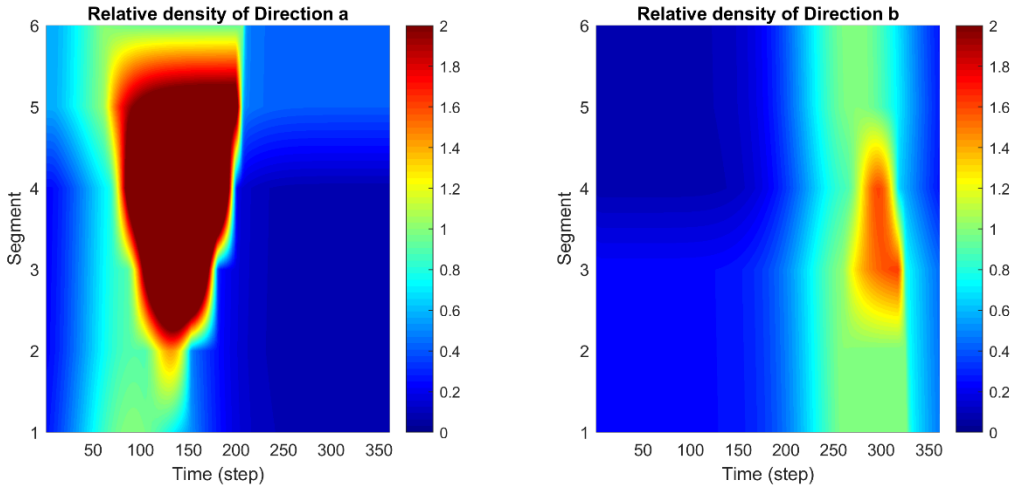


Figure 2.7:Uncongested scenario: Relative density of the two directions in the no-control case

Figure 2.7 shows that, as expected from the analysis with Fig. 2.6, heavy congestion is created in section 5 for direction a due to the strong ramp inflow, in combination with the increased mainstream demand, at around $k = 60$. The congestion tail propagates backwards, reaching up to section 2, and is dissolved at around $k = 200$, thanks to the rapid decrease of the mainstream demand (Fig.2.5a). In direction b, we have also a congestion being triggered by the increasing mainstream demand, in combination with the on-ramp flow, in section 3 at around $k = 250$. Due to lower on-ramp flow, this congestion is smaller than in direction a; it spills back up to section 5 and dissolves at around $k = 330$.

It should be noted that the results displayed in Fig. 2.7 were obtained using the CTM equations with capacity drop, and the corresponding value of TTS is reported in Table 1. When the option of creating capacity drop is de-activated, then the space-time extent of the created congestions is reduced. The corresponding diagrams are omitted for space economy, but the resulting, lower TTS value is also reported in Table 1.

2.4.2.3 Control case

Based on the analysis with Fig. 2.6, there exist infinitely many internal-boundary trajectories that may accommodate the bi-directional demand in all sections such that the assigned capacity in each direction is never exceeded. In fact, the mentioned green curves in the diagrams of Fig. 2.6 reflect the obtained solution for this scenario, and it may be seen that this curve does not intersect with the projected demands in any section and direction. As a consequence, the resulting traffic conditions are expected to be under-critical everywhere, and this is indeed confirmed by the spatio-temporal evolution of the relative densities depicted in Fig. 2.8.

Figures 2.9, 2.10, 2.11 display more detailed information for this case. Specifically, each figure holds the results of two respective sections; for each section, we provide three diagrams:

- The first diagram shows the two traffic densities (in veh/km), for directions a and b, and the corresponding two critical densities, which are changing according to the sharing factor in the section.
- The second diagram shows the two traffic flows, for directions a and b, and the corresponding two capacities, which are changing according to the sharing factor in the section. In addition, the sum of both flows is also displayed (yellow curve).
- The third diagram shows the two sharing factors $\varepsilon^a(k_c)$ and $\varepsilon^b(k_c)$, for directions a and b, respectively. Note that the time axis in this case displays the control time steps k_c.

The displayed results confirm that densities (flows) are always lower than the respective critical densities (capacities) in all sections and in both directions; hence traffic conditions are always and everywhere under-critical. In fact, the total-flow curve (for both directions) does not reach the total carriageway capacity (of 12,000 veh/h) at any time anywhere. Also, the margins of densities (flows) to the respective critical densities (capacities) are seen to be sufficiently balanced for the two directions at all sections for all time steps, taking into account that space-time smoothness of the sharing factors is also considered. In short, congestion is utterly avoided and any occurring delays in the no-control case are now utterly nullified, while all operational sub-objectives are sufficiently accounted.

Indeed, the sharing factor trajectories of the sections reveal that this excellent outcome is enabled via a smooth swapping of assigned capacity to the two directions, whereby more capacity is assigned to direction a during the first half of the time horizon and vice-versa for the second half, so as to accommodate the changing respective demands and their peaks.

The reported results were obtained with activation of capacity drop, which however, has no impact, as there is not the least congestion in the QP-problem solution. For the same reason, virtually no holding-back is observed in the solution. All related TTS values are given in Table 1,

indicating improvements of 29 % and 21 % over the no-control case with and without capacity drop activation, respectively.

Finally, Fig. 2.11a displays the space-time diagram of the control input (sharing factors) which demonstrates that it is a smooth function in space and time. Clearly, this shape may be further influenced by appropriate changes in the utilized weights of the cost criterion.

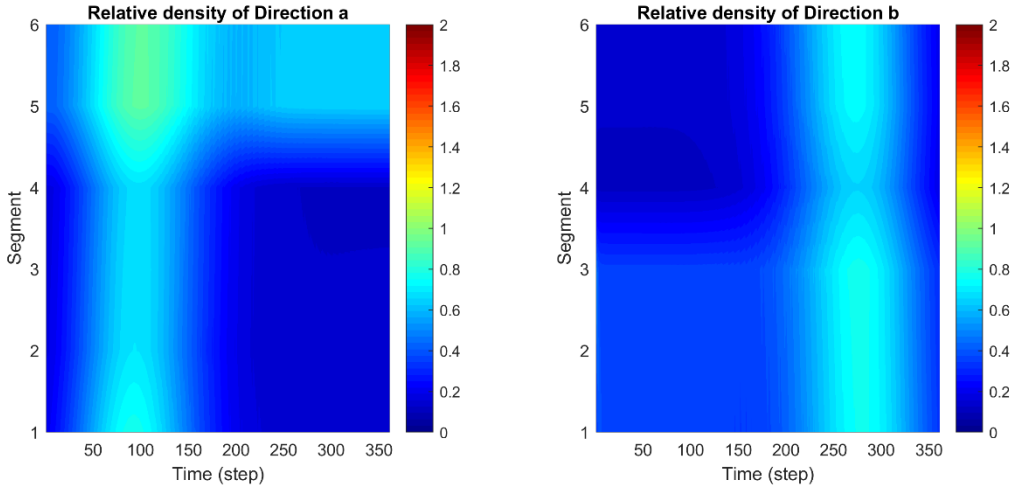


Figure 2.8: Uncongested scenario: Relative density of the two directions in the control case

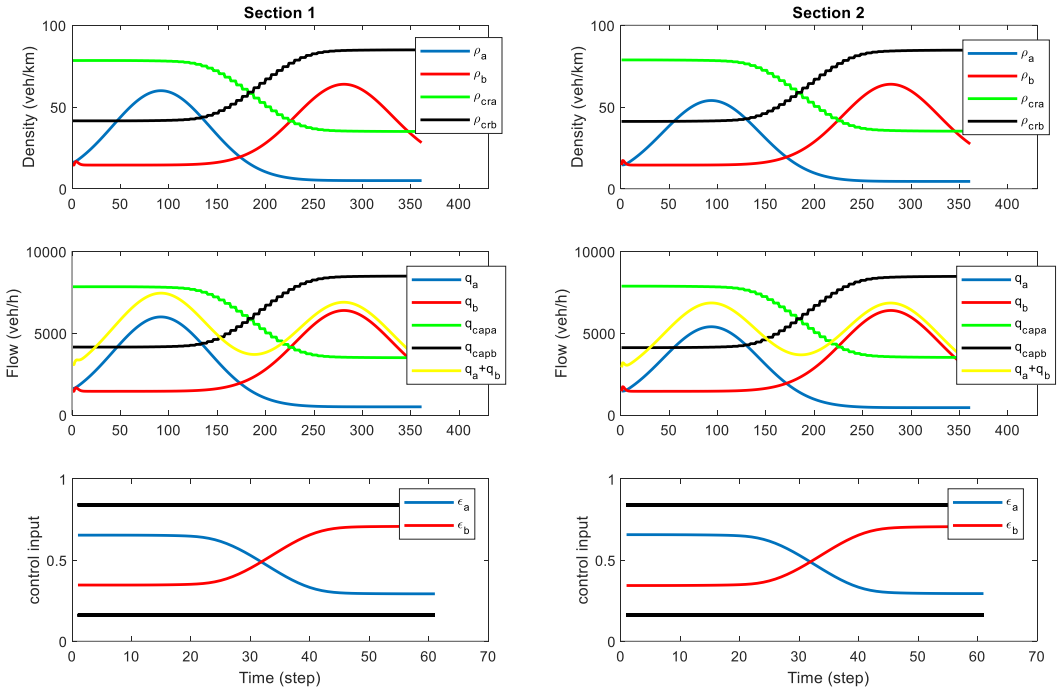


Figure 2.9: Uncongested scenario: Density, flow and control trajectories in the control case (sections 1 and 2)

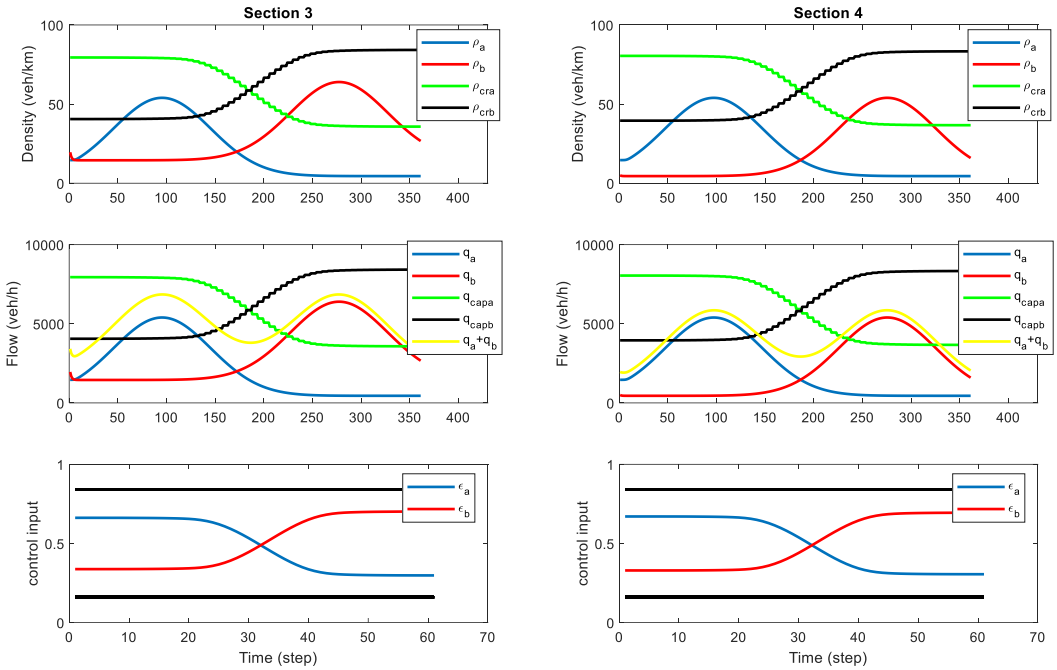


Figure 2.10: Uncongested scenario: Density, flow and control trajectories in the control case (sections 3 and 4)

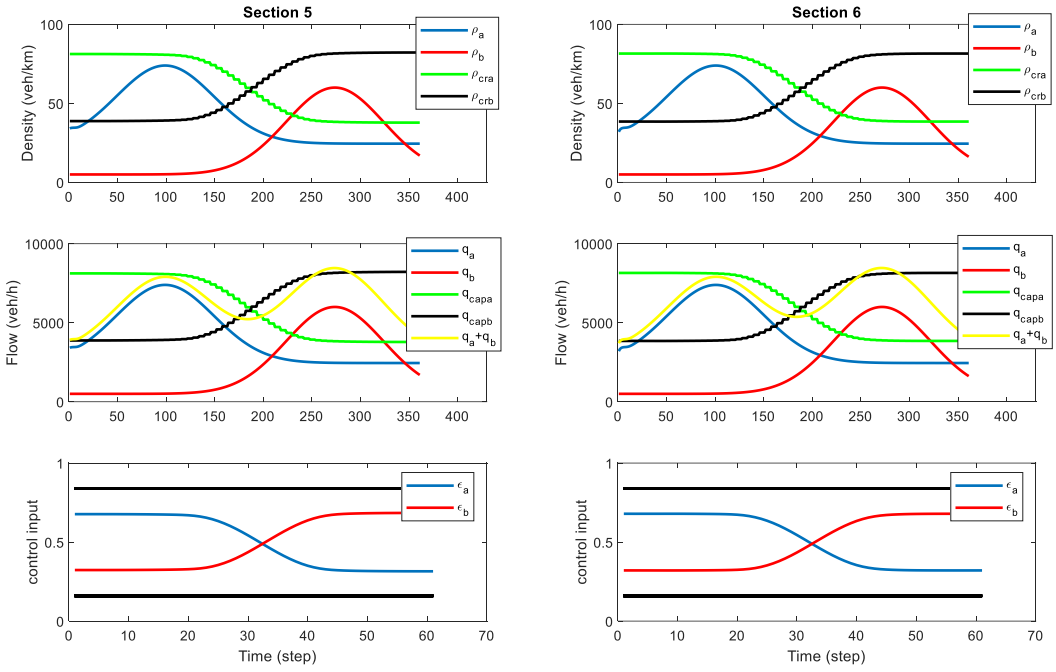


Figure 2.11:Uncongested scenario: Density, flow and control trajectories in the control case (sections 5 and 6)

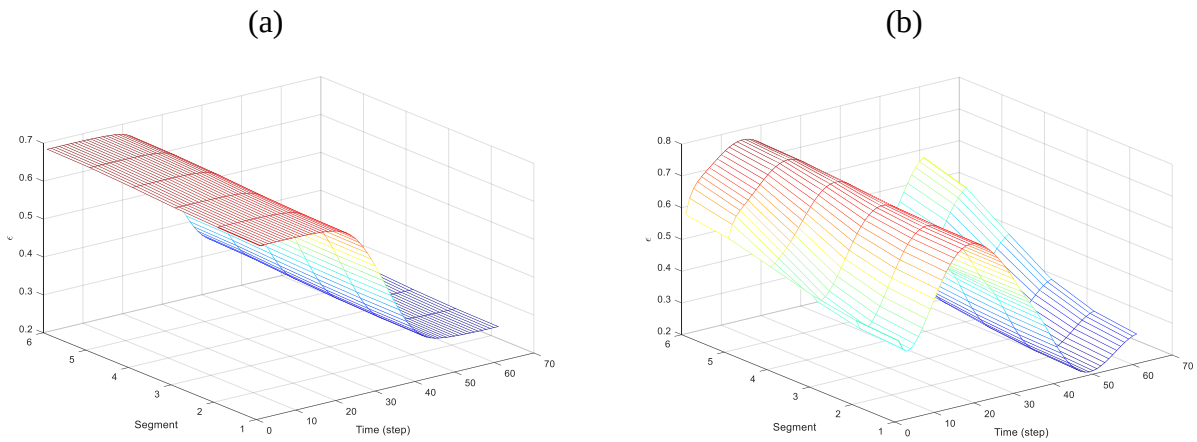


Figure 2.12: 3-D space-time diagram of the control input (sharing factors) for: (a) uncongested scenario; (b) congested scenario

2.4.3 Congested scenario

2.4.3.1 Scenario description

The demand flows for this scenario are displayed in Fig. 2.5b for both directions. Essentially, all external flows are similar as in the uncongested scenario, with the noticeable difference that the two mainstream demands have been moved closer to each other. This leads to a longer overlapping period with strong flows on both directions, which implies capacity problems even in presence of internal boundary control. The on-ramp demands are constant at the same level as in the uncongested scenario.

The demand and supply situation for the congested scenario can be seen in Fig. 2.13, which displays the same kind of information as Fig. 2.6. It may be seen that the no-control fixed capacity line of 6,000 veh/h intersects with the projected demand curves. Specifically, the projected demand in direction a exceeds the fixed capacity first at section 5 at around $k = 120$; while in direction b, the projected demand exceeds capacity first in section 3 at around $k = 200$. Again, exceeding of capacity is due to the presence of on-ramps in the respective sections, and, obviously, these sections and time periods are candidates for congestion forming in the no-control case. In contrast to the uncongested scenario, we now see that the curves of the two projected demands also interfere slightly in sections 5 and 6, starting at around $k = 200$. This implies that, even with optimal boundary control, formation of congestion is unavoidable at these sections and time.

It is very interesting to emphasize here that, in contrast to conventional traffic where bottlenecks may be present in either direction independently, the application of internal boundary control implies that bottlenecks concern both traffic directions simultaneously. In other words, for a bottleneck to be present in a section, the total bi-directional projected demand must exceed the total carriageway capacity, and this obviously involves both traffic directions. Note though that a

bottleneck could also occur in some cases at low demand in one traffic direction, simply because the minimum-width constraint (2.5) has been activated, hence no additional road width can be assigned to the opposite high-demand direction.

Figure 2.13 also displays a possible set of sharing factor trajectories (green) that lead to congestion avoidance anywhere except at the location and time of the identified bottleneck. These trajectories are seen to balance approximately the capacity reserves in the two directions wherever possible (i.e. at times and locations without a bottleneck).

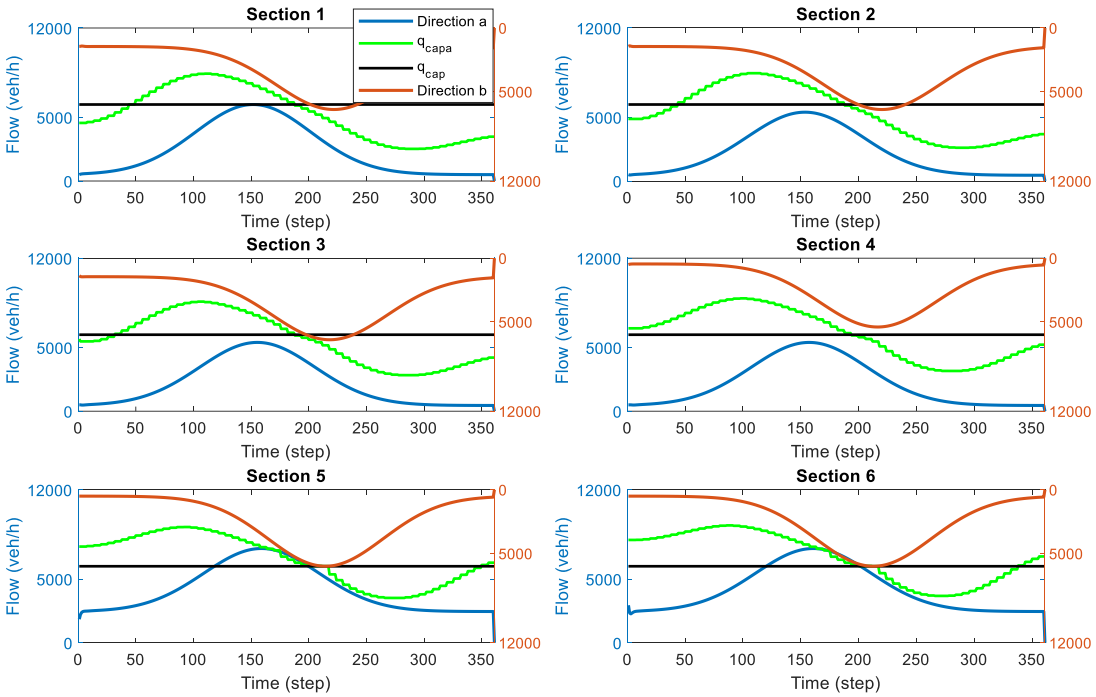


Figure 2.13: Demand-supply analysis for the congested scenario.

2.4.3.2 No control case

Using the entering flows of the congested scenario in the CTM equations of Section 2.2.1 with constant internal boundary at for all sections, we obtain the simulation results of the no- control case. Figure 2.14 displays the corresponding spatio-temporal relative density evolution, where, as expected from the analysis with Fig. 2.13, heavy congestion is created in section 5 for direction a due to the strong ramp inflow, in combination with the increased mainstream demand, at around $k = 120$. The congestion tail propagates backwards, reaching up to section 2, and is dissolved at around $k = 250$, thanks to the rapid decrease of the mainstream demand (Fig.2.5b). In direction b, we have also a congestion being triggered by the increasing mainstream demand, in combination with the on-ramp flow, in section 3 at around $k = 200$. Due to lower on-ramp flow, this congestion is smaller than in direction a; it spills back up to section 5 and dissolves at around $k = 270$.

The results displayed in Fig. 2.14 were obtained using the CTM equations with capacity drop, and the corresponding value of TTS is reported in Table 1. When the option of creating capacity drop is de-activated, then the space-time extent of the created congestions is reduced. The corresponding diagrams are omitted for space economy, but the resulting, lower TTS value is also reported in Table 1. Note that, compared with the uncongested scenario, the congestion extent and TTS values are not much different, simply because the total demand per direction did not change much, while the capacity of each direction is constant in the no-control case.

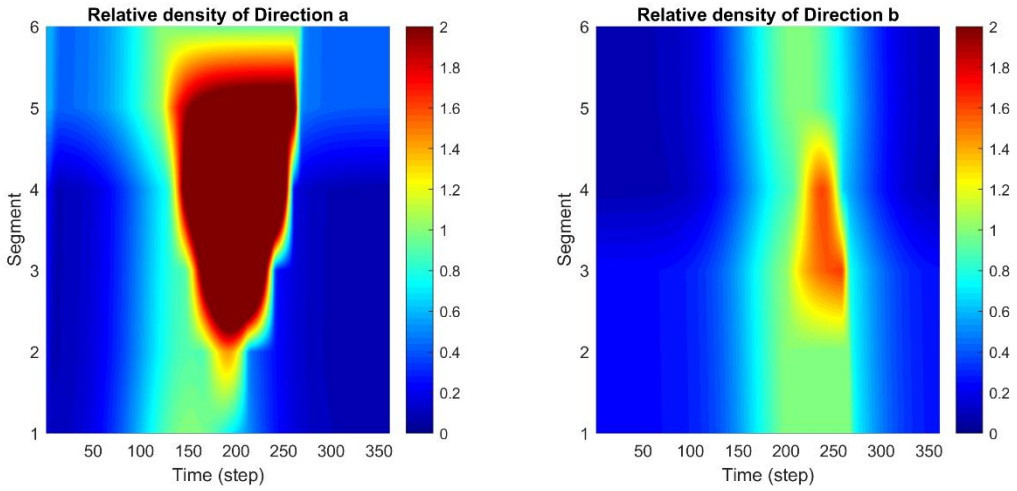


Figure 2.14: Congested scenario: Relative density of the two directions in the no-control case

2.4.3.3 Control case

Based on the analysis with Fig. 2.13, there exist infinitely many internal-boundary trajectories that may accommodate the projected demand so that the assigned capacity in each direction is not exceeded, with one exception, which concerns the bottleneck location and time period, as mentioned earlier. The bottleneck is common for both directions, and the sharing factors at sections 5 and 6 during the identified period should be such that the unavoidable congestion(s) affect as little as possible the resulting TTS value. In fact, the mentioned green curves in the diagrams of Fig. 2.13 reflect the obtained optimal solution for this scenario, and it may be seen that this curve intersects with the projected demands only at the bottleneck location and time period, while it exhibits balanced capacity margins elsewhere. As a consequence, the resulting traffic conditions are expected to be over-critical at the bottleneck location and time period, although any congestion formed there may of course propagate upstream and cover further upstream sections, depending on the amount of excess demand.

Figure 2.15 displays the spatio-temporal evolution of the relative densities in the control case; while Figures 2.16, 2.17, 2.18 display more detailed information for this case, as for the uncongested scenario. The displayed results indicate that densities (flows) are always lower than or equal to the respective critical densities (capacities) in all sections of direction b. In direction a, congestion may be observed, mainly in section 5 and to much lesser extent in section 6, starting at time $k=170$ and lasting up to $k=220$. It is important to emphasize that the total-

flow curve (for both directions) is seen to reach and remain close to the total carriageway capacity (of 12.000 veh/h) at section 6, which is indeed the bottleneck of this scenario. Thus, full exploitation of the carriageway capacity (both directions) is indeed enabled at the bottleneck for the duration of the critical period, so as to minimize congestion and delays. Finally, the margins of densities (flows) to the critical densities (capacities) are seen to be sufficiently balanced for the two directions at any time, except for the bottleneck location and time period.

The sharing factor trajectories are slightly more varying in this scenario, compared with the uncongested scenario, so as to assign the necessary share of capacity where and when needed. Figure 2.12b displays the space-time diagram of the control input (sharing factors), which illustrates that it is a slightly more complex, but still smooth function in space and time.

The reported results were obtained based on the CTM equations, which were fed with the QP-optimal sharing factor trajectories. As a matter of fact, the QP problem solution contains some limited flow holding back (at section 4), which enables a slightly lower TTS value, compared with the one resulting from the CTM equations. All TTS values, namely with/without capacity drop activation and with/without holding-back, are reported in Table 1, where it may be seen that the respective differences are minor. TTS improvements over the no-control case are 28 % and 20 % with and without capacity drop activation, respectively.

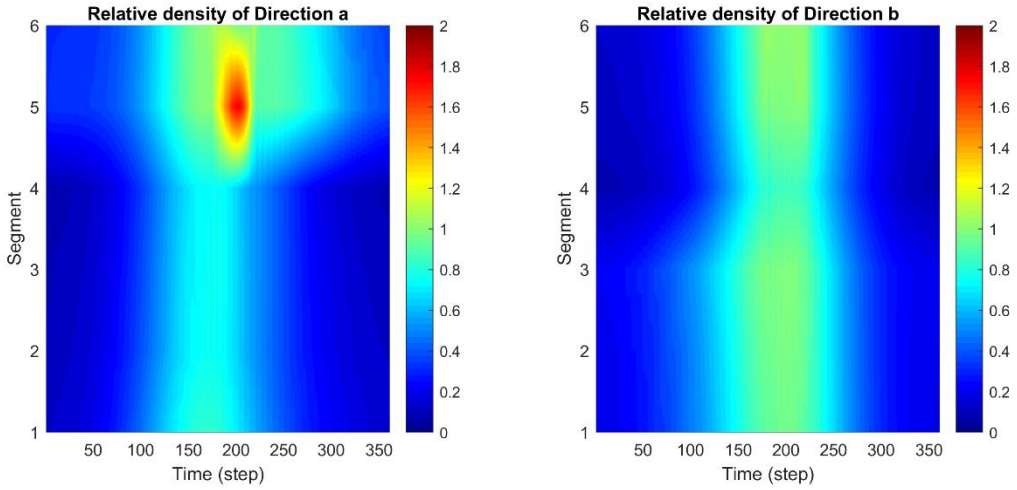


Figure 2.15: Congested scenario: Relative density of the two directions in the control case

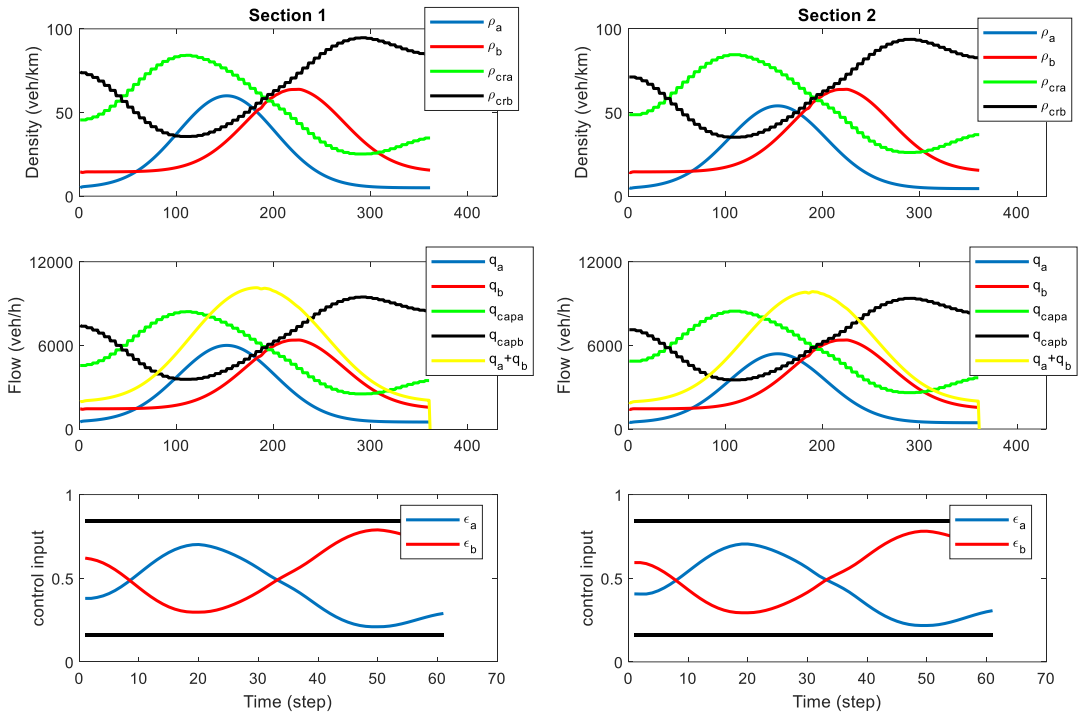


Figure 2.16: Congested scenario: Density, flow and control trajectories in the control case (sections 1 and 2)

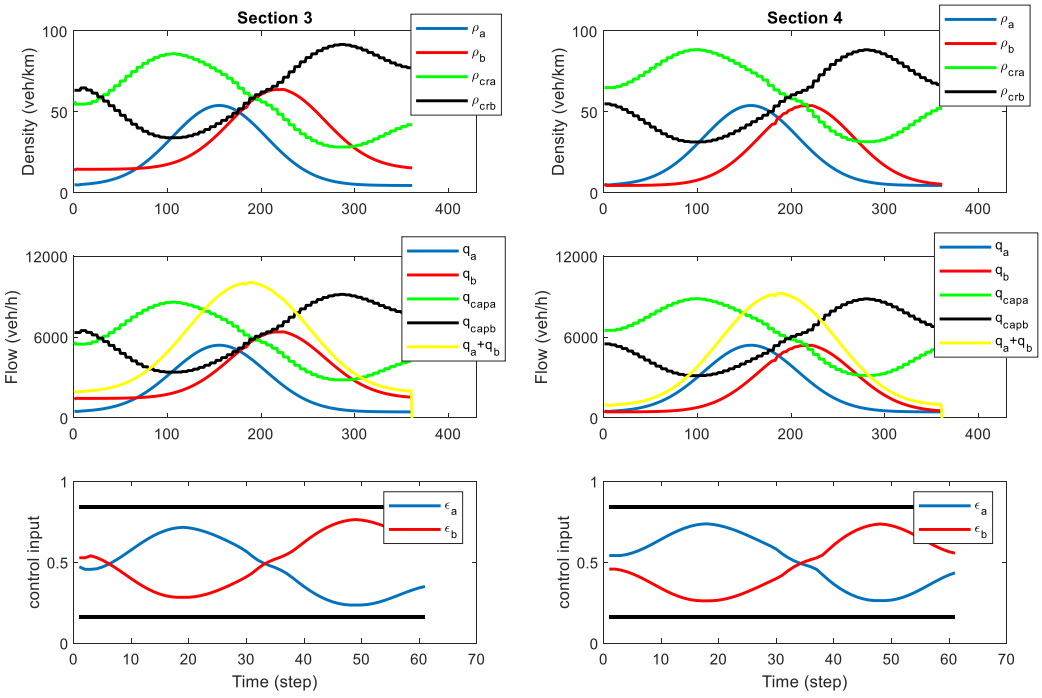


Figure 2.17: Congested scenario: Density, flow and control trajectories in the control case (sections 3 and 4)

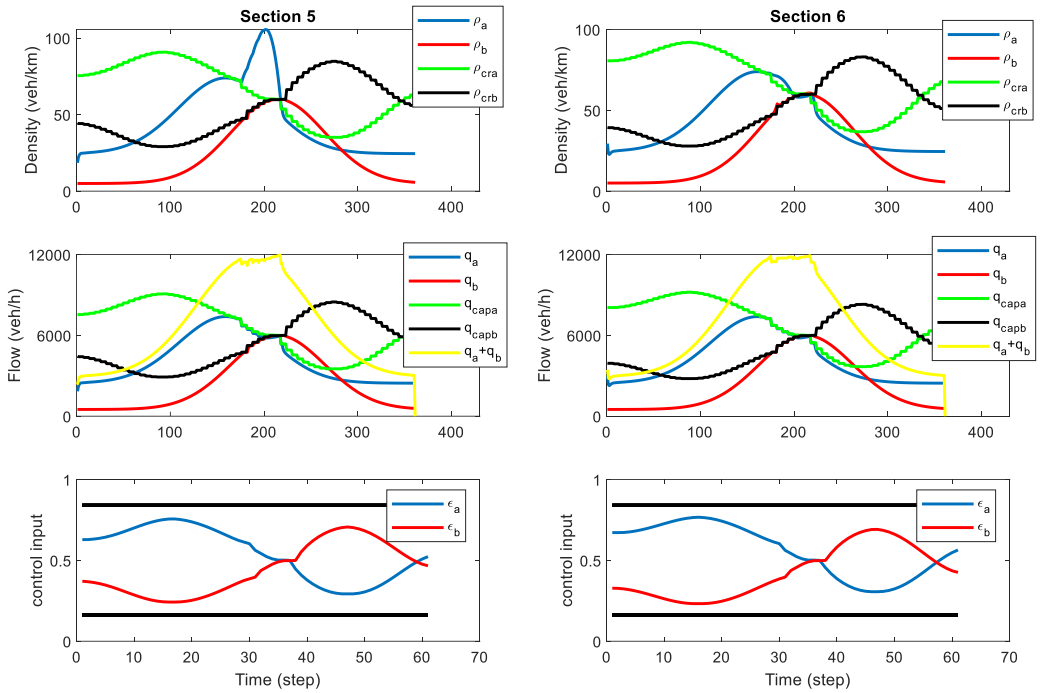


Figure 2.18: Congested scenario: Density, flow and control trajectories in the control case (sections 5 and 6)

Table 2.1. The value of TTS (veh·h) and related improvement (%) over the no-control case in different scenarios

Scenario	No-control case	QP results	CTM simulation with QP control
Uncongested with capacity drop	231.9	164.9 (-28.9 %)	164.9 (-28.9 %)
Uncongested without capacity drop	209.8	164.8 (-21.4 %)	164.9 (-21.4 %)
Congested with capacity drop	236.0	170.4 (-27.8 %)	171.0 (-27.5 %)
Congested without capacity drop	213.9	170.1 (-20.5 %)	170.9 (-20.1 %)

2.5 Conclusions

In this study, a new traffic control concept called internal boundary control has been presented, which is applicable in lane-free CAV traffic. Based on this concept, the capacity flow for each traffic direction is not constant, but can be flexibly adjusted according to the bi-directional demand and congestion conditions. For demonstration of the new concept, a quadratic programming formulation was adopted, which is fast enough for employment in a model predictive control (MPC) frame for real-time control.

Simulation case studies were designed for two scenarios that are representative for many others. Specifically, an uncongested scenario was considered first, where congestion appearing with fixed internal boundary (no-control case) can be utterly lifted with flexible boundary. This is likely the most common case in current highways and arterials, and the potential usage of the method would lead to full delay elimination in such cases. Secondly, a congested scenario was considered, where, due to strong bi-directional demand, congestion cannot be fully eliminated with optimal internal boundary control, but can be strongly mitigated with substantial benefits.

To achieve these results, the new concept necessitated the elaboration of novel definitions, notions and developments, including:

- The sharing factor, which is a novel traffic control input extending over space (sections) and time; and its inclusion into macroscopic traffic flow models, so as to reflect the corresponding traffic impact.
- The relative density, to help distinguishing under-critical from over-critical traffic conditions;
- The novel incurred notion of bottleneck, which, in contrast to conventional traffic, now refers to both traffic directions simultaneously.

As mentioned earlier, Duell et al. (2016) considered a lane reversal approach, incorporated in a system-optimal route guidance problem, whereby integer variables were needed to account for segment-wise whole-lane swapping to one or the other traffic direction. The incremental internal boundary control problem for highway stretches presented in this paper provides a good basis for considering management or planning tasks at the network level, such as route guidance, dynamic traffic assignment or network design problems. The significant advantage in such an approach would be that all capacity share variables (sharing factors) would be real-valued, something that would strongly simplify the related mathematical problems and accelerate solution algorithms. In addition, internal boundary control, as a real-time traffic management measure, may be integrated with other traffic management tools, such as ramp metering and speed control, for further increase of traffic flow efficiency where possible.

3 LQR for IBC of lane-free automated vehicle traffic

3.1 Abstract

Lane-free vehicle movement has been recently proposed for connected automated vehicles (CAV) due to various potential advantages. One such advantage stems from the fact that incremental changes of the road width in lane-free traffic lead to corresponding incremental changes of the traffic flow capacity. Based on this property, the concept of internal boundary control was recently introduced to flexibly share the total road width and capacity among the two traffic directions of a highway in real-time, in response to the prevailing traffic conditions, so as to maximize the cross-road (both directions) infrastructure utilization. Feedback-based Linear-Quadratic regulators with or without Integral action (LQI and LQ regulators) are appropriately developed in this paper to efficiently address the internal boundary control problem. Simulation investigations, involving a realistic highway stretch and different demand scenarios, demonstrate that the proposed simple regulators are robust and similarly efficient as an open-loop nonlinear constrained optimal control solution, while circumventing the need for accurate modelling and external demand prediction.

Keywords: lane-free traffic, internal boundary control, linear-quadratic regulator, capacity sharing

3.2 Introduction

Recurrent traffic congestion on urban freeways, highways and arterials is an increasingly serious problem for most big cities around the world, causing substantial delays, increased fuel consumption, excessive environmental pollution and reduced traffic safety. Traffic management measures, utilizing conventional means, are valuable (Papageorgiou et al., 2003; Kurzhanskiy and Varaiya, 2010) and, in some cases, able to delay or even avoid the onset of congestion. However, they are not always sufficient to tackle heavily congested traffic conditions. Gradually emerging and future ground-breaking vehicle automation and communication systems should be exploited to develop innovative solutions that can be applied within a smart road infrastructure.

Vehicle automation systems range from different kinds of driver support systems (e.g. ACC and lane-assist systems that are already commercially available with most car manufacturers) to highly or fully automated driving (i.e. SAE levels 4 and 5 vehicles); while vehicle communication enables vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication that may support various potential applications. Most vehicle manufacturers and, lately, some information-technology companies (like Waymo/Google), as well as research institutions, have been developing and testing in real traffic conditions high-automation or virtually driverless autonomous vehicles that monitor their environment and make sensible driving decisions based on appropriate decision and control algorithms (Ardelt et al., 2012; Aeberhard et al., 2015; Kamal et al., 2016; Makantasis and Papageorgiou, 2018).

Recently, Papageorgiou et al. (2021) launched the TrafficFluid concept, a novel paradigm for vehicular traffic that is applicable at high penetration rates of vehicles equipped with high levels

of vehicle automation and communication systems. The TrafficFluid concept suggests: (1) lane-free traffic, whereby vehicles are not bound to fixed traffic lanes, as in conventional traffic; (2) vehicle nudging, whereby vehicles may exert a "nudging" effect on, i.e. influence the movement of vehicles in front of them. Vehicles in a lane-free environment do not necessarily align to form lanes, but are self-organizing into dynamically changing 2-D clusters, depending on the vehicle sizes, their desired speeds, the employed vehicle movement strategies and the prevailing density, so as to maximize the available infrastructure utilization. Thus, if the road width is increased or decreased by some amount, the vehicles, driven by their movement strategy, spread accordingly to cover the changed 2-D road space, and new possible clusters may form. As demonstrated by Papageorgiou et al. (2021), if the road is widened by some amount, vehicles immediately move laterally to cover the free space. This increases the average inter-vehicle spacing, thus allowing for higher vehicle speeds and hence higher flow and capacity. In this context, the internal boundary control concept, introduced by Malekzadeh et al. (2021), exploits the lane-free principle of TrafficFluid, specifically the property that the road capacity may exhibit incremental (increasing or decreasing) changes in response to corresponding incremental (widening or narrowing) changes of the road width. This is in contrast to lane-based roads and traffic, where capacity changes may only occur if the road width is changed by a lane or a lane-multiple.

Consider a road with two opposite traffic directions serving connected automated vehicles (CAVs). The total available cross-road capacity (for both directions) may be shared between the two directions in a flexible way, according to the prevailing demand per direction, so as to maximize the infrastructure exploitation. Flexible capacity sharing may be achieved by virtually moving the internal boundary, which separates the two traffic directions, and communicating this decision to CAVs, so that they respect the changed road boundary. This way, the road width portion (and total capacity share) assigned to each direction can be changed in space and time (subject to constraints) according to an appropriate real-time control strategy, as illustrated in Fig 3.1, so as to maximize the total traffic efficiency of the overall system.

The idea of sharing the total cross-road capacity among the two traffic directions is not new and has been occasionally employed for conventional lane-based traffic, typically with manual interventions (see Wolshon and Lambert, 2006; Ampountolas et al. (2020) and references therein for a review of systems that exist around the globe). The measure is known as tidal flow (or contra-flow or reversible lanes) control, and its main principle is to adapt the total available cross-road supply to the demand per direction. Its most basic form is the steady allocation of one (or more) lanes of one direction to the other direction for a period of time in the aim of addressing abnormal traffic supply or demand in one traffic direction, e.g. at work zones or at big events, holiday departure or return, evacuation etc. More advanced reversible lane control systems may operate in real time, e.g. to balance delays on both sides of a known bottleneck (e.g. bridge, tunnel) by assigning a lane to one of the two directions in alternation in response to the prevailing traffic conditions. To this end, optimal control or feedback control algorithms of various types were proposed by Xue and Dong (2000), Frejo et al. (2015) and Ampountolas et al. (2020).

Reversible lanes have also been considered in connection with lane-based CAV driving. Duell et al. (2015) use the system optimal dynamic traffic assignment models formulated by

Ziliaskopoulos (2000) for a single destination and by Li et al. (2003) for more general networks, using the Cell Transmission Model (CTM) by Daganzo (1994). Lanes are introduced as integer variables, and the problem is formulated as a mixed integer linear programming (MILP) problem that has, however, high (exponential) complexity due to the many integers variables involved. Levin and Boyles (2016) use this model for a single link and utilize stochastic demand as a Markov decision process. The MILP problem is solved using a heuristic and is incorporated within a UE routing problem.

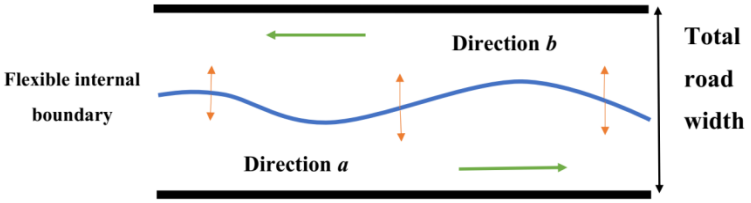


Figure 3.1: Space-time flexible internal road boundary.

The use of tidal flow control systems in lane-based traffic is not widespread for a number of reasons (see Malekzadeh et al. (2021)), such as: harsh resolution of infrastructure sharing (only by lane quanta) among the two traffic directions; serious counter-problems due to frequent merging or diverging traffic at lane-drop or lane-gain areas; safety-induced time-delays after each lane switch. These serious difficulties entail very limited capacity sharing flexibility in space and time and hinder reversible lane control from being a major traffic management measure. Even in the future CAV traffic, some of the mentioned difficulties would persist in lane-based conditions, notably the low capacity sharing resolution and the merging nuisance.

In contrast, in a lane-free CAV traffic environment, the mentioned difficulties are largely mitigated (Malekzadeh et al. (2021)). More specifically:

- The resolution of road-width sharing among the two directions can be high, still leading to corresponding intended capacity changes for the two opposite traffic directions.
- Assuming smooth CAV driving on a lane-free road surface, the internal boundary may be a smooth space-function, as illustrated in Fig 3.1, which may be smoothly changed in real time in response to the prevailing traffic conditions.
- Assuming moderate changes of the internal boundary over time and space, the aforementioned safety-induced time-delay may be very small.

Thanks to these characteristics, real-time internal boundary control for lane-free CAV traffic may be broadly applicable to the high number of arterial or highway infrastructures that feature unbalanced demands during the day in the two traffic directions, so as to strongly mitigate or even utterly avoid congestion. Even for infrastructures experiencing strong demand in both directions quasi-simultaneously, real-time internal boundary control may intensify the road utilization and lead to sensible improvements.

Malekzadeh et al. (2021) analyzed the internal boundary control problem and demonstrated its high improvement potential by formulating and solving a model-based open-loop nonlinear

constrained optimal control problem, which requires accurate modelling and external demand prediction, in the form of a convex Quadratic Programming (QP) problem. That approach considers explicitly the optimization of a physical measure of performance Total Time Spent (TTS). Although the application of such a controller appears difficult in the field, it was developed in order to be able to know the upper limit of performance improvement that can be achieved for any scenario in a macroscopic simulation environment. In fact, the open-loop approach may be used within a Model Predictive Control (MPC) frame with updated external demand predictions for real-time application; however, simpler real-time algorithms, that do not call for demand prediction and reach a similar efficiency level as the mentioned QP approach, are preferable. This paper develops and investigates the application of Linear-Quadratic regulators with or without Integral action (LQI and LQ regulators) for the internal boundary control problem. The well-known CTM is used, after linearization, for controller design; while its full nonlinear form is used for simulation testing of the developed controllers. Using the same highway stretch as in (Malekzadeh et al., 2021), a couple of well-designed and challenging demand scenarios are considered, and the performance of the LQ regulators is compared to each other, to the no-control case, as well as to the optimal results obtained using the QP formulation by Malekzadeh et al. (2021) with perfect prediction (upper limit of achievable performance improvement).

Section 3.3 presents some background issues, specifically the CTM equations and an outline of the QP problem formulation by Malekzadeh et al. (2021). Section 3.4 presents the design of the LQ and LQI regulators. Section 3.5 presents the simulation investigations, while conclusions are drawn in Section 3.6.

3.3 Background issues

3.3.1 Sharing factor

The internal boundary control problem should be designed in a macroscopic setting, so as to account for the traffic conditions on the highway and act accordingly. Lane-free traffic is not expected to give rise to structural changes of existing macroscopic traffic flow models. It is reasonable to assume, as also supported by results in (Bhavathrathan and Mallikarjuna, 2012; Asaithambi et al., 2016; Munigety et al., 2016; Papageorgiou et al., 2021), that notions and concepts like the conservation equation, the Fundamental Diagram (FD), as well as moving traffic waves will continue to characterize macroscopic traffic flow modelling in the case of CAV lane-free traffic. Additionally, specific physical traffic parameters, such as free speed, critical density, flow capacity, jam density, are also relevant for lane-free traffic, but may of course take different values than in lane-based traffic. The exact values that these parameters will take in CAV lane-free traffic are of minor importance for the presented demonstration of the novel internal boundary control measure.

Let us call the two opposite traffic directions, presented in Fig 3.1, directions a and b, respectively. We assume that, at specific road sections, each direction is assigned a respective road width $w^a = \varepsilon \cdot w$ and $w^b = (1-\varepsilon) \cdot w$, where $0 \le \varepsilon \le 1$ is the sharing factor, to be specified in real time as a control input by the internal boundary controller, and w is the total road width (both directions).

Let $Q(\rho)$, where ρ is the traffic density in veh/km, be the FD of a road section, which would apply if the whole road width would be assigned to only one of the two opposite traffic directions (i.e. for ε equal 0 or 1), with total critical density ρ_{cr}, total capacity q_{cap} (in veh/h) and total jam density $\rho_{\max}$. Let us now consider the case of partial road sharing, i.e. $\varepsilon_{\min} \le \varepsilon \le \varepsilon_{\max}$, where $\varepsilon_{\min}, \varepsilon_{\max} \in (0,1)$ are appropriate bounds aiming to suppress utter closure of either direction. As shown by Malekzadeh et al. (2021), the FDs for the two directions are functions of ε given by

$$\begin{aligned} Q^a(\rho^a, \varepsilon) &= \varepsilon \cdot Q(\rho^a / \varepsilon) \\ Q^b(\rho^b, \varepsilon) &= (1-\varepsilon) \cdot Q(\rho^b / (1-\varepsilon)) \end{aligned} \tag{3.1}$$

where ρ^a and ρ^b (in veh/km) are the respective densities of the two directions. The critical density $\rho_{cr}^a(\varepsilon)$, capacity $q_{cap}^a(\varepsilon)$, and jam density $\rho_{\max}^a(\varepsilon)$ for direction a are thus functions of the sharing factor ε and are given (Malekzadeh et al., 2021) by $\rho_{cr}^a(\varepsilon) = \varepsilon \cdot \rho_{cr}$, $q_{cap}^a(\varepsilon) = \varepsilon \cdot q_{cap}$ and $\rho_{\max}^a(\varepsilon) = \varepsilon \cdot \rho_{\max}$, i.e. the sharing factor ε scales density (veh/km) and flow (veh/h), leaving unaffected the speed (km/h). The corresponding relations for direction b are $\rho_{cr}^b(\varepsilon) = (1-\varepsilon) \cdot \rho_{cr}$, $q_{cap}^b(\varepsilon) = (1-\varepsilon) \cdot q_{cap}$ and $\rho_{\max}^b(\varepsilon) = (1-\varepsilon) \cdot \rho_{\max}$. All relations for the FD parameters hold for FDs that may be smooth or non-differentiable at ρ_{cr}, e.g. the triangular FD; see Figure 3.2 for illustration; see also (Ampountolas et al., 2020).

For controller design, a dynamic traffic flow model must be used. A simple but realistic option is CTM. CTM is a first-order dynamic traffic flow model with a triangular FD, which attains a space-time discretized form by application of the Godunov numerical scheme. The following section presents CTM and appropriate adjustments that have been introduced to incorporate capacity drop and the effect of the sharing factor ε.

3.3.2 Extended CTM

Consider a highway stretch with two reverse traffic directions a (from left to right) and b (from right to left). The stretch is subdivided into n sections, with lengths L_i, $i = 1, 2, \ldots, n$. As explained in the previous section, the total road width, which is assumed constant over all sections for simplicity, can be flexibly shared among the two directions in real time. As the sharing may be different for every section, we have corresponding sharing factors ε_i, $i = 1, 2, \ldots, n$; and (3.1) applies to each section. As a consequence, the total section capacity, as well as the critical density and jam density, are shared among traffic directions a and b according to

$$\begin{aligned} & q_{i,cap}^a(\varepsilon_i) = \varepsilon_i \cdot q_{cap}, \; q_{i,cap}^b(\varepsilon_i) = (1-\varepsilon_i) \cdot q_{cap} \\ & \rho_{i,cr}^a(\varepsilon_i) = \varepsilon_i \cdot \rho_{cr}, \; \rho_{i,cr}^b(\varepsilon_i) = (1-\varepsilon_i) \cdot \rho_{cr} \\ & \rho_{i,\max}^a(\varepsilon_i) = \varepsilon_i \cdot \rho_{\max}, \; \rho_{i,\max}^b(\varepsilon_i) = (1-\varepsilon_i) \cdot \rho_{\max}. \end{aligned} \tag{3.2}$$

The corresponding changes of the triangular FD that may occur at each section and traffic direction are illustrated in Figure 3.2. More specifically, when the value of the sharing factor is 0.5, i.e., the flow capacities of the two directions are equal, their FDs are "nominal" (blue line with $(.)^N$ parameters); when the sharing factor is different than 0.5, we have two FDs: the

extended one (green line with $(.)^E$ parameters) applies to the direction that is assigned more width and hence more flow capacity, and the reduced, complementary FD (orange line with $(.)^R$ parameters) applies to the other direction that is assigned less width and flow capacity. Based on (3.2), all FD parameters of a section change, whenever it is decided to change the corresponding sharing factor in real time. The exact values of the specific physical traffic parameters used in the CTM model, such as free speed, critical density, flow capacity, jam density, will depend on the vehicles' physical dimensions (width and length) and desired speed distribution, the total width of the road and the vehicle moving strategy employed for lane-free traffic.

As stated earlier, the sharing factors take values $\varepsilon_i \in [0,1]$. However, for the internal boundary control problem, we would like to disallow the utter closure of either direction; hence, the assigned road width in either direction should never be smaller than the widest vehicles driving on the road. This requirement gives rise to stricter constraints for the sharing factors as follows

$$0 < \varepsilon_{i,\min} \le \varepsilon_i \le \varepsilon_{i,\max} < 1 \tag{3.3}$$

where $\varepsilon_{i,\min} \cdot w$ and $(1-\varepsilon_{i,\max}) \cdot w$ are the minimum admissible widths to be assigned to directions a and b, respectively.

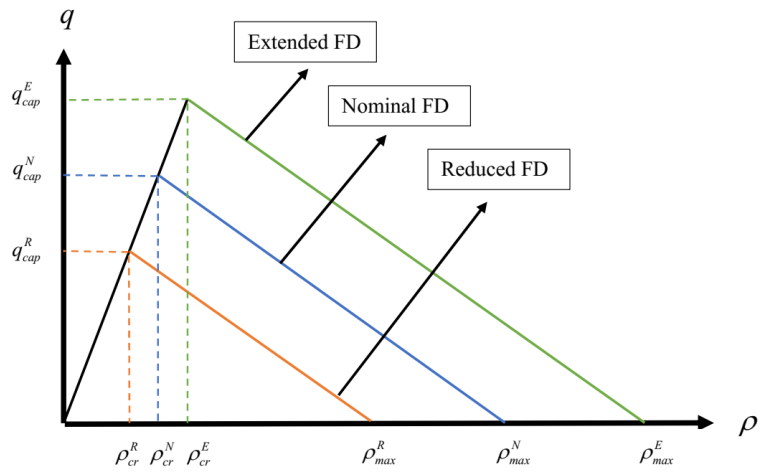


Figure 3.2: The triangular fundamental diagram with flexible internal boundary.

Another restriction to be applied to the sharing factors concerns the time-delay needed to evacuate traffic on the direction that receives a restricted width, compared with the previous control time-step. As discussed earlier, this time-delay is small in lane-free CAV traffic without physical barrier among the two traffic directions and with moderate changes of the sharing factors applied to short sections, but needs nevertheless to be considered. Clearly, the time-delay should apply only to the traffic direction that is being widened, compared to the previous control interval; while the direction that is restricted should promptly apply the smaller width, so that CAVs therein move out of the reduced-width zone. Assume that the required time-delay is smaller than or equal to the control time interval T_c; then, the time-delay requirement is

automatically fulfilled for each section i, if the sharing factors that are actually applied to the two directions, i.e. ε_i^a and ε_i^b, respectively, are calculated as follows

$$\varepsilon_i^a(k_c) = \min\left\{\varepsilon_i(k_c), \varepsilon_i(k_c - 1)\right\} \tag{3.4}$$

$$\varepsilon_i^b(k_c) = \min\left\{1 - \varepsilon_i(k_c), 1 - \varepsilon_i(k_c - 1)\right\} \tag{3.5}$$

where $k_c = 0,1,\ldots$ is the discrete control time index. It is noted that the notation $\varepsilon_i^a(k_c)$ and $\varepsilon_i^b(k_c)$ indicates that the sharing factors are applied for the duration of the control time interval $[k_c \cdot T_c, (k_c + 1) \cdot T_c)$. The above equations may be readily extended if the required time-delay is a multiple of the control time interval T_c.

Traffic flows from section 1 to section n in direction *a*; and from section n to section 1 in direction b (see Figure 3.3 as an example). We denote ρ_i^a, $i = 1,2,\ldots,n$, the traffic density of section *i*, direction *a*; and ρ_i^b, $i = 1,2,\ldots,n$, the traffic density of section *i*, direction b. Similarly, we denote q_i^a, $i = 1,2,\ldots,n$, and q_i^b, $i = 1,2,\ldots,n$, the mainstream exit flows of section i for directions a and b, respectively. Thus, q_0^a is the feeding upstream mainstream inflow for direction *a*; and q_{n+1}^b is the feeding upstream mainstream inflow for direction b. Every section, except for the most upstream in each direction, may have an on-ramp or an off-ramp at its upstream boundary. The on-ramp flows (if any) at section i are denoted r_i^a for direction *a*, and r_i^b for direction *b*. The off-ramp flow (if any) of section *i*, direction *a*, is calculated based on known exit rates β_i^a multiplied with the upstream-section flow, i.e. $\beta_i^a q_{i-1}^a$; and the off-ramp flow (if any) of section *i*, direction b, is calculated based on known exit rates β_i^b multiplied with the upstream-section flow, i.e. $\beta_i^b q_{i+1}^b$.

The conservation equations for the sections of direction a are:

$$\begin{aligned}
\rho_1^a(k+1) &= \rho_1^a(k) + \frac{T}{L_i}\left(q_0^a(k) - q_1^a(k)\right) \\
\rho_i^a(k+1) &= \rho_i^a(k) + \frac{T}{L_i}\left((1-\beta_i^a)q_{i-1}^a(k) - q_i^a(k) + r_i^a(k)\right),\ i = 2,3,\ldots,n
\end{aligned} \tag{3.6}$$

where T is the model time-step, typically set equal to 5 – 10 s for section lengths of some 500 m in length, and $k = 0,1,\ldots$ is the corresponding discrete time index of the model.

According to CTM, traffic flow is obtained as the minimum of demand and supply functions, except for the last section, where only the demand function is considered, assuming that the downstream traffic conditions are uncongested. Clearly, when writing the demand and supply functions Q_D and Q_S, respectively, for the case of the internal boundary control problem, we need to consider the impact of the respective sharing factors $\varepsilon_i^a(k_c)$ on the FDs. Thus we have

$$\begin{aligned}
q_i^a(k) &= \min\left\{Q_D(\rho_i^a(k), \varepsilon_i^a(k_c)), \frac{Q_S(\rho_{i+1}^a(k), \varepsilon_{i+1}^a(k_c))}{(1-\beta_{i+1}^a)} - \lambda_r r_{i+1}^a(k)\right\},\ i = 1,2,\ldots,n-1 \\
q_n^a(k) &= Q_D(\rho_n^a(k), \varepsilon_i^a(k_c)).
\end{aligned} \tag{3.7}$$

The demand and supply functions are given by the following respective equations

$$
\begin{aligned}
Q_D(\rho,\varepsilon) &= \min\left\{\varepsilon q_{cap} + \lambda_d q_{cap}\frac{\rho-\varepsilon\rho_{cr}}{\rho_{cr}-\rho_{\max}}, v_f\rho\right\},\\
Q_S(\rho,\varepsilon) &= \min\left\{\varepsilon q_{cap}, w_s(\varepsilon\rho_{\max}-\rho)\right\},
\end{aligned}
\tag{3.8}
$$

where v_f is the free speed (which is assumed equal for all sections for simplicity) and w_s is the back-wave speed. The control time-step T_c does not need to be equal to the model time-step, but is assumed to be a multiple of T, in which case the control time index is given by $k_c = \lfloor kT/T_c \rfloor$, where $\lfloor . \rfloor$ is the integer part notation.

It is well-known that CTM does not reproduce the capacity drop phenomenon, i.e. the empirical finding that, at the head of congestion, the observed flow in real traffic is reduced compared to the road capacity. Capacity drop is deemed to occur in conventional traffic due to bounded and differing accelerations of different vehicles (Yuan et al., 2015). Recently, CTM has been extended in a number of possible ways to enable the reproduction of capacity drop, see (Kontorinaki et al., 2017) for an overview and comparison. The presence of capacity drop in conventional freeway traffic is a major reason for infrastructure degradation and for the need of introducing traffic control measures to restore capacity (Papageorgiou et al., 2003). In contrast, in the present context of internal boundary control, the presence of capacity drop marks a secondary source of amelioration of the traffic conditions, because the potential benefits achievable via opportune capacity sharing are expected to be much higher than those resulting from capacity drop avoidance. In fact, it is unknown at the moment, if and to what extent capacity drop may occur in lane-free CAV traffic. To be able to investigate the impact of possible capacity drop, we have incorporated in the above equations the option of introducing capacity drop trough appropriate terms according to Kontorinaki et al. (2017). More specifically, this option is enabled via the parameters λ_r and λ_d in equations (3.7) and (3.8). If these parameters are set to $\lambda_r = 1$ and $\lambda_d = 0$, no capacity drop is reproduced, as typical for CTM; if these values are set between 0 and 1, a corresponding level of capacity drop is produced by the model.

The equations for direction b are analogous to those of direction a, with few necessary index modifications. Section numbers in direction b are descending, hence we have

$$
\begin{aligned}
\rho_i^b(k+1) &= \rho_i^b(k) + \frac{T}{L_i}\left((1-\beta_i^b)q_{i+1}^b(k) - q_i^b(k) + r_i^b(k)\right),\ i = 1,2,\ldots,n-1\\
\rho_n^b(k+1) &= \rho_n^b(k) + \frac{T}{L_i}\left(q_{n+1}^b(k) - q_n^b(k)\right)
\end{aligned}
\tag{3.9}
$$

and the flows are given by

$$
\begin{aligned}
q_1^b(k) &= Q_D(\rho_1^b(k), \varepsilon_i^b(k_c))\\
q_i^b(k) &= \min\left\{Q_D(\rho_i^b(k), \varepsilon_i^b(k_c)), \frac{Q_S(\rho_{i-1}^b(k), \varepsilon_{i-1}^b(k_c))}{(1-\beta_{i-1}^b)} - \lambda_r r_{i-1}^b(k)\right\},\ i = 2,3,\ldots,n
\end{aligned}
\tag{3.10}
$$

3.3.3 The QP problem formulation

To be able to assess the level of performance of the LQ regulators to be developed, we will compare its results with those obtained with a nonlinear constrained optimal control scheme developed for the same problem by Malekzadeh et al. (2021). That scheme is based on the formulation and solution of a convex Quadratic Programming (QP) problem that is outlined here for completeness.

The CTM, presented in Section 3.3.2, is utilized in the problem formulation. Due to the presence of the min-operators in (3.7), (3.8) and (3.10), CTM is a nonlinear model. As proposed by Papageorgiou (1995) and practiced in most previous utilizations of CTM for optimal control (Ziliaskopoulos, 2000; Gomes and Horowitz, 2006; Roncoli et al., 2015), such nonlinearities may be transformed to linear inequalities by requesting the left-hand side of the equation, where the min-operator appears, to be smaller than or equal to each one of the terms included in the min-operator. As a result, the above mentioned equations of CTM are replaced in this formulation by respective groups of linear inequalities for each section and direction. Similarly, each one of the equations (3.4) and (3.5) is replaced by two linear inequalities. Finally, the constraints (3.3) must be considered at each discrete point to appropriately limit the sharing factors.

All these inequalities, presented in detail by Malekzadeh et al. (2021), transform the CTM equations, including the changing sharing factors and their constraints, to a set of linear equalities and inequalities, which thus constitute a convex feasible region for an optimization problem. Combining this model with a convex quadratic objective function, to be minimized, results in a QP formulation that can be solved numerically using very efficient available codes. The objective function is as follows:

$$\begin{aligned} J_{QP} = {} & T\sum_{k=1}^{K}\sum_{i=1}^{n}\left(L_i\rho_i^a(k)+L_i\rho_i^b(k)\right) - w_1\sum_{k_c=0}^{K_c-1}\sum_{i=1}^{n}\left(\varepsilon_i^a(k_c)+\varepsilon_i^b(k_c)\right) \\ & + w_2\sum_{k_c=1}^{K_c-1}\sum_{i=1}^{n}\left(\varepsilon_i(k_c)-\varepsilon_i(k_c-1)\right)^2 + w_3\sum_{k_c=0}^{K_c-1}\sum_{i=2}^{n}\left(\varepsilon_i(k_c)-\varepsilon_{i-1}(k_c)\right)^2 \\ & + w_4\sum_{k_c=0}^{K_c-1}\sum_{i=1}^{n}\left(\frac{\varepsilon_i(k_c)^2}{\hat{d}_i^a(k_c)}+\frac{(1-\varepsilon_i(k_c))^2}{\hat{d}_i^b(k_c)}\right). \end{aligned} \tag{3.11}$$

The cost function extends over a time horizon of K model time-steps or K_c control time-steps; it includes five terms, the first two being linear and the rest of them quadratic. The first term is the most important one and represents the Total Time Spent (TTS) that determines the traffic efficiency resulting from the proposed control actions. The second term ensures that one of the two inequalities introduced to replace equation (3.4), and one of the two inequalities introduced to replace equation (3.5) will be activated and, as a result, one of the two terms included in the min-operators of equations (3.4) and (3.5) will actually materialize. The following three terms are quadratic and reflect secondary operational and policy objectives. The first of these quadratic terms penalizes the variation of the control input in consecutive time-steps, so that changes of the internal boundary of each section from one control time-step to the next remain small. The

second quadratic term penalizes the space variation of the control input from section to section, so as to mitigate strong changes in the road width assigned to each direction within a short distance. The last quadratic term is policy related and attempts to assign to the two directions respective capacity shares that balance the respective capacity reserves for each section. To this end, we use $\hat{d}_i^a(k_c)$ and $\hat{d}_i^b(k_c)$, $i=1,..,n,$ which are the projected demand trajectories for each section of the two respective directions. The weights w_1, w_2, w_3, w_4 were tuned using various demand scenarios in order to guarantee that the minimum achievable TTS value is always achieved. More details are provided by Malekzadeh et al. (2021).

3.4 Design of the LQ regulators

3.4.1 Relative densities

In conventional traffic management, traffic densities have a central role, as they reflect unequivocally the state of traffic, e.g. on the FD. In the novel internal boundary control setting, however, the traffic density variables ρ_i^a and ρ_i^b (in veh/km) in the two opposite directions of each section i are not directly indicating the traffic conditions (e.g. under-critical or congested) encountered. This is because the critical density for each direction is a function of the sharing factor and is changing according to the applied control action; thus, the same density value may mark an under-critical or congested traffic state, depending on the value of the critical density; or, equivalently, depending on the share of road width among the two directions.

Therefore, we proceed with the definition of new variables that are appropriate to reflect the actual traffic conditions in the internal boundary control context. The new variables are the *relative densities* (dimensionless) and are defined per section and per direction as the ordinary densities. The relative density of section i and direction a or b is obtained by dividing the corresponding traffic density with the corresponding critical density, which, on its turn, depends on the sharing factor prevailing during the last time-step. Considering (3.2), we get the following relations for the relative densities in the two directions

$$\tilde{\rho}_i^a(k)=\frac{\rho_i^a(k)}{\varepsilon_i(k-1)\rho_{cr}},\ \tilde{\rho}_i^b(k)=\frac{\rho_i^b(k)}{(1-\varepsilon_i(k-1))\rho_{cr}},\ i=1,2,\ldots,n. \tag{3.12}$$

Note that, according to the respective variable definitions, which are typical in discrete-time modelling, the sharing factor $\varepsilon_i(k-1)$ applies during the time interval $[(k-1)T, kT)$; while a density $\rho_i^a(k)$ or $\rho_i^b(k)$ is an instantaneous value at time instant kT; and these definitions explain the choice of the discrete-time arguments in (3.12).

As evidenced with equations (3.4) and (3.5), a control time delay must be applied only to the direction that is being widened, not to the direction that is being narrowed. However, this distinction is not possible in the design of the LQ regulators, as (3.4) and (3.5) cannot be explicitly considered in the linearized control design model. Thus, we have the option of either applying the time delay to both traffic directions or to none. Both options were tried and resulted in similar results. Therefore, given that the first option complicates the design model, we opted for the second option in (3.12). Please note that the correct control time delays are always

applied, according to equations (3.4) and (3.5), in the simulation model while testing the regulators.

The relative densities reflect unequivocally the state of traffic in the internal boundary control context. Specifically, the relative density of a section and direction reflects under-critical traffic conditions if it is in the range $[0,1)$; capacity flow if it equals 1; and over-critical traffic conditions if it is bigger than 1.

3.4.2 Linearized model

In order to derive the LQ regulators, we need to include in our problem a discrete-time linearized system. To this end, we use as a basis the CTM equations delivered in the previous section. Replacing (3.12) in (3.6), i.e. in the conservation equation for direction a, we get

$$\varepsilon_i(k)\rho_{cr}\tilde{\rho}_i^a(k+1) = \varepsilon_i(k-1)\rho_{cr}\tilde{\rho}_i^a(k) + \frac{T}{L_i}((1-\beta_i^a)q_{i-1}^a(k) - q_i^a(k) + r_i^a(k)),\ i = 2,3,\ldots,n$$
$$\varepsilon_1(k)\rho_{cr}\tilde{\rho}_1^a(k+1) = \varepsilon_1(k-1)\rho_{cr}\tilde{\rho}_1^a(k) + \frac{T}{L_i}(q_0^a(k) - q_1^a(k)) \tag{3.13}$$

and by dividing with $\varepsilon_i(k-1)\rho_{cr}$ we get

$$\frac{\varepsilon_i(k)}{\varepsilon_i(k-1)}\tilde{\rho}_i^a(k+1) = \tilde{\rho}_i^a(k) + \frac{T}{L_i\rho_{cr}}\left((1-\beta_i^a)\frac{q_{i-1}^a(k)}{\varepsilon_i(k-1)} - \frac{q_i^a(k)}{\varepsilon_i(k-1)} + \frac{r_i^a(k)}{\varepsilon_i(k-1)}\right),\ i = 2,3,\ldots,n$$
$$\frac{\varepsilon_1(k)}{\varepsilon_1(k-1)}\tilde{\rho}_1^a(k+1) = \tilde{\rho}_1^a(k) + \frac{T}{L_i\rho_{cr}}\left(\frac{q_0^a(k)}{\varepsilon_1(k-1)} - \frac{q_1^a(k)}{\varepsilon_1(k-1)}\right). \tag{3.14}$$

It is now assumed that the values of the control inputs do not change significantly in consecutive sample times. In fact, as we will see at the end of this section, the control time-step is a multiple M of the model time-step, hence the control input changes only every M model time-steps. In addition, due to the quadratic penalization of the control inputs within the objective function (3.29), that is defined later, the control input changes moderately over time (at each control interval). As a result, it is reasonable to assume that $\varepsilon_i(k)/\varepsilon_i(k-1) \simeq 1,\ i = 1,2,\ldots,n.$ This assumption simplifies significantly the equations and decreases the burden of derivative calculations necessary for the linearization of the model, without really affecting the control results achieved, as found in corresponding result comparisons. Defining also the one-step retarded control input as a new state variable according to

$$\gamma_i(k+1) = \varepsilon_i(k),\ i = 1,2,\ldots,n, \tag{3.15}$$

we finally get the state equations that replace the conservation equations for direction a

$$\tilde{\rho}_i^a(k+1) = \tilde{\rho}_i^a(k) + \frac{T}{L_i \rho_{cr}} \left((1-\beta_i^a) \frac{q_{i-1}^a(k)}{\gamma_i(k)} - \frac{q_i^a(k)}{\gamma_i(k)} + \frac{r_i^a(k)}{\gamma_i(k)} \right), \; i = 2,3,\ldots,n$$
$$\tilde{\rho}_1^a(k+1) = \tilde{\rho}_1^a(k) + \frac{T}{L_1 \rho_{cr}} \left(\frac{q_0^a(k)}{\gamma_1(k)} - \frac{q_1^a(k)}{\gamma_1(k)} \right). \tag{3.16}$$

Similarly, replacing (3.12) in (3.6), i.e. in the conservation equation for direction b, and making the same assumption for the value of the control input in consecutive sample times, we get the state equations that replace the conservation equations for direction b

$$\tilde{\rho}_i^b(k+1) = \tilde{\rho}_i^b(k) + \frac{T}{L_i \rho_{cr}} \left((1-\beta_i^b) \frac{q_{i+1}^b(k)}{1-\gamma_i(k)} - \frac{q_i^b(k)}{1-\gamma_i(k)} + \frac{r_i^b(k)}{1-\gamma_i(k)} \right), \; i = 1,2,...,n-1$$
$$\tilde{\rho}_n^b(k+1) = \tilde{\rho}_n^b(k) + \frac{T}{L_n \rho_{cr}} \left(\frac{q_{n+1}^b(k)}{1-\gamma_n(k)} - \frac{q_n^b(k)}{1-\gamma_n(k)} \right). \tag{3.17}$$

According to CTM, traffic flow is obtained as the minimum of the demand and supply functions. Starting from CTM, we need to come up with an approximate linearized system to allow for the derivation of feedback controllers by use of the LQ regulator methodology. This is a procedure with a long history in traffic management and control, see e.g. (Isaksen and Payne, 1973; Papageorgiou, 1984; Papageorgiou et al., 1990; Diakaki and Papageorgiou, 1994; Aboudolas and Geroliminis, 2013), to name just a few. In such works, the nominal linearization state (density) is taken to be uncongested and close to the critical density. Following this procedure in our case, we assume that traffic flow is nominally operating around capacity and is determined only by the demand function; which however is also obtained as the minimum of two terms according to (3.8), namely as the minimum of the flow capacity (without the capacity-drop term) and the FD flow-density relation. In conventional traffic control, the flow capacity is constant; hence a linear or linearized flow-density relation is taken from the FD to characterize the linearized flow dynamics. In contrast, in the internal boundary control case, the capacity is directly proportional to the control input (sharing factor), as evidenced from (3.8), hence this term is deemed more significant in the linearized approximation of the system dynamics.

To confirm these arguments, we have considered, in preliminary simulation testing, three cases: (i) Use of the FD flow-density relation only; (ii) Use of the capacity term only (omitting the capacity drop term); and (iii) Use of a convex combination of the two terms included in the min-operator of the demand function. The investigation confirmed that case (iii) is indeed most efficient and robust for all traffic conditions.

In conclusion, to proceed with the linearized approximation of system dynamics, $q_i^a(k)$ and $q_i^b(k)$ are given by

$$q_i^a(k) = \sigma \varepsilon_i(k) q_{cap} + (1-\sigma) v_f \rho_i^a(k)$$
$$q_i^b(k) = \sigma (1-\varepsilon_i(k)) q_{cap} + (1-\sigma) v_f \rho_i^b(k) \tag{3.18}$$

where $0 \le \sigma \le 1$ is the convex combination parameter used. Taking into account (3.12) and (3.15), (3.18) can be rewritten as

$$q_i^a(k) = \sigma\varepsilon_i(k)q_{cap} + (1-\sigma)v_f\tilde{\rho}_i^a(k)\gamma_i(k)\rho_{cr}$$
$$q_i^b(k) = \sigma(1-\varepsilon_i(k))q_{cap} + (1-\sigma)v_f\tilde{\rho}_i^b(k)(1-\gamma_i(k))\rho_{cr} \quad (3.19)$$

Considering these equations, the system of dynamic equations (3.15)-(3.17) can be rewritten in the following state-space form:

$$\tilde{\rho}_i^a(k+1) = f_i^a\left(\tilde{\rho}_i^a(k), \tilde{\rho}_{i-1}^a(k), \varepsilon_i(k), \varepsilon_{i-1}(k), \gamma_i(k), \gamma_{i-1}(k), d_i^a(k)\right), i = 2,3,\ldots,n$$
$$\tilde{\rho}_1^a(k+1) = f_1^a\left(\tilde{\rho}_1^a(k), \varepsilon_1(k), \gamma_1(k), d_1^a(k)\right) \quad (3.20)$$

$$\tilde{\rho}_i^b(k+1) = f_i^b\left(\tilde{\rho}_i^b(k), \tilde{\rho}_{i+1}^b(k), \varepsilon_i(k), \varepsilon_{i+1}(k), \gamma_i(k), \gamma_{i+1}(k), d_i^b(k)\right), i = 1,2,\ldots,n-1$$
$$\tilde{\rho}_n^b(k+1) = f_n^b\left(\tilde{\rho}_n^b(k), \varepsilon_n(k), \gamma_n(k), d_n^b(k)\right) \quad (3.21)$$

$$\gamma_i(k+1) = g_i(\varepsilon_i(k)), i = 1,\ldots,n \quad (3.22)$$

where the upstream mainstream inflow, as well as the on-ramp flows of each direction, which are external inputs (disturbances), have been included in the respective demand vectors $\mathbf{d}^a(k) = \left[q_0^a(k), r_2^a(k), \ldots, r_n^a(k)\right]^T$ and $\mathbf{d}^b(k) = \left[r_1^b(k), \ldots, r_{n-1}^b(k), q_{n+1}^b(k)\right]^T$.

Linearization of the system (3.20)-(3.22) around a nominal point, denoted by superscript N, yields

$$\Delta\tilde{\rho}_i^a(k+1) = \left.\frac{\partial f_i^a}{\partial\tilde{\rho}_i^a}\right|_N \Delta\tilde{\rho}_i^a(k) + \left.\frac{\partial f_i^a}{\partial\tilde{\rho}_{i-1}^a}\right|_N \Delta\tilde{\rho}_{i-1}^a(k) + \left.\frac{\partial f_i^a}{\partial\varepsilon_i}\right|_N \Delta\varepsilon_i(k) + \left.\frac{\partial f_i^a}{\partial\varepsilon_{i-1}}\right|_N \Delta\varepsilon_{i-1}(k)$$
$$+ \left.\frac{\partial f_i^a}{\partial\gamma_i}\right|_N \Delta\gamma_i(k) + \left.\frac{\partial f_i^a}{\partial\gamma_{i-1}}\right|_N \Delta\gamma_{i-1}(k), i = 2,3,\ldots,n \quad (3.23)$$
$$\Delta\tilde{\rho}_1^a(k+1) = \left.\frac{\partial f_1^a}{\partial\tilde{\rho}_1^a}\right|_N \Delta\tilde{\rho}_1^a(k) + \left.\frac{\partial f_1^a}{\partial\varepsilon_1}\right|_N \Delta\varepsilon_1(k) + \left.\frac{\partial f_1^a}{\partial\gamma_1}\right|_N \Delta\gamma_1(k)$$

$$\Delta\tilde{\rho}_i^b(k+1) = \left.\frac{\partial f_i^b}{\partial\tilde{\rho}_i^b}\right|_N \Delta\tilde{\rho}_i^b(k) + \left.\frac{\partial f_i^b}{\partial\tilde{\rho}_{i+1}^b}\right|_N \Delta\tilde{\rho}_{i+1}^b(k) + \left.\frac{\partial f_i^b}{\partial\varepsilon_i}\right|_N \Delta\varepsilon_i(k) + \left.\frac{\partial f_i^b}{\partial\varepsilon_{i+1}}\right|_N \Delta\varepsilon_{i+1}(k)$$
$$+ \left.\frac{\partial f_i^a}{\partial\gamma_i}\right|_N \Delta\gamma_i(k) + \left.\frac{\partial f_i^a}{\partial\gamma_{i+1}}\right|_N \Delta\gamma_{i+1}(k), i = 1,2,\ldots,n-1 \quad (3.24)$$
$$\Delta\tilde{\rho}_n^b(k+1) = \left.\frac{\partial f_n^b}{\partial\tilde{\rho}_n^b}\right|_N \Delta\tilde{\rho}_n^b(k) + \left.\frac{\partial f_n^b}{\partial\varepsilon_n}\right|_N \Delta\varepsilon_n(k) + \left.\frac{\partial f_n^a}{\partial\gamma_n}\right|_N \Delta\gamma_n(k)$$

$$\Delta\gamma_i(k+1) = \left.\frac{\partial g_i}{\partial\varepsilon_i}\right|_N \Delta\varepsilon_i(k), i = 1,2,\ldots,n \quad (3.25)$$

where $\Delta(.)(k) = (.)(k) - (.)^N$. Additionally, it has been assumed that $\Delta(.)(k) = 0$ for all disturbances, i.e. for the elements of the demands $\mathbf{d}^a(k)$ and $\mathbf{d}^b(k)$. The derivation of the partial derivatives necessary in (3.23)-(3.25), is presented in Appendix A.

Merging all the above, we obtain a linear state-space model

$$\boldsymbol{\Delta}\mathbf{x}(k+1) = \mathbf{A}\boldsymbol{\Delta}\mathbf{x}(k) + \mathbf{B}\boldsymbol{\Delta}\mathbf{u}(k) \tag{3.26}$$

where $\boldsymbol{\Delta}\mathbf{x}(k) = [\boldsymbol{\Delta}\tilde{\boldsymbol{\rho}}^a(k)^T, \boldsymbol{\Delta}\tilde{\boldsymbol{\rho}}^b(k)^T, \boldsymbol{\Delta}\boldsymbol{\gamma}(k)^T]^T$ is the state vector and $\boldsymbol{\Delta}\mathbf{u}(k) = \boldsymbol{\Delta}\boldsymbol{\varepsilon}(k)$ is the control input vector, whereby $\boldsymbol{\Delta}\tilde{\boldsymbol{\rho}}^a(k) = [\Delta\tilde{\rho}_1^a(k), \ldots, \Delta\tilde{\rho}_n^a(k)]^T$, $\boldsymbol{\Delta}\tilde{\boldsymbol{\rho}}^b(k) = [\Delta\tilde{\rho}_1^b(k), \ldots, \Delta\tilde{\rho}_n^b(k)]^T$, $\boldsymbol{\Delta}\boldsymbol{\gamma}(k) = [\Delta\gamma_1(k), \ldots, \Delta\gamma_n(k)]^T$ and $\boldsymbol{\Delta}\boldsymbol{\varepsilon}(k) = [\Delta\varepsilon_1(k), \ldots, \Delta\varepsilon_n(k)]^T$. $\mathbf{A} \in \mathbb{R}^{3n\times 3n}$ and $\mathbf{B} \in \mathbb{R}^{3n\times n}$ are the state and input matrices, respectively.

If the control time-step is defined as a multiple of the model time-step, i.e. $T_c = MT$, where M is an integer, then we have $k_c = \lfloor kT/T_c \rfloor = \lfloor k/M \rfloor$. Thus, the linear state-space equation may be changed as follows, in order to be based on the control time-step T_c,

$$\boldsymbol{\Delta}\mathbf{x}(k_c+1) = \mathbf{A}^M\boldsymbol{\Delta}\mathbf{x}(k_c) + \underbrace{(\mathbf{A}^{M-1} + \mathbf{A}^{M-2} + \ldots. + \mathbf{I})\mathbf{B}}_{\hat{\mathbf{B}}}\boldsymbol{\Delta}\mathbf{u}(k_c). \tag{3.27}$$

3.4.3 Quadratic cost function and integration states

The QP problem formulation of Section 3.3.3 provides sufficient freedom for considering, in the cost objective, the main physical traffic efficiency metric (TTS), as well as several secondary operational sub-objectives. In contrast, when employing the LQ methodology, the objective function is restricted to be a weighted quadratic norm of the states and of the control inputs. Nevertheless, we wish to achieve with the designed LQ regulators similarly efficient control results in terms of TTS, although TTS is not an explicit formal objective in the LQ problem formulation.

The state vector in (3.27) includes as state variables $\Delta\tilde{\rho}_i^a(k)$ and $\Delta\tilde{\rho}_i^b(k)$. By setting the nominal value of relative densities equal to 1, the squares of those state variables, to be included in the quadratic objective function, will motivate the resulting controller to operate the system near capacity, which is good for traffic efficiency. In addition, if capacity flow is not feasible (due to lack of demand or due to excessive demand), then minimizing a sum of squares has the tendency to balance deviations from the nominal values, something that is conform with the secondary operational sub-objective of balancing the margin to capacity across sections.

The quadratic objective also includes the sum of squares of the control inputs, i.e. of $\Delta\varepsilon_i(k)$. Using 0.5 as a nominal value for the sharing factors, minimization will have two effects: (i) to mitigate the deviations of the sharing factors from 0.5; and (ii) to balance these deviations in space (sections) and time, which are also secondary operational sub-objectives.

A final specification concerns the possible inclusion of integral terms in the regulator (LQI regulator), as often practiced in traffic control, see (Papageorgiou et al., 1990; Diakaki and Papageorgiou, 1994, Aboudolas and Geroliminis, 2013). The LQI regulator offers the possibility to drive to zero, in the steady state, a number of linear combinations of state variables, which cannot be higher than the number of control variables. In our case, the number of control variables equals the number of highway sections, and a good choice for the linear combinations to be considered could be $\Delta\tilde{\rho}_i^a(k) - \Delta\tilde{\rho}_i^b(k)$, which equals $\tilde{\rho}_i^a(k) - \tilde{\rho}_i^b(k)$ for nominal state values equal to 1. Striving for the same value of the relative densities on the two directions at

each section seems equitable, and also seems to foster further the aforementioned balancing of capacity margins.

In view of the above, we may now proceed with the formal design of an LQI regulator. To enable the inclusion of integral terms in the regulator, we consider the state equation (3.27) augmented by use of the integrators

$$\mathbf{y}(k_c+1)=\mathbf{y}(k_c)+\underbrace{\left[\mathbf{I}_n,-\mathbf{I}_n,\mathbf{0}_{n\times n}\right]}_{\mathbf{H}}\underbrace{\begin{bmatrix}\Delta\tilde{\boldsymbol{\rho}}^{\mathbf{a}}(k_c)\\ \Delta\tilde{\boldsymbol{\rho}}^{\mathbf{b}}(k_c)\\ \Delta\boldsymbol{\gamma}(k_c)\end{bmatrix}}_{\Delta\mathbf{x}(k_c)}. \tag{3.28}$$

We recall that, with this choice of integration state variables, we achieve a steady-state error equal to zero for the differences between the relative densities per direction for each section.

The control goal is to minimize the quadratic criterion

$$J_{LQI}=\frac{1}{2}\sum_{k_c=0}^{\infty}\left[\left\|\Delta\mathbf{x}(k_c)\right\|_{\mathbf{Q}}^2+\left\|\mathbf{y}(k_c)\right\|_{\mathbf{S}}^2+\left\|\Delta\mathbf{u}(k_c)\right\|_{\mathbf{R}}^2\right] \tag{3.29}$$

where $\mathbf{Q}\in\mathbb{R}^{3n\times 3n}$, $\mathbf{S}\in\mathbb{R}^{n\times n}$, $\mathbf{R}\in\mathbb{R}^{n\times n}$ are symmetric positive definite matrices, which will be chosen in Section 3.5 to be diagonal for simplicity of tuning. The first term penalizes deviations of the elements of the state variable $\Delta\mathbf{x}$ from zero, i.e. deviations of $\tilde{\rho}_i^a(k)$, $\tilde{\rho}_i^b(k)$, $\gamma_i(k)$, $i=1,2,\ldots,n$, from their desired nominal values. The second term penalizes deviations of the elements of the state variable $\mathbf{y}$ from zero, i.e. deviations of the integral of the differences between the relative densities per direction for each section from zero. Finally, the third term penalizes deviations of the control inputs from the nominal values (i.e. from 0.5).

Considering the discrete-time linear system (3.27), (3.28) and the quadratic criterion (3.29), the following augmented problem matrices are defined

$$\tilde{\mathbf{A}}=\begin{bmatrix}\mathbf{A}^M & \mathbf{0}_{3n\times n}\\ \mathbf{H} & \mathbf{I}_n\end{bmatrix},\ \tilde{\mathbf{B}}=\begin{bmatrix}\hat{\mathbf{B}}\\ \mathbf{0}_{n\times n}\end{bmatrix},\ \tilde{\mathbf{Q}}=\begin{bmatrix}\mathbf{Q} & \mathbf{0}_{3n\times n}\\ \mathbf{0}_{n\times 3n} & \mathbf{S}\end{bmatrix},\ \tilde{\mathbf{R}}=\mathbf{R}. \tag{3.30}$$

It is noted that, if the weighting matrix $\mathbf{S}$ is set equal to a zero matrix, then the solution of the above problem with state $\Delta\mathbf{x}$ and control $\Delta\mathbf{u}$ yields the solution of the LQ problem without integration terms, i.e. an LQ regulator (without integral parts).

3.4.4 Derivation of the LQ regulators

It can be shown that, for a convex combination parameter $0\le\sigma<1$ used in (3.19), the linear system used is controllable. Given the above specifications, the optimal controller minimizing the criterion is given by

$$\Delta\mathbf{u}(k_c)=-\mathbf{K}\begin{bmatrix}\Delta\mathbf{x}(k_c)\\ \mathbf{y}(k_c)\end{bmatrix} \tag{3.31}$$

where $\mathbf{K} \in \mathbb{R}^{n\times 3n}$ is the constant gain matrix given by

$$\mathbf{K} = (\tilde{\mathbf{R}} + \tilde{\mathbf{B}}^T \mathbf{P} \tilde{\mathbf{B}})^{-1} \tilde{\mathbf{B}}^T \mathbf{P} \tilde{\mathbf{A}} \tag{3.32}$$

and $\mathbf{P}$ is a unique positive semidefinite solution of the discrete-time matrix algebraic Riccati equation

$$\mathbf{P} = \tilde{\mathbf{Q}} + \tilde{\mathbf{A}}^T (\mathbf{P} - \mathbf{P}\tilde{\mathbf{B}}(\tilde{\mathbf{R}} + \tilde{\mathbf{B}}^T \mathbf{P} \tilde{\mathbf{B}})^{-1} \tilde{\mathbf{B}}^T \mathbf{P}) \tilde{\mathbf{A}}. \tag{3.33}$$

One way to calculate $\mathbf{P}$ is by iterating backwards in time (starting from any terminal positive semidefinite condition, e.g. $\mathbf{P}(K_c) = \tilde{\mathbf{Q}}$ or $\mathbf{P}(K_c) = \mathbf{0}$) the dynamic Riccati equation of the finite-horizon case

$$\mathbf{P}(k-1) = \tilde{\mathbf{Q}} + \tilde{\mathbf{A}}^T (\mathbf{P}(k) - \mathbf{P}(k)\tilde{\mathbf{B}}(\tilde{\mathbf{R}} + \tilde{\mathbf{B}}^T \mathbf{P}(k) \tilde{\mathbf{B}})^{-1} \tilde{\mathbf{B}}^T \mathbf{P}(k)) \tilde{\mathbf{A}} \tag{3.34}$$

until it converges to a stationary value. This is possible only for $0 \le \sigma < 1$ because then the system is controllable. However, if one selects $\sigma = 1$, then again gain matrix $\mathbf{K}$ can be calculated by iterating backwards in time the dynamic Riccati equation of the finite-horizon case until $\mathbf{K}$ converges to a stationary value, as the system is then not controllable, but stabilizable.

Decomposing the gain matrix $\mathbf{K} = [\mathbf{K}_1, \mathbf{K}_2]$, we have from (3.31)

$$\Delta\mathbf{u}(k_c) = -\mathbf{K}_1 \Delta\mathbf{x}(k_c) - \mathbf{K}_2 \mathbf{y}(k_c) \tag{3.35}$$

and, after some algebra, the final form of the Linear-Quadratic regulator with Integral action (LQI regulator) for the internal boundary control problem is given by the following differential form

$$\boldsymbol{\varepsilon}(k_c) = \boldsymbol{\varepsilon}(k_c - 1) - \mathbf{K}_P [\mathbf{x}(k_c) - \mathbf{x}(k_c - 1)] - \mathbf{K}_I \left[\tilde{\boldsymbol{\rho}}^a(k_c) - \tilde{\boldsymbol{\rho}}^b(k_c) \right] \tag{3.36}$$

where $\mathbf{K}_P = \mathbf{K}_1 - \mathbf{K}_2 \mathbf{H}$ and $\mathbf{K}_I = \mathbf{K}_2$ are the proportional and integral gain matrices, respectively, while $\boldsymbol{\varepsilon}(k_c) = [\varepsilon_1(k_c), \ldots, \varepsilon_n(k_c)]^T$, $\mathbf{x}(k_c) = [\tilde{\boldsymbol{\rho}}^a(k_c)^T, \tilde{\boldsymbol{\rho}}^b(k_c)^T, \boldsymbol{\gamma}(k_c)^T]^T$, $\tilde{\boldsymbol{\rho}}^a(k_c) = [\tilde{\rho}_1^a(k_c), \ldots, \tilde{\rho}_n^a(k_c)]^T$, $\tilde{\boldsymbol{\rho}}^b(k_c) = [\tilde{\rho}_1^b(k_c), \ldots, \tilde{\rho}_n^b(k_c)]^T$ and $\boldsymbol{\gamma}(k_c) = [\gamma_1(k_c), \ldots, \gamma_n(k_c)]^T$.

Some comments should be given regarding this regulator:

- No nominal values are explicitly required in the regulator.
- If the weighting matrix $\mathbf{S}$ is set equal to a zero matrix, then $\mathbf{K}_2$ becomes also a zero matrix, i.e. (3.35) becomes an LQ regulator without integral terms.
- It is immediately seen that, at a steady state, the last term of the regulator must equal zero, and, since $\mathbf{K}_2$ is full rank, this corresponds to equal relative densities on the two directions of every section, as specified.
- The regulator is utterly reactive, i.e. it makes use only of real-time density measurements in each section and direction. In particular, there is no need for external demand prediction, as necessary in the QP problem of Section 3.3.3.
- The values obtained for each one of the control variables, i.e. the sharing factors per section, must be truncated before application in order to satisfy (3.3). These truncated

values are used as $\boldsymbol{\varepsilon}(k_c - 1)$ in (3.36) in the next time-step to avoid the well-known wind-up effect of regulators with integral terms. In case $\mathbf{S}$ is set equal to a zero matrix, i.e. for an LQ regulator, the differential form (3.36) is equivalent to (3.35) only if no truncation is necessary.

- Finally, (2.6) and (2.7) are applied before actual implementation of the sharing factors to ensure safe evacuation of the direction being restricted by the last control decision.

3.5 Case studies

3.5.1 Simulation set-up

The performance of the proposed feedback-based controllers is investigated using two scenarios for a bi-directional highway stretch depicted in Figure 3.3. These scenarios have been first considered by Malekzadeh et al. (2021) to check the performance of the optimal QP methodology proposed therein. Using the same scenarios here, allows us to compare the efficiency improvements achieved by the two LQ control methods. The two demand scenarios were carefully selected to challenge the controller more than in real traffic, as they feature very strongly rising and falling inflow values that require accordingly fast controller reaction. The two demand scenarios belong to the following respective classes (see (Malekzadeh et al., 2021) for a more in-depth discussion):

- *Uncongested scenario*, where congestion is created in one or both directions without internal boundary control (no-control case); but congestion can be utterly avoided with activation of internal boundary control. Such situations are likely to constitute the majority of real congestion cases on highways.
- *Congested scenario*, where no-control congestion can be mitigated with internal boundary control, but cannot be utterly suppressed due to strong and strongly overlapping bi-directional external demands.

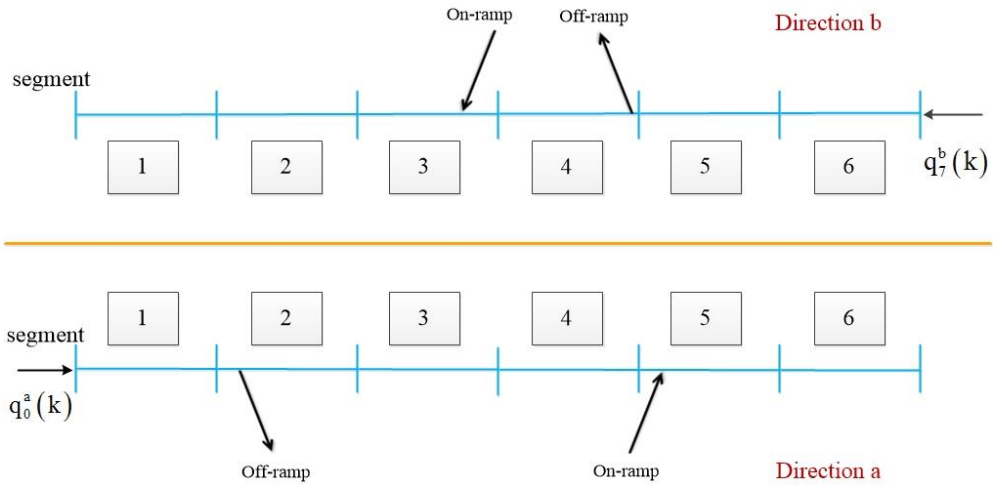


Figure 3.3: The considered highway stretch

The considered highway stretch has a length of 3 km and is subdivided in 6 sections of 0.5 km each. For direction a, there is an off-ramp in section 2 and an on-ramp in section 5, while for direction b, there is an off-ramp in section 4 and an on-ramp in section 3. The modelling time-step, T, is set to 10 s and the considered time horizon is 1 h, hence $K = 360$.

While a linearization of CTM was used for controller design, the full nonlinear CTM is used to represent the emulated ground truth in this section. The model parameters used are $v_f = 100$ km/h and $w_s = 12$ km/h; while the total cross-road capacity to be shared among the two directions is $q_{cap} = 12{,}000$ veh/h. Based on the above and the triangular FD, we can calculate $\rho_{cr} = 120$ veh/km and $\rho_{\max} = 1120$ veh/km. The parameter values used to enable capacity drop, when this possibility is mentioned to be activated in the scenarios, are $\lambda_r = 0.7$ and $\lambda_d = 0.4$; when no capacity drop is activated in the model, these parameters are $\lambda_r = 1$ and $\lambda_d = 0$. The exit rates for the two off-ramps are both assumed equal to 0.1. For both scenarios, the initial density values used are $\rho^a(0) = [5.0, 5.0, 5.0, 5.0, 18.5, 29.4]$ veh/km, $\rho^b(0) = [14.4, 14.4, 14, 5.0, 5.0, 5.0]$ veh/km. For each scenario, the simulation results of the no-control case are presented first for completeness and comparison purposes, although they are already included in (Malekzadeh et al., 2021). The no-control cases are followed by the results obtained when using the LQ and LQI regulators. The efficiency improvements achieved are also presented and compared to those obtained when using the optimal control solution reported by Malekzadeh et al. (2021).

3.5.2 Regulator design

In order to apply the feedback-based LQ and LQI regulators (3.36) developed in Section 3.4, we need to calculate off-line the static gain matrix $\mathbf{K} = [\mathbf{K}_1, \mathbf{K}_2]$. The convex combination parameter is selected to be $\sigma = 0.95$. A nominal point of operation is selected for the calculation, via (3.37)-(3.39), of the matrices $\mathbf{A}$ and $\mathbf{B}$ used in the linear model (3.26). The controller has to be designed to perform well around the critical area of the fundamental diagram for each direction. To this end, it is assume that the nominal point for the relative densities is equal to 1. The selection of the nominal demand values is done such that the resulting mainstream flows are roughly around the capacity of the nominal system, i.e. the system with sharing factors equal to 0.5. Under these conditions, the exact nominal demand values did not have any noticeable impact on the control results. In the simulations performed, we have strong variations of the demand over time, which do not have a negative effect on the performance of the LQ and LQI controllers that is approaching the performance of optimal control. In this sense, the selected nominal values are $q_0^{aN} = q_7^{bN} = 5000$ veh/h, $r_5^{aN} = r_3^{bN} = 1000$ veh/h, $\tilde{\rho}_i^{aN} = \tilde{\rho}_i^{bN} = 1$ and $\varepsilon_i^N = 0.5$, $i = 1, 2, \ldots, 6$. From (3.15), we have that $\gamma_i^N = 0.5$, $i = 1, 2, \ldots, 6$. The control time-step, T_c, is set to 60 s, hence $K_c = 60$ and $M = 6$. Based on (3.30) and the $\mathbf{A}$ and $\mathbf{B}$ matrices, we can calculate the $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ matrices for the augmented linear system. The weighing matrices used in the objective function (3.29) are selected to be $\mathbf{Q} = [\mathbf{I}_{2n\times 2n}, \mathbf{0}_{n\times n}; \mathbf{0}_{n\times 3n}]$, $\mathbf{S} = 10^{p_1}\mathbf{I}_{n\times n}$ and $\mathbf{R} = 10^{p_2}\mathbf{I}_{n\times n}$, where the values for the parameters p_1 and p_2, appearing as exponents in the above specification of the weighing matrices $\mathbf{S}$ and $\mathbf{R}$, respectively, will be used in Sections 0 and 3.5.4.3 to investigate the robustness of the controller. The selection of the $\mathbf{Q}$ matrix is such that the control variables are not considered (indirectly, via the states $\boldsymbol{\Delta\gamma}$) twice in the objective

function. The above specifications allow for the calculation of $\tilde{\mathbf{Q}}$ and $\tilde{\mathbf{R}}$. Then, starting from the terminal condition $\mathbf{P}(K_c)=\mathbf{0}$, we iterate backwards in time the dynamic Riccati equation of the finite-horizon case (3.34) until it converges to a stationary value $\mathbf{P}$ that is used in (3.32) to get the gain matrix $\mathbf{K}$.

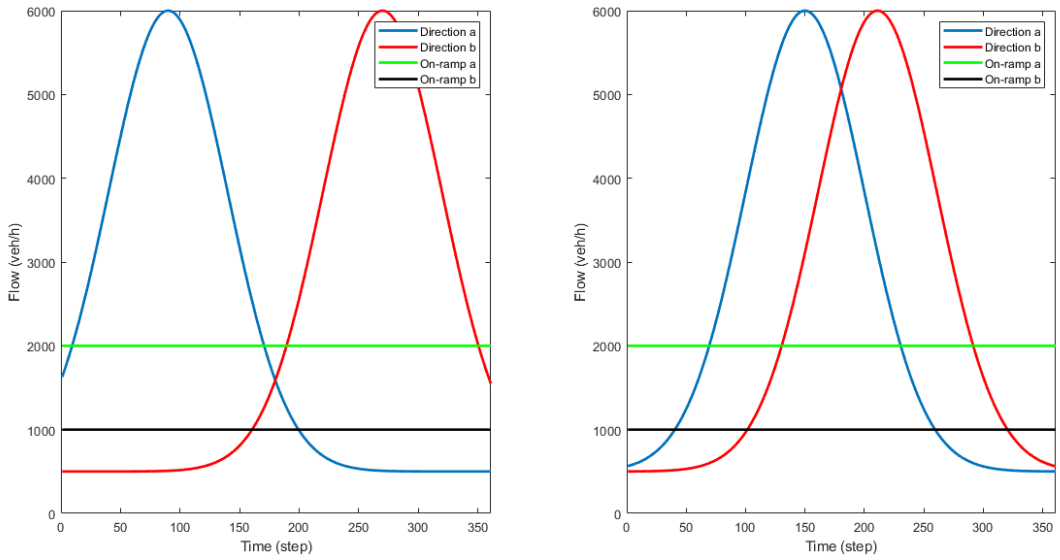
The LQI regulator (3.36) starts with initial control input values set equal to the nominal values, i.e. equal to 0.5 for all sections. The upper and lower bounds for the sharing factors, used to avoid utter blocking of any of the two directions, are equal for all sections $i=1,2,\ldots,6$ and are given the values $\varepsilon_{i,\min}=0.16$ and $\varepsilon_{i,\max}=0.84$.

With these settings, the regulator is operated in a closed-loop mode, receiving in emulated real time all section density values per direction from the CTM model equations; and responding with the sharing factors, which are calculated according to (3.36) and are eventually truncated (if necessary) and delayed as mentioned earlier. This is repeated every $T_c=60\text{s}$.

3.5.3 Uncongested scenario

3.5.3.1 Scenario description

The mainstream and on-ramp demand flows per direction for this scenario are presented in Figure 3.4a. Compared to real freeways, the demand rising and falling profiles applied are quite extreme for such a short time period. It may be seen that the two directions feature respective peaks in their mainstream demands that are slightly overlapping. In addition, the on-ramp demands are constant, with the on-ramp demand in direction *a* being higher than in direction *b*.



(a) (b)

Figure 3.4: Demand flows per direction and on-ramp for (a) the uncongested scenario; and (b) the congested scenario

3.5.3.2 No-control case

Using the demand flows of the uncongested scenario in the CTM equations of Section 3.3.2 with constant sharing factors at $\varepsilon_i = 0.5$ for all sections, we obtain the simulation results of the no-control case. Figure 3.5 displays the corresponding spatio-temporal evolution of the relative density defined in (3.12). According to the definition, relative density values lower than 1 refer to uncongested traffic; while values higher than 1 refer to congested traffic; when the relative density equals 1, and the downstream section is uncongested, we have capacity flow at the corresponding section.

Figure 3.5 shows that heavy congestion is created in section 5 for direction a due to the strong ramp inflow, in combination with the increased mainstream demand, at around $k = 60$. The congestion propagates upstream, reaching up to section 2, and dissolves at around $k = 200$, due to the rapid decrease of the mainstream demand for this direction (Figure 3.4a). In direction b, we have also a congestion being triggered in section 3 by the increasing mainstream demand, in combination with the on-ramp flow, at around $k = 250$. Due to lower on-ramp flow, the congestion extent is smaller than in direction a; the congestion propagates upstream to section 5 and dissolves at around $k = 330$.

It should be noted that the results displayed in Figure 3.5 were obtained using the CTM equations with the capacity drop terms. The TTS value obtained for this case is reported in Table 3.1. When the capacity drop terms are de-activated, then the space-time extent of the created congestions is slightly reduced. The corresponding diagrams are omitted for space economy, but the resulting, lower TTS value is also reported in Table 1.

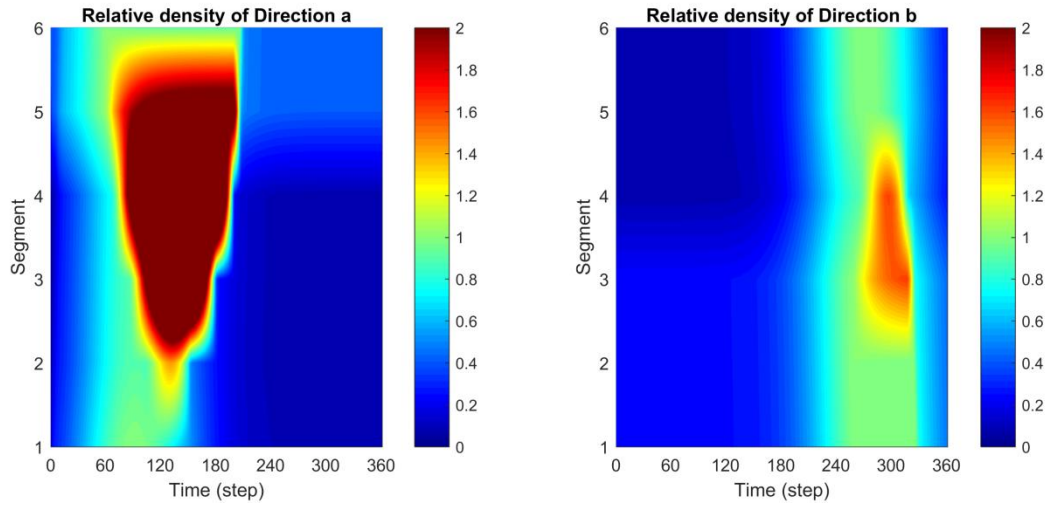


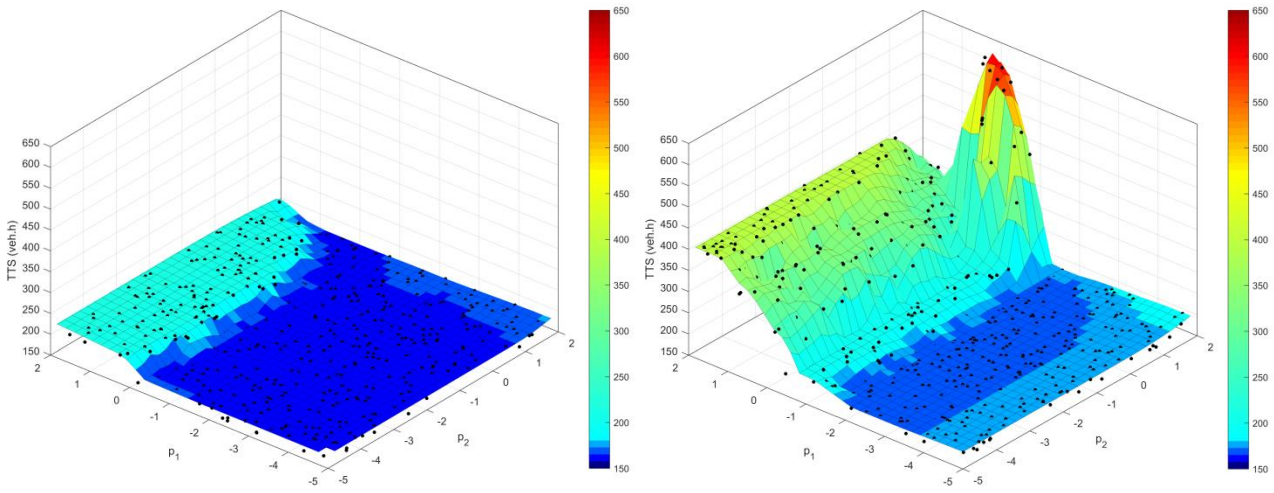
Figure 3.5: Uncongested scenario: Relative density for the two directions in the no-control case

Table 3.1: TTS (veh·h) and related improvement (%) over the no-control case for different scenarios

Scenarios	**No-control**	**LQI with** $(p_1, p_2) = (-2.5, -3.0)$	**LQ with** $(p_1, p_2) = (-\infty, -3.0)$	**QP**
Uncongested with capacity drop	231.9	164.9 (28.9)	164.9 (28.9)	164.9 (28.9)
Uncongested without capacity drop	209.8	164.9 (21.4)	164.9 (21.4)	164.9 (21.4)
Congested with capacity drop	236.0	170.3 (27.8)	175.6 (25.6)	171.0 (27.5)
Congested without capacity drop	213.9	170.0 (20.5)	172.7 (19.3)	170.9 (20.1)

3.5.3.3 Control case

In order to investigate the robustness properties of the LQ regulators, in particular the influence of the weighting matrices in the quadratic objective (3.29) on control performance, we have calculated 1000 different values of the gain matrix $\mathbf{K}$, following the procedure of Section 3.5.2, with randomly selected values for the parameters p_1 and p_2 within the range $[-5, 2]$. The corresponding values for the gain matrix $\mathbf{K}$, reflecting different combinations of the weighting matrices, were then used to run a set of 1000 corresponding control experiments for the uncongested scenario. Each such experiment delivered a corresponding TTS value from the simulator, and the TTS values for all (p_1, p_2) combinations are depicted, marked as black dots, in Figure 3.6a. The same plot includes a fitted function.



(a) (b)

Figure 3.6: 3-D (p_1, p_2)-diagram of the TTS values for (a) the uncongested scenario; and (b) the congested scenario

Note that the 3-D diagram in Figure 3.6a actually refers to the LQI controller results. On the other hand, the LQ regulator is obtained for weight matrix $\mathbf{S} = \mathbf{0}$, i.e. for $p_1 \to -\infty$, in the objective function, but the displayed results for $p_1 = -5$ are virtually equivalent to those obtained with $\mathbf{S} = \mathbf{0}$. Thus, the robustness diagram for the LQ regulator without integral parts is 2-D (depending on p_2 only) and is obtained from the 3-D diagram of Figure 3.6a as the 2-D plane cut at $p_1 = -5$. This 2-D diagram is visible in Figure 3.6a, at the front-right edge of the 3-D diagram (for $p_1 = -5$).

Based on the fitted function, it can be safely concluded that the LQ and LQI regulators are robust with respect to variations in the values of the selected weighing matrices within a very large area, that is seen to remain flat, featuring similarly low TTS values.

In the following we have selected to use the LQI regulator with a parameter set $(p_1, p_2) = (-2.5, -3.0)$, which is within the good-performance area of Figure 3.6a; and the LQ regulator with the same weight p_2, i.e. $(p_1, p_2) = (-\infty, -3.0)$. The resulting traffic conditions for both regulators are virtually the same and under-critical anywhere anytime, i.e. congestion observed in both traffic directions in the no-control case is utterly avoided. The TTS values for both regulators are exactly the same, and the activation of the capacity drop terms has no impact, as there is no congestion. All related TTS values are given in Table 3.1 indicating improvements of 29 % and 21 % over the corresponding no-control cases with and without capacity drop activation, respectively. The TTS value obtained using the LQI or the LQ regulators is, in fact,

also equal to the value that is achieved when applying the optimal control resulting from the QP problem formulation. Thus, despite their simple feedback character, where no model or demand predictions are involved, the LQI and LQ regulators achieve highest efficiency, even though they do not explicitly consider TTS in their objective.

Since the performance of the two regulators is equivalent, we use only the LQ regulator results for the following detailed discussion of the control-case for the uncongested scenario. The spatio-temporal evolution of the relative densities is depicted in Figure 3.7. Figure 3.8, Figure 3.9 and Figure 3.10 display more detailed information for this case. Specifically, each figure holds two columns with the results of two respective sections; for each section (column), we provide three diagrams (rows):

- The first diagram shows the two traffic densities (in veh/km), for directions a and b, and the corresponding critical densities, which are changing according to the sharing factor in the section.
- The second diagram shows the two traffic flows, for directions a and b, and the corresponding capacities, which are changing according to the sharing factor in the section. In addition, the sum of both flows is also displayed (cyan curve).
- The third diagram shows the value of the control input, i.e. the sharing factor applied, as well as the constant bounds (black curves), which may lead to possible truncation of the control input.

The displayed results confirm that densities (flows) are always lower than the respective (controlled) critical densities (capacities) in all sections and in both directions; hence traffic conditions are always and everywhere under-critical. In fact, the total-flow curve (for both directions) does not reach the total road capacity (of 12,000 veh/h) at any time anywhere. In short, congestion is utterly avoided and any occurring delays in the no-control case do not exist anymore.

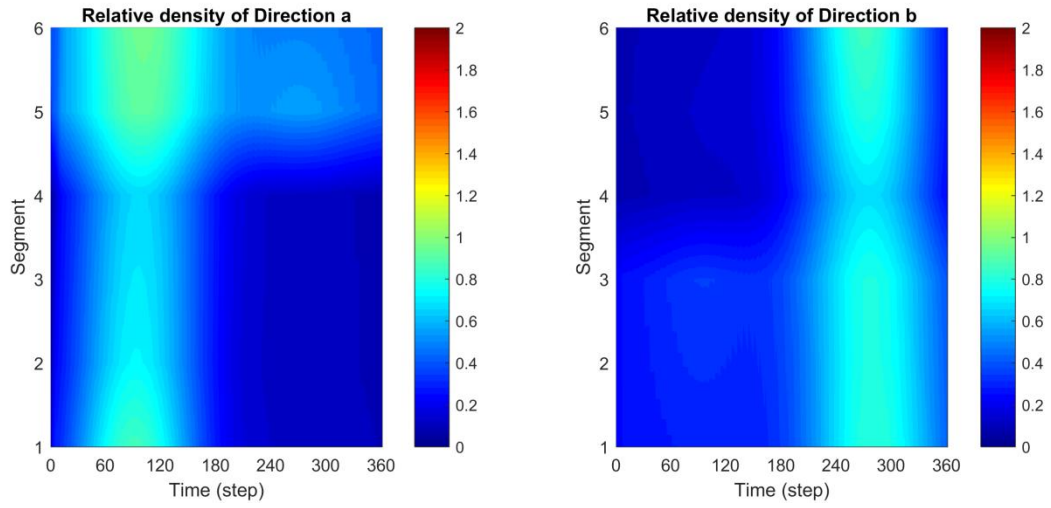


Figure 3.7: Uncongested scenario: Relative density for the two directions in the control case (LQ)

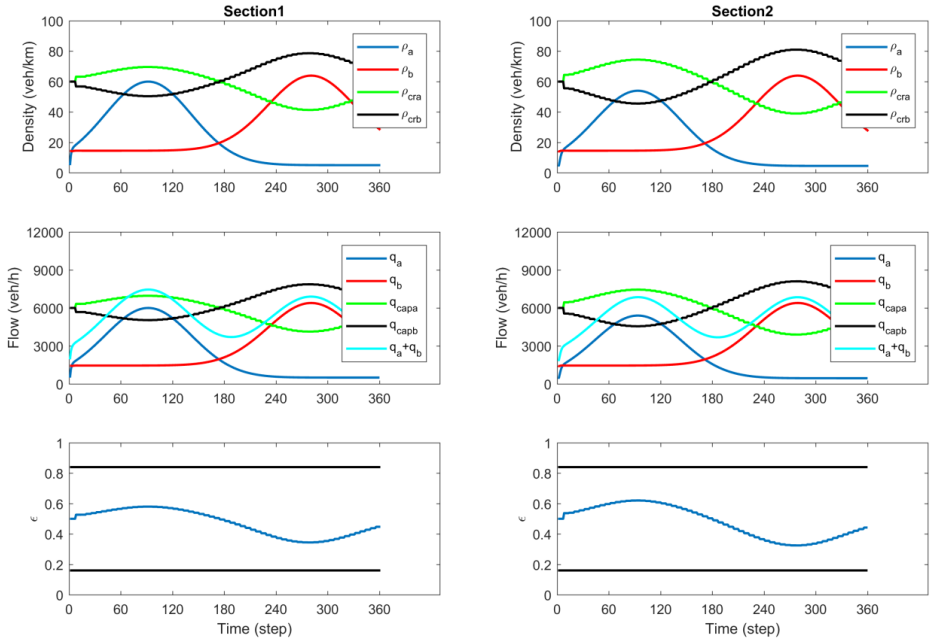


Figure 3.8: Uncongested scenario: Density, flow and control trajectories in the control case (LQ) (sections 1 and 2)

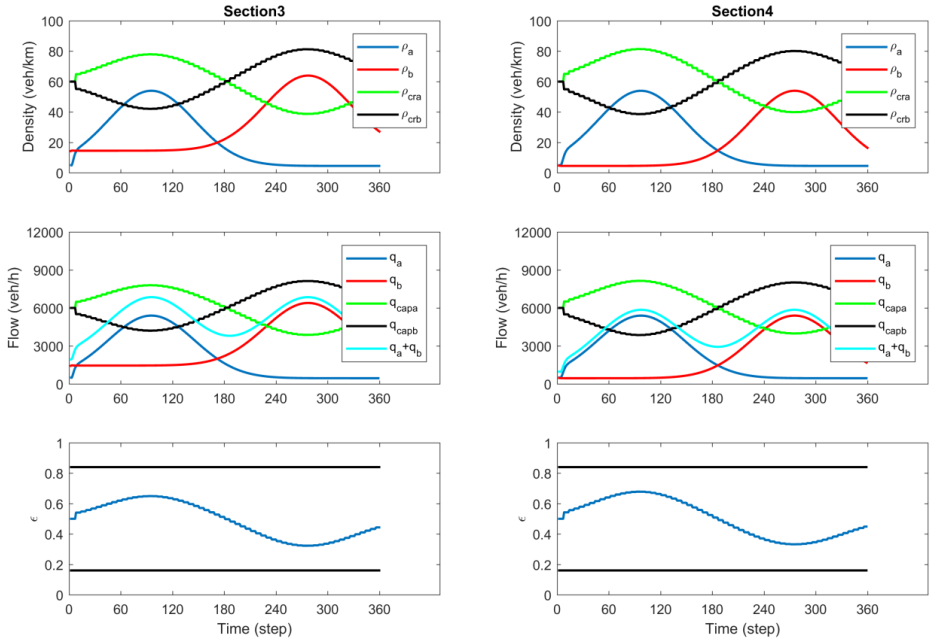


Figure 3.9: Uncongested scenario: Density, flow and control trajectories in the control case (LQ) (sections 3 and 4)

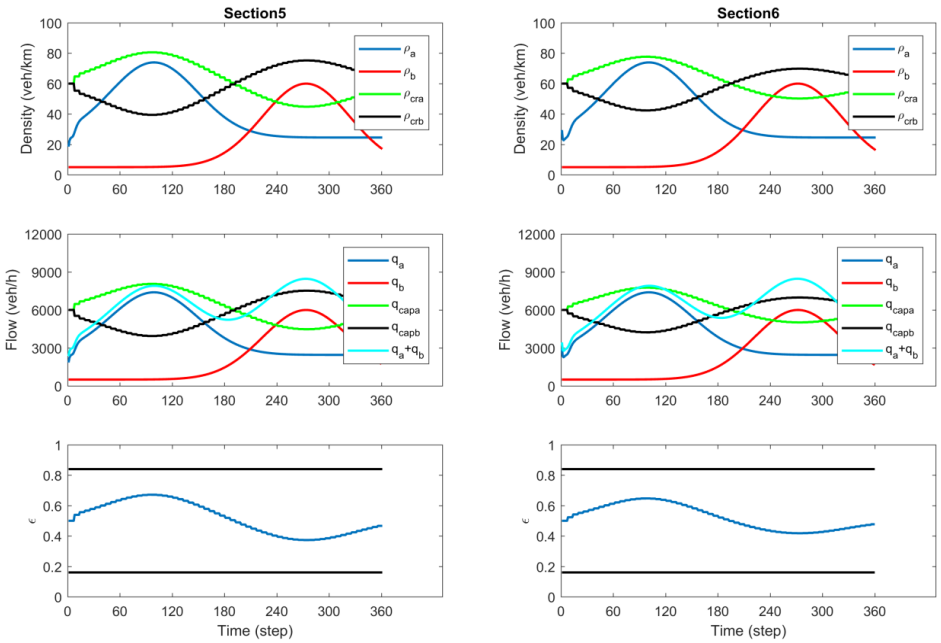
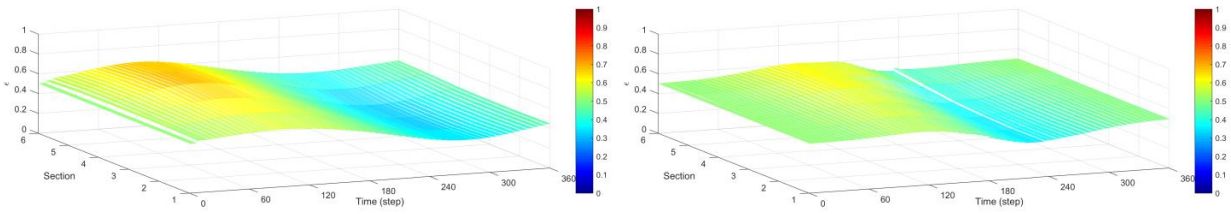


Figure 3.10: Uncongested scenario: Density, flow and control trajectories in the control case (LQ) (sections 5 and 6)

The sharing factor trajectories of the sections reveal that this excellent outcome is enabled via a smooth swapping of assigned capacity to the two directions, whereby more capacity is assigned to direction a during the first half of the time horizon and vice-versa for the second half, in response to the traffic (density) changes caused by the changing respective demands and their peaks. It is interesting to notice that the value of the control input (sharing factor) is never saturated and, as a result, the use of the differential form (3.36) for the LQ controller is equivalent to the use of (3.35). Finally, Figure 3.11a displays the space-time diagram of the control input which demonstrates that it is a smooth function in space and time.



(a) (b)

Figure 3.11: 3-D space-time diagram of the LQ control input for (a) the uncongested scenario; and (b) the congested scenario

3.5.4 Congested scenario

3.5.4.1 Scenario description

The demand flows for this scenario are displayed in Figure 3.4b for both directions. Essentially, all external flows are similar as in the uncongested scenario, with the notable difference that the two mainstream demand profiles have been shifted closer to each other. This leads to a longer overlapping period with strong flows on both directions, which implies capacity problems even in presence of internal boundary control. The on-ramp demands are constant at the same respective levels as in the uncongested scenario.

3.5.4.2 No-control case

Using the demand flows of the congested scenario in the CTM equations of Section 3.3.2 with constant internal boundary at $\varepsilon_i = 0.5$ for all sections, we obtain the simulation results of the no-control case. Figure 3.12 displays the corresponding spatio-temporal relative density evolution, where heavy congestion is created in section 5 for direction a due to the strong ramp inflow, in combination with the increased mainstream demand, at around $k = 120$. The congestion propagates upstream, reaching up to section 2, and is dissolved at around $k = 250$, thanks to the rapid decrease of the mainstream demand for this direction (Figure 3.4b). In direction b, we have also a congestion being triggered in section 3 by the increasing mainstream demand, in combination with the on-ramp flow, at around $k = 200$. Due to lower on-ramp flow, the congestion extent is smaller than in direction a. The congestion propagates upstream up to section 5 and dissolves at around $k = 270$.

The results displayed in Figure 3.12 were obtained using the CTM equations with the capacity drop terms, and the corresponding value of TTS is reported in Table 3.1. When the capacity drop terms are de-activated, then the space-time extent of the created congestions is reduced. The corresponding diagrams are omitted for space economy, but the resulting, lower TTS value is also reported in Table 3.1. Note that, compared with the uncongested scenario, the congestion extent and TTS values are not much different, simply because the total demand per direction did not change much, while the capacity of each direction is constant in the no-control case. In other words, any observed congestion phenomena in the two directions are independent of each other and only dependent on the corresponding one-directional demand and capacity, which are both the same in the two scenarios for the no-control case.

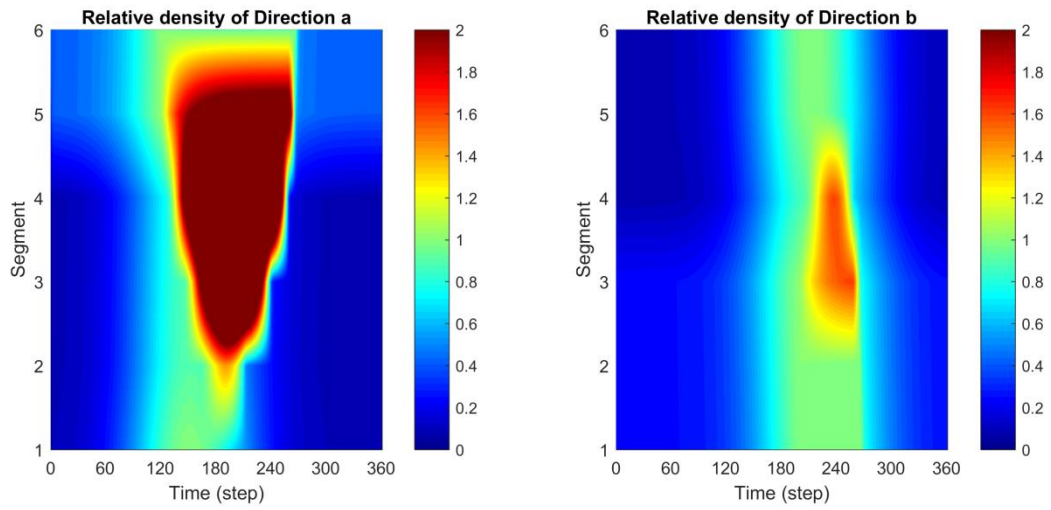


Figure 3.12: Congested scenario: Relative density for the two directions in the no-control case

3.5.4.3 Control case

In order to further investigate the robustness properties of the LQI and LQ regulators with respect to the objective criterion weights, we used the 1000 different values of the gain matrix $\mathbf{K}$, calculated in Section 0, to run a set of 1000 corresponding control experiments, now for the congested scenario. Each such experiment delivered a corresponding TTS value from the simulator, and the TTS values for all (p_1, p_2) combinations are depicted, marked as black dots, in Figure 3.6b. Again, the same plot includes a fitted function. Recall that the 2-D diagram for the LQ regulator appears in Figure 3.6b at the front-right edge of the 3-D diagram (for $p_1 = -5$).

Comparing the two diagrams of Figure 3.6, we observe that:

- The good-performance area has slightly shrunk in the congested scenario compared to the uncongested one, but remains very large, extending over several orders of magnitude of the weights p_1 and p_2.
- The good-performance area for the congested scenario is a subset of its counterpart of the uncongested scenario. This implies that the same weights may be used for high-performance regulation in either case.
- The best control performance of the LQI regulator is slightly better compared to the LQ regulator.
- In case of inappropriate weight selection, the worst performance of the LQ regulator (obtained for $p_2 \to \infty$) is equivalent to no control. In contrast, inappropriate weight selection for the LQI regulator may lead to results much worse than no control.

Further scenario investigations with stronger congestion confirmed that both regulators perform excellently (i.e. deliver TTS values that are not more than 5 % higher than the optimal ones) under all scenarios, for $p_1 < -2$ and $p_2 < 0$.

In fact, for the present demand scenario, the LQI regulator features similarly low TTS values, virtually equal to the one achieved when applying the optimal control resulting from the QP problem formulation, within a broad area of (p_1, p_2) combinations, marked blue in Figure 3.6b; thus, the selection of any pair within this area, e.g. $(p_1, p_2) = (-2.5, -3.0)$, leads to optimal results. On the other hand, for $p_1 < -4$ the matrix $\mathbf{S} = 10^{p_1}\mathbf{I}_{n\times n}$ is approaching the zero matrix, i.e. the LQI regulator is approaching the LQ regulator, and we observe that for $p_2 < 0$ (and $p_1 = -5$) the LQ regulator is producing close to optimal results. It must be emphasized that this excellent outcome of the feedback regulators is achieved without assuming a perfect model and perfect demand prediction, as for the optimal QP problem.

Activation of the capacity drop terms has only a minor impact, as there is no heavy congestion in the control results. The TTS value obtained using the LQI regulator is even slightly lower than the value resulting from the optimal control of the QP problem of Section 3.3.3. This may sound paradoxical, but is explained by the fact that the optimal QP problem solution takes into account, apart from TTS, also other operational sub-objectives in (3.11), which may slightly affect the resulting TTS value upon minimization.

The results reported in the following figures were obtained with activation of the capacity drop terms. Figure 3.13 displays the spatio-temporal evolution of the relative densities in the control case, utilizing the LQI regulator with $(p_1, p_2) = (-2.5, -3.0)$, while Figure 3.14 displays the same output when utilizing the LQ regulator with $(p_1, p_2) = (-\infty, -3.0)$. It can be observed that in both cases a short-lasting congestion exists in section 6 for direction b, while the main difference is the congestion created in section 5 for direction a when the LQ regulator is used, which is absent in the LQI regulator results.

Figure 3.15, Figure 3.16 and Figure 3.17 display more detailed information for the case of the LQ regulator, as for the uncongested scenario. The displayed results indicate that in direction a congestion appears in section 5, starting at time $k = 150$ and lasting up to $k = 225$. In direction b, congestion is observed in section 6, starting at time $k = 175$ and lasting up to $k = 225$. It is important to emphasize that the total-flow curve (for both directions) reaches and remains close to the total road capacity (of 12,000 veh/h) at sections 5 and 6 during this peak period. In fact, sections 5 and 6 form a bottleneck for this scenario. As explained in (Malekzadeh et al., 2021), bottlenecks in the internal boundary control context, under an efficient control strategy, refer to both traffic directions simultaneously, which is contrast to independent uni-directional bottlenecks appearing in conventional traffic control. Thus, full exploitation of the road capacity (both directions) is indeed enabled at the bottleneck by the LQ regulator for the duration of the critical period, so as to minimize congestion and delays.

The sharing factor trajectories are smooth in this scenario as well (see Figure 3.11b), achieving the efficient assignment of the necessary share of capacity where and when needed. No saturation is observed for the control input (sharing factors) at any section.

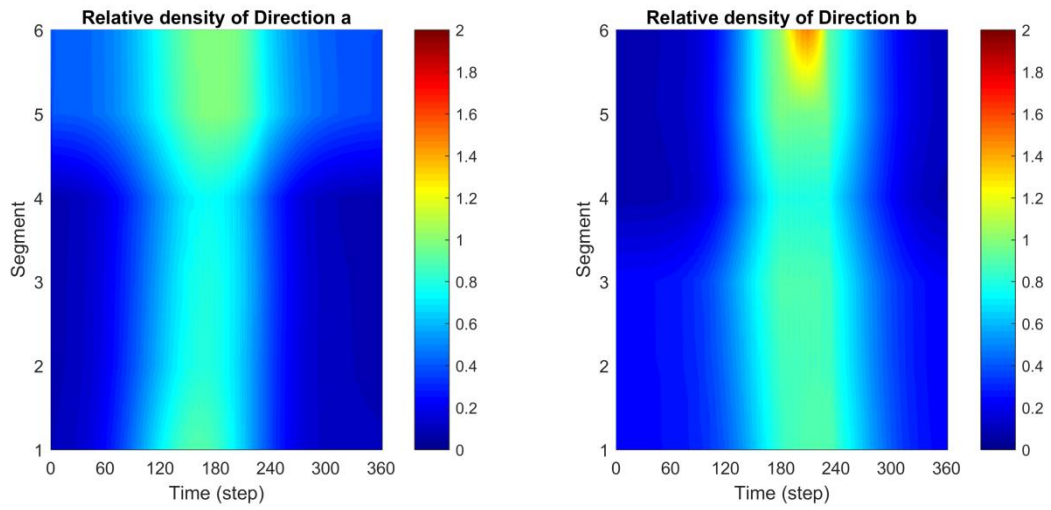


Figure 3.13: Congested scenario: Relative density for the two directions in the control case (LQI)

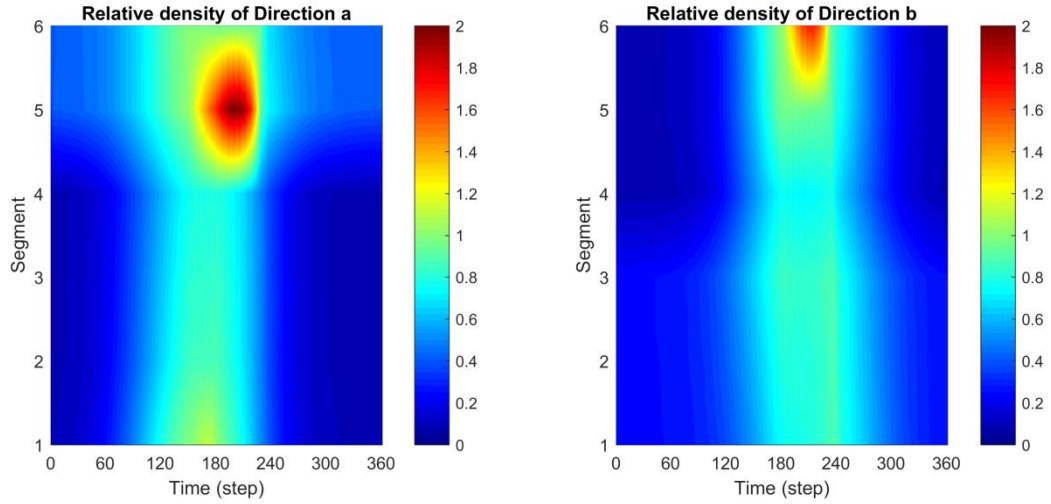


Figure 3.14: Congested scenario: Relative density for the two directions in the control case (LQ)

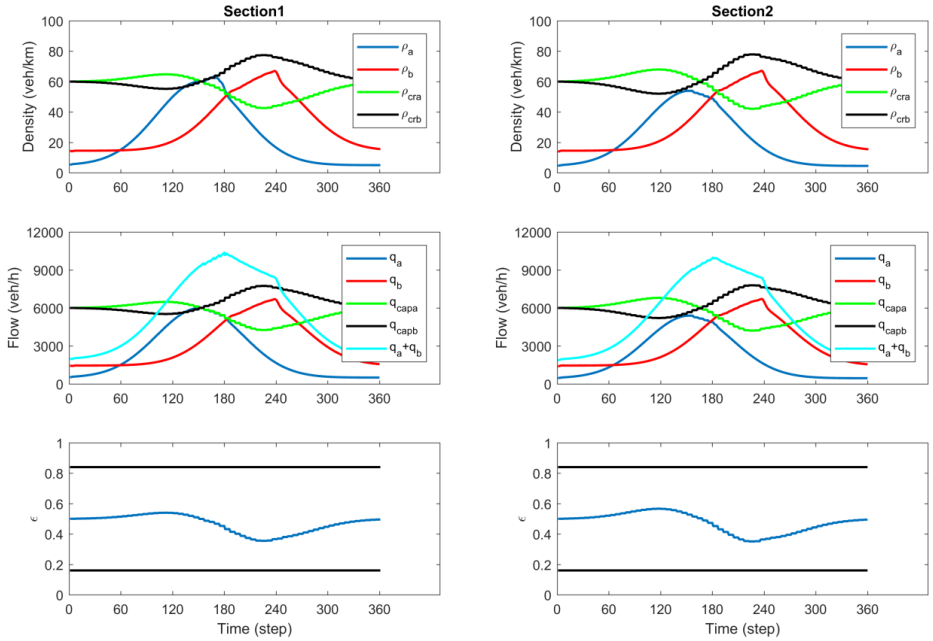


Figure 3.15: Congested scenario: Density, flow and control trajectories in the control case (LQ) (sections 1 and 2)

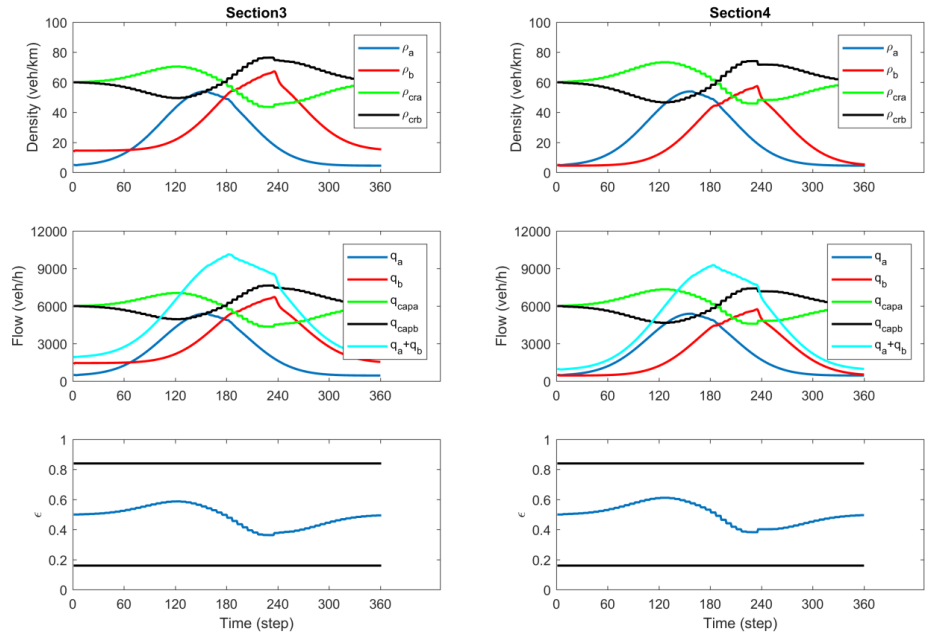


Figure 3.16: Congested scenario: Density, flow and control trajectories in the control case (LQ) (sections 3 and 4)

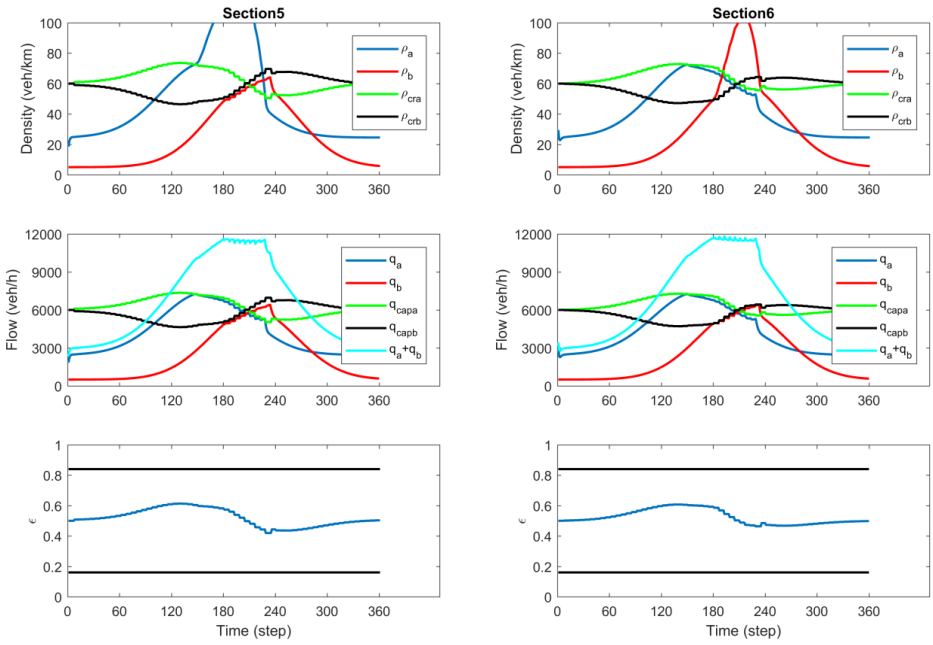


Figure 3.17: Congested scenario: Density, flow and control trajectories in the control case (LQ) (sections 5 and 6)

3.5.5 Uncongested scenario with late control activation

In the scenarios considered already, the LQ regulator is able to either completely avoid congestion (when possible) or minimize the extent of congestion in space and time (when congestion is unavoidable). In the considered scenarios, the regulator is applied for the whole simulation period, starting from under-critical initial values of the traffic state. In order to test and demonstrate that the regulator is successful even in cases where congestion has already occurred, we simulate the system for the uncongested scenario without applying any control action for an initial period of time, which includes the appearance of congestion due to lack of control, as in the no-control case. Then, at $k=72,$ the LQ regulator is activated and remains active till the end of the simulation.

The resulting spatio-temporal evolution of the relative densities is depicted in Figure 3.18. As expected, up to $k=72,$ traffic conditions are exactly the same with those of the no-control case described in Section 3.5.3.2 and depicted in Figure 3.5. The onset of congestion takes place at around $k=60$ in section 5 of direction a and propagates upstream to section 3. At $k=72,$ thanks to the activation of the controller, the congestion propagation is mitigated, and its extent starts soon after to decrease rapidly.

Figure 3.19, Figure 3.20 and Figure 3.21 display more detailed information for this case. The actions of the controller lead to an increase of the assigned capacity for direction $a,$ immediately after the activation of the controller, in reaction to the accumulated demand in the congested area. Congestion is dissolved at around $k=170,$ much earlier compared to the case of no control

activation. Despite this action, direction b remains uncongested, as, in the second half of the simulation, more capacity is assigned to this direction to accommodate its own rising demand. The TTS value achieved with the late activation of the LQ regulator is 180.6 veh·h, which is a 22% improvement compared to the case of no control activation reported in Table 3.1.

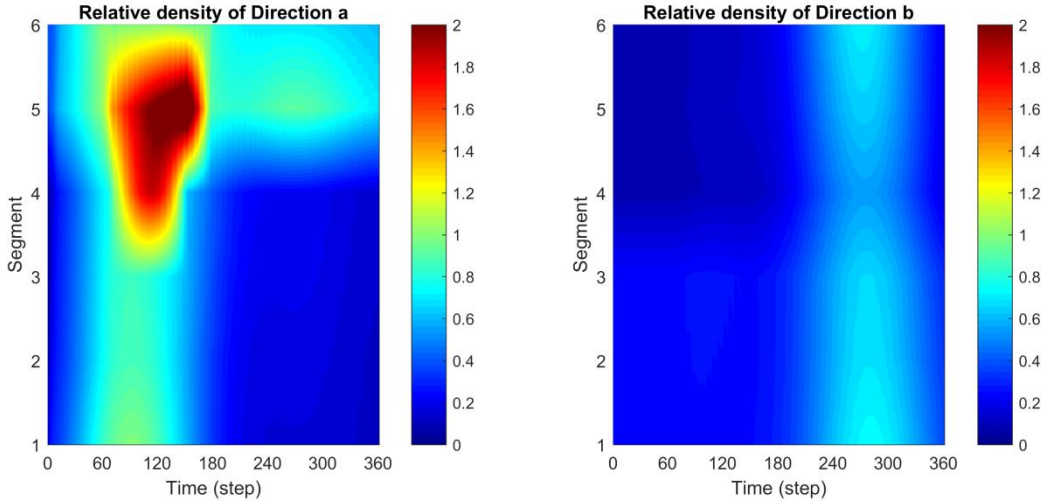


Figure 3.18: Uncongested scenario with late control (LQ) activation: Relative density for the two directions

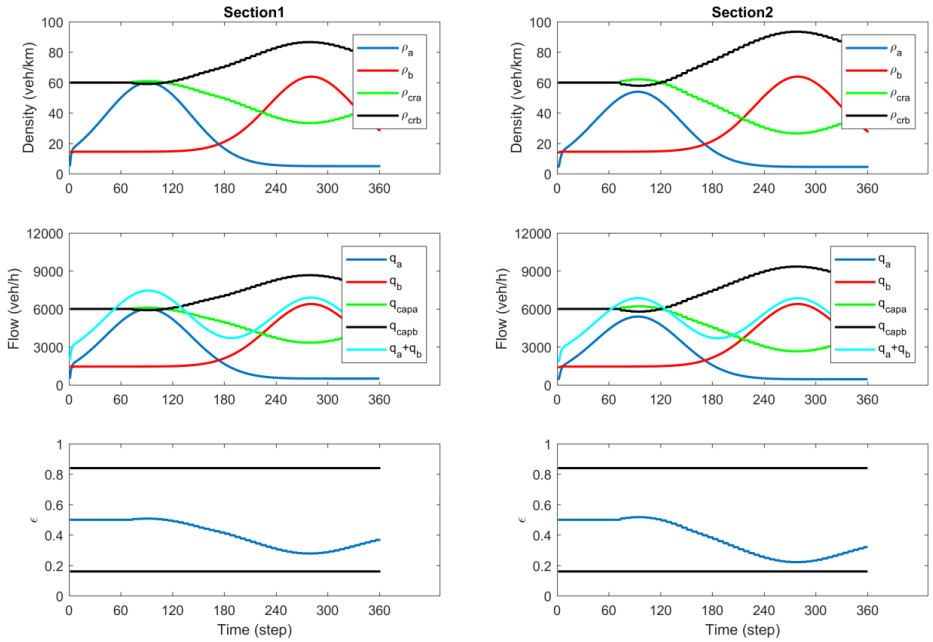


Figure 3.19: Uncongested scenario with late control (LQ) activation: Density, flow and control trajectories (sections 1 and 2)

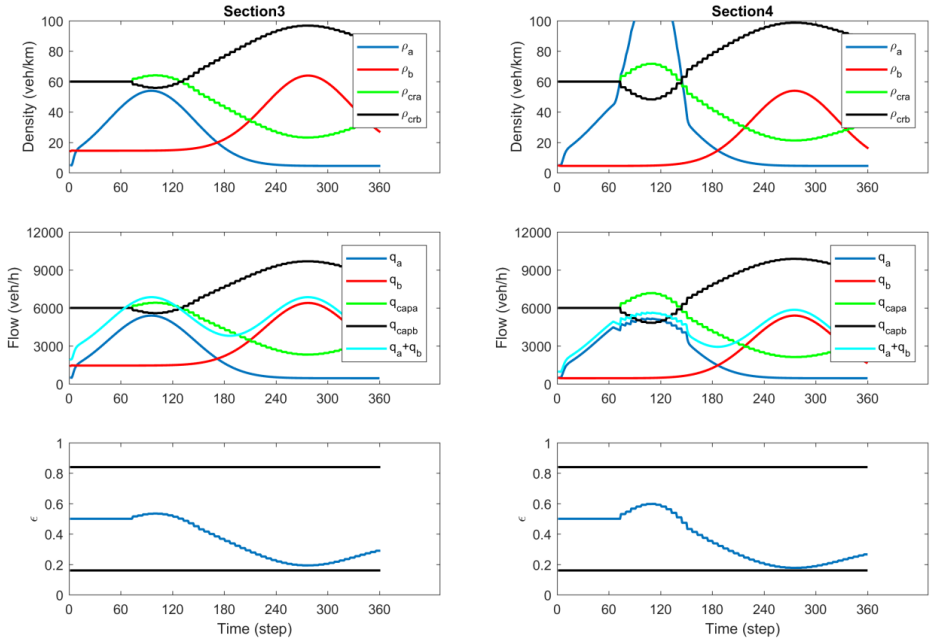


Figure 3.20: Uncongested scenario with late control (LQ) activation: Density, flow and control trajectories (sections 3 and 4)

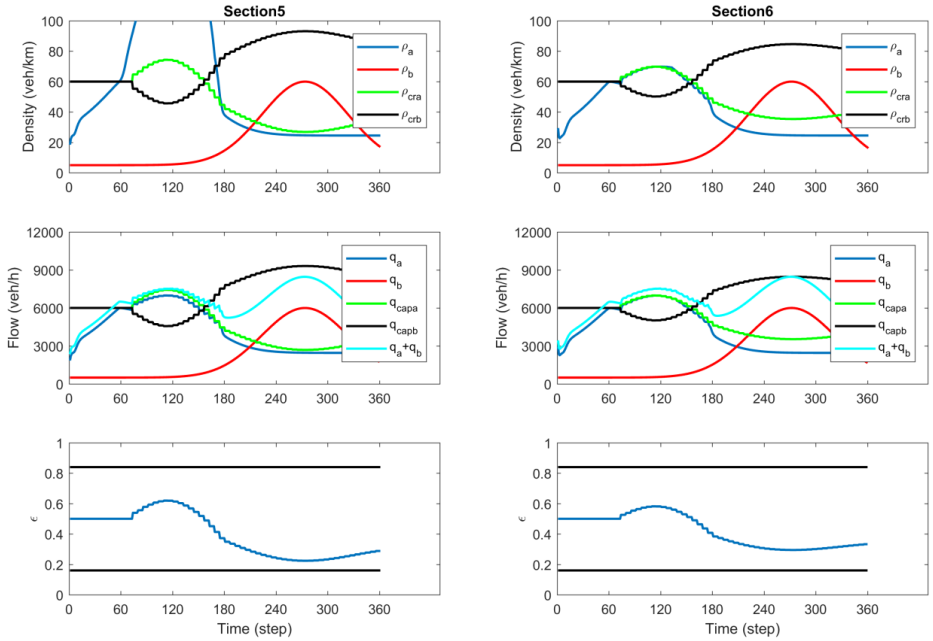


Figure 3.21: Uncongested scenario with late control (LQ) activation: Density, flow and control trajectories (sections 5 and 6)

3.6 Conclusions

The concept of internal boundary control, introduced by Malekzadeh et al. (2021), has been revisited in this study by use of a different control approach. The well-known CTM, appropriately adjusted to introduce the effect of the sharing factors, has been utilized for the design of feedback-based LQ and LQI regulators for the internal boundary control problem. Using these controllers, the total road width and capacity are shared in each section in real-time among the two directions of the road in response to the prevailing traffic conditions, without the need for model and demand predictions. The regulators were found to behave adequately for a broad range of weight values used in their quadratic objective and for different traffic scenarios, in which congestion may be avoidable or unavoidable.

The LQ and LQI regulators are easy to design and implement (feedback-based) and robust to disturbances (no need to predict the arriving demands). Simulation investigations demonstrate, for carefully selected scenarios, that the regulators (with appropriate selection of the weighting matrices) are equally efficient as an open-loop optimal control solution (with perfect model and demand predictions) developed for the same problem by Malekzadeh et al. (2021) using a convex QP problem formulation.

Appendix A

This Appendix presents the derivation of the partial derivatives necessary in (3.23)-(3.25). For direction a, the partial derivatives necessary for (3.23) are as follows:

$$
\begin{aligned}
&\left.\frac{\partial f_i^a}{\partial \tilde{\rho}_i^a}\right|_N = 1+\frac{T}{L_i\rho_{cr}}\left[-(1-\sigma)v_f\gamma_i\,\rho_{cr}\right]\Big|_N,\ i=1,2,\ldots,n\\
&\left.\frac{\partial f_i^a}{\partial \tilde{\rho}_{i-1}^a}\right|_N = \frac{T}{L_i\rho_{cr}}\left[(1-\beta_i^a)\frac{(1-\sigma)v_f\gamma_{i-1}\rho_{cr}}{\gamma_i}\right]\Big|_N,\ i=2,3,\ldots,n\\
&\left.\frac{\partial f_i^a}{\partial \varepsilon_i}\right|_N = \frac{T}{L_i\rho_{cr}}\left[-\frac{\sigma q_{cap}}{\gamma_i}\right]\Big|_N,\ i=1,2,\ldots,n\\
&\left.\frac{\partial f_i^a}{\partial \varepsilon_{i-1}}\right|_N = \frac{T}{L_i\rho_{cr}}\left[(1-\beta_i^a)\frac{\sigma q_{cap}}{\gamma_i}\right]\Big|_N,\ i=2,3,\ldots,n\\
&\left.\frac{\partial f_i^a}{\partial \gamma_i}\right|_N = \frac{T}{L_i\rho_{cr}}\left[-(1-\beta_i^a)\frac{\sigma\varepsilon_{i-1}q_{cap}+(1-\sigma)v_f\tilde{\rho}_{i-1}^a\gamma_{i-1}\rho_{cr}}{\gamma_i^2}+\frac{\sigma\varepsilon q_{cap}}{\gamma_i^2}-\frac{d_i^a}{\gamma_i^2}\right]\Big|_N,\ i=2,3,\ldots,n\\
&\left.\frac{\partial f_i^a}{\partial \gamma_{i-1}}\right|_N = \frac{T}{L_i\rho_{cr}}\left[(1-\beta_i^a)\frac{(1-\sigma)v_f\tilde{\rho}_{i-1}^a\rho_{cr}}{\gamma_i}\right]\Big|_N,\ i=2,3,\ldots,n\\
&\left.\frac{\partial f_1^a}{\partial \gamma_1}\right|_N = \frac{T}{L_1\rho_{cr}}\left[-\frac{d_1^a}{\gamma_1^2}+\frac{\sigma\varepsilon_1 q_{cap}}{\gamma_1^2}\right]\Big|_N
\end{aligned}
\tag{3.37}
$$

Similarly for direction b, the partial derivatives necessary for (3.24) are as follows:

$$\left.\frac{\partial f_i^b}{\partial \tilde{\rho}_i^b}\right|_N = 1 + \frac{T}{L_i \rho_{cr}}\left[-(1-\sigma) v_f \rho_{cr}\right]\Big|_N,\ i = 1, 2, \ldots, n$$

$$\left.\frac{\partial f_i^b}{\partial \tilde{\rho}_{i+1}^b}\right|_N = \frac{T}{L_i \rho_{cr}}\left[(1-\beta_i^b)\frac{(1-\sigma) v_f (1-\gamma_{i+1}) \rho_{cr}}{1-\gamma_i}\right]\Big|_N,\ i = 1, 2, \ldots, n-1$$

$$\left.\frac{\partial f_i^b}{\partial \varepsilon_i}\right|_N = \frac{T}{L_i \rho_{cr}}\left[\frac{\sigma q_{cap}}{1-\gamma_i}\right]\Big|_N,\ i = 1, 2, \ldots, n$$

$$\left.\frac{\partial f_i^b}{\partial \varepsilon_{i+1}}\right|_N = \frac{T}{L_i \rho_{cr}}\left[-(1-\beta_i^b)\frac{\sigma q_{cap}}{1-\gamma_i}\right]\Big|_N, i = 1, 2, \ldots, n-1$$

$$\left.\frac{\partial f_i^b}{\partial \gamma_i}\right|_N = \frac{T}{L_i \rho_{cr}}\left[(1-\beta_i^b)\frac{\sigma(1-\varepsilon_{i+1}) q_{cap} + (1-\sigma) v_f \tilde{\rho}_{i+1}^b (1-\gamma_{i+1}) \rho_{cr}}{(1-\gamma_i)^2} - \frac{\sigma(1-\varepsilon_i) q_{cap}}{(1-\gamma_i)^2} + \frac{d_i^b}{(1-\gamma_i)^2}\right]\Big|_N,$$

$$i = 1, 2, \ldots, n-1$$

$$\left.\frac{\partial f_i^b}{\partial \gamma_{i+1}}\right|_N = \frac{T}{L_i \rho_{cr}}\left[-(1-\beta_i^b)\frac{(1-\sigma) v_f \tilde{\rho}_{i+1}^b \rho_{cr}}{1-\gamma_i}\right]\Big|_N, i = 1, 2, \ldots, n-1$$

$$\left.\frac{\partial f_n^b}{\partial \gamma_n}\right|_N = \frac{T}{L_n \rho_{cr}}\left[\frac{d_n^b}{(1-\gamma_n)^2} - \frac{\sigma(1-\varepsilon_n) q_{cap}}{(1-\gamma_n)^2}\right]\Big|_N$$

(3.38)

Moreover, we have simply

$$\left.\frac{\partial g_i}{\partial \varepsilon_i}\right|_N = 1,\ i = 1, 2, \ldots, n. \tag{3.39}$$

4 Overlapping IBC of lane-free automated vehicle traffic

4.1 Abstract

Lane-free vehicle driving has been recently proposed for connected automated vehicles on highways or arterials. Lane-free traffic implies that incremental changes of the road width lead to corresponding incremental changes of traffic flows and capacity. In these conditions, internal boundary control (IBC) was proposed for flexible sharing of the total road width and capacity among the two traffic directions of the highway in real-time in order to maximize the cross-road infrastructure utilization and minimize congestion and delays. Centralized solutions , e.g. LQ (linear quadratic) regulators, requiring information from the whole highway stretch under consideration, have already been proposed, which, however, may be cumbersome for long highways with respect to the offline controller design effort; the extent of real-time communications; and the physical system architecture. This paper investigates two different overlapping decentralized control schemes for IBC in lane-free automated vehicle traffic. The first approach is based on a contractible controller, which is designed and employed in a decomposed way, for each subsystem of an extended highway system. In the second approach, the overall discrete-time LQ control problem is transformed into a linear matrix inequalities (LMI) problem. After selecting the overlapping feedback structure and solving the overall LMI problem, the correspondingly structured gain matrix is obtained, enabling a decentralized overlapping control deployment in real-time. Simulation investigations, involving a realistic highway stretch and demand scenarios, demonstrate that the proposed decentralized regulators are almost as efficient as the centralized control solutions for the problem at hand.

Keywords: lane-free traffic, internal boundary control, linear-quadratic regulator, overlapping control, linear matrix inequalities (LMI).

4.2 Introduction

Road and highway traffic congestion is one of the most serious problems for big cities around the world, causing extensive delays, increased fuel consumption, excessive environmental pollution, and reduced traffic safety. Conventional traffic control measures are valuable (Papageorgiou et al. 2003; Kurzhanskiy and Varaiya, 2010), but not always sufficient to tackle the heavy traffic demand. Gradually emerging and future ground-breaking capabilities of connected automated vehicles (CAV) should be exploited to develop innovative solutions that can be applied to smart road infrastructures. Recently, there has been a strong interest in the scientific community in developing a variety of methods attempting to exploit current and future vehicle automation and communication systems for improved real-time traffic management, see e.g. (Diakaki et al., 2015; Papamichail et al., 2019) and references therein.

A novel paradigm for vehicular traffic, which is appropriate for high penetration rates of vehicles equipped with high-level automation and communication systems, was recently launched with the TrafficFluid concept (Papageorgiou et al., 2021), suggesting to abandon, in the CAV era, two main conventional restrictions that reflect physical limitations of human drivers. Specifically, the

TrafficFluid concept suggests: (1) lane-free traffic, whereby vehicles are not bound to fixed traffic lanes, as in conventional traffic; (2) vehicle nudging, whereby vehicles may exert a "nudging" effect on, i.e. influence the movement of, vehicles in front of them.

In this context, it becomes possible to employ internal boundary control (IBC), a promising and innovative control measure aiming to achieve an unprecedented exploitation of the available road infrastructure (Malekzadeh et al. 2021a). It should be noted that IBC is facilitated by the first TrafficFluid feature, lane-free traffic, while vehicle nudging is not a prerequisite for IBC. More specifically, IBC relies on the fact that, in lane-free traffic, the traffic flow and capacity may exhibit incremental changes in response to corresponding incremental changes of the road width. Thus, on a highway or arterial with two opposite traffic directions, the total cross-road capacity (for both directions) may be shared between the two directions in real-time, according to the prevailing demand per direction, by virtually moving the internal boundary that separates the two traffic directions and communicating this decision to CAVs, so that they respect the changed road boundary. This calls for an appropriate real-time control strategy that changes the internal road boundary in space and time, as illustrated in Fig. 4.1, in response to the current traffic conditions, so as to maximize the traffic efficiency in both directions.

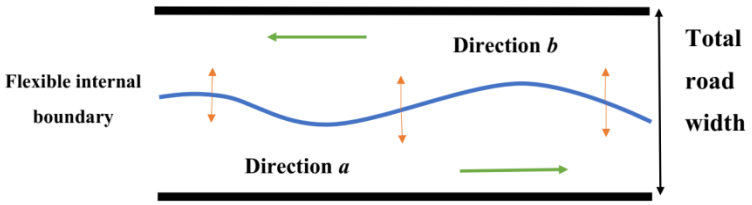


Figure 4.1: Space-time flexible internal road boundary.

Real-time IBC for lane-free CAV traffic may be broadly applicable to the high number of arterial or highway infrastructures that feature unbalanced demands during the day in the two traffic directions, to strongly mitigate or even utterly suppress congestion. Even for infrastructures experiencing strong demand in both directions quasi-simultaneously, real-time IBC may intensify the road utilization and lead to sensible improvements.

The characteristics of IBC are analyzed by Malekzadeh et al. (2021a), where its high improvement potential is demonstrated by formulating and solving an open-loop optimal control problem, in the form of a convex Quadratic Programming (QP) problem, aiming at minimizing a physical criterion, the total time spent by all vehicles. That approach may be used within a Model Predictive Control (MPC) frame, with online demand prediction, for real-time application. However, simpler real-time approaches with similar efficiency, but without the need for online demand prediction, are preferable. Therefore, a feedback-based Linear-Quadratic Regulator (LQR) was developed by Malekzadeh et al. (2021b) for IBC, aiming at balancing the relative densities in the two directions; and have been demonstrated to be robust and similarly efficient as the open-loop optimal control solution, while avoiding the need for accurate modelling and external demand prediction. It is noteworthy that, for the needs of the present investigation, the two previously developed centralized strategies, QP and LQR, are applied here to different and,

most importantly, more sizable highway systems than in (Malekzadeh, et al. 2021a) and (Malekzadeh, et al. 2021b).

However, for very long highways, the centralized LQR applying to the whole highway may have to be designed with hundreds of state variables; also, its application requires real-time information from the whole highway under consideration, which may be problematic for very long highways with respect to the required communications and physical system architecture. In other words, to calculate the control input (internal boundary position) at a specific highway location, the centralized LQR requires formally real-time information about the traffic conditions in the whole highway, i.e., even from highway parts that may be dozens of kilometers away, something that is obviously not necessary for efficient IBC operations.

Generally speaking, control decentralization in large-scale and geographically extended systems is a desirable and reasonable control structure, as it is simpler to monitor and maintain and more reliable in cases of device failures. Such features appear, for example, in the control of interconnected power systems (Patil et al. 2019), water systems (Hafeez et al. 2021), as well as in space or traffic systems (Pasquale et al. 2020), where analysis, synthesis and implementation is distributed to interacting subsystems (Siljak., 2011) under the motto "think global, act local". Such systems are typically composed of various physically or functionally local, possibly interacting control sub-stations, where each sub-station is in charge only of the operation of a part of the overall system, and the overall control system objective is reached by the actions of the local control subsystems.

Various methodologies were proposed in the past decades for the design of such decentralized control systems, the structure of which may be dictated by the structure of the process under control; or may have to be defined based on entailed performance, maintenance and communication requirements, reliability and other properties. In the latter case, there are four well-known system decomposition structures: disjoint decomposition, overlapping decomposition, border block diagonal decomposition, and epsilon decomposition (Bakule, 2014). The disjoint decomposition of the overall system is not reasonable if the subsystems are strongly interconnected, in which case an overlapping control structure may be employed, where the overlapping parts create connections between subsystems to enhance the local and overall control performance. Several overlapping control methods and applications have been proposed, see for example (Trudnowski and Pierre, 1992), (Bakule et al., 2006), (Hug-Glanzmann and Andersson, 2009), (Ahmadi and Aldeen, 2017).

A highway is an one-dimensional structure, where traffic flow is modelled on the basis of a tandem of sections (each about 500 m long), each of them interacting with its two adjacent sections. More specifically, the information flow in the highway traffic process points downstream in free-flow conditions; and upstream in congested conditions. The research question investigated in this paper is whether the centralized IBC-LQR may be replaced, without significant loss of efficiency, by a tandem of local sub-controllers, each dealing with a corresponding subsystem that consists of a few subsequent highway sections. Given the interactions among adjacent sections, the subsystems should be overlapping, i.e., adjacent subsystems have one or more sections in common. Such a decentralization may reduce the

regulator design effort in cases of very long highways; and would certainly improve on the communication requirements and reliability of the overall control system.

This paper introduces two overlapping decentralized control schemes for IBC of lane-free automated vehicle traffic. The first approach is based on a contractible controller developed for an extended model that exploits the overlapping structure of the system. This approach allows for decomposed control design, separately for each subsystem, which enhances scalability of design and expandability of the controlled system when considering additional highway parts for control, since the additional subsystem control can be designed independently of the existing control system. The method was first introduced by Ikeda et al. (1981) and was extended by İftar and Özgüner (1990) to consider input inclusion, additionally to state inclusion. In the second approach, the discrete-time LQR problem is first transformed into a linear matrix inequalities (LMI) problem (Zečević and Šiljak (2005). The overlapping control structure used is selected, the LMI problem is solved, and the accordingly structured gain matrix is obtained, enabling a decentralized overlapping control scheme. This second approach calls for a centralized design procedure (LMI problem), but may, for this reason, lead to better control results under the same decentralized overlapping control structure, compared to the first approach.

The well-known Cell Transmission Model (CTM) (Daganzo, 1994) is used, after linearization, for controller design; and, in its full nonlinear form, for simulation testing. The next section presents some background issues and the appropriately adjusted CTM equations, while Section 4.3 provides the design of the proposed control schemes. Simulation investigations are discussed in Section 4.4 while conclusions are given in Section 4.5.

4.3 Background

Some background information on IBC modelling and centralized LQR control design, which is necessary for understanding the overlapping controller design and application, is reported in this section for completeness; see (Malekzadeh et al. 2021a; Malekzadeh et al. 2021b) for more details.

4.3.1 CTM for Internal Boundary Control

Lane-free traffic is not expected to give rise to structural changes of the basics of existing macroscopic traffic flow models. As also supported by results presented by Papageorgiou et al. (2021), Yanumula et al. (2021), and Malekzadeh et al. (2022), basic notions and concepts, like the conservation equation, the stationary Fundamental Diagram (FD), as well as limited-capacity bottlenecks activating congested-traffic waves, continue to characterize macroscopic traffic flow modelling in the case of lane-free automated vehicle traffic. An extended version of CTM, a first-order dynamic traffic flow model with a triangular FD, is considered for control design (in linearized form) and for simulation testing (in full nonlinear form). Note that a triangular FD may be characterized by three parameters, e.g., the capacity q_{cap},, the critical density ρ_{cr} (at which capacity flow occurs) and the jam (maximum) density $\rho_{\max}$, see (Daganzo, 1994). The present study is focusing on the macroscopic level to test and demonstrate the unprecedented improvements enabled by IBC. The CTM is used in this context as a proxy for traffic flow modelling, under the reasonable assumption that AVs are not changing the mentioned

fundamental properties of traffic, but mainly quantitative aspects (free speed, capacity, critical density values). In fact, recent, yet unpublished, CTM calibration results, using microscopic AV lane-traffic simulation data, confirm this assumption

Let us call the two opposite traffic directions, presented in Fig. 4.1, directions a (from left to right) and b (from right to left). The stretch is subdivided in n sections, with respective lengths L_i, $i=1,2,\ldots,n$. The total road width (both directions) w, which is assumed constant over all sections for simplicity, can be shared in real-time, independently for each section, between the two directions. As a result, each direction is assigned a corresponding road width $w_i^a = \varepsilon_i \cdot w$ and $w_i^b = (1-\varepsilon_i)\cdot w$, where $0 \le \varepsilon_i \le 1$ is the *sharing factor* per section $i=1,2,\ldots,n$, to be specified in real-time as a control input by the internal boundary controller. The total (both directions) section capacity q_{cap}, as well as the total critical density ρ_{cr} and the total jam density $\rho_{\max}$, are accordingly shared between the two traffic directions a and b. Based on the derivation presented by Malekzadeh (2021a), these quantities are given by direction by

$$\begin{aligned} &q_{i,cap}^a(\varepsilon_i) = \varepsilon_i \cdot q_{cap},\ q_{i,cap}^b(\varepsilon_i) = (1-\varepsilon_i)\cdot q_{cap} \\ &\rho_{i,cr}^a(\varepsilon_i) = \varepsilon_i \cdot \rho_{cr},\ \rho_{i,cr}^b(\varepsilon_i) = (1-\varepsilon_i)\cdot \rho_{cr} \\ &\rho_{i,\max}^a(\varepsilon_i) = \varepsilon_i \cdot \rho_{\max},\ \rho_{i,\max}^b(\varepsilon_i) = (1-\varepsilon_i)\cdot \rho_{\max} \end{aligned} \tag{4.1}$$

For the IBC problem, the complete closure of either direction is disallowed; to this end, the assigned road width in either direction should never be less than the widest vehicles driving on the road. This requirement gives rise to stricter constraints for the sharing factors as follows

$$0 < \varepsilon_{i,\min} \le \varepsilon_i \le \varepsilon_{i,\max} < 1 \tag{4.2}$$

where $\varepsilon_{i,\min}\cdot w$ and $(1-\varepsilon_{i,\max})\cdot w$ are the minimum admissible widths to be assigned to directions a and b, respectively. Another restriction to be applied to the sharing factors concerns the time-delay needed to evacuate traffic on the direction that receives a restricted width, compared to the previous control time-step. This time-delay is small in lane-free CAV traffic, where no physical barrier exists among the two traffic directions, for moderate changes of the sharing factors applied to short sections, but needs nevertheless to be considered. Clearly, the time-delay should apply only to the traffic direction that is being widened, compared to the previous control interval; while the direction that is restricted should promptly apply the smaller width, so that CAVs therein move out of the reduced-width zone. Assume that the required time-delay is smaller than or equal to the control time interval T_c; then, the time-delay requirement is automatically fulfilled for each section i, if the sharing factors that are actually applied to the two directions, i.e. ε_i^a and ε_i^b, respectively, are calculated as follows

$$\begin{aligned} &\varepsilon_i^a(k_c) = \min\left\{\varepsilon_i(k_c), \varepsilon_i(k_c-1)\right\} \\ &\varepsilon_i^b(k_c) = \min\left\{1-\varepsilon_i(k_c), 1-\varepsilon_i(k_c-1)\right\} \end{aligned} \tag{4.3}$$

where $k_c = 0,1,\ldots$ is the discrete control time index. It is noted that the notation $\varepsilon_i^a(k_c)$ and $\varepsilon_i^b(k_c)$ indicates that the sharing factors are applied for the duration of the control time interval

$[k_c \cdot T_c, (k_c+1)\cdot T_c)$. The above equations may be readily extended if the required time-delay is a multiple of the control time interval T_c.

Traffic flows from section 1 to section n in direction a; and from section n to section 1 in direction b (see Fig. 4.3 as an example). It is defined ρ_i^a, $i=1,2,\ldots,n$, to be the traffic density (in veh/km) of section i, direction a; and ρ_i^b, $i=1,2,\ldots,n$, the traffic density of section i, direction b. Similarly, q_i^a and q_i^b, $i=1,2,\ldots,n$, are defined to be the mainstream exit flows (in veh/h) of section i for directions a and b, respectively. Thus, q_0^a is the feeding upstream mainstream inflow for direction a; and q_{n+1}^b is the feeding upstream mainstream inflow for direction b. Every section may have an on-ramp or an off-ramp at its upstream boundary. The on-ramp flows (if any) at section i are denoted r_i^a for direction a, and r_i^b for direction b. The off-ramp flow (if any) of section i, direction a, is calculated based on known exit rates β_i^a multiplied with the upstream-section flow, i.e. $\beta_i^a q_{i-1}^a$; and the off-ramp flow (if any) of section i, direction b, is calculated based on known exit rates β_i^b multiplied with the upstream-section flow, i.e. $\beta_i^b q_{i+1}^b$. The conservation equations for the sections of direction a are:

$$\begin{aligned}
\rho_1^a(k+1) &= \rho_1^a(k) + \frac{T}{L_i}(q_0^a(k) - q_1^a(k)) \\
\rho_i^a(k+1) &= \rho_i^a(k) + \frac{T}{L_i}((1-\beta_i^a)q_{i-1}^a(k) - q_i^a(k) + r_i^a(k)),\ i = 2,3,\ldots,n
\end{aligned} \tag{4.4}$$

where T is the model time-step, typically set equal to 5–10 s for section lengths of some 500 m, and $k=0,1,\ldots$ is the corresponding discrete-time index of the model.

According to CTM, traffic flow is obtained as the minimum of demand and supply functions, except for the last section, where only the demand function is considered, assuming that the downstream traffic conditions are uncongested. Considering the impact of the respective sharing factors on the FDs, the flow is obtained as follows (see (Malekzadeh et al. 2021a))

$$\begin{aligned}
q_i^a(k) &= \min\left\{ Q_D(\rho_i^a(k), \varepsilon_i^a(k_c)), \frac{Q_S(\rho_{i+1}^a(k), \varepsilon_{i+1}^a(k_c))}{(1-\beta_{i+1}^a)} - \lambda_r r_{i+1}^a(k) \right\},\ i = 1,2,\ldots,n-1 \\
q_n^a(k) &= Q_D(\rho_n^a(k), \varepsilon_i^a(k_c)).
\end{aligned} \tag{4.5}$$

The demand and supply functions are given by the following respective equations

$$\begin{aligned}
Q_D(\rho,\varepsilon) &= \min\left\{ \varepsilon q_{cap} + \lambda_d q_{cap} \frac{\rho - \varepsilon\rho_{cr}}{\rho_{cr} - \rho_{\max}}, v_f \rho \right\}, \\
Q_S(\rho,\varepsilon) &= \min\left\{ \varepsilon q_{cap}, w_s(\varepsilon\rho_{\max} - \rho) \right\},
\end{aligned} \tag{4.6}$$

where v_f is the free speed (which is assumed equal for all sections for simplicity) and w_s is the back-wave speed. Note that, for a triangular FD, these two parameters may be deduced from q_{cap}, ρ_{cr}, and $\rho_{\max}$,; note also that the two parameters are invariant to the road width. The control time-step T_c does not need to be equal to the model time-step, but is assumed to be a multiple of

T, in which case the control time index is given by $k_c = \lfloor kT/T_c \rfloor$, where $\lfloor . \rfloor$ is the integer part notation.

It is well-known that CTM does not reproduce the capacity drop phenomenon, i.e., the empirical finding that, at the head of congestion, the observed flow in real traffic is reduced, compared to the road capacity. Capacity drop is deemed to occur in conventional traffic due to bounded and differing accelerations of different vehicles (Yuan et al., 2015). Recently, CTM has been extended in a number of possible ways to enable the reproduction of capacity drop, see (Kontorinaki et al., 2017) for an overview and comparison. The presence of capacity drop in conventional freeway traffic is a major reason for infrastructure degradation and for the need of introducing traffic control measures to restore capacity (Papageorgiou et al., 2003). In contrast, in the present context of internal boundary control, the presence of capacity drop marks a secondary source of amelioration of the traffic conditions, because the potential benefits achievable via opportune capacity sharing are expected to be much higher than those resulting from capacity drop avoidance. In fact, it is unknown at the moment, if and to what extent capacity drop may occur in lane-free CAV traffic. To be able to investigate the impact of possible capacity drop, the option of introducing capacity drop has been incorporated in the above equations trough appropriate terms according to Kontorinaki et al. (2017). More specifically, this option is enabled via the parameters λ_r and λ_d in equations (5) and (6). If these parameters are set to $\lambda_r = 1$ and $\lambda_d = 0$, no capacity drop is reproduced, as typical for CTM; if these values are set between 0 and 1, a corresponding level of capacity drop is produced by the model.

The equations for direction b are analogous to those of direction a, with few necessary index modifications. Section numbers in direction b are descending, hence

$$\begin{aligned} \rho_i^b(k+1) &= \rho_i^b(k) + \frac{T}{L_i}((1-\beta_i^b)q_{i+1}^b(k) - q_i^b(k) + r_i^b(k)),\ i = 1,2,\ldots,n-1 \\ \rho_n^b(k+1) &= \rho_n^b(k) + \frac{T}{L_i}(q_{n+1}^b(k) - q_n^b(k)) \end{aligned} \tag{4.7}$$

and the flows are given by

$$\begin{aligned} q_1^b(k) &= Q_D(\rho_1^b(k), \varepsilon_i^b(k_c)) \\ q_i^b(k) &= \min\left\{ Q_D(\rho_i^b(k), \varepsilon_i^b(k_c)), \frac{Q_S(\rho_{i-1}^b(k), \varepsilon_{i-1}^b(k_c))}{(1-\beta_{i-1}^b)} - \lambda_r r_{i-1}^b(k) \right\},\ i = 2,3,\ldots,n \end{aligned} \tag{4.8}$$

In conventional traffic management, traffic densities (in veh/km) characterize clearly the state of traffic, depending on their value versus the critical density: free traffic (when density is lower than critical density), critical traffic (when density is around critical density) or congested traffic (when density is higher than critical density). However, in the proposed IBC measure, the critical density for each direction and section is not constant, but a function of the sharing factor (see (3.27)), and is changing according to the applied control action. Therefore, the density value by itself is not sufficient, in the IBC context, to characterize the traffic situation in a section.

To address this issue, the following relations define the relative densities (dimensionless) per section and per direction. The relative density of section i and direction a or b is obtained by dividing the corresponding traffic density with the corresponding critical density, which, on its turn, depends (via (3.27)) on the sharing factor prevailing during the last time-step, as follows

$$\tilde{\rho}_i^a(k) = \frac{\rho_i^a(k)}{\varepsilon_i(k-1)\rho_{cr}},\ \tilde{\rho}_i^b(k) = \frac{\rho_i^b(k)}{(1-\varepsilon_i(k-1))\rho_{cr}} \tag{4.9}$$

The relative densities reflect clearly the state of the traffic in the IBC context. Specifically, if the relative density of a section and direction is less than 1, it reflects under-critical (free-flow) traffic conditions; if it is around 1, it reflects capacity flow; and if it is greater than 1, it reflects over-critical (congested) traffic conditions.

4.3.2 Centralised open-loop optimisation for Internal Boundary Control

Malekzadeh et al. (2021a) first addressed the IBC problem in lane-free traffic and demonstrated its high improvement potential by formulating and solving an open-loop optimal control problem, in the form of a convex Quadratic Programming (QP) problem. The QP problem minimizes a physical performance criterion, the total time spent by all vehicles over a time horizon, along with some operational sub-objectives. The full CTM model is considered in the optimization problem, albeit with an appropriate transformation of the piecewise linear (triangular) FD, so as to involve only linear equality and inequality constraints and enable the usage of efficient QP solution algorithms. Availability of demands over the considered time horizon is necessary. That approach may be used within a Model Predictive Control (MPC) frame, with online demand prediction, for real-time application. In an open-loop setting, the approach delivers the maximum achievable performance, as it is based on the full CTM equations, considers all constraints and external inputs (demands) and minimizes a physical performance criterion. Therefore, in this paper, the QP approach is used as a benchmark, delivering an upper bound of achievable performance, against which other, simpler approaches may be compared.

4.3.3 Centralised LQR design for Internal Boundary Control

The QP approach is efficient, but simpler real-time approaches with similar efficiency, but without the need for online demand prediction, are preferable. Therefore, feedback-based Linear-Quadratic Regulators (LQR) have been developed by Malekzadeh et al. (2021b) for IBC, aiming at balancing the relative densities in the two directions; and have been demonstrated to be robust and similarly efficient as the open-loop QP solution, while avoiding the need for accurate modelling and external demand prediction.

Linearization of the CTM dynamic equations around a nominal point was presented analytically by Malekzadeh (2021b). To achieve this, the one-step retarded control input was defined as a new state variable according to $\gamma_i(k+1) = \varepsilon_i(k),\ \ i = 1, 2, \ldots, n.$ Following the linearization procedure by Malekzadeh (2021b), the linearized state-space model is

$$\mathbf{x}(k+1) = \hat{\mathbf{A}}\mathbf{x}(k) + \hat{\mathbf{B}}\mathbf{u}(k) \tag{4.10}$$

where $\mathbf{x}(k) = [\Delta\tilde{\rho}_1^a(k), \Delta\tilde{\rho}_1^b(k), \Delta\gamma_1(k), ..., \Delta\tilde{\rho}_n^a(k), \Delta\tilde{\rho}_n^b(k), \Delta\gamma_n(k)]^T$ is the state vector and $\mathbf{u}(k) = \boldsymbol{\Delta\varepsilon}(k)$ is the control vector, whereby $\boldsymbol{\Delta\varepsilon}(k) = [\Delta\varepsilon_1(k), \ldots, \Delta\varepsilon_n(k)]^T$. Also, $\Delta(.)(k) = (.)(k) - (.)^N$, the superscript N denoting the nominal values, while it has been assumed that $\Delta(.)(k) = 0$ for all disturbances (upstream mainstream inflows, as well as the on-ramp flows, of each direction). $\hat{\mathbf{A}} \in \mathbb{R}^{3n \times 3n}$ and $\hat{\mathbf{B}} \in \mathbb{R}^{3n \times n}$ are the time-invariant state and input matrices, respectively, while $\mathbf{x} \in \mathbb{R}^{3n}$ and $\mathbf{u} \in \mathbb{R}^n$.

If the control time-step is defined as a multiple of the model time-step, i.e. $T_c = MT$, where M is an integer, then the discrete control time index is $k_c = \lfloor kT/T_c \rfloor$. Thus, the linear state-space equation may be changed as follows, in order to be based on the control time-step T_c,

$$\varphi: \mathbf{x}(k_c + 1) = \mathbf{A}\mathbf{x}(k_c) + \mathbf{B}\mathbf{u}(k_c) \tag{4.11}$$

where $\mathbf{A} = \hat{\mathbf{A}}^M$, and $\mathbf{B} = (\hat{\mathbf{A}}^{M-1} + \hat{\mathbf{A}}^{M-2} + ... + \mathbf{I})\hat{\mathbf{B}}$.

When employing the LQR methodology, as done by Malekzadeh (2021b), the control goal is the minimization of the quadratic criterion

$$J = \frac{1}{2} \sum_{k_c=0}^{\infty} \left[\mathbf{x}^T(k_c)\mathbf{Q}\mathbf{x}(k_c) + \mathbf{u}^T(k_c)\mathbf{R}\mathbf{u}(k_c) \right] \tag{4.12}$$

where $\mathbf{Q} \in \mathbb{R}^{3n \times 3n}$ is a diagonal positive semidefinite matrix and $\mathbf{R} \in \mathbb{R}^{n \times n}$ is a diagonal positive definite matrix. The first term penalizes deviations of the state variables from zero, i.e. deviations of $\tilde{\rho}_i^a(k_c)$, $\tilde{\rho}_i^b(k_c)$, $\gamma_i(k_c)$, $i = 1, 2, \ldots, n$, from their respective desired nominal values. The second term penalizes deviations of the control inputs from the nominal values.

The nominal value of relative densities on both directions is set equal to 1 (with positive, under-critical derivative), so that the controller is motivated to operate the system near capacity, which is good for traffic efficiency. In particular, due to the quadratic penalty terms, the controller tends to mitigate strong density departures from the critical density at specific sections, i.e., mitigate traffic congestion. In addition, if capacity flow is not feasible (e.g. due to lack of demand), then minimizing a sum of squares has the tendency to balance deviations from the nominal values at different sections and directions, something that is conform with the secondary operational sub-objective of balancing the margin to capacity across sections and directions. On the other hand, the nominal value for the sharing factors is set to 0.5, so as to have smooth and moderate internal boundary changes. Thus, minimization of the second term in (4.12) mitigates deviations of the sharing factors from 0.5 and balances these deviations in space and time, which is a secondary operational sub-objective, as unnecessarily strong internal boundary changes over space and time should be avoided.

The optimal controller minimizing the criterion (4.12) subject to the model (4.11) is given by a linear state-feedback control law of the form $\mathbf{u}(k_c) = \mathbf{K}\mathbf{x}(k_c)$, where $\mathbf{K} \in \mathbb{R}^{n \times 3n}$ is a constant gain matrix given by

$$\mathbf{K} = (\mathbf{R} + \mathbf{B}^T\mathbf{P}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{P}\mathbf{A} \tag{4.13}$$

and $\mathbf{P}$ is a unique positive semidefinite solution of the discrete-time algebraic Riccati equation.

4.4 Overlapping control schemes

Summarizing from above, each highway section (of some 500 m length) has three state variables and one control input, which is the corresponding sharing factor. For very long highways, designing the centralized LQR, proposed by Malekzadeh (2021b) and outlined above, calls for the solution of an accordingly large-scale Riccati equation. More importantly, the centralized LQR requires real-time information for all the states of the system even very remote ones, to compute each control input. This requirement may be problematic for long highways due to the need to transfer data from the whole highway in order to compute each sharing factor. Lastly, the traffic situation may be quite different at different parts of a long highway; then, balancing relative densities and sharing factors over the whole highway may not be fully relevant, as it is reasonable to address this sub-objective primarily at local levels.

For the above reasons, there is an interest in developing a decentralized control scheme, which may reduce the LQR design complexity; reduce the burden of real-time data transferring; and increase the system reliability in cases of device failure. Due to strong dependencies between consecutive sections in traffic flow, developing a fully decentralized control scheme (with disjoint subsystems) may reduce the control efficiency compared to the centralized case, as was indeed confirmed in preliminary investigations. Therefore, in this study, overlapping control schemes are adopted and tested for IBC. The approaches rely on the separation of the highway in a number of subsequent subsystems (highway stretches) which are overlapping, i.e., adjacent subsystems have some sections in common; hence, adjacent subsystems share some states and some control inputs. This approach delivers overlapping decentralized controllers, where, thanks to the overlapping structure, subsystems remain partly aware about the inter-relations with adjacent subsystems. Overlapping controllers offer less communication and equipment requirements, less monitoring and maintenance effort and higher reliability versus the centralized LQR, for the price of a minor performance reduction. It is hardly possible to predict appropriate trade-offs of these features versus performance, therefore the paper investigates the issue empirically, using different overlapping structures

The control inputs of the non-overlapping sections are computed based only on local subsystem information; while control inputs in the overlapping sections use information from both adjacent subsystems. The following sections present the details of the employed overlapping control approaches. The structure of the overlapping control scheme is illustrated in Fig. 4.2 for a two-subsystem case with one overlapping section. The overlapping section 4 delivers real-time measurements to both adjacent subsystem controllers; and its control input is specified with contributions from both adjacent subsystem controllers.

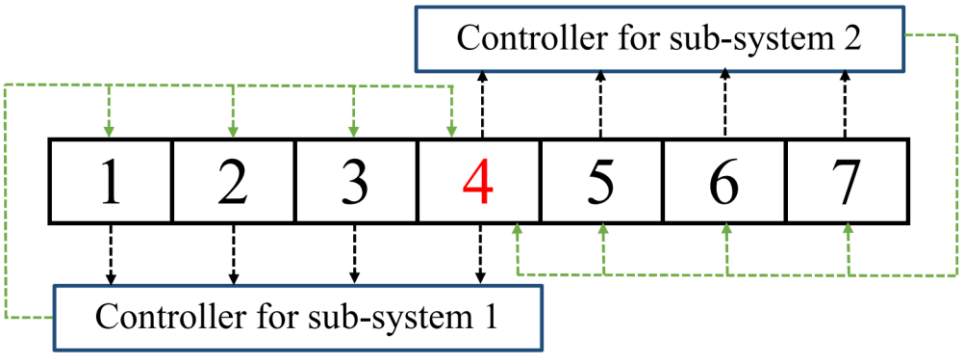


Fig. 4.2: A block diagram of overlapping control scheme for two subsystems.

4.4.1 Contractible controller design with state and input inclusion

Consider the following linear time-invariant system

$$\tilde{\varphi}:\ \tilde{\mathbf{x}}(k_c+1)=\tilde{\mathbf{A}}\tilde{\mathbf{x}}(k_c)+\tilde{\mathbf{B}}\tilde{\mathbf{u}}(k_c) \tag{4.14}$$

where $\tilde{\mathbf{x}}\in\mathbb{R}^{\tilde{n}}$ and $\tilde{\mathbf{u}}\in\mathbb{R}^{\tilde{m}}$. In the following, the system φ in (4.11) is referred to as the original system, and $\tilde{\varphi}$ is referred to as the expanded system. It is assumed that $\tilde{n}\geq 3n$ and $\tilde{m}\geq n$, as the expanded system $\tilde{\varphi}$ has typically more states and more inputs than the original system φ due to the expansion. More specifically, the goal of the expansion is to create overlapping areas that belong to both corresponding adjacent subsystems, and this leads to higher state and control dimensions for the expanded system.

The state and the input of the original system are partitioned as

$$\begin{aligned}\mathbf{x}&=(\mathbf{x}_1^T,\mathbf{x}_2^T,\ldots,\mathbf{x}_P^T),\quad \mathbf{x}_i\in\mathbb{R}^{n_i}\\ \mathbf{u}&=(\mathbf{u}_1^T,\mathbf{u}_2^T,\ldots,\mathbf{u}_P^T),\quad \mathbf{u}_i\in\mathbb{R}^{m_i}\end{aligned} \tag{4.15}$$

where $P\geq 3$ is the odd number of partitions (subsequent highway stretches). Here it is assumed that, for all even indexes i, $\mathbf{x}_i$ and $\mathbf{u}_i$ correspond to the overlapping parts of the state and input vectors, respectively, while for all odd indexes i, $\mathbf{x}_i$ and $\mathbf{u}_i$ correspond to the non-overlapping parts of the state and input spaces.

Now consider that the matrices $\mathbf{A}$ and $\mathbf{B}$ for the system φ are also partitioned compatibly, as follows:

$$\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1P}\\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2P}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{A}_{P1} & \mathbf{A}_{P2} & \ldots & \mathbf{A}_{PP}\end{bmatrix},\quad \begin{bmatrix}\mathbf{B}_{11} & \mathbf{B}_{12} & \cdots & \mathbf{B}_{1P}\\ \mathbf{B}_{21} & \mathbf{B}_{22} & \cdots & \mathbf{B}_{2P}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{B}_{P1} & \mathbf{B}_{P2} & \cdots & \mathbf{B}_{PP}\end{bmatrix}. \tag{4.16}$$

An expanded system $\tilde{\varphi}$ can be obtained from the original system φ, as will be explained below. If the expanded system satisfies certain conditions presented by İftar and Özgüner (1990), such an expansion is called an extension and can be leveraged in order to design the controller for the

original system based on the design of a controller for the extended system. For the case of overlapping areas belonging to exactly two corresponding adjacent subsystems, as in the IBC application, the expanded system matrices $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ can be easily obtained from the original matrices $\mathbf{A}$ and $\mathbf{B}$ by simply doubling the state equations of the overlapping partitions. This expansion may be readily shown to satisfy the conditions required in (İftar and Özgüner, 1990) for an extension of the original system.

A highway stretch that is subdivided in 10 sections (see Fig. 4.3) is considered as an example. This system may be separated in two subsystems with a single overlapping area of two sections, i.e. based on three partitions ($P=3$) of the original system. As can be seen in Fig. 4.3, two subsystems with six sections each are introduced; while sections 5 and 6 are considered as an overlapping area. The two subsystems are strongly connected through the overlapping area. The matrices for the extended system with $P=3$ are obtained by doubling the state equations of the overlapping sections:

$$\tilde{\mathbf{A}} = \left[\begin{array}{cc:cc} \mathbf{A}_{11} & \mathbf{A}_{12} & 0 & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & 0 & \mathbf{A}_{23} \\ \hdashline \mathbf{A}_{21} & 0 & \mathbf{A}_{22} & \mathbf{A}_{23} \\ \mathbf{A}_{31} & 0 & \mathbf{A}_{32} & \mathbf{A}_{33} \end{array}\right] \triangleq \left[\begin{array}{cc} \tilde{\mathbf{A}}_1 & \tilde{\mathbf{A}}_{12} \\ \tilde{\mathbf{A}}_{21} & \tilde{\mathbf{A}}_2 \end{array}\right] \tag{4.17}$$

$$\tilde{\mathbf{B}} = \left[\begin{array}{cc:cc} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{22} & \mathbf{B}_{23} \\ \hdashline \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{22} & \mathbf{B}_{23} \\ \mathbf{B}_{31} & \mathbf{B}_{32} & \mathbf{B}_{32} & \mathbf{B}_{33} \end{array}\right] \triangleq \left[\begin{array}{cc} \tilde{\mathbf{B}}_1 & \tilde{\mathbf{B}}_{12} \\ \tilde{\mathbf{B}}_{21} & \tilde{\mathbf{B}}_2 \end{array}\right] \tag{4.18}$$

Note that in the specific IBC application, the non-zero coupling matrices $\tilde{\mathbf{A}}_{12}$, $\tilde{\mathbf{A}}_{21}$, $\tilde{\mathbf{B}}_{12}$ and $\tilde{\mathbf{B}}_{21}$ have most of their elements equal to zero.

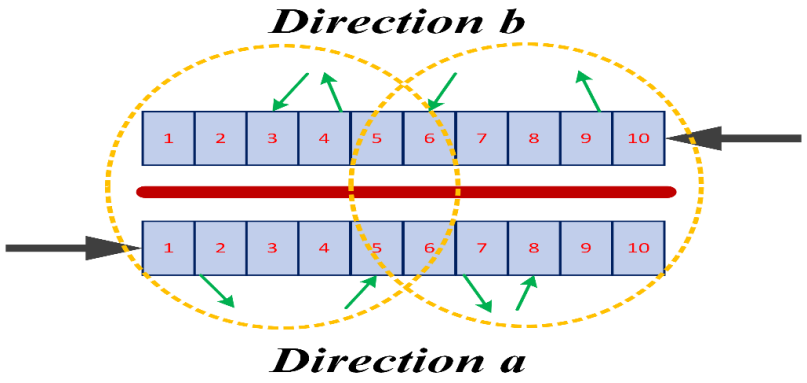


Fig. 4.3: A highway stretch example.

For the general case, consider $n_D = (P+1)/2$ decoupled subsystems:

$$\varphi_i^D : \tilde{\mathbf{x}}_i(k_c+1) = \tilde{\mathbf{A}}_i \tilde{\mathbf{x}}_i(k_c) + \tilde{\mathbf{B}}_i \tilde{\mathbf{u}}_i(k_c), \; i = 1, \ldots n_D \tag{4.19}$$

where

$$\tilde{\mathbf{A}}_1 = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \quad \tilde{\mathbf{A}}_{n_D} = \begin{bmatrix} \mathbf{A}_{(P-1)(P-1)} & \mathbf{A}_{(P-1)P} \\ \mathbf{A}_{P(P-1)} & \mathbf{A}_{PP} \end{bmatrix},$$
$$\tilde{\mathbf{A}}_i = \begin{bmatrix} \mathbf{A}_{(2i-2)(2i-2)} & \mathbf{A}_{(2i-2)(2i-1)} & \mathbf{A}_{(2i-2)(2i)} \\ \mathbf{A}_{(2i-1)(2i-2)} & \mathbf{A}_{(2i-1)(2i-1)} & \mathbf{A}_{(2i-1)(2i)} \\ \mathbf{A}_{(2i)(2i-2)} & \mathbf{A}_{(2i)(2i-1)} & \mathbf{A}_{(2i)(2i)} \end{bmatrix}, \quad i = 2, \ldots, n_D - 1 \tag{4.20}$$

and

$$\tilde{\mathbf{B}}_1 = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}, \tilde{\mathbf{B}}_{n_D} = \begin{bmatrix} \mathbf{B}_{(P-1)(P-1)} & \mathbf{B}_{(P-1)P} \\ \mathbf{B}_{P(P-1)} & \mathbf{B}_{PP} \end{bmatrix},$$
$$\tilde{\mathbf{B}}_i = \begin{bmatrix} \mathbf{B}_{(2i-2)(2i-2)} & \mathbf{B}_{(2i-2)(2i-1)} & \mathbf{B}_{(2i-2)(2i)} \\ \mathbf{B}_{(2i-1)(2i-2)} & \mathbf{B}_{(2i-1)(2i-1)} & \mathbf{B}_{(2i-1)(2i)} \\ \mathbf{B}_{(2i)(2i-2)} & \mathbf{B}_{(2i)(2i-1)} & \mathbf{B}_{(2i)(2i)} \end{bmatrix}, \quad i = 2, \ldots, n_D - 1. \tag{4.21}$$

The gains can be calculated for the feedback law for each decoupled subsystem $\tilde{\mathbf{u}}_i(k_c) = \tilde{\mathbf{K}}^i \tilde{\mathbf{x}}_i(k_c)$ using the LQR method for objectives similar to (4.12). The resulting subsystem gains $\tilde{\mathbf{K}}^i$ are partitioned as follows:

$$\tilde{\mathbf{K}}^i = \begin{bmatrix} \tilde{\mathbf{K}}^i_{11} & \tilde{\mathbf{K}}^i_{12} \\ \tilde{\mathbf{K}}^i_{21} & \tilde{\mathbf{K}}^i_{22} \end{bmatrix}, \quad i = 1, n_D$$
$$\tilde{\mathbf{K}}^i = \begin{bmatrix} \tilde{\mathbf{K}}^i_{11} & \tilde{\mathbf{K}}^i_{12} & \tilde{\mathbf{K}}^i_{13} \\ \tilde{\mathbf{K}}^i_{21} & \tilde{\mathbf{K}}^i_{22} & \tilde{\mathbf{K}}^i_{23} \\ \tilde{\mathbf{K}}^i_{31} & \tilde{\mathbf{K}}^i_{32} & \tilde{\mathbf{K}}^i_{33} \end{bmatrix}, \quad i = 2, \ldots, n_D - 1. \tag{4.22}$$

The overall gain for the feedback control law of the extended system $\tilde{\mathbf{u}}(k_c) = \tilde{\mathbf{K}}\tilde{\mathbf{x}}(k_c)$ is given by

$$\tilde{\mathbf{K}} = \text{blockdiag}(\tilde{\mathbf{K}}^1, \ldots, \tilde{\mathbf{K}}^{n_D}). \tag{4.23}$$

If $\tilde{\varphi}$ is an extension of φ, then the gain value $\mathbf{K}$ for the equivalent feedback control law of the original system, i.e. $\mathbf{u}(k_c) = \mathbf{K}\mathbf{x}(k_c)$, can be constructed using the gain value $\tilde{\mathbf{K}}$ from the feedback law of the extended system $\tilde{\mathbf{u}}(k_c) = \tilde{\mathbf{K}}\tilde{\mathbf{x}}(k_c)$ (İftar and Özgüner, 1990). For the case of the overlapping systems considered here, the gain is as follows:

$$\mathbf{K}=\begin{bmatrix}
\tilde{\mathbf{K}}_{11}^{1} & \tilde{\mathbf{K}}_{12}^{1} & 0 & 0 & 0 & 0 & \cdots & 0 \\
\tilde{\mathbf{K}}_{21}^{1} & \tilde{\mathbf{K}}_{22}^{1}+\tilde{\mathbf{K}}_{11}^{2} & \tilde{\mathbf{K}}_{12}^{2} & \tilde{\mathbf{K}}_{13}^{2} & 0 & 0 & \cdots & 0 \\
0 & \tilde{\mathbf{K}}_{21}^{2} & \tilde{\mathbf{K}}_{22}^{2} & \tilde{\mathbf{K}}_{23}^{2} & 0 & 0 & \cdots & 0 \\
0 & \tilde{\mathbf{K}}_{31}^{2} & \tilde{\mathbf{K}}_{32}^{2} & \tilde{\mathbf{K}}_{33}^{2}+\tilde{\mathbf{K}}_{11}^{3} & \tilde{\mathbf{K}}_{12}^{3} & \tilde{\mathbf{K}}_{13}^{3} & \cdots & 0 \\
0 & 0 & 0 & \tilde{\mathbf{K}}_{21}^{3} & \tilde{\mathbf{K}}_{22}^{3} & \tilde{\mathbf{K}}_{23}^{3} & \cdots & 0 \\
0 & 0 & 0 & \tilde{\mathbf{K}}_{31}^{3} & \tilde{\mathbf{K}}_{32}^{3} & \tilde{\mathbf{K}}_{33}^{3}+\tilde{\mathbf{K}}_{11}^{4} & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & \ldots & \tilde{\mathbf{K}}_{22}^{n_D}
\end{bmatrix} \tag{4.24}$$

and for the specific case of two subsystems:

$$\mathbf{K}=\begin{bmatrix}
\tilde{\mathbf{K}}_{11}^{1} & \tilde{\mathbf{K}}_{12}^{1} & 0 \\
\tilde{\mathbf{K}}_{21}^{1} & \tilde{\mathbf{K}}_{22}^{1}+\tilde{\mathbf{K}}_{11}^{2} & \tilde{\mathbf{K}}_{12}^{2} \\
0 & \tilde{\mathbf{K}}_{21}^{2} & \tilde{\mathbf{K}}_{22}^{2}
\end{bmatrix}. \tag{4.25}$$

The approach presented above has advantages as well as some disadvantages. Traffic conditions on a long highway are usually inhomogeneous, i.e. they may be simultaneously free-flowing, critical or congested at different highway parts. In such circumstances, the mentioned balancing of relative density and sharing factor deviations from their respective nominal values may not be fully appropriate when applied to all sections of the long highway, as in the case of the centralized LQR control approach, due to strong differences in the prevailing traffic conditions. It appears more reasonable to consider subsystems with corresponding sub-objectives as done here. Additionally, this method may enhance the extendibility to new parts of a highway, as this may be done without the need for re-designing the controllers for the whole highway. On the other hand, it is evident that the decoupled, independently designed control systems cannot capture fully the dynamics of the extended system due to only partial consideration of the interconnections among subsystems. Therefore, the proposed decentralized overlapping controller (called overlapping LQR or OLQR when presenting the results) is expected to be sub-optimal compared to the centralized LQR control approach.

4.4.2 Overlapping control scheme using Linear Matrix Inequalities (LMIs)

Rewriting the quadratic criterion (4.12) yields

$$\begin{aligned} J &= \frac{1}{2}\sum_{k_c=0}^{\infty}\left[\mathbf{x}^T(k_c)\mathbf{Q}\mathbf{x}(k_c)+\mathbf{u}^T(k_c)\mathbf{R}\mathbf{u}(k_c)\right] \\ &= \frac{1}{2}\sum_{k_c=0}^{\infty}\left[(\mathbf{Q}^{1/2}\mathbf{x}(k_c))^T\mathbf{Q}^{1/2}\mathbf{x}(k_c)+(\mathbf{R}^{1/2}\mathbf{u}(k_c))^T\mathbf{R}^{1/2}\mathbf{u}(k_c)\right] \end{aligned} \tag{4.26}$$

where, as already stated above, $\mathbf{Q}\in\mathbb{R}^{3n\times 3n}$ is for IBC a diagonal positive semidefinite matrix and $\mathbf{R}\in\mathbb{R}^{n\times n}$ is a diagonal positive definite matrix. Let

$$\mathbf{C} = \begin{bmatrix} \left(\frac{1}{2}\mathbf{Q}\right)^{1/2} \\ 0_{n\times 3n} \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 0_{3n\times n} \\ \left(\frac{1}{2}\mathbf{R}\right)^{1/2} \end{bmatrix} \tag{4.27}$$

and $\mathbf{Z}(k_c) = \mathbf{C}\mathbf{x}(k_c) + \mathbf{D}\mathbf{u}(k_c)$. Using the above, the objective function (4.26) becomes

$$J = \sum_{k=0}^{\infty} \mathbf{Z}^T(k_c)\mathbf{Z}(k_c). \tag{4.28}$$

Assuming the use of a linear state-feedback control law

$$\mathbf{u}(k_c) = \mathbf{K}\mathbf{x}(k_c) \tag{4.29}$$

leads to $\mathbf{Z}(k_c) = (\mathbf{C} + \mathbf{D}\mathbf{K})\mathbf{x}(k_c)$, and (4.28) becomes

$$J = \sum_{k=0}^{\infty} \left[\mathbf{x}^T(k_c)(\mathbf{C} + \mathbf{D}\mathbf{K})^T(\mathbf{C} + \mathbf{D}\mathbf{K})\mathbf{x}(k_c) \right]. \tag{4.30}$$

Using the trace operator $\mathbf{Tr}(.)$, (4.30) becomes

$$J = \sum_{k=0}^{\infty} \mathbf{Tr}\left((\mathbf{C} + \mathbf{D}\mathbf{K})\mathbf{x}(k_c)\mathbf{x}^T(k_c)(\mathbf{C} + \mathbf{D}\mathbf{K})^T \right) = \mathbf{Tr}\left((\mathbf{C} + \mathbf{D}\mathbf{K})\mathbf{P}(\mathbf{C} + \mathbf{D}\mathbf{K})^T \right) \tag{4.31}$$

where $\mathbf{P} = \sum_{k=0}^{\infty} \mathbf{x}(k_c)\mathbf{x}^T(k_c)$ is a symmetric positive definite matrix satisfying the following Lyapunov equation (Datta, 2004) for the linear state-space model (4.11) after using the feedback law (4.29)

$$\mathbf{P} - (\mathbf{A} + \mathbf{B}\mathbf{K})\mathbf{P}(\mathbf{A} + \mathbf{B}\mathbf{K})^T - \mathbf{x}_0\mathbf{x}_0^T = \mathbf{0} \tag{4.32}$$

with $\mathbf{x}_0 = \mathbf{x}(0)$. Then, defining the parameter $\gamma \in (0,1)$, the following is true

$$\mathbf{P} - (\mathbf{A} + \mathbf{B}\mathbf{K})\mathbf{P}(\mathbf{A} + \mathbf{B}\mathbf{K})^T - \gamma^2\mathbf{x}_0\mathbf{x}_0^T > \mathbf{0}. \tag{4.33}$$

Considering a matrix $\mathbf{F} \in \mathbb{R}^{n\times 3n}$ given by $\mathbf{F} = \mathbf{K}\mathbf{P}$, the objective function (4.31) becomes

$$J = \mathbf{Tr}\left((\mathbf{C} + \mathbf{D}\mathbf{F}\mathbf{P}^{-1})\mathbf{P}(\mathbf{C} + \mathbf{D}\mathbf{F}\mathbf{P}^{-1})^T \right) = \mathbf{Tr}\left((\mathbf{C}\mathbf{P} + \mathbf{D}\mathbf{F})\mathbf{P}^{-1}(\mathbf{C}\mathbf{P} + \mathbf{D}\mathbf{F})^T \right) \tag{4.34}$$

and (4.33) becomes,

$$\mathbf{P} - (\mathbf{A}\mathbf{P} + \mathbf{B}\mathbf{F})\mathbf{P}^{-1}(\mathbf{A}\mathbf{P} + \mathbf{B}\mathbf{F})^T - \gamma^2\mathbf{x}_0\mathbf{x}_0^T > \mathbf{0}. \tag{4.35}$$

Using the Schur complement lemma (Van Antwerp and Braatz, 2000) for an auxiliary variable $\mathbf{W} > (\mathbf{C}\mathbf{P} + \mathbf{D}\mathbf{F})\mathbf{P}^{-1}(\mathbf{C}\mathbf{P} + \mathbf{D}\mathbf{F})^T$ and for (4.35), the following optimization problem can be formulated

$$
\begin{aligned}
&\min \mathbf{Tr}(\mathbf{W}) \\
&\text{s.t.} \\
&\begin{bmatrix} \mathbf{W} & \mathbf{CP}+\mathbf{DF} \\ (\mathbf{CP}+\mathbf{DF})^T & \mathbf{P} \end{bmatrix} > \mathbf{0} \\
&\begin{bmatrix} \mathbf{P} & \mathbf{AP}+\mathbf{BF} & \gamma \mathbf{x}_0 \\ (\mathbf{AP}+\mathbf{BF})^T & \mathbf{P} & 0 \\ \gamma \mathbf{x}_0^T & 0 & \mathbf{I} \end{bmatrix} > \mathbf{0}
\end{aligned}
\tag{4.36}
$$

The discrete-time LQR problem has now been transformed into an LMI problem that can be solved using infeasible path-following algorithms for solving standard semidefinite programs (Toh et al., 1999) in order to get the optimal matrices $\mathbf{F}$ and $\mathbf{P}$. The difference compared to the discrete-time LQR problem is that the structure of the gain matrix $\mathbf{K}$ can be pre-specified, before solving the LMI problem, by specifying the structure of the matrices $\mathbf{F}$ and $\mathbf{P}$. Thus, the design procedure remains a centralized one, but it delivers an overlapping decentralized control scheme for IBC of lane-free automated vehicle traffic. The overlapping control structure used can be selected appropriately for the control system and the physical system at hand, something that may call for experimentation with different possible overlapping control structures.

As an example, in the case depicted in Fig. 4.3, where two subsystems with a single overlapping area of two sections (sections 5 and 6) are inter-connected, the structure of the gain matrix $\mathbf{K}$ needed is the following:

$$
\mathbf{K}_{10\times 30} = \begin{bmatrix} \mathbf{K}^1_{4\times 18} & \mathbf{0} \\ \mathbf{K}^{1,2}_{2\times 30} & \\ \mathbf{0} & \mathbf{K}^2_{4\times 18} \end{bmatrix}. \tag{4.37}
$$

Consequently, the matrices $\mathbf{F}$ and $\mathbf{P}$ can be selected to have the following structure:

$$
\mathbf{F}_{10\times 30} = \begin{bmatrix} \mathbf{F}^1_{4\times 18} & \mathbf{0} \\ \mathbf{F}^{1,2}_{2\times 30} & \\ \mathbf{0} & \mathbf{F}^2_{4\times 18} \end{bmatrix}, \ \mathbf{P} = \begin{bmatrix} P_1 & 0 & \dots & \dots & 0 \\ 0 & P_2 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & P_{30} \end{bmatrix}. \tag{4.38}
$$

The LMI approach results in an overlapping decentralized control scheme, i.e. some of the gain matrix elements have been set equal to zero (see e.g. (4.37)) and, as a result, the optimal value of the objective function in (4.36) will be at best equal to the optimal value of the objective function (4.12) for the original centralized LQR problem. Nevertheless, the whole system dynamics have been considered in the design procedure, as opposed to the case of the overlapping control design procedure presented in Section 0. Therefore, this strategy may lead to better performance in case of complicated scenarios with stronger interconnections.

4.5 Simulation investigations

4.5.1 Simulation set-up

The performance of the proposed overlapping control schemes is investigated using the large bi-directional highway stretch depicted in Fig. 4.4. The considered highway stretch has a length of 20 km and is subdivided in 40 sections of 0.5 km each. There are two aspects to be considered when selecting the length of a section. First, sections are used for modelling within the discrete CTM, where the considered length of 0.5 km is appropriate, see (Kontorinaki et al., 2017). Second, the sections are used for IBC, to determine the range of each sharing factor. For the latter, too long cells would reduce the flexibility of IBC; while too short cells would call for frequent vehicle maneuvering due to frequent road width changing. Given these, the employed cell length of 0.5 km is deemed appropriate, although other cell lengths in this order may also be employed, depending on the road infrastructure and the typical traffic conditions.

The considered highway contains eight on-ramps and off-ramps (Fig. 4.4) in each direction, which create realistically inhomogeneous traffic conditions in space and time. The modelling time-step, T, is set to 10 s, and the considered time horizon is 1 h. While a linearization of CTM was used for the controller design, the full nonlinear extended CTM is used to represent the emulated ground truth in this section.

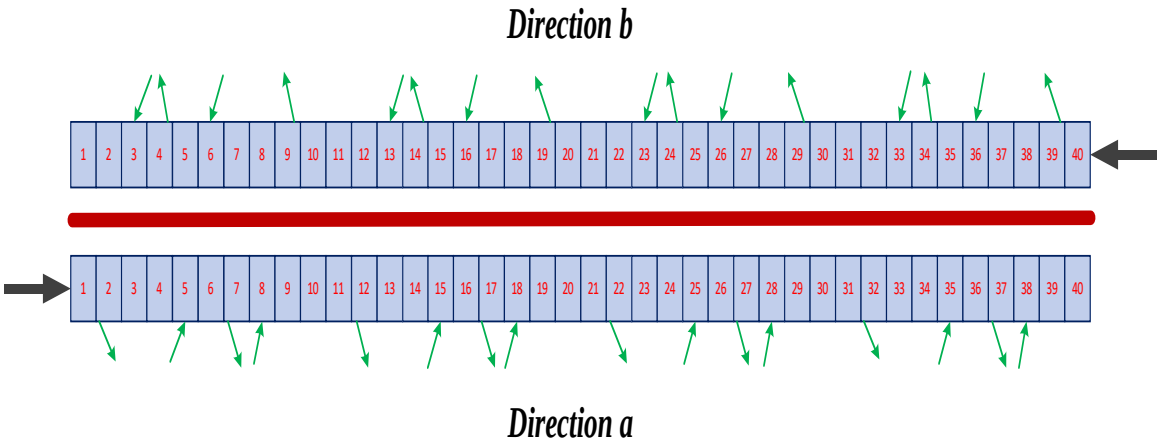


Fig. 4.4: The considered highway stretch.

The model parameters used in the simulation are $v_f = 100\,\text{km/h}$ and $w_s = 12\,\text{km/h}$; while the total cross-road capacity to be shared among the two directions is $q_{cap} = 12{,}000\,\text{veh/h}$. Based on the above and the triangular FD, $\rho_{cr} = 120$ veh/km and $\rho_{\max} = 1120$ veh/km can be calculated. The parameter values used to enable capacity drop are $\lambda_r = 0.7$ and $\lambda_d = 0.4$. The exit rates for all the off-ramps are set equal to 0.1.

The mainstream and on-ramp demand flows per direction are presented in Fig. 4.5 including two demand scenarios. The uncongested scenario causes congestion in the no-control case, but, which, by activation of the IBC, can be dissolved entirely. In the congested scenario, the no-control congestion can be mitigated by application of IBC, but cannot be utterly dissolved due to stronger overlapping between the two mainstream demands. In such cases, smart and flexible

capacity sharing via IBC is still efficient, but not sufficient to utterly suppress traffic congestion. The on-ramp demands are assumed equal for all on-ramps of both directions and scenarios, as depicted in Fig. 4.5. It should be noted that interconnections in traffic flow are due to two main effects (Ferrara et al., 2018): (i) In free-flowing traffic, upstream traffic states (flow) influence downstream states; (ii) In congested traffic, downstream traffic states (mounting congestion tail) influence upstream states. Thus, in our considered scenarios, both effects are included to test the performance of the overlapping strategies under various levels of both mentioned effects.

As performance metrics for comparison of different approaches, two criteria are considered: (i) the Total Time Spent (TTS) by all vehicles in the highway; and (ii) the Total Delay (TD), which is the total time spent by all vehicles in excess of what would be needed if they would drive at free speed. Both metrics are expressed in (veh·h) and may be readily computed based on the macroscopic variables of the simulation model, see (Papageorgiou, et al. 2003).

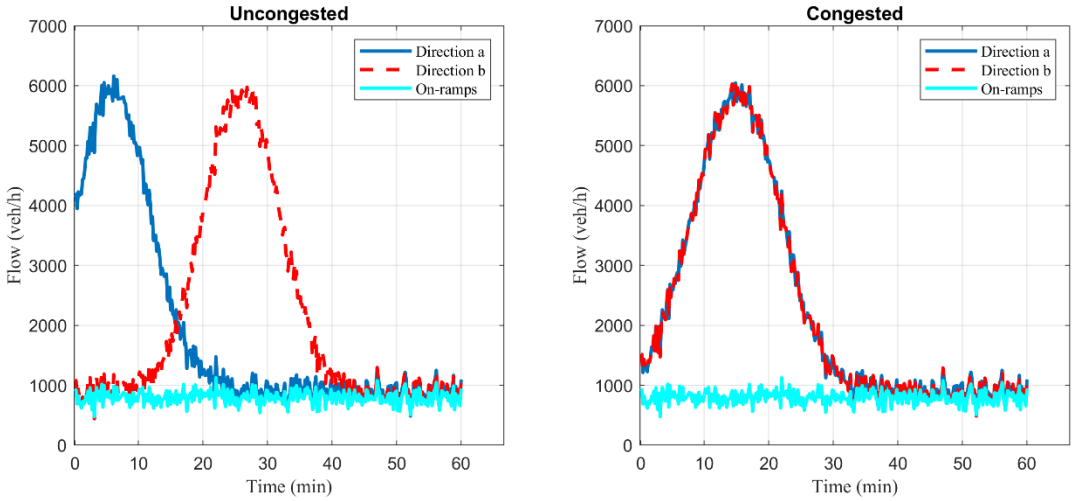


Fig. 4.5: Demand flows per direction and on-ramps for the uncongested scenario (left) and the congested scenario (right).

4.5.2 Uncongested scenario

4.5.2.1 No-control case

When no IBC is applied, the total width of the highway stretch is equally shared by the two directions, i.e. the sharing factors ε_i are constant and equal to 0.5 for all sections. Feeding the demand profiles, presented in Fig. 4.5 for the uncongested scenario, to the nonlinear CTM model with $\varepsilon_i = 0.5$, the simulation results for the no-control case are obtained. Figure 4.6 displays the corresponding spatio-temporal evolution of the relative density defined in (4.9). According to the definition, relative density values lower than 1 refer to uncongested traffic; while values higher than 1 refer to congested traffic; when the relative density equals 1, and the downstream section is uncongested, the corresponding section is at capacity flow.

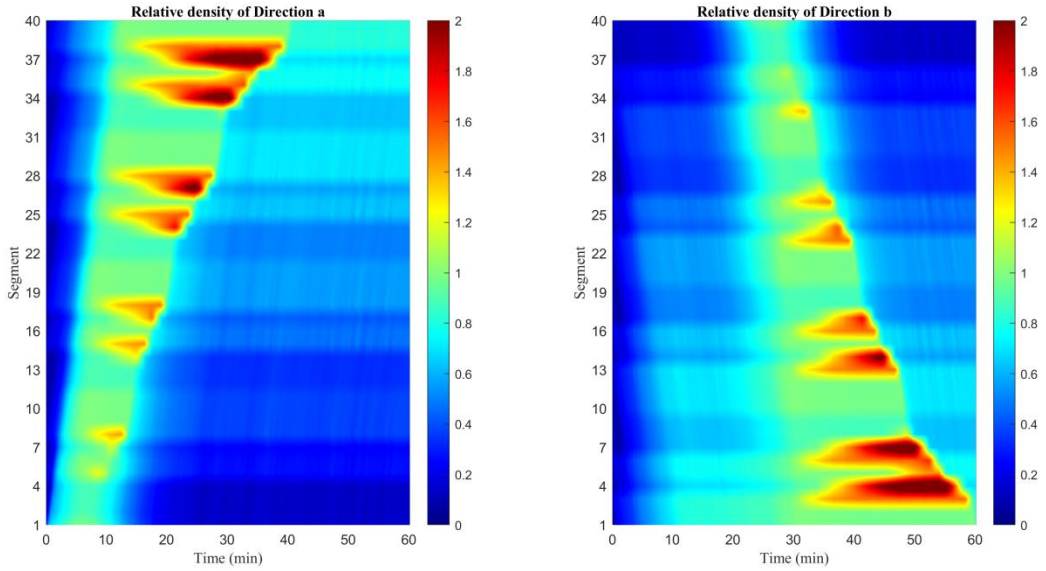


Fig. 4.6: Uncongested scenario: Relative density for the two directions in the no-control case.

Figure 4.6 shows that several congestions are created on both directions due to the high mainstream demand, in combination with the ramp inflows. As can be seen, the congestion phenomena appear around the on-ramp merge segments and, as moving downstream on each direction, the phenomena last more and cover longer areas of the highway stretch. The TTS as well as the TD performance metrics for this case are reported in Table 4.1.

4.5.2.2 Control case

To compare the performance of the proposed overlapping control schemes with previously developed methods, two centralized control schemes are considered first. The first is the QP approach (Malekzadeh et al., 2021a) and the second is the LQR approach (Malekzadeh et al., 2021b), which were outlined in Section 3.

The nominal values for linearization that is necessary for the LQR case are $q_0^a\big|_N = q_{40}^b\big|_N = 5000$ veh/h, $r_i^a\big|_N = 1000$ veh/h, $i = 5,8,15,18,25,28,35,38$, $r_i^b\big|_N = 1000$ veh/h, $i = 3,6,13,16,23,26,33,36$, $\tilde{\rho}_i^a\big|_N = \tilde{\rho}_i^b\big|_N = 1$ and $\varepsilon_i\big|_N = 0.5$, $i = 1,2,\ldots,40$. The control time-step, T_c, is set to 60 s, hence $M = 6$. The weighing matrices used in the objective function are selected to be $\mathbf{Q} = \mathrm{diag}(\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_n)$ where $\mathbf{S} = [\mathbf{I}_2, \mathbf{0}_{2\times 1}; \mathbf{0}_{1\times 3}]$, $\mathbf{R} = 10^{-1}\mathbf{I}_{n\times n}$. The upper and lower bounds for the sharing factors, used to avoid utter blocking of any of the two directions, are equal for all sections $i = 1,2,\ldots,40$ and are given the values $\varepsilon_{i,\min} = 0.16$ and $\varepsilon_{i,\max} = 0.84$. These values are used to truncate the LQR outcome.

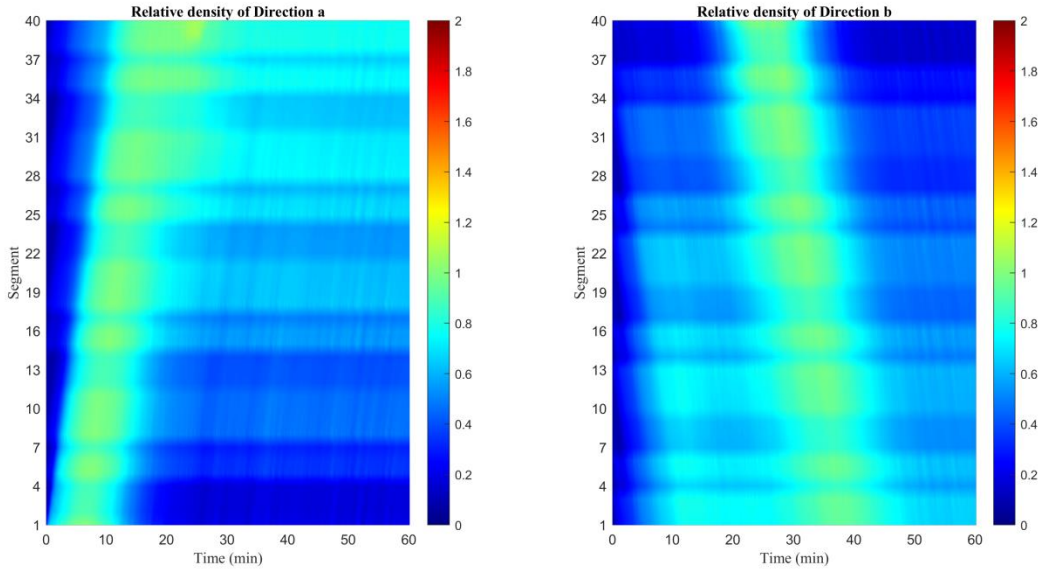
The simulation results applying QP and LQR for the considered highway are depicted in Fig. 4.7. As can be seen, congestion is utterly avoided by employing open-loop QP, and the results obtained via LQR are close to that optimal solution. Table 4.1 figures demonstrate the effectiveness of these strategies in terms of the TTS and TD performance metrics. Note that the TD index is virtually nullified, i.e. any delays caused by traffic congestion are almost fully

rejected by IBC. It is worth mentioning that in the QP problem, the smoothed versions of the demand profiles are considered as the predicted demand required by the method; while the simulations delivering the reported performance values are performed using CTM with noisy demands, as presented in Fig. 4.5. This is why the achieved TD improvement using the QP approach is slightly lower than 100% in this uncongested scenario.

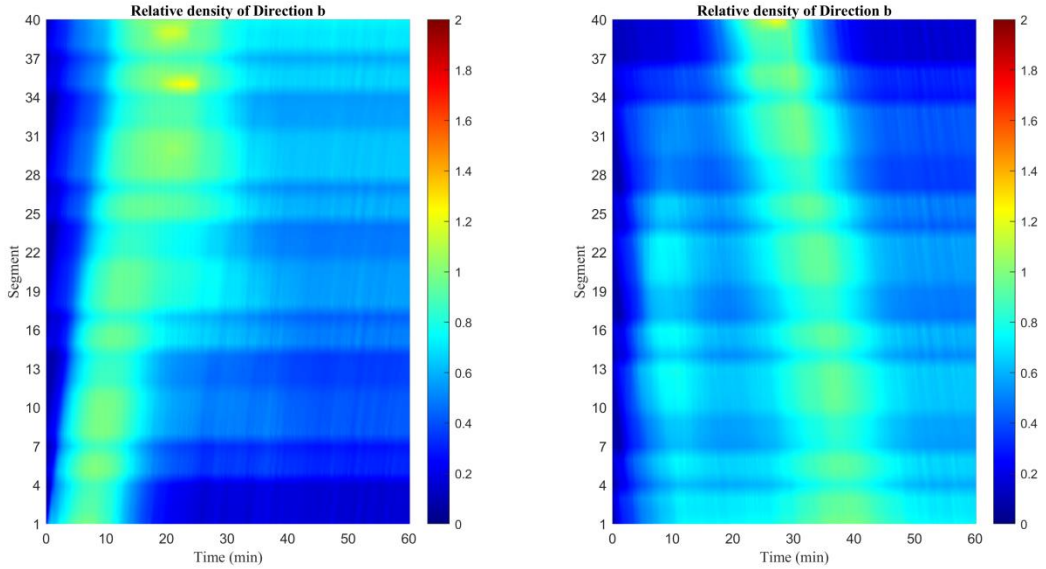
Table 4.1: Uncongested scenario: TTS and TD values with related improvement (%) over the no-control case for different scenarios.

	TTS (veh·h)		**TD (veh·h)**	
No-control	1553		110.8	
Centralized	**QP**	**LQR**	**QP**	**LQR**
	1446 (6.8%)	1447 (6.8%)	3.6 (97%)	4.7 (96%)

The IBC action, while achieving the remarkable reported results, can be better understood in Fig. 4.8. The figure depicts flow-versus-time windows for nine selected sections, where time extends over the time-horizon of the scenario. The height of each window equals the total road capacity that must be shared among the two directions at each section. The lower displayed curve (blue) is the projected demand in direction a for each section; and the upper curve (red), starting from the upper window edge, is the projected demand in direction b. Projected demand is the flow expected in those sections in case of sufficiently high capacities to accommodate all the demands. The fact that these two curves do not intersect in the presented (and in any other) section(s), indicates that flexible capacity sharing may be applied, so as to avoid any congestion forming in either direction. Furthermore, the black line at 6000 veh/h indicates the capacity of each direction in the no-control case. Evidently, when the blue or red curve exceeds the black line, it means that the flow is higher than the capacity and a congestion is expected in the corresponding section, direction and time. The green line indicates a trajectory of capacity sharing (resulting from the corresponding sharing factor), here produced by the centralized LQR approach, which is accommodated among the blue and red curves, without ever crossing either of them. Such a sharing factor trajectory may be enabled via IBC to avoid (in this demand scenario) the otherwise unavoidable congestion. However, this analysis also reveals that it is possible, for other demand scenarios, that IBC cannot utterly avoid the congestion, like in the next scenario, where the blue and red lines intersect; hence there is no possibility for the capacity sharing curve to separate them. In such cases, IBC will be mitigating the congestion at an extent that depends on the specific demand scenario.



(a)



(b)

Fig. 4.7: Uncongested scenario: Relative density for the two directions in the centralized control cases: (a) QP; (b) LQR

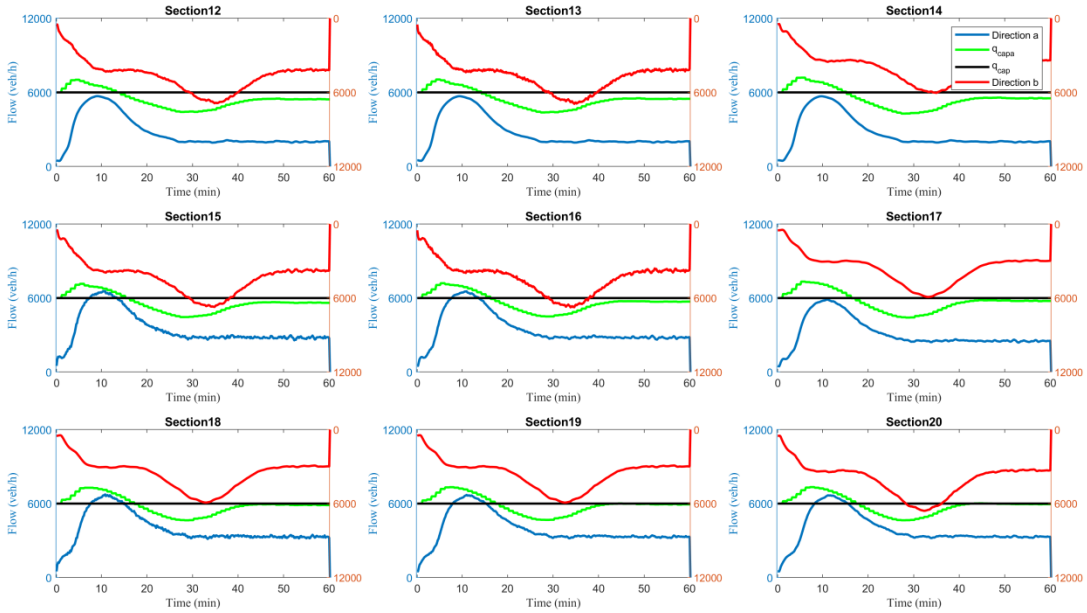


Fig. 4.8: Demand-supply analysis for the uncongested scenario applying centralized LQR.

In the following, the overlapping control approaches will be applied. The main issue that one must consider before applying these approaches is the arrangement of the subsystems. There are many different possibilities to choose from; the arrangement may depend on the position of the on-ramps and off-ramps, it may be symmetric or not, and may include different numbers of overlapping sections. In order to investigate some combinations, a number of different schemes with three, four or five subsystems are considered, while for each scheme the number of overlapping sections ranges from one to three. First, the highway stretch is partitioned into $P = 2n_S - 1$ partitions, where n_S is the number of subsystems considered. The partitions with an even index form the overlapping areas, while the partitions with an odd index form the non-overlapping areas. The employed overlapping control methods enable a variety of possible overlapping structures, and this calls for empirical investigations to test the performance of different such structures under realistic traffic conditions. Tables 2, 3 and 4 present the different schemes considered and the corresponding partitions. As an example, Scheme 1-2 has 5 partitions that form 3 subsystems and 2 overlapping areas composed of 2 sections each. The first subsystem includes section 1 to 14; the second subsystem includes sections 13 to 28; and the third subsystem includes sections 27 to 40.

Table 4.2. Distribution of sections per partition for schemes with three subsystems.

	Partitions				
	1	2	3	4	5
Scheme 1-1	1-13	14	15-26	27	28-40
Scheme 1-2	1-12	13-14	15-26	27-28	29-40
Scheme 1-3	1-11	12-14	15-25	26-28	29-40

Table 4.3. Distribution of sections per partition for schemes with four subsystems.

	Partitions						
	1	2	3	4	5	6	7
Scheme 2-1	1-9	10	11-19	20	21-29	30	31-40
Scheme 2-2	1-9	10-11	12-19	20-21	22-29	30-31	32-40
Scheme 2-3	1-8	9-11	12-19	20-22	23-29	30-32	33-40

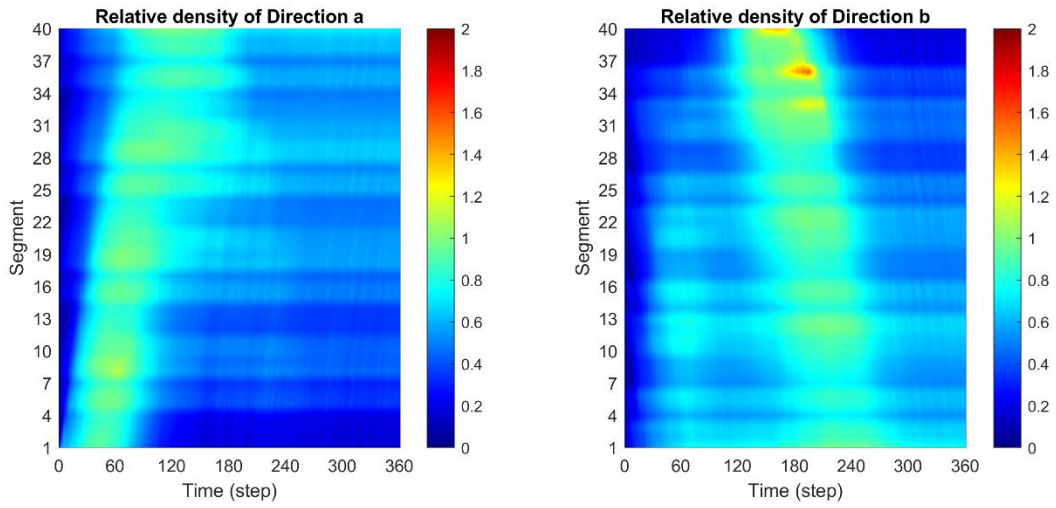
Table 4.4. Distribution of sections per partition for schemes with five subsystems.

	Partitions								
	1	2	3	4	5	6	7	8	9
Scheme 3-1	1-7	8	9-15	16	17-23	24	25-31	32	33-40
Scheme 3-2	1-7	8-9	10-15	16-17	18-23	24-25	26-31	32-33	34-40
Scheme 3-3	1-6	7-9	10-14	15-17	18-22	23-25	26-30	31-33	34-40

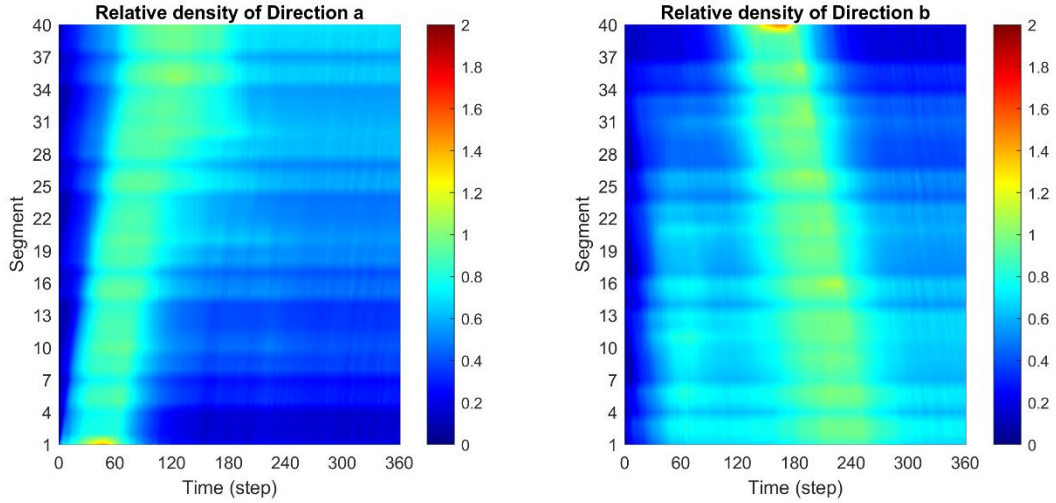
The TTS and TD values related to the aforementioned schemes are presented in Table 5 for the uncongested scenario. Both approaches perform very well under all examined partitions, with the LMI approach achieving TTS and TD values closer to the centralized approaches. The results indicate that the control performance in this demand scenario is not related to the number of subsystems considered. In contrast, for all subsystem cases, the control performance is getting better when the overlapping areas include more than one section, something that underlines the necessity of considering the interconnections among subsystems for more pertinent control. As an example, the results related to scheme (2-2) for the decentralized case are depicted in Fig. 4.9. When compared with the results presented in Fig. 4.7, it can be observed that the relative density heat plots are similar as also expected from the similar efficiency improvement.

Table 4.5: Uncongested scenario: TTS and TD values with related improvement (%) over the no-control case for different scenarios.

	TTS (veh·h)		TD (veh·h)	
Decentralized	**LMI**	**OLQR**	**LMI**	**OLQR**
Scheme 1-1	1450 (6.7%)	1461 (5.9%)	7.7 (93.1%)	18.4 (83.4%)
Scheme 1-2	1448 (6.8%)	1457 (6.2%)	6.0 (94.6%)	15.3 (86.2%)
Scheme 1-3	1448 (6.8%)	1453 (6.5%)	6.3 (94.3%)	11.2 (89.9%)
Scheme 2-1	1451 (6.6%)	1452 (6.5%)	8.8 (92.1%)	9.4 (91.5%)
Scheme 2-2	1449 (6.7%)	1450 (6.7%)	6.6 (94.0%)	8.0 (92.8%)
Scheme 2-3	1448 (6.8%)	1450 (6.7%)	6.3 (94.3%)	8.0 (92.8%)
Scheme 3-1	1457 (6.2%)	1460 (6.0%)	14.5 (86.9%)	18.1 (83.7%)
Scheme 3-2	1453 (6.5%)	1457 (6.2%)	10.9 (90.2%)	14.8 (86.6%)
Scheme 3-3	1450 (6.7%)	1453 (6.5%)	7.4 (93.3%)	11.3 (89.8%)



(a)



(b)

Figure 4.9:Uncongested scenario: Relative density for the two directions in the decentralized control (scheme 2-2) cases: (a) LMI; (b) OLQR

4.5.3 Congested scenario

4.5.3.1 No-control case

Using the demand profiles, presented in Fig. 4.5 for the congested scenario, to feed the nonlinear CTM model with $\varepsilon_i = 0.5$, the simulation results for the no-control case are obtained. Essentially, all external flows are similar to those in the uncongested scenario, with the notable difference that the two mainstream demand profiles have been shifted so that they completely overlap with each other in time. This gives rise to strong flows on both directions at similar

times, which implies limited opportunities for IBC and, indeed, capacity problems even in presence of IBC.

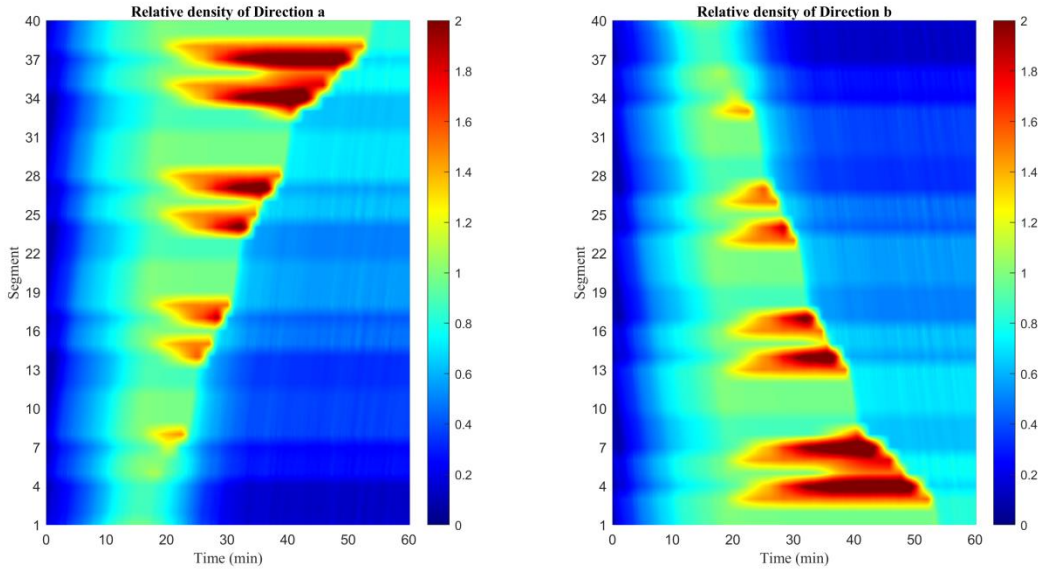
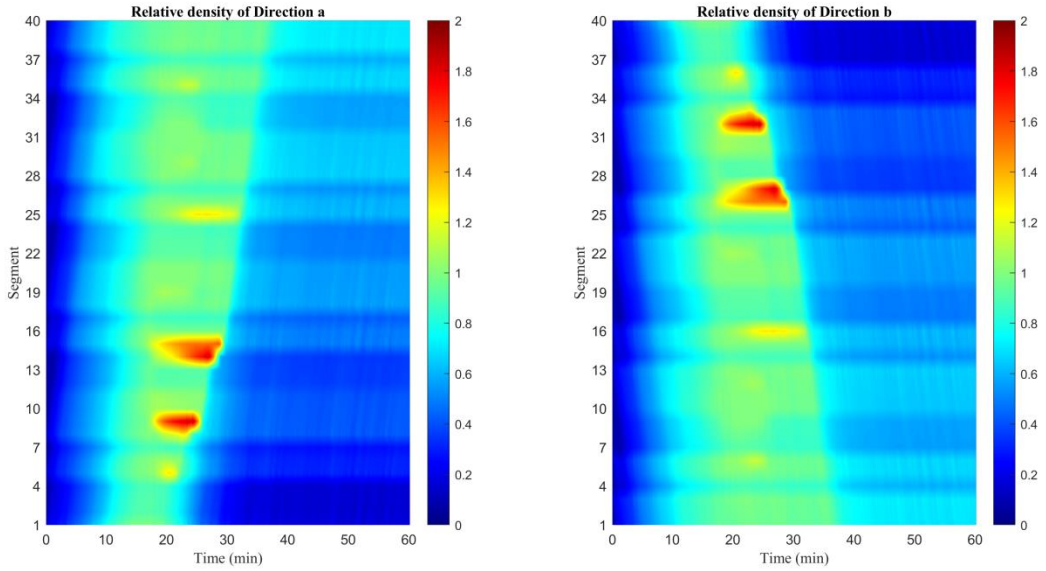


Figure 4.10: Congested scenario: Relative density for the two directions in the no-control case.

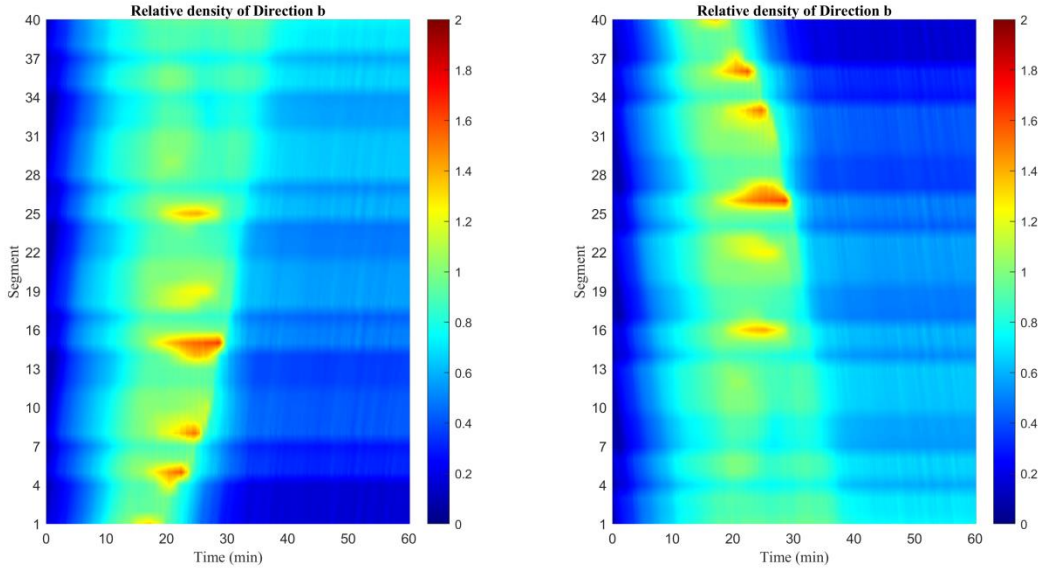
Figure 4.10 shows that congestion areas are created on both directions due to the increased mainstream demand, in combination with the ramp inflows. As in the case of the uncongested scenario considered in Section 4.5.2.1, the congestion phenomena appear around the on-ramp merge segments, and, as moving further downstream on each direction, the phenomena last longer and cover longer areas of the highway stretch. The main difference here is that the phenomena are shifted in time due to the corresponding shift that has been applied to the demand profiles. The Total Time Spent (TTS) in the system as well as the Total Delay (TD) for this case are reported in Table 6.

4.5.3.2 Control case

The simulation results applying the centralized open-loop QP and LQR strategies for the considered highway are depicted in Fig. 4.11. As can be seen, congestion cannot be avoided completely with IBC, but it is strongly mitigated via flexible internal boundary moving by the control strategies. Table 4.6 figures demonstrate the effectiveness of these strategies with respect to TTS and TD values. The results obtained via the centralized LQR approach are, also for this scenario, very close to the open-loop optimal solution. Notice the remarkable reduction of the TD metric, despite the challenging demand scenario with two simultaneous strong mainstream demands.



(a)



(b)

Fig. 4.11*:* Congested scenario: Relative density for the two directions in the centralized control case: (a) QP; (b) LQR

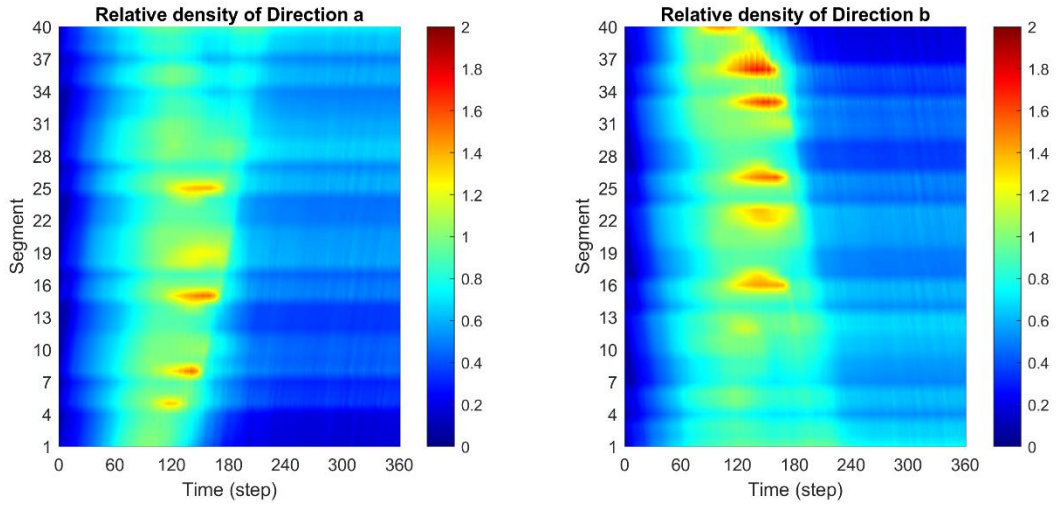
The overlapping control approaches are applied next for all the different schemes presented in Tables 4.2, 4.3 and 4.4. Table 4.6 includes the corresponding TTS and TD values for the congested scenario. Both approaches perform very well, with the LMI approach achieving TTS and TD values closer to the centralized approaches. In this scenario, more performance variations are observed, particularly for the OLQR approach, but no consistent relation of the

control performance to the number of subsystems considered or to the number of sections in the overlapping areas can be deduced.

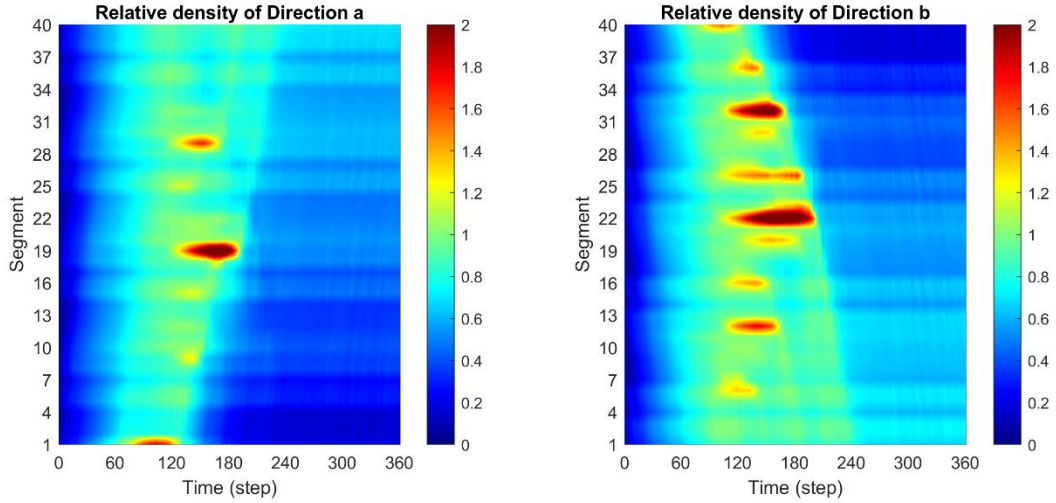
Table 4.6: Congested scenario: TTS and TD values with related improvement (%) over the no-control case for different scenarios.

	TTS (veh·h)		**TD (veh·h)**	
No-control	1681		173.2	
Centralized	**QP**	**LQR**	**QP**	**LQR**
	1536 (8.7%)	1537(8.6%)	27.9 (83.9%)	29.4 (83.0%)
Decentralized	**LMI**	**OLQR**	**LMI**	**OLQR**
Scheme 1-1	1546 (8.1%)	1565 (6.9%)	37.8 (78.2%)	56.8 (67.2%)
Scheme 1-2	1542 (8.3%)	1579 (6.1%)	34.1 (80.3%)	70.9 (59.1%)
Scheme 1-3	1547 (8.0%)	1621 (3.6%)	39.0 (77.5%)	113.2 (34.6%)
Scheme 2-1	1540 (8.4%)	1587 (5.6%)	32.2 (81.4%)	78.7 (54.6%)
Scheme 2-2	1540 (8.4%)	1568 (6.8%)	32.5 (81.2%)	59.6 (65.6%)
Scheme 2-3	1541 (8.4%)	1569 (6.7%)	33.3 (80.8%)	61.3 (64.6%)
Scheme 3-1	1542 (8.3%)	1606 (4.5%)	33.9 (80.4%)	97.8 (43.5%)
Scheme 3-2	1542 (8.3%)	1614 (4.0%)	34.2 (80.3%)	107.8 (37.8%)
Scheme 3-3	1543 (8.2%)	1614 (4.0%)	34.7 (80.0%)	105.7 (39.0%)

As an example, the results related to scheme (2-2) for the decentralized control are depicted in Fig. 4.12. When compared with the results presented in Fig. 4.11, it can be observed that the relative density heat plots are quite similar, with the congestion phenomena appearing at nearby locations.



(a)



(b)

Figure 4.12: Congested scenario: Relative density for the two directions in the decentralized control (scheme 2-2) cases: (a) LMI; (b) OLQR

4.6 Conclusions

The concept of internal boundary control, introduced by Malekzadeh et al. (2021a), has been revisited in this study by use of different feedback control approaches with an overlapping structure for a sizable highway system. According to the IBC concept, the total road width and capacity are shared in each section in real-time among the two directions of the road in response to the prevailing traffic conditions. As in previous studies, the well-known CTM, extended to include the capacity drop and appropriately adjusted to introduce the effect of the sharing factors, has been utilized (in linearized form) for the development of the overlapping decentralized regulators for the IBC problem, as well as (in full nonlinear form) for simulation testing and comparison. Two overlapping decentralized control design methods were adopted:

(i) a contractible controller with state and input inclusion developed for an extended system (İftar and Özgüner, 1990) and

(ii)a linear matrix inequalities approach (Zečević and Šiljak, 2005).

Approach (i) may be designed separately for each considered subsystem, which is favorable for scalability and extendibility in cases of long highways. In contrast, approach (ii) calls for centralized design based on a pre-selected control structure. Both overlapping regulator schemes reduce the communication, monitoring and maintainance requirements and increase control system reliability in cases of device failures, which is beneficial in case of long highways.

Simulation investigations first demonstrate the adequacy of previously developed centralized IBC approaches in sizable highway scenarios. Regarding the novel overlapping decentralized control schemes, it was found that, for the case of the uncongested demand scenario (where congestion is avoidable via IBC), the overlapping control schemes are as efficient as an open-loop optimal control solution (with perfect model knowledge and demand prediction) developed for the same problem by Malekzadeh et al. (2021a) using a convex QP problem formulation; and as efficient as the centralized LQR. On the other hand, for the more challenging case of the congested scenario (where congestion may be mitigated, but cannot be fully avoided via IBC), the LMI-based approach (ii) performs better than the contractible controller (i), as it takes into account the whole system in the design phase.

5 Microscopic Simulation of IBC in Lane-free Automated Vehicle Traffic

5.1 Abstract

The recently introduced *TrafficFluid* concept proposes that automated vehicles drive lane-free, thus enabling capacity sharing between the two opposite road directions via real-time Internal Boundary Control (IBC). This novel traffic control measure was demonstrated, using macroscopic traffic flow models, to deliver unprecedented improvements of traffic flow efficiency. The present study completes and validates the IBC concept in a much more realistic way via microscopic simulation and active internal boundary moving, using the SUMO-based *TrafficFluid-Sim* simulation tool. To effectuate IBC, a Linear Quadratic Regulator (LQR), which is a feedback control scheme, is employed. In addition, to enhance the performance of the LQR controller, a feedforward term, accounting for external disturbances, i.e. entering flow and on-ramp flows, is also designed, leading to an augmented LQR-FF control scheme. The LQR and LQR-FF controllers are tested and compared in the created realistic environment, demonstrating how IBC may operate in practice to combat traffic congestion on highways.

Keyword- Internal Boundary control; LQR; Feedforward; TrafficFluid-Sim; Microscopic simulation.

5.2 Introduction

Despite the development and partial deployment of traffic management systems (Papageorgiou et al., 2003; Kurzhanskiy and Varaiya, 2010), traffic congestion on freeway and urban road networks continues to be a major problem for society. Existing and emerging vehicle automation and communication systems enable innovative traffic management actions that may further mitigate traffic congestion (Diakaki et al., 2015; Papamichail et al., 2019). Further advances facilitate a number of prototype vehicles to keep track of their surroundings and make reliable and efficient autonomous driving decisions. Furthermore, vehicle communication enables vehicles to communicate with each other and with the infrastructure, which opens the door for a variety of enhanced or additional applications.

Recently, Papageorgiou et al. (2021) put forward the *TrafficFluid* concept, a forward-thinking paradigm for vehicular traffic that is suitable for traffic with a high percentage of vehicles outfitted with high-level automation and communication systems. A distinguishing feature of TrafficFluid is lane-free traffic of connected and automated vehicles (CAV), which allows for an improved exploitation of the road infrastructure. In this context, it becomes possible to introduce Internal Boundary Control (IBC), a novel and highly promising control measure, which relies on the fact that, in lane-free traffic, the traffic flow and capacity exhibit incremental changes in accordance with corresponding incremental changes of the road width, as demonstrated in (Papamichail et al., 2023). Thus, by virtually shifting the internal boundary that separates the two directions of traffic on highways or arterials, the total crossroad capacity can be flexibly shared between them in real-time, depending on the prevailing traffic conditions. The boundary-shifting

information is sent to CAV so that they can drive in accordance with the changed internal road boundary. Real-time IBC could potentially be widely used for lane-free CAV traffic on the numerous arterial or highway infrastructures that feature different demand levels in each direction, in time or in space, so as to drastically reduce or even eliminate congestion.

Malekzadeh et al. (2021a) studied the features of IBC, illustrating the underlying mechanism and improvement potential through formulating and resolving an open-loop optimal control problem in form of a convex Quadratic Programming (QP) problem, aiming to minimize the total time spent by all vehicles in both directions of the highway. This strategy can be cast in a Model Predictive Control (MPC) frame, along with a demand prediction module for real-time operation. Though, it is preferable to use simpler real-time methods which achieve similar levels of efficacy, but do not require online traffic flow modelling and demand prediction. To this end, Malekzadeh et al. (2021b) developed a feedback-based Linear-Quadratic Regulator (LQR) for IBC, which strives to keep the relative density of the two traffic directions near their nominal value (close to capacity). This closed-loop LQR approach was demonstrated, using macroscopic simulation, to be almost as efficient as the open-loop optimal control solution, without requiring an accurate model or demand forecasts (Malekzadeh et al., 2021b; Malekzadeh et al., 2023).

As already practiced in other applications (Marinaki and Papageorgiou, 2005; Papageorgiou, et al., 2015), LQR performance can be enhanced in the presence of external disturbances by adding a feedforward term to the basic feedback controller. In the IBC context, such an addition is considered in this paper to mitigate the effect of entering and on-ramp flows on the controller performance. The stationary LQR feedback and feedforward gains are designed offline; then, in real-time operation, the control input is computed as the sum of: (i) the state measurements multiplied with the feedback gain and (ii) the inflow measurements multiplied with the feedforward gain. Following this procedure, LQR-Feedforward (LQR-FF) is employed, tested and compared for IBC in this study.

Up to this point, the mentioned IBC control schemes were applied and tested using macroscopic traffic flow models. Microscopic implementation is a crucial step to investigate, beyond the basic macroscopic concept, the suitability, feasibility, robustness and efficiency of the proposed IBC measure and the related control strategies in realistically emulated vehicle behavior and emerging traffic conditions. The investigation addresses open questions, such as: Can lane-free driving CAV timely and safely evacuate internal road zones assigned to the other traffic direction in real-time? Can CAV efficiently exploit additional road zones assigned to their direction? Can CAV drive safely on a highway where the internal boundary is changing over space? How do flow and capacity of a traffic direction relate to narrowing or widening of the road infrastructure? What is an appropriate space resolution for IBC application, and how should different sharing decisions be implemented in space? What is an appropriate time-step for real-time IBC application? Is the achievable efficiency increase as high as indicated by the afore-mentioned macroscopic investigations? Is the macroscopically derived controller robust to the inevitable subtleties and system “noise” introduced by a realistic microscopic environment?

In this regard, a lane-free extension of the well-known microscopic simulator SUMO (Lopes et al., 2018), called *TrafficFluid-Sim* (Troullinos et al., 2021), has been developed for use in various

TrafficFluid-related investigations. Additional functionalities had to be specifically developed to facilitate the use of changing internal boundaries, some of which are outlined as part of future work by Troullinos et al. (2022). Also, in the microscopic environment, the vehicles must follow a lane-free movement strategy. In this context, the ad-hoc model suggested by Malekzadeh et al. (2022) for a ring-road is modified here to incorporate policies for on- and off-ramps in highway scenarios.

In summary, the primary objectives of this paper are:

- To complete the IBC application scheme and control loop with the detailed internal boundary design and changing in real-time.
- To present a thorough exploration of the microscopic implementation and testing of IBC.
- To design, test and compare LQR-FF versus LQR in the realistic microscopic environment.

The remainder of the paper is organized as follows: Section 5.2 covers some background for understanding the IBC concept and LQR; and presents the novel feedforward term design. Information on TrafficFluid-Sim and its extension to accommodate IBC implementation are given in Section 5.3. Section 5.4 outlines the extended ad-hoc vehicle moving strategy. Section 5.5 presents the simulation results, while Section 5.6 concludes the paper.

5.3 Internal Boundary Control Background and Controller Design

This section provides an overview of the IBC macroscopic modelling, the LQR and the novel LQR-FF control design, which are essential for proper understanding of the IBC application. For more details, see (Malekzadeh et al., 2021a; Malekzadeh et al., 2021b).

5.3.1 Macroscopic Modelling for IBC

Macroscopic IBC modelling was based on the expectation, to be verified in this paper, that lane-free traffic does not entail any significant modifications to the fundamental components of existing macroscopic traffic flow models. An extended version of CTM (Cell Transmission Model), a first-order dynamic traffic flow model with a triangular fundamental diagram (Daganzo, 1994), was considered for control design (Malekzadeh et al., 2021b).

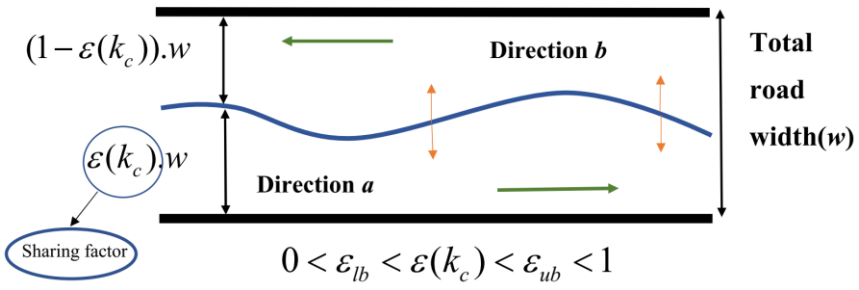


Figure 5.1: Space-time flexible internal road boundary.

Figure 5.1 shows a road stretch with two directions of traffic: direction a from left to right; and direction b from right to left. IBC implies that the internal boundary separating the two directions may be changed flexibly in space and time according to the traffic conditions. For macroscopic modelling, rather than using a space-continuous separation curve as in Figure 5.1, the road stretch is partitioned into n sections, with respective lengths L_i (of some 500 m), and the total road width (both directions) w, which is assumed constant over all sections for simplicity, can be shared in real-time, independently for each section, between the two traffic directions. As a result, each direction is assigned a corresponding road width $w_i^a = \varepsilon_i \cdot w$ and $w_i^b = (1-\varepsilon_i)\cdot w$, where $0 \le \varepsilon_i \le 1$ is the *sharing factor* per section $i = 1, 2, \ldots, n$, to be specified in real-time as a control input by the internal boundary controller. The total (both directions) section capacity q_{cap}, as well as the total critical density ρ_{cr} and the total jam density $\rho_{\max}$, are accordingly shared at each section between the two traffic directions a and b. Based on the derivation presented by Malekzadeh, et al. (2021a), these quantities are given by direction from the following relations

$$\begin{aligned} &q_{i,cap}^a(\varepsilon_i) = \varepsilon_i \cdot q_{cap},\ q_{i,cap}^b(\varepsilon_i) = (1-\varepsilon_i)\cdot q_{cap} \\ &\rho_{i,cr}^a(\varepsilon_i) = \varepsilon_i \cdot \rho_{cr},\ \rho_{i,cr}^b(\varepsilon_i) = (1-\varepsilon_i)\cdot \rho_{cr} \\ &\rho_{i,\max}^a(\varepsilon_i) = \varepsilon_i \cdot \rho_{\max},\ \rho_{i,\max}^b(\varepsilon_i) = (1-\varepsilon_i)\cdot \rho_{\max} \end{aligned} \tag{5.1}$$

While applying IBC, it is not permissible to fully close any of the two directions; hence, the designated road width in any direction must never be narrower than the greatest width of vehicles travelling on the road, leading to stricter constraints on the sharing factors, as follows

$$0 < \varepsilon_{lb} \le \varepsilon_i \le \varepsilon_{ub} < 1 \tag{5.2}$$

where $\varepsilon_{lb} \cdot w$ and $(1-\varepsilon_{ub})\cdot w$ are the minimum admissible widths to be assigned to directions a and b, respectively.

Another restriction to be applied to the sharing factors concerns the time-delay needed to evacuate traffic on the direction that receives a restricted width, compared to the previous control time-step. Obviously, the time-delay in applying the new sharing factor should only take place for the direction of traffic that is being widened in comparison to the previous control interval. In contrast, the restricted direction should promptly be provided with the narrower width, so that CAV therein can evacuate the limited-width region removed. Assume that the required time-delay is smaller than or equal to the control time interval T_c; then, the time-delay requirement is automatically fulfilled for each section i, if the sharing factors that are actually applied to the two directions, i.e. ε_i^a and ε_i^b, respectively, are calculated as follows

$$\begin{aligned} &\varepsilon_i^a(k_c) = \min\left\{\varepsilon_i(k_c), \varepsilon_i(k_c-1)\right\} \\ &\varepsilon_i^b(k_c) = \min\left\{1-\varepsilon_i(k_c), 1-\varepsilon_i(k_c-1)\right\} \end{aligned} \tag{5.3}$$

where $k_c = 0, 1, \ldots$ is the discrete control time index. It should be mentioned that the notation $\varepsilon_i^a(k_c)$ and $\varepsilon_i^b(k_c)$ indicates that the sharing factors are applied for the duration of the control

time interval $[k_c \cdot T_c, (k_c+1)\cdot T_c)$. The above equations may be readily extended if the required time-delay is a multiple of the control time interval T_c.

In conventional traffic management, the traffic density (in veh/km) gives a clear indication of the state of traffic, depending on its value compared to the critical density: free-flow traffic (when density is lower than critical density), critical traffic (when density is around critical density) or congested traffic (when density is higher than critical density). However, as opposed to having a constant critical density for each direction and section, IBC has the critical density determined by the sharing factor (see (1)); hence altering dynamically according to the control action. Consequently, the sole use of the density value does not provide a consistent view of the traffic conditions in a given section in the presence of IBC.

To tackle this problem, the following relations define the relative densities (dimensionless) per section and per direction. The relative density of section i and direction a or b is obtained by dividing the corresponding traffic density with the corresponding critical density, which, on its turn, depends on the sharing factor (see (1)) prevailing during the last time-step, as follows

$$\tilde{\rho}_i^a(k) = \frac{\rho_i^a(k)}{\varepsilon_i(k-1)\rho_{cr}},\ \tilde{\rho}_i^b(k) = \frac{\rho_i^b(k)}{(1-\varepsilon_i(k-1))\rho_{cr}}. \tag{5.4}$$

The relative densities enable consistent assessment of the traffic conditions in the IBC context. Specifically, if the relative density of a section and direction is less than 1, it reflects under-critical (free-flow) traffic conditions; if it is around 1, it reflects capacity flow; and if it is greater than 1, it reflects over-critical (congested) traffic conditions.

5.3.2 LQR Design for IBC

CTM contains one state equation per section and direction, which is the conservation equation describing the evolution of traffic density. Although the conservation equation is linear, the flow variables included on its right-hand side are nonlinear functions of density. Therefore, Malekzadeh et al. (2021b) proffered linearization of the CTM dynamic equations around a nominal point. In addition, as densities are replaced by relative densities according to (5.4) as state variables, the one-step retarded control input was attached as a new state variable via $\gamma_i(k+1) = \varepsilon_i(k),\ i = 1, 2, \ldots, n.$ This way, the following linearized state-space model was derived in (Malekzadeh, et al., 2021b)

$$\Delta\mathbf{x}(k+1) = \hat{\mathbf{A}}\Delta\mathbf{x}(k) + \hat{\mathbf{B}}\Delta\mathbf{u}(k) \tag{5.5}$$

where $\Delta\mathbf{x}(k) = [\Delta\tilde{\rho}_1^a(k), \Delta\tilde{\rho}_1^b(k), \Delta\gamma_1(k), ..., \Delta\tilde{\rho}_n^a(k), \Delta\tilde{\rho}_n^b(k), \Delta\gamma_n(k)]^T$ is the state vector and $\Delta\mathbf{u}(k) = \Delta\boldsymbol{\varepsilon}(k)$ is the control vector, whereby $\Delta\boldsymbol{\varepsilon}(k) = [\Delta\varepsilon_1(k), \ldots, \Delta\varepsilon_n(k)]^T$. Also, $\Delta(.)(k) = (.)(k) - (.)^N$, the superscript N denoting the nominal values, while it was assumed for simplicity that $\Delta(.)(k) = 0$ for all external disturbances (upstream mainstream inflows, as well as the on-ramp flows, of each direction). $\hat{\mathbf{A}} \in \mathbb{R}^{3n\times 3n}$ and $\hat{\mathbf{B}} \in \mathbb{R}^{3n\times n}$ are the time-invariant state and input matrices, respectively, which result from the linearization, while $\Delta\mathbf{x} \in \mathbb{R}^{3n}$ and $\Delta\mathbf{u} \in \mathbb{R}^n$.

As is well-known, the time-step of system (5.5) must be determined according to the section lengths for numerical stability (e.g., $T = 10\,s$ for $L = 0.5\,km$). If the control time-step is taken

as a multiple of the model time-step, i.e. $T_c = MT$, where M is an integer, then the discrete control time index is $k_c = \lfloor kT/T_c \rfloor$. Thus, the linear state-space equation (5.5) may be changed as follows, in order to be based on the control time-step T_c,

$$\Delta\mathbf{x}(k_c+1) = \mathbf{A}\Delta\mathbf{x}(k_c) + \mathbf{B}\Delta\mathbf{u}(k_c) \tag{5.6}$$

where $\mathbf{A} = \hat{\mathbf{A}}^M$, and $\mathbf{B} = (\hat{\mathbf{A}}^{M-1} + \hat{\mathbf{A}}^{M-2} + ... + \mathbf{I})\hat{\mathbf{B}}$.

To enable the derivation of a time-invariant LQR, Malekzadeh et al. (2021b) proposed minimization of the quadratic criterion

$$J = \frac{1}{2}\sum_{k_c=0}^{\infty}\left[\Delta\mathbf{x}^T(k_c)\mathbf{Q}\Delta\mathbf{x}(k_c) + \Delta\mathbf{u}^T(k_c)\mathbf{R}\Delta\mathbf{u}(k_c)\right] \tag{5.7}$$

where $\mathbf{Q} \in \mathbb{R}^{3n\times 3n}$ is a diagonal positive semidefinite matrix, and $\mathbf{R} \in \mathbb{R}^{n\times n}$ is a diagonal positive definite matrix. The first term penalizes deviations of the state variables from zero, i.e. deviations of $\tilde{\rho}_i^a(k_c)$, $\tilde{\rho}_i^b(k_c)$, $\gamma_i(k_c)$, $i = 1, 2, \ldots, n$, from their respective nominal values. The second term penalizes deviations of the control inputs from the nominal values.

The controller is set to operate the system near capacity by setting the nominal value of relative densities on both directions to 1, which is beneficial for traffic efficiency. Due to the quadratic penalty terms, the controller attempts to reduce the strong deviation of density from the critical density at certain sections, which works to alleviate traffic congestion. On the other hand, by setting the nominal value for the sharing factors to 0.5, the internal boundaries are allowed to shift moderately if density deviations are moderate. Thus, minimization of the second term in (5.7) mitigates deviations of the sharing factors from 0.5 and balances these deviations in space and time, which is a secondary operational sub-objective, as unnecessarily strong internal boundary changes over space and time should be suppressed.

The optimal controller minimizing the criterion (5.7) subject to the model (5.6) is given by a linear state-feedback control law of the form $\Delta\mathbf{u}(k_c) = \mathbf{K}\Delta\mathbf{x}(k_c)$, where $\mathbf{K} \in \mathbb{R}^{n\times 3n}$ is a constant gain matrix given by

$$\mathbf{K} = (\mathbf{R} + \mathbf{B}^T\mathbf{P}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{P}\mathbf{A} \tag{5.8}$$

and $\mathbf{P}$ is the unique positive semidefinite solution of the discrete-time algebraic Riccati equation. To drop the nominal point from the controller, the differential form of the controller is given as follows

$$\mathbf{u}(k_c) = \mathbf{u}(k_c - 1) + \mathbf{K}(\mathbf{x}(k_c) - \mathbf{x}(k_c - 1)). \tag{5.9}$$

The control input is truncated if it exceeds the bounds (5.2). This LQR was applied in (Malekzadeh, et al., 2021b) to the nonlinear extended CTM as ground truth, with control inputs $\varepsilon_i^a(k_c)$ and $\varepsilon_i^b(k_c)$ according to (3) being directly applied to the nonlinear macroscopic model. This first successful test opened the way for a more detailed and realistic implementation and evaluation, using lane-free microscopic simulation, where the sharing factors must first be transformed to a continuous internal boundary curve (as in Figure 5.1); while lane-free driving vehicles are called to respect the internal boundary that changes at every time-step T_c, according to the control decisions. These developments are presented in the following sections.

5.3.3 LQR-Feedforward (LQR-FF) Design for IBC

The LQR above was derived under the assumption that $\Delta(.)(k) = 0$ for all external disturbances (mainstream inflows and on-ramp flows of each direction); i.e., that the disturbances remain constant at their nominal values. Clearly, the external inflows are not constant in highway traffic, but feature strong variations during the day (peak and off-peak periods). LQR may nevertheless be efficient thanks to its feedback structure that reacts to the changes of traffic density caused by the inflows, without explicitly accounting for the actual inflow values. Having said that, it is worth investigating whether an extended LQR, namely LQR-FF, that includes, beyond the state feedback, also a feedforward term to explicitly account for the current inflow values, might provide additional benefits. Further details on the design and applications of LQR-FF may be found in (Marinaki and Papageorgiou, 2005; Papageorgiou, et al., 2015).

To derive such an extended LQR-FF, we assume that, also for the disturbances, $\Delta(.)(k)$ are non-zero, hence we have the following augmented linearized system equation instead of (5.5)

$$\Delta\mathbf{x}(k+1) = \hat{\mathbf{A}}\Delta\mathbf{x}(k) + \hat{\mathbf{B}}\Delta\mathbf{u}(k) + \hat{\mathbf{D}}\Delta\mathbf{d}(k) \tag{5.10}$$

where $\Delta\mathbf{d}(k) = [\Delta q_0^a(k), \Delta r_1^b(k), \Delta r_2^a(k), \Delta r_2^b(k), ..., \Delta r_n^a(k), \Delta q_{n+1}^b(k)]^T$. q_0^a and q_{n+1}^b are the upstream entering flows for the directions a and b, respectively, while r_i^a and r_i^b denote the on-ramp flows for section i (if any) for the directions a and b, respectively. Thus, the state-space equation (4.11) is also changed as follows, in order to be based on the control time-step T_c,

$$\Delta\mathbf{x}(k_c+1) = \mathbf{A}\Delta\mathbf{x}(k_c) + \mathbf{B}\Delta\mathbf{u}(k_c) + \mathbf{D}\Delta\mathbf{d}(k_c) \tag{5.11}$$

where $\mathbf{D} = (\hat{\mathbf{A}}^{M-1} + \hat{\mathbf{A}}^{M-2} + ... + \mathbf{I})\hat{\mathbf{D}}$.

To derive a time-invariant controller, we will start by assuming that the disturbance is constant, i.e., that $\Delta\mathbf{d}(k_c) = \Delta\bar{\mathbf{d}}$. Then, considering minimization of the same objective criterion (7) with the extended system (11) yields (Marinaki and Papageorgiou, 2005) the optimal controller

$$\Delta\mathbf{u}(k_c) = \mathbf{K}\Delta\mathbf{x}(k_c) - \mathbf{F}\Delta\bar{\mathbf{d}} \tag{5.12}$$

where $\mathbf{F} = \mathbf{\Psi}(\mathbf{I} - \mathbf{Z})^{-1}\mathbf{P}\mathbf{D}$, and $\mathbf{\Psi} = [\mathbf{R} + \mathbf{B}^\mathbf{T}\mathbf{P}\mathbf{B}]^{-1}\mathbf{B}^\mathbf{T}$, $\mathbf{Z} = \mathbf{A}^\mathbf{T}\left[\mathbf{I} - \mathbf{P}\mathbf{B}\mathbf{\Psi}\right]$, where $\mathbf{F}$ is a constant feedforward gain matrix.

Assuming now that the disturbance is not constant but slowly changing, to account for the time-variations of the inflows, we may render the feedforward term in (5.12) time-varying by replacing $\Delta\bar{\mathbf{d}}$ in (12) by $\Delta\mathbf{d}(k_c)$, which yields the LQR-FF as follows

$$\Delta\mathbf{u}(k_c) = \mathbf{K}\Delta\mathbf{x}(k_c) - \mathbf{F}\Delta\mathbf{d}(k_c). \tag{5.13}$$

Eventually, LQR-FF may be taken in differential form

$$\mathbf{u}(k_c) = \mathbf{u}(k_c - 1) + \mathbf{K}(\mathbf{x}(k_c) - \mathbf{x}(k_c - 1)) - \mathbf{F}(\Delta\mathbf{d}(k_c) - \Delta\mathbf{d}(k_c - 1)). \tag{5.14}$$

Clearly, LQR-FF requires, beyond state measurements, also external inflow measurements for real-time operation.

5.4 Lane-Free Microscopic Implementation and Simulation of IBC

As mentioned, the IBC concept was tested in previous works macroscopically, whereby the sharing factors ε_i were implemented directly in (5.1) as control inputs. However, the real physical control input in practice is the space-dependent internal boundary curve (blue line in Fig. 5.1), which needs to be specified appropriately, based on the controller decisions, i.e., the sharing factors ε_i by section. It should be noted that the internal boundary curve should be smooth, to allow for lane-free driving vehicles to smoothly and safely adapt to internal boundary changes encountered on their way. The same requirements apply to the microscopic simulator used for IBC testing; hence, microscopic IBC evaluation is deemed to be close to real implementation. Experimental evaluation is performed with TrafficFluid-Sim (Troullinos et al., 2021), a lane-free extension of the well-known microscopic simulator SUMO (Lopes et al., 2018). For IBC, additional functionalities had to be developed in order to facilitate the possibility of changing internal boundaries. Some of these extensions are outlined as part of future work by Troullinos et al. (2022).

5.4.1 Lateral Boundaries in Lane-Free Environments

Consider a network with multiple routes, e.g., a highway with on-ramps and off-ramps. Each vehicle entering the network (at the upstream boundary or at an on-ramp) is assigned a specific itinerary according to its destination (downstream boundary or an off-ramp). In conventional lane-based traffic, the vehicle would need to undertake appropriate lane-changing operations in order to follow its itinerary. In a lane-free environment, realization of the prescribed itinerary calls for appropriate considerations within the vehicle movement strategy, to be detailed in Section 4. In addition, a corresponding admissible movement corridor, delineated by smooth left and right boundaries, is designed that limits the vehicle's lateral placement appropriately. For example, a vehicle entering from an on-ramp needs to first merge appropriately in the main highway. To this end, its movement boundaries are designed with the use of sigmoid functions that specify the laterally admissible positions for the associated itinerary as a function of longitudinal space x (Fig. 5.2).

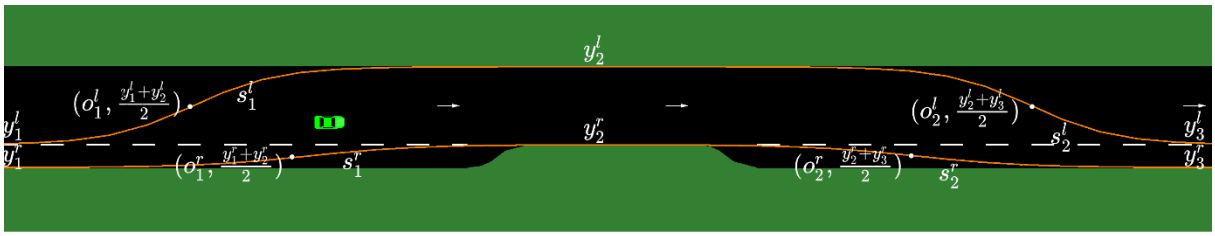


Figure 5.2: Illustration of the boundaries for vehicles entering from an on-ramp and scheduled to exit from an off-ramp

For every itinerary, the boundaries (left and right) are designed generically with a set of parameters that dictate their shape using hyperbolic tangent functions. In Fig. 5.2, we showcase the boundaries in TrafficFluid-Sim and illustrate the design procedure through relevant parameters for an itinerary where vehicles enter from an on-ramp; drive on the main road for a

while; and then exit through an off-ramp located downstream. The l, r superscripts on the parameters indicate elements relevant to the left or right boundary, respectively, and both boundaries are designed following the same principle.

To clarify the design process, we focus on the left boundary from the perspective of the vehicle, which is partitioned into 3 distinct phases corresponding to changes in the lateral availability of the road. First, we define the 3 distinct lateral levels y_1^l, y_2^l, y_3^l as shown in the figure, according to the itinerary that the vehicles shall follow; i.e., y_1^l is set according to the left boundary of the on-ramp's acceleration lane; y_2^l is set according to the left boundary of the main road; and y_3^l is set according to the left boundary of the off-ramp's deceleration lane. These three levels need to be connected via a smooth line to form the complete left boundary of the itinerary to be considered by the corresponding vehicles, as they move longitudinally through the road network. To connect the three levels, longitudinal offset locations are specified, imposing where (longitudinally) the smooth change between two consecutive lateral levels takes place. For instance, the first offset o_1^l in the figure determines the merging area into the main highway. In addition, the slope value s_1^l at the offset point calibrates how steeply the boundary moves from y_1^l to y_2^l.

In general, for each boundary of an itinerary, the user needs to provide a set of $P+1$ lateral level values $\mathbf{y} \in \mathbb{R}^{P+1}$, along with a set of P corresponding offset values $\mathbf{o} \in \mathbb{R}^P$ and slope values $\mathbf{s} \in \mathbb{R}^P$. As mentioned, the hyperbolic tangent function enables gradual change of the boundary from each y_p to y_{p+1}, with steepness at offset o_p being regulated by the slope value s_p. Using these sets of parameters, each vehicle can obtain the boundary's (left or right) lateral placement $y_b(x)$ at any longitudinal position x and incorporate this into its movement strategy (see Section 4). The calculation of $y_b(x)$ follows the equation below:

$$y_b(x) = \sum_{p=1}^{P} \left[\frac{y_{p+1} - y_p}{2} \tanh\left(s_p (x - o_p)\right) \right] + \frac{y_1 + y_{P+1}}{2} \tag{5.15}$$

where the summation involves all provided elements $\forall p \in [1, P]$, while the constant term shifts the resulting sum to yield the proper lateral position of the boundary.

Equation (5.15) can be better understood through the following two examples. Considering a boundary with P offset points, we look for the boundary's position $y_b(x)$ for $o_m < x < o_{m+1}$, where $m \in [1, P]$ is an offset index. When x is not too close to any offset, we have $\tanh(s_p(x - o_p)) \approx 1, \quad \forall p \in [1, m]$, and $\tanh(s_p(x - o_p)) \approx -1, \quad \forall p \in [m+1, P]$. As such, the equation above results in:

$$y_b(x) \approx \left(\frac{y_2 - y_1}{2} + \ldots + \frac{y_{m+1} - y_m}{2} - \frac{y_{m+2} - y_{m+1}}{2} - \ldots - \frac{y_{P+1} - y_P}{2}\right) + \frac{y_1 + y_{P+1}}{2} \tag{5.16}$$

which yields $y_b(x) \approx y_{m+1}$, meaning that, when sufficiently far from offset points, the boundary corresponds to level y_{m+1}, see, e.g., the straight road level y_2^l in Fig. 5.2. Another indicative case refers to the value of $y_b(x)$ when a vehicle crosses an offset, i.e., at $x = o_m$. Assuming that the nearby offsets are not too close, then $\tanh(s_p(x - o_m)) \approx 0$ while the tangent functions for all other $p \neq m$ are close to either $+1$ or -1, as in the previous example. Following this, we get $y_b(x_m) \approx (y_m + y_{m+1})/2$, i.e., the boundary's lateral position at offsets is approximately in the middle of the two adjacent levels, see Fig 5.2. Note that offset points are inflection points of the formed curve $y_b(x)$ in (5.15).

If needed, it is possible, within TrafficFluid-Sim, to design asymmetric lateral level changes via selection of yet another set of parameters $\mathbf{n} \in \mathbb{R}^P$. Then, at offset o_p, the boundary is placed approximately at $n_p \cdot (y_p + y_{p+1})$, where $n_p \in (0,1)$. Asymmetric changes for the lateral levels may be useful when designing boundaries within limited space. Note that for such asymmetric lateral changes $(n_p \neq 0.5)$, equation (5.15) delivering $y_b(x)$ is accordingly generalized.

5.4.2 Microscopic IBC Loop

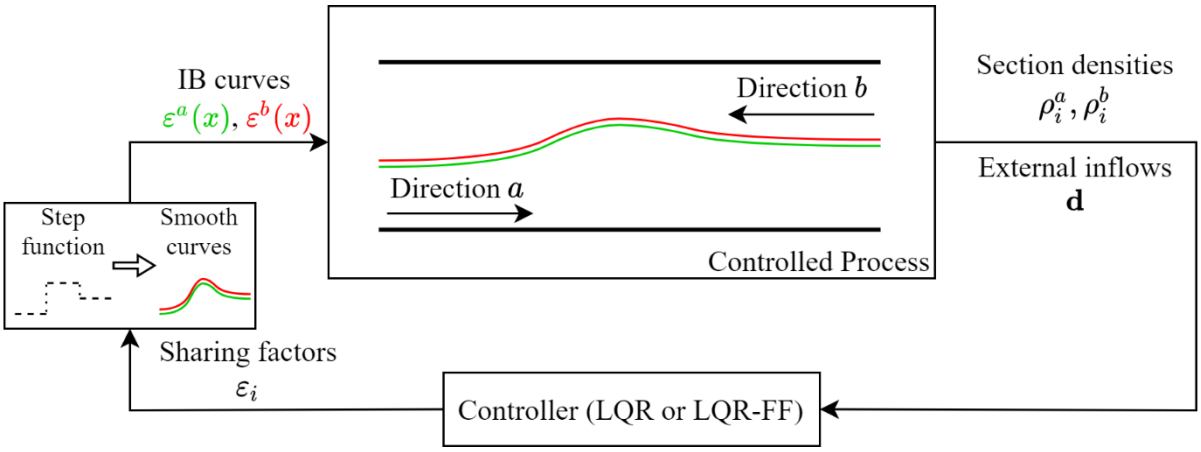


Figure 5.3: Overview of the IBC closed-loop system

The control loop that involves the controlled traffic process and the controller (LQR or LQR-FF) for IBC is provided in Fig. 5.3, where the loop is updated with time interval T_c. The controlled process in this study is the microscopic simulation environment which reflects a potential real-world system. In the microscopic context, the controller output needs to be transformed to a smooth Internal Boundary (IB) curve for the traffic process, while the latter provides real-time measurements of density and external inflows (only for LQR-FF) to the controller.

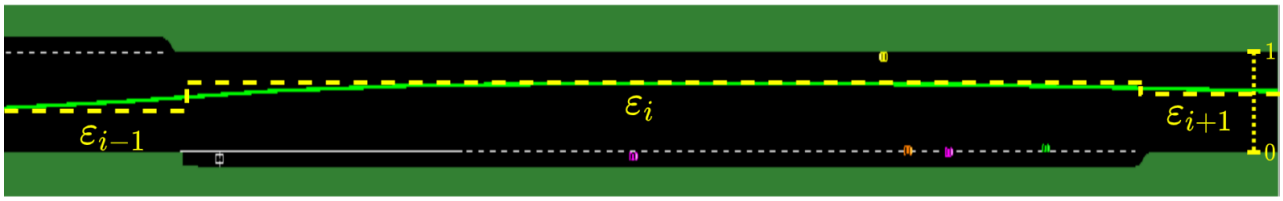


Figure 5.4: Transformation of IBC controller output to the smooth internal boundary.

As discussed in Section 2, the controller (eq. (5.9) or (5.14)) outputs sharing factor values ε_i per highway section i, i.e., a step-function of longitudinal space x. For both the real system and our microscopic simulation, this step-function must be transformed to a continuous and smooth IB curve of space x, as illustrated in Fig. 5.4 for a section i, where the provided sharing factors ε_{i-1}, ε_i, ε_{i+1} are seen to form a step-function (indicated with a dashed yellow line) since each of them is associated with a corresponding section. To transform this step-function to a smooth one (green line), we use the same sigmoid function of Section 3.1, whereby the offsets points are

naturally set at the borders of the highway sections used for the controller design (see Section 2); while the lateral levels y_i are updated according to the values of the sharing factors ε_i, based on the road width.

Since the controller is applied in a microscopic environment, some safety-related aspects emerge and need to be considered while converting sharing factors to IB curves. First, a small lateral distance margin (0.5 m) is placed between the two traffic directions, so that vehicles travelling near the IB retain a safety distance from the opposite traffic direction. As a result, two IB curves $\varepsilon^a(x), \varepsilon^b(x)$, colored in Fig. 5.3 green and red, respectively, are produced, corresponding to the two directions a, b. In addition, while the controller updates its output according to the control time interval T_c (e.g. every 1 min), a time delay in applying the sharing factor is introduced, in every section i, for the direction that is widened, so as to leave time to the vehicles of the narrowed direction to evacuate the shifted IB zone. This was already mentioned in Section 2.1 and led to the equations (5.3), where the time-delay was set equal to T_c. In the microscopic simulation environment, it was observed that lane-free driving CAV subtly and safely evacuate internal road zones assigned to the other traffic direction, while, on the widened road side, CAV efficiently exploit additional road zones assigned to their direction. In fact, a smaller delay is sufficient for vehicles travelling on the narrowed direction to comply with the reduced lateral space, hence a time-delay of $T_c/2$ was adopted.

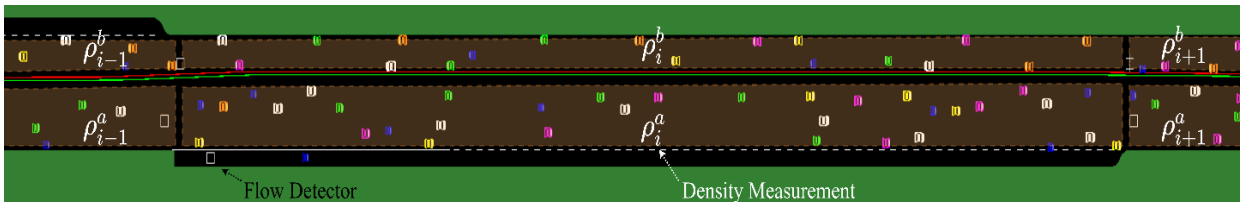


Figure 5.5: Density measurements at two sections of the highway

Finally, as shown in Fig. 5.3 as well as in (5.9) and (5.14), the controller requires real-time information regarding density (veh/km) per section i for each direction a, b. Such density measurements ρ_i^a, ρ_i^b, shown in Fig. 5.5, may be picked from the simulator by counting the corresponding numbers of vehicles in directions a, b and dividing by the section length; and may be readily obtained in a real environment based on appropriate vehicle messages or estimators, see e.g. (Papadopoulou, et al., 2018). The relative desities $\tilde{\rho}_i^a(k)$, $\tilde{\rho}_i^b(k)$ that are relevant to the controller may then be computed according to equation (4) using the last sharing factor values. For LQR-FF, the external inflows $\mathbf{d}(k)$ are measured using loop detectors at the network's entries, e.g. at the on-ramp in Fig. 5.5.

5.5 Vehicle Movement Strategy

Developing a safe and effective approach for the movement of vehicles in a lane-free traffic environment is a complex problem (Yanumula et al., 2023, Naderi et al., 2022). In order to quickly demonstrate and verify the TrafficFluid concept, an ad-hoc vehicle movement strategy for lane-free roads was developed and demonstrated for a ring-road by Malekzadeh et al. (2022).

This section outlines briefly this scheme and the necessary changes made for operation when on-/off-ramps are present.

A double-double-integrator (DDI) model is used for the vehicle motion dynamics, consisting of the following two-dimensional kinematic equations that describe the vehicle's longitudinal and lateral position and speed

$$\begin{aligned} x(t+T_v) &= x(t) + v_x(t)T_v + 0.5\, f_x(t)T_v^2 \\ v_x(t+T_v) &= v_x(t) + f_x(t)T_v \\ y(t+T_v) &= y(t) + v_y(t)T_v + 0.5\, f_y(t)T_v^2 \\ v_y(t+T_v) &= v_y(t) + f_y(t)T_v \end{aligned} \tag{5.17}$$

for $t = 0, T_v, 2T_v, \ldots$ Thus, at the beginning of each time-step t of length T_v, each vehicle departs from position $(x(t), y(t))$, that is actually the center of the vehicle, with longitudinal (lateral) speed $v_x(t)$ $(v_y(t))$; and, moving with constant longitudinal (lateral) acceleration $f_x(t)$ $(f_y(t))$, it reaches its updated state (positions and speeds) at time $t+T_v$. In the above equations, the acceleration value is considered as a control input to be specified by the moving strategy. The DDI model is deemed appropriate for vehicle movement that does not involve strong turnings, as is typically the case on highways.

In the employed vehicle movement strategy, longitudinal (x-direction) and lateral (y-direction) accelerations for each CAV at time t are computed via the following equations:

$$\begin{aligned} f_x(t) &= \sigma_x^{ts} f_x^{ts}(t) + f_x^{rp}(t) + \sigma^{ng}\gamma_x f_x^{ng}(t) \\ f_y(t) &= f_y^{ts}(t) + f_y^{rp}(t) + \sigma^{ng}\gamma_y f_y^{ng}(t) \end{aligned} \tag{5.18}$$

A brief account of the role of each "force" on the right-hand side of (5.18) is provided here:

- The *target-speed forces* f_x^{ts} and f_y^{ts} strive for the vehicle to attain its respective longitudinal and lateral target speeds. f_x^{ts} is not active all the time. Whether this term is activated or not is contingent on the vehicle's relative position to other vehicles in its vicinity. Activation or deactivation of f_x^{ts} is performed using the binary variable $\sigma_x^{ts} \in \{0,1\}$.
- The *repulsive forces* f_x^{rp} and f_y^{rp} are present to prevent any crashes from happening with vehicles ahead (downstream).
- The *nudging forces* f_x^{ng} and f_y^{ng} are present due to vehicles behind (upstream) and may facilitate advancement of faster vehicles behind, but have also an impact on the macroscopic characteristics of the emerging traffic flow (Karafyllis et al., 2022). The coefficients $\gamma_x, \gamma_y \in [0,1]$ regulate the influence of nudging forces versus repulsive forces, enabling attenuation of nudging forces if desired. The nudging forces f_x^{ng} and f_y^{ng} may under circumstances have to be de-activated, and this is performed using the binary variable $\sigma^{ng} \in \{0,1\}$.

Fig. 5.6 visualizes repulsive and nudging forces between two adjacent vehicles i and j in lane-free traffic. A repulsive force f_j^{rp} is created for each couple of adjacent vehicles, with a direction along the line connecting both vehicle centers. The repulsive force indicates towards, and applies

to, the vehicle i, whose center is farther upstream. Its magnitude depends on the distance, relative speed and lateral displacement of both vehicles and is partly inspired by lane-based car-following ACC (Adaptive Cruise Control) laws. The downstream vehicle j receives a nudging force f_i^{ng} with opposite direction and equal magnitude to the repulsive force f_j^{rp}, which, however, can be moderated by the choice of the aforementioned parameters γ_x, γ_y.

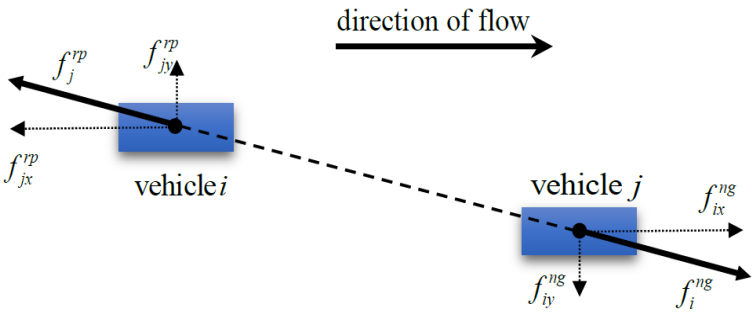


Figure 5.6: Visualization of repulsive and nudging forces between two vehicles.

The total repulsive force f^{rp}, acting on a reference vehicle, equals the sum of the four strongest-in-magnitude individual repulsive forces due to vehicles within a downstream longitudinal distance. Similarly, the nudging force f^{ng}, acting on a reference vehicle, equals the sum of two strongest-in-magnitude individual nudging forces due to vehicles within an upstream longitudinal distance. After f_x and f_y have been produced from the respective sums of target-speed, repulsive and nudging forces via (5.18), a constraining mechanism is applied before they are used as accelerations. In particular, in order to keep a vehicle within the boundaries of the highway, lateral acceleration bounds are computed through a regulation problem, whereby a (left or right) road boundary is considered as a reference lateral position, and a lateral acceleration of the vehicle is obtained via a boundary feedback controller, such that the vehicle moves asymptotically towards the boundary, or remains on the boundary, without violating it. Consequently, such an acceleration can be used as a lateral acceleration bound to ensure that the vehicle will never cross the corresponding road boundary. Other bounds are applied as well, e.g. to avoid the backward moving of vehicles, to avoid speeds that exceed excessively the desired (target) speeds, etc.

A comprehensive explanation of the relations leading to the aforementioned forces can be found in (Malekzadeh et al., 2022). The simulation results presented in that work considered a ring-road without any on-ramps or off-ramps. However, to study the effect of congestion creation due to merging and the related spillback caused, the present work considers the presence of on-ramps and off-ramps, and therefore, the moving strategy was appropriately extended compared to (Malekzadeh et al., 2022). Specifically, while vehicles travelling on the mainstream have zero lateral target (desired) speed, vehicles travelling towards a deceleration lane that leads to an off-

ramp; or vehicles travelling on an on-ramp and trying to merge to the mainstream through an acceleration lane are assigned a negative or a positive desired lateral speed, respectively, to enable corresponding smooth exiting or entering maneuvers. Also, exiting or entering vehicles have extended road boundaries near the acceleration and deceleration lanes (according to Section 3.1, see also Fig. 5.2).

5.6 Simulation Investigations

5.6.1 Simulation Set-Up, Control Parameters and Demand Scenario

The bi-directional highway stretch displayed in Fig. 5.7 is used to test and analyse the IBC control scheme proposed. This highway is 3 km long and was divided into six sections, each being 0.5 km long, for LQR design and IBC application according to Section 2. When deciding on the section lengths, two points need to be taken into account. Firstly, sections are utilized for modelling within the discrete CTM, in which the length of 0.5 km is deemed to be suitable for the purpose, see (Kontorinaki et al., 2017). Secondly, the sections are utilized for IBC, to specify the scope of every sharing factor. If the sections are too long, the flexibility of IBC might be affected; while shorter sections would necessitate accordingly more frequent vehicular maneuvers to adapt to the changing IB. In light of these factors, the decision was made to use a section length of 0.5 km, but other section lengths in this order may also be used, e.g. to adapt to the roadway geometry (for example the position of on-ramps and off-ramps in each direction or changes in total road width) and the typical traffic situations. The investigated highway comprises one on-ramp and one off-ramp per direction (Fig. 5.5), yielding an inhomogeneous traffic situation in both space and time.

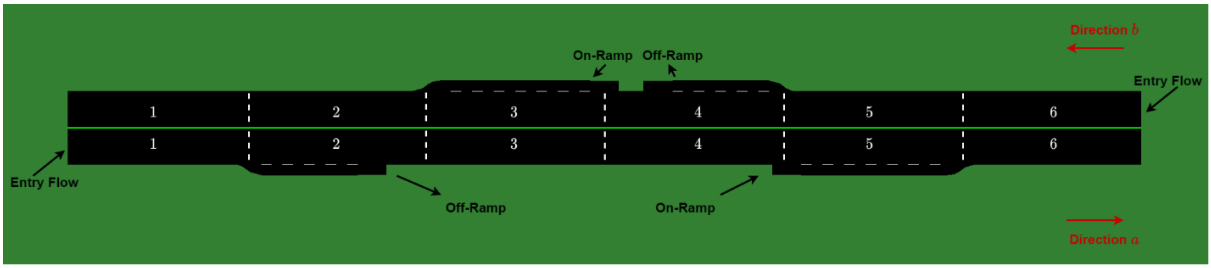


Figure 5.7: The considered stretch of highway

With regard to the microscopic simulation, the vehicle movement strategy is employed with a time step of 0.2 s. The dimensions of the vehicles are determined by choosing randomly (with uniform distribution) one out of the six "dimension classes" reported in Table 5.1. Also, the desired longitudinal speed assigned to a vehicle is chosen randomly (with uniform distribution) within the range [25, 35] m/s = [90, 126] km/h.

Table 5.1: The different dimension classes of vehicles applied in the simulation

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
Length (m)	3.20	3.90	4.25	4.55	4.60	5.15
Width (m)	1.60	1.70	1.80	1.82	1.77	1.84

As outlined in Section 2, the relative density is a key factor in IBC design and scope. According to (4), the relative density depends on the total critical density of the road. Based on the CTM calibration reported in (Papamichail, et al. 2023) with a triangular fundamental diagram, the model parameters used for control design have the following values: $q_{cap} = 28{,}000\,\text{veh/h}$, $\rho_{cr} = 360\ \text{veh/km}$, $v_f = 95\,\text{km/h}$.

Initializing the LQR (5.9) or (5.14) requires the control input (sharing factor) values to be set to the nominal values, i.e. 0.5 for all sections. The highest and lowest admissible values for the sharing factors, used to prevent complete blocking in either direction, are identical throughout $i = 1, 2, \ldots, 6$ and given the values $\varepsilon_{lb} = 0.1$ and $\varepsilon_{ub} = 0.9$. Sharing factors delivered by the IB controller are truncated if they exceed these bounds. The weighing matrices of the objective function are chosen $\mathbf{Q} = \text{diag}(\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_n)$ where $\mathbf{S} = [\mathbf{I}_2, \mathbf{0}_{2\times1}; \mathbf{0}_{1\times3}], \mathbf{R} = \mathbf{I}_{n\times n}$. These weights were specified empirically, but sensitivity of the control performance with respect to the weight values is low, see also (Malekzadeh, et al. 2021b). The used control sample time is $T_c = 60\,\text{s}$.

The employed mainstream and on-ramp demand trajectories for each direction are presented in Fig. 5.8, for a time horizon of 90 min. While the on-ramp demands are relatively low, the mainstream demand feature strong peaks, which are displaced in time for the two directions, so as to highlight the working and benefits of IBC. Notice the high-frequency fluctuation, which is natural in highway demands and flows.

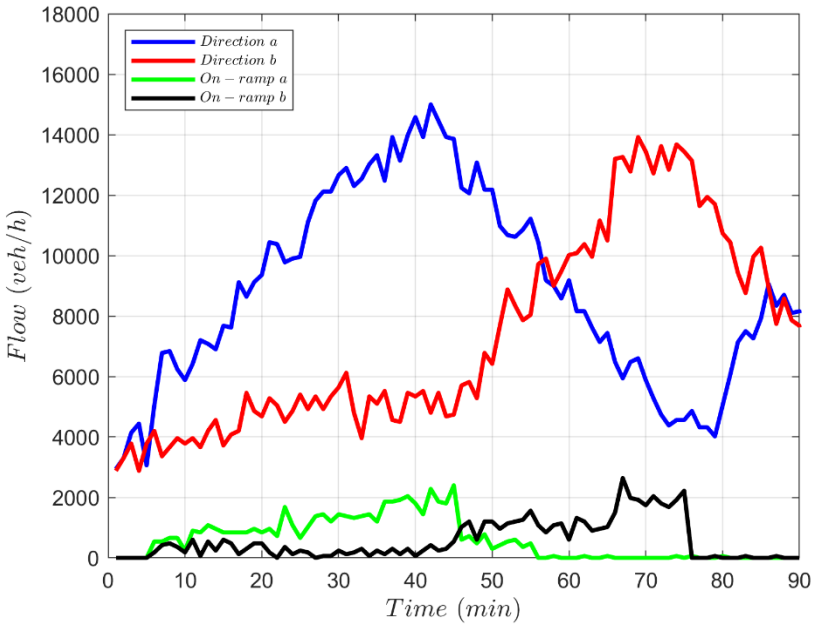


Figure 5.8: Upstream and on-ramp demands per direction

When IBC is activated, the controller uses the relative densities to compute in real-time the IB curves. As a result of the high-frequency demand fluctuation, which is transferred also to the density measurements and relative density values, the control input resulting from the feedback loop may also exhibit some fluctuation, which would not be conducive to IBC application. To address this situation, the collected measurement fluctuations are mitigated via an exponential filter

$$\mathbf{x}^{sm}(k_c) = \alpha\,\mathbf{x}(k_c) + (1-\alpha)\mathbf{x}^{sm}(k_c - 1) \tag{5.19}$$

where $0 \le \alpha \le 1$, before using them for feedback control. For the same reason, exponential filtering is also applied to the external demands used in the feedforward term of LQR-FF

$$\mathbf{d}^{sm}(k_c) = \alpha\,\mathbf{d}(k_c) + (1-\alpha)\mathbf{d}^{sm}(k_c - 1)\,. \tag{5.20}$$

The value $\alpha = 0.5$ was used in the reported investigations.

5.6.2 Investigation Results

When no IBC is applied, hence the total mainstream width of the road is shared equally among the two traffic directions, congestion is created in both of them, displaced in time according to the demand peaks of Fig. 5.8. Congestion in each direction is created at the respective on-ramp sections, when the increasing mainstream demands of either direction plus the respective on-ramp flows exceed the respective (equal) road capacities, i.e. yield densities higher than the respective (equal) critical densities. A graphical representation of the spatio-temporal changes of the relative density in the two directions can be seen in Fig. 5.9. Recall that relative density values lower than 1 are indicative of non-congested traffic; whereas values exceeding 1 signify

congested traffic. If the relative density is around 1, and the downstream section is free-flowing, then we have capacity flow in the relevant section.

Figure 5.9 illustrates a high level of congestion in section 5 for direction a, caused by the merging on-ramp inflow with the increased mainstream demand, at around $t = 40\,\text{min}$. The tail of the created congestion travels upstream, until it ultimately reaches section 1. Eventually, the congestion dissolves at around $t = 55\,\text{min}$, owing to the reduction of the mainstream demand for this direction (Fig. 5.8). In direction b, the combination of the on-ramp flow and the increasing mainstream demand leads also to a congestion in section 3 at around $t = 70\,\text{min}$. The congestion propagates upstream to section 5 and dissolves at around $t = 80\,\text{min}$.

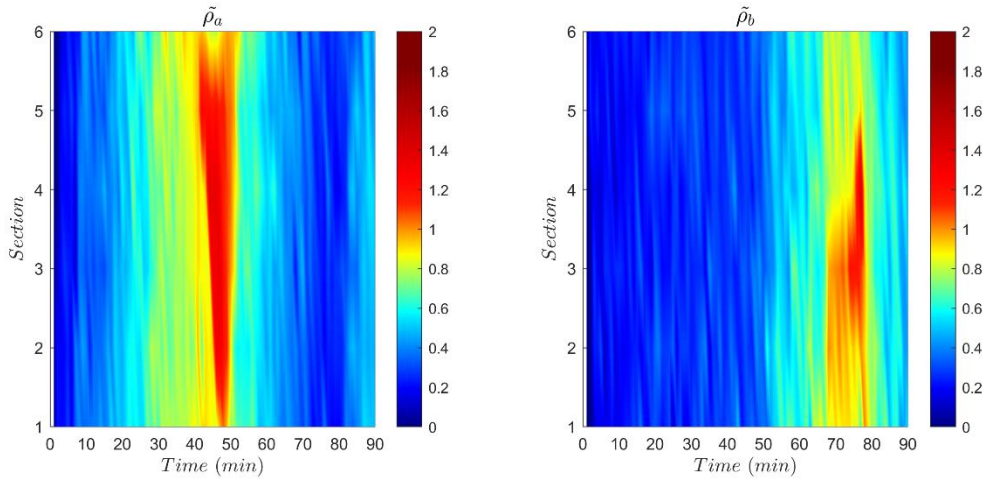


Figure 5.9: Relative density for the two directions in the no-control case

Turning to the control cases, our investigation has two main aims: First, to test and demonstrate the previously developed LQR (5.9) for IBC in a realistic microscopic environment; second, to investigate on possible benefits provided by its extension, introduced in this paper, with a feedforward term (LQR-FF) in (5.14).

Figures 5.10 and 5.11 depict the spatio-temporal evolution of the relative densities with LQR and LQR-FF control, respectively. Further insights are provided from Figs. 5.12 and 5.13, which refer to the merging sections 3 and 5, respectively, where congestions appear in the no-control case. Each figure comprises four diagrams, namely for the respective flows, relative densities, mean speeds and sharing factors (control inputs), the latter along with their upper and lower bounds (black lines). Each diagram (except for the sharing factors) displays the mentioned information for both directions a and b, and for both LQR and LQR-FF. As a first observation, the sharing factor trajectories exhibit only minor time-fluctuations, compared to the demand (Fig. 5.8) and relative density trajectories that feed the controllers, thanks to the smoothing (5.19), (5.20).

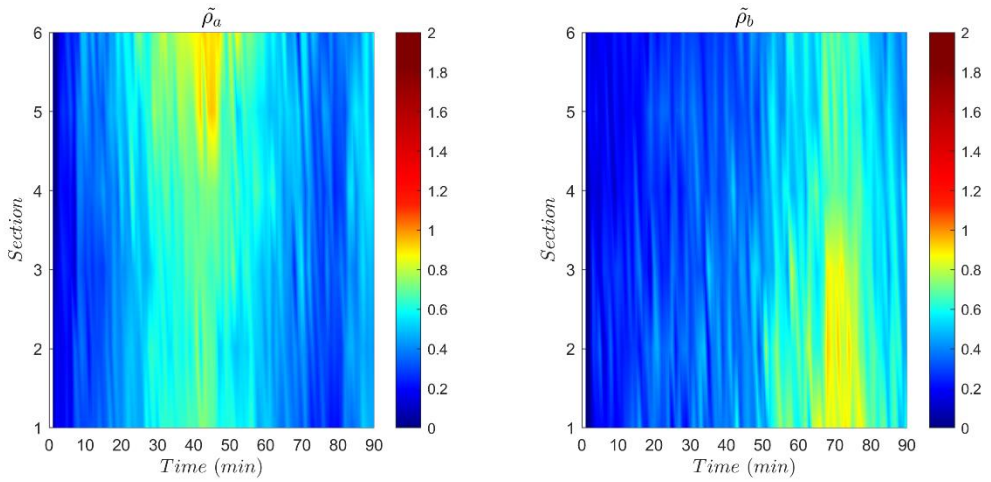


Figure 5.10: Relative density for the two directions in the LQR control case

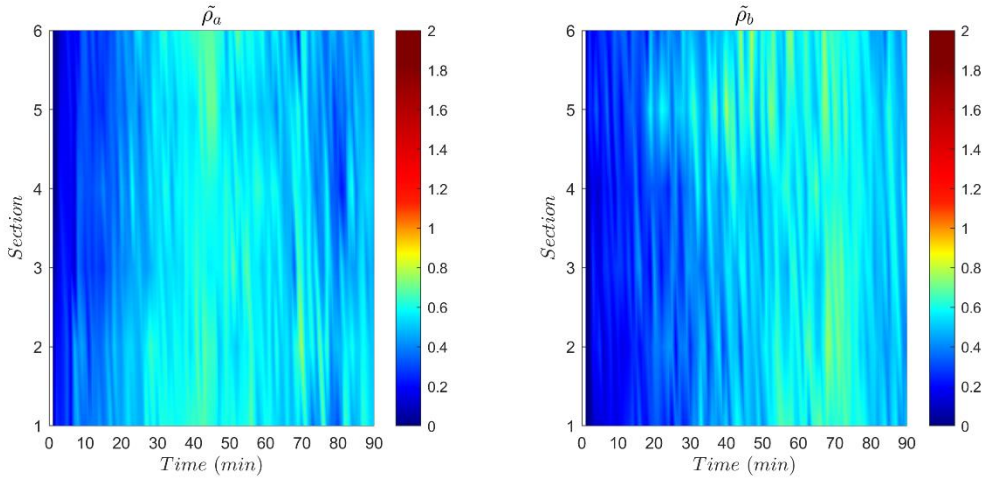


Figure 5.11: Relative density for the two directions in the LQR-FF control case

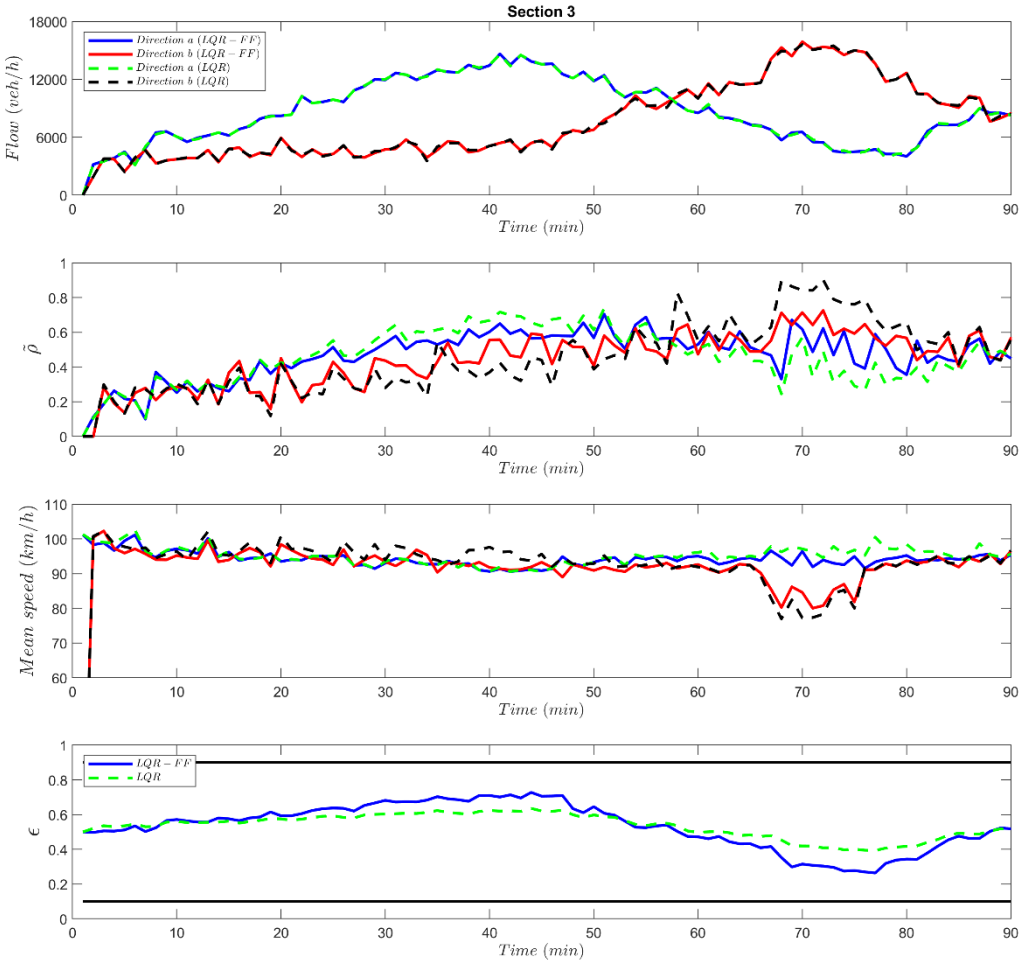


Figure 5.12: Flow, relative density, mean speed and control trajectories in the control case (section 3)

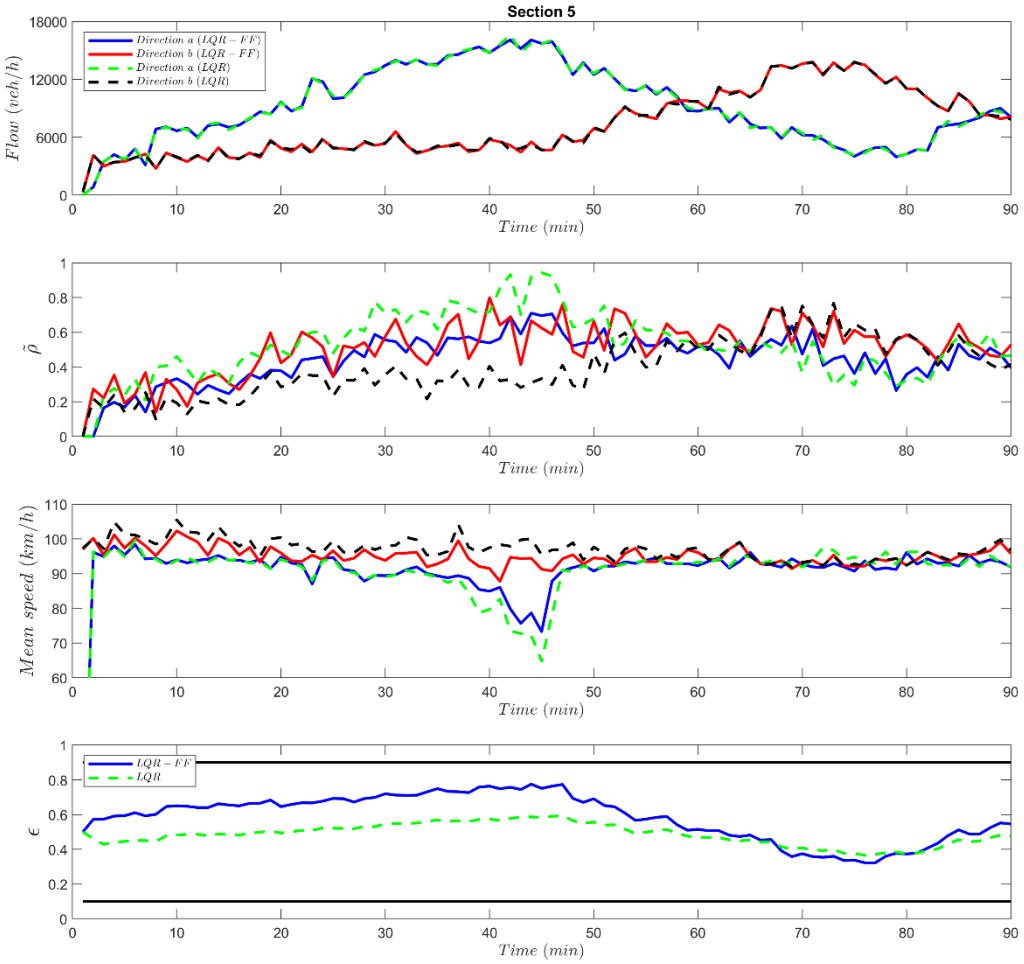


Figure 5.13: Flow, relative density, mean speed and control trajectories in the control case (section 5)

All relative densities in all sections are lower than 1 for both controllers; hence traffic conditions are always and everywhere under-critical, i.e. congestion is utterly avoided, as also evidenced by the mean speed trajectories in Figs. 5.12 and 5.13. Thus, both LQR and LQR-FF appear equally successful in this respect for the considered scenario. It is worth noting that the avoidance of a no-control congestion in one direction is achieved via appropriate IB shifting towards the other direction, which yields a small RD increase in the latter, compared to the no-control case, as evidenced in Figs. 5.9 – 5.11.

Taking a closer look at the sharing factor diagrams in Figs. 5.12 and 5.13, LQR-FF is more reactive than LQR, i.e. the occurring ε-differences from the nominal value (0.5) are stronger. As a result, the two relative densities (RDs) in the two directions of every section are closer to each other (see Figs. 5.12, 5.13), i.e. LQR-FF is better in balancing the RDs of the two directions. This may be further assessed with Table 5.2. The first three columns of Table 5.2 indicate that

the space-time averages of RD are highest for no control, followed by LQR and LQR-FF, but the differences are limited due to two reasons: First, the created congestions in the no-control case are short-lived and contribute only moderately to the averages; second, as mentioned earlier, congestion avoidance in a traffic direction via IB shifting increases RD in the opposite direction. In contrast, clear differences are observed in the fourth column, which displays the space-time average of the absolute value of the RD difference in the two directions, i.e. of $\left|\tilde{\rho}_i^a(k) - \tilde{\rho}_i^b(k)\right|$. This column highlights, in quantitative terms, the superiority of LQR-FF in balancing the RDs in both directions, something that is the direct goal of IBC according to the objective criterion (7).

Another useful assessment criterion for control performance is the RD margin-to-congestion $m(k)$, which is defined, for each time k, as the smallest distance to 1 among all section RDs of both directions. Clearly, the bigger the margin, the higher the "distance" from a congested state; and $m(k) < 0$ indicates that a congestion has already occurred in at least one section and direction. Figure 5.14 displays $m(k)$ for the three examined cases. It is seen that for the no-control case, $m(k) < 0$ over the two periods of congestion in directions a and b, respectively. It is also seen that $m(k)$ is mostly higher for LQR-FF than for LQR, hence, in case of stronger demand, LQR-FF has larger margin to the critical density and congestion creation. The smallest occurring value of m(k) is also depicted in the last column of Table 5.2, confirming the improved behaviour of LQR-FF in this respect.

Table 5.2: Averages of the relative density for no-control, LQR and LQR-FF cases

	Space-time average RD (direction a)	Space-time average RD (direction b)	Space-time average RD (directions a and b)	Space-time average $\left\|\tilde{\rho}_i^a(k) - \tilde{\rho}_i^b(k)\right\|$	Time average RD-margin	Minimum RD-margin
No-control	0.52	0.43	0.47	0.33	0.23	-0.45
LQR	0.48	0.43	0.45	0.21	0.31	0.06
LQR-FF	0.47	0.45	0.46	0.11	0.4	0.2

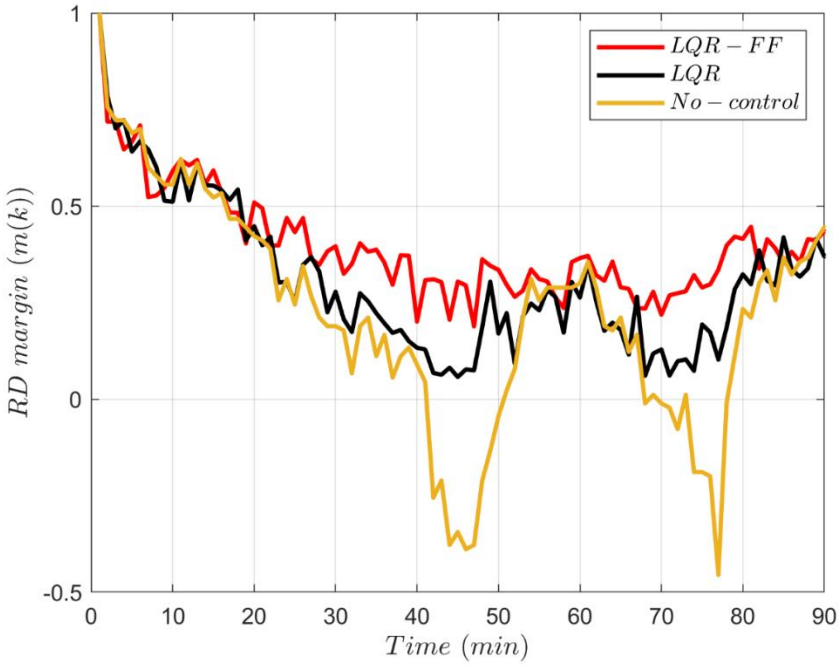


Figure 5.14: The smallest (among all sections of both directions) margin of RD from 1 over time

In the macroscopic IBC testing in (Malekzadeh et al. 2021a; 2021b) by use of CTM, the mean speed is constant (free speed) for any under-critical RD due to the triangular fundamental diagram. In contrast, in the present microscopic evaluation, every increase in RD, even in the under-critical range, leads to a corresponding small decrease of the mean speed, as evidenced by Figs. 5.12 and 13, and also confirmed with Fig. 5.15 that depicts the space-average value of mean speeds for each direction and each control case over time. Clearly, in the no-control case, there is a substantial speed reduction at the period of congestion in both directions. But also for the two control cases, a small speed reduction is observed at times of increased, though under-critical, RDs. Hence, increasing the capacity for one direction (to avoid congestion) comes at a small cost for the other direction due to slightly increased speeds. For example, consider in Fig. 5.15 the speeds in direction b in the time window $40-50\,\text{min}$, i.e. at the period of no-control congestion in direction a. During this period, LQR and LQR-FF restrict direction b in order to handle the increased demand on direction a. The impact on direction b is a small speed decrease, compared to the no-control case, which is more pronounced for LQR-FF due to stronger ε-deviations. The same holds true also when the controllers restrict direction a within the time-interval $70-80\,\text{min}$ to avoid congestion in direction b.

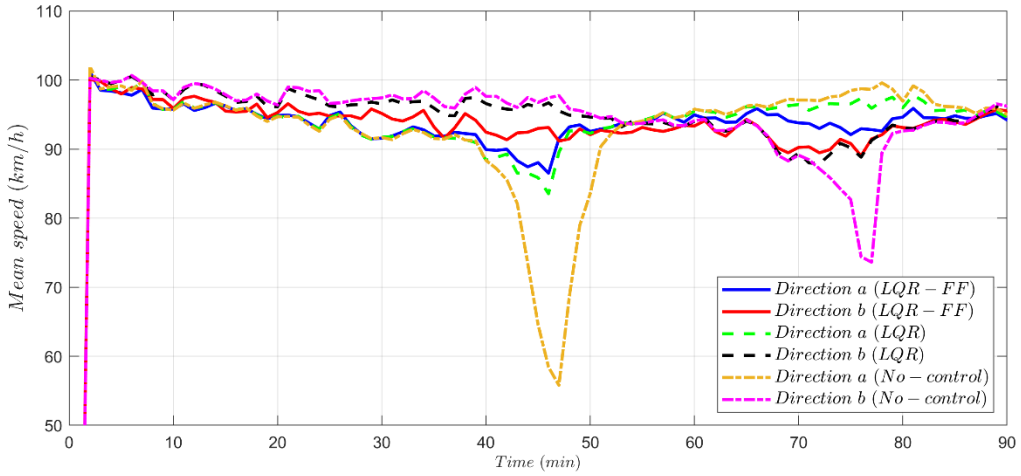


Figure 5.15: Space-average of mean speeds for each direction over time for no-control, LQR and LQR-FF.

We also evaluate IBC with a well-known measure of efficiency, namely the Time Delay (TD) metric, which is depicted in Fig. 5.16 for 20-min time-intervals, for each direction as well as in total. TD is computed for every vehicle when it crosses the exit point and is the time difference between the actual time the vehicle spent in the network and the ideal time it would have accomplished if it was travelling at its desired speed. In the related diagrams, we provide the average value of TD (in seconds) for all vehicles that have exited the network during the associated time-interval.

We again compare LQR with LQR-FF and the no-control case, and the results are consistent with the comparisons on mean speed presented earlier. LQR and LQR-FF exhibit very similar performance, while the no-control results are comparable only when RDs in either direction have under-critical values. Thus, during the third time-interval for direction a and the fourth time-interval for direction b, when respective congestions occur, the no-control TD is clearly higher. In no-congestion periods, TD values may vary among the three cases in one or the other direction, depending on the respective values of the sharing factors. However, a closer look at the total TD diagram during such periods indicates a very slight increase of TD for LQR, compared to no control, and slightly more for LQR-FF. The main reason is that LQR-FF uses stronger control to better equalises the RDs of two opposite directions. The time delay associated with each change of IB has likely also an impact on this, as parts of the roadway are temporarily neutralised and remain unexplored when IBC is applied. In a nutshell, the strongly improved performance of the controllers versus no control, but also the better behaviour of LQR-FF versus LQR in balancing RDs in the two directions, come at a very low price of slightly decreased mean speeds and increased delays when traffic conditions are clearly under-critical. For this, but also for operational reasons, it is reasonable that IBC is only switched on when a threshold of RD in either direction is exceeded, something that is practiced in many other traffic management measures as well.

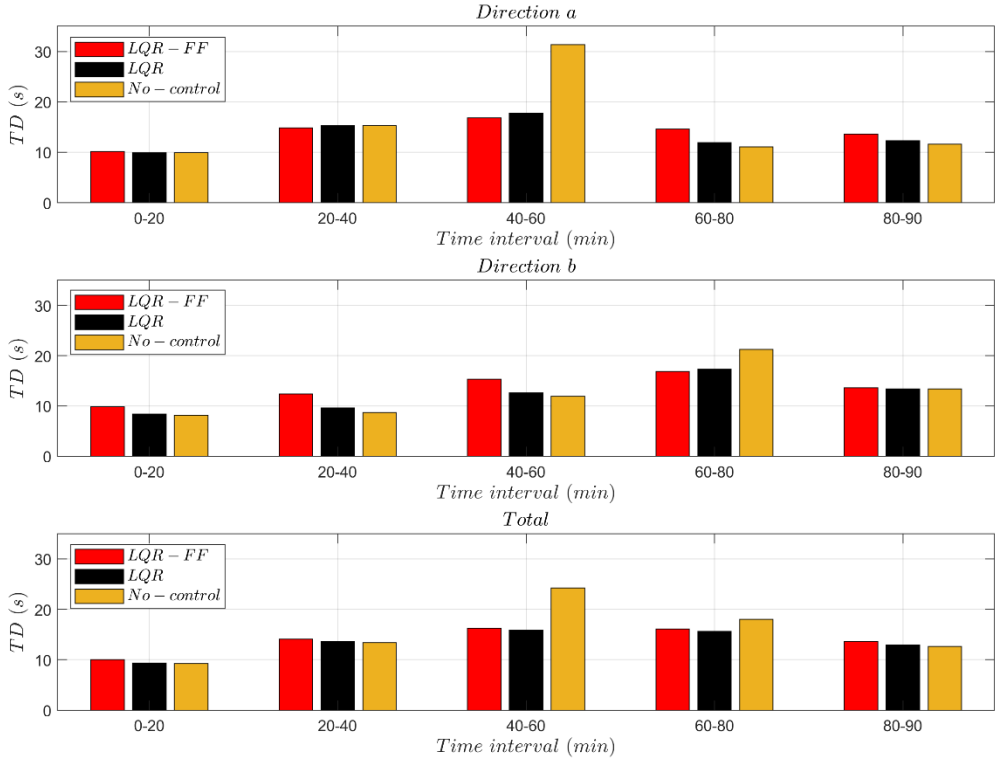


Figure 5.16: Average Time Delay (TD) for 20-min periods

Finally, Fig. 5.17 displays the space-time diagram of the control input, illustrating that the behaviour is consistently smooth over both space and time. Thus, vehicles driving in either direction encounter only small changes of IB on their way that may be easily accommodated (by vehicles close or on the IB) with smooth maneuvers.

5.7 Conclusion

This study has explored the potential use of internal boundary control for lane-free traffic on highways or arterials through microscopic simulation. The IBC scheme has been implemented by use of a previously designed LQR. In addition, a novel, extended controller, LQR-FF, incorporating a feedforward term into the LQR, was introduced, tested and compared. LQR uses real-time measurements of the relative densities, while the feedforward term uses real-time inflow measurements at the mainstream entrance of the highway and at on-ramps. The simulation results confirm and extend previous insights and demonstrate the efficiency and feasibility of IBC in a very realistic microscopic environment. Both controllers were found to act adequately to mitigate or avoid congestion creation in cases of unbalanced demands. LQR-FF was found to offer benefits versus LQR in better balancing the relative densities in both traffic directions, thus enlarging the margin to congestion creation in case of stronger demands.

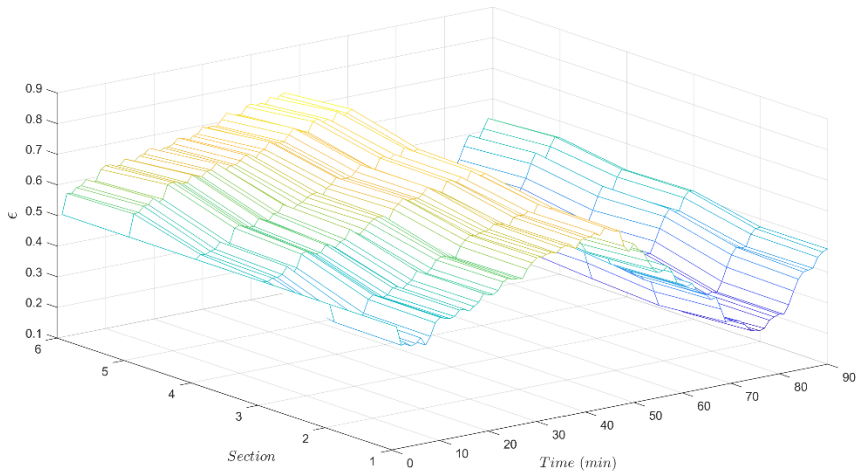


Figure 5.17: Space-time diagram of the control input (LQR-FF)

6 Conclusions and future research directions

Through this thesis, we have delved into a detailed and thorough examination of traffic management within the intelligent transportation system in the lane-free paradigm. We endeavored to address the ever-changing challenges of contemporary mobility by rethinking the traditional limitations of lanes and introducing flexible, adaptable approaches to traffic management. The journey involved a systematic approach that employed different methodologies and solutions to create a smart traffic management system. This was achieved by introducing Internal Boundary Control (IBC) as an innovative solution. When it comes to the IBC, the capacity available for two opposite directions can be dynamically shared without any traffic constraints implied by the presence of lanes.

We commenced our investigation by applying Quadratic Programming (QP) as an open-loop solution. This initial foray allowed us to establish a baseline understanding of optimizing traffic flow while working within the confines of predefined road configurations, as well as to dispose of optimal solutions that were useful in judging the level of suboptimality of other approaches. Despite the benefits of open-loop optimal solutions, we acknowledged their limitations and the importance of implementing more adaptive and responsive control strategies.

In response to this imperative, we turned to the principles of feedback control, employing the Linear Quadratic Regulator (LQR) as a key element of our approach. By harnessing the power of real-time feedback and control theory, we ventured into a more dynamic realm of traffic management, one capable of adapting to changing conditions, traffic patterns, and unexpected events. This pivotal transition marked a significant step toward the realization of IBC.

Recognizing the challenges posed by long highways, we further refined our approach by presenting innovative decentralized solutions. These solutions, rooted in principles of distributed control, not only enhance scalability but also imbue our system with a robust resilience, capable of withstanding localized disruptions while maintaining efficient traffic flow across extensive highway networks. In this context, two different methods, including overlapping LQR and Linear Matrix Inequality (LMI) have been employed. It has been confirmed by the simulation results that the proposed approaches have the potential to serve as a good alternative for decentralization.

In the final phase of our journey, we leveraged microscopic simulation to validate IBC. Microscopic simulation served as a critical proving ground, enabling us to test our hypotheses, refine our algorithms, and gain valuable insights into the intricate dynamics of traffic under the lane-free paradigm. The IBC scheme has been implemented by use of a previously designed LQR. In addition, a novel, extended controller, LQR-FF, incorporating a feedforward term into the LQR, was introduced, tested and compared. LQR uses real-time measurements of the relative densities, while the feedforward term uses real-time inflow measurements at the mainstream entrance of the highway and at on-ramps. The simulation results confirm and extend previous insights and demonstrate the efficiency and feasibility of IBC in a very realistic microscopic environment. Both controllers were found to act adequately to mitigate or avoid congestion creation in cases of unbalanced demands. LQR-FF was found to offer benefits versus LQR in

better balancing the relative densities in both traffic directions, thus enlarging the margin to congestion creation in case of stronger demands.

As we conclude this thesis, we emphasize the significance of our research in the broader context of intelligent transportation. Our work not only challenges conventional notions of traffic management but also provides a roadmap for the future of mobility. By embracing adaptability, decentralization, and feedback-driven control, we offer a promising vision of a transportation ecosystem that can alleviate congestion, enhance safety, and promote sustainability.

The lane-free paradigm is not merely a theoretical concept; it represents a tangible evolution in the way we conceive and manage transportation systems. As we move forward, it is imperative to continue refining and implementing the principles and strategies outlined in this thesis. Collaborations between researchers, policymakers, and industry stakeholders will be crucial in translating these innovations into real-world solutions that benefit society.

In conclusion, the lane-free paradigm, coupled with intelligent traffic management techniques, holds immense potential for revolutionizing mobility. As we bid adieu to traditional lanes, we usher in a new era of adaptive, efficient, and sustainable transportation systems, with the promise of enhancing the quality of life for citizens around the world.

Future work could focus on enhancing the IBC outcome through the combination with ramp metering and variable speed limit control, which have the potential to significantly improve traffic flow in cases where IBC alone may not suffice, e;g due to strong demands in both traffic directions. Furthermore, IBC may be also applied to urban road networks, in addition to urban intersection control, to deliver similar performance surge in cases of unbalanced demands, which are similarly frequent in urban road networks as in highways. Finally, the utilization of Model Predictive Control (MPC), e.g. on the basis of the developed QP problem formulation, could potentially be an effective solution to this problem and should be taken into consideration.

Bibliography

Aboudolas, K. and Geroliminis, N., 2013. Perimeter and boundary flow control in multi-reservoir heterogeneous networks. Transportation Research Part B: Methodological, 55, pp. 265-281.

Aeberhard, M., Rauch, S., Bahram, M., Tanzmeister, G., Thomas, J., Pilat, Y., Homm, F., Huber, W. and Kaempchen, N., 2015. Experience, results and lessons learned from automated driving on Germany's highways. IEEE Intelligent transportation systems magazine, 7(1), pp. 42-57.

Ahmadi, A. and Aldeen, M., (2017). Robust overlapping load frequency output feedback control of multi-area interconnected power systems. International Journal of Electrical Power & Energy Systems, 89, pp.156-172.

Ampountolas, K., dos Santos, J.A. and Carlson, R.C., 2020. Motorway tidal flow lane control. *IEEE Transactions on Intelligent Transportation Systems*, 21(4), pp. 1687-1696.

Ardelt, M., Coester, C. and Kaempchen, N., 2012. Highly automated driving on freeways in real traffic using a probabilistic framework. IEEE Transactions on Intelligent Transportation Systems, 13(4), pp. 1576-1585.

Asaithambi, G., Kanagaraj, V. and Toledo, T., 2016. Driving behaviors: Models and challenges for non-lane based mixed traffic. Transportation in Developing Economies, 2, no.19.

Bakule, L., Rodellar, J. and Rossell, J.M., (2006). Robust overlapping guaranteed cost control of uncertain state-delay discrete-time systems. IEEE Transactions on Automatic Control, 51(12), pp.1943-1950.

Bakule, L., (2014). Decentralized control: Status and outlook. Annual Reviews in Control, 38(1), pp.71-80.

Bhavathrathan, B. and Mallikarjuna, C., 2012. Evolution of macroscopic models for modeling the heterogeneous traffic: an Indian perspective. Transportation Letters, 4(1), pp. 29-39.

Burger, M., van den Berg M., Hegyi, A., De Schutter, B., Hellendoorn, J., 2013. Considerations for model-based traffic control. Transportation Research Part C: Emerging Technologies, 35, pp. 1-19.

Chu, K.F., Lam, A.Y. and Li, V.O., 2019. Dynamic lane reversal routing and scheduling for connected and autonomous vehicles: Formulation and distributed algorithm. IEEE Transactions on Intelligent Transportation Systems, 21(6), pp.2557-2570.

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.

Datta, B. (2004). Numerical methods for linear control systems (Vol. 1). Academic Press.

Diakaki, C. and Papageorgiou, M., 1994. Design and Simulation Test of Coordinated Ramp Metering Control (METALINE) for A10-West in Amsterdam. Internal Report No. 1994-2, Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania, Greece.

Diakaki, C., Papageorgiou, M., Papamichail, I. and Nikolos, I., 2015. Overview and analysis of vehicle automation and communication systems from a motorway traffic management perspective. Transportation Research Part A: Policy and Practice, 75, pp. 147-165.

Doan, K. and Ukkusuri, S.V., 2012. On the holding-back problem in the cell transmission

based dynamic traffic assignment models. Transportation Research Part B: Methodological, 46(9), pp. 1218-1238.

Duell, M., Levin, M.W., Boyles, S.D. and Waller, S.T., 2015. System optimal dynamic lane reversal for autonomous vehicles, *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pp. 1825-1830.

Duell, M., Levin, M.W., Boyles, S.D. and Waller, S.T., 2016. Impact of autonomous vehicles on traffic management: Case of dynamic lane reversal. Transportation Research Record, 2567, pp. 87-94.

Ferrara, A., Sacone, S. and Siri, S., (2018). Freeway traffic modelling and control. Berlin: Springer.

Frejo, J.R.D., Papamichail, I., Papageorgiou, M. and Camacho, E.F., 2015. Macroscopic modeling and control of reversible lanes on freeways. IEEE Transactions on Intelligent Transportation Systems, 17(4), pp. 948-959.

Gomes, G. and Horowitz, R., 2006. Optimal freeway ramp metering using the asymmetric cell transmission model. Transportation Research Part C: Emerging Technologies, 14(4), pp. 244-262.

Hafeez, A., Shamair, Z., Shezad, N., Javed, F., Fazal, T., ur Rehman, S., Bazmi, A.A. and Rehman, F., 2021. Solar powered decentralized water systems: a cleaner solution of the industrial wastewater treatment and clean drinking water supply challenges. Journal of Cleaner Production, 289, p.125717.

Han, Y. Hegyi, A., Yuan, Y., Hoogendoorn, S., Papageorgiou, M. and Roncoli, C., 2017: Resolving freeway jam waves by discrete first-order model-based predictive control of variable speed limits. Transportation Research Part C: Emerging Technologies, 77, pp. 405-420.

Hausknecht, M., Au, T.-C., Stone, P., Fajardo, D. and Waller, T., 2011. Dynamic lane reversal in traffic management. *2011 14th International IEEE Conference on Intelligent Transportation Systems*, pp.1929-1934.

Hug-Glanzmann, G. and Andersson, G., (2009). Decentralized optimal power flow control for overlapping areas in power systems. IEEE Transactions on Power Systems, 24(1), pp.327-336.

İftar, A. & Özgüner, Ü. (1990). Contractible controller design and optimal control with state and input inclusion. Automatica, 26(3), 593–597.

Ikeda, M., Šiljak, D.D. & White, D.E. (1981). Decentralized control with overlapping information sets. Journal of Optimization Theory and Applications, 34(2), 279–310.

Isaksen, L., Payne, H.J., 1973. Suboptimal control of linear systems by augmentation with application to freeway traffic regulation. IEEE Transactions on Automatic Control, 18, pp. 210-219.

Kamal, M.A.S., Taguchi, S. and Yoshimura, T., 2016. Efficient driving on multilane roads under a connected vehicle environment. IEEE Transactions on Intelligent Transportation Systems, 17(9), pp. 2541-2551.

Kanagaraj, V. and Treiber, M., 2018. Self-driven particle model for mixed traffic and other disordered flows. Physica A: Statistical Mechanics and its Applications, 509, pp. 1-11.

Karafyllis, I., Theodosis, D. and Papageorgiou, M., 2022. Analysis and control of a non-local PDE traffic flow model. International Journal of Control, 95(3), pp. 660-678.

Kontorinaki, M., Spiliopoulou, A., Roncoli, C. and Papageorgiou, M., 2017. First-order traffic

flow models incorporating capacity drop: Overview and real-data validation. Transportation Research Part B: Methodological, 106, pp. 52-75.

Kurzhanskiy, A.A. and Varaiya, P., 2010. Active traffic management on road networks: a macroscopic approach. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 368(1928), pp. 4607-4626.

Levin, M.W. and Boyles, S.D., 2016. A cell transmission model for dynamic lane reversal with autonomous vehicles. Transportation Research Part C: Emerging Technologies, 68, pp. 126-143.

Li, Y., Waller, S,T., and Ziliaskopoulos, A.K., 2003. A decomposition scheme for system optimal dynamic traffic assignment models. Networks and Spatial Economics, 3(4), pp. 441-455.

Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves II: a theory of traffic flow on long crowded roads. Royal Soc. Lond. A 229, pp. 317-345.

Lopez, P.A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P. and Wießner, E., 2018. Microscopic traffic simulation using sumo. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 2575-2582). IEEE.

Ma, Y., Song, C. and Wen, C., 2018. Dynamic Traffic Control and Direction Switch of Reversible Lanes for Continuous Flow Considering V2I. In *CICTP 2017: Transportation Reform and Change—Equity, Inclusiveness, Sharing, and Innovation* (pp.505-514). Reston, VA: American Society of Civil Engineers.

Makantasis, K. and Papageorgiou, M., 2018. Motorway path planning for automated road vehicles based on optimal control methods. Transportation Research Record, 2672(19), pp. 112-123.

Malekzadeh, M., Papamichail, I., Papageorgiou, M. and Bogenberger, K., 2021a. Optimal internal boundary control of lane-free automated vehicle traffic. Transportation Research Part C: Emerging Technologies, 126, Article 103060.

Malekzadeh, M., Papamichail, I. and Papageorgiou, M., 2021b. Linear–quadratic regulators for internal boundary control of lane-free automated vehicle traffic. Control Engineering Practice, 115, Article 104912.

Malekzadeh, M., Manolis, D., Papamichail, I. and Papageorgiou, M., 2022. Empirical investigation of properties of lane-free automated vehicle traffic. In 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC) (pp. 2393-2400). IEEE.

Malekzadeh, M., Yanumula, V.K., Papamichail, I. and Papageorgiou, M., 2023. Overlapping internal boundary control of lane-free automated vehicle traffic. Control Engineering Practice, 133, Article105435.

Manjunatha, P., Vortisch, P. and Mathew, T.V., 2013, January. Methodology for the Calibration of VISSIM in Mixed Traffic. In 92nd Transportation Research Board Annual Meeting, Washington, DC, United States.

Marinaki, M. and Papageorgiou, M., 2005. Optimal real-time control of sewer networks. London: Springer.

Mulla, A.K., Joshi, A., Chavan, R., Chakraborty, D. and Manjunath, D., 2018. A microscopic model for lane-less traffic. IEEE Transactions on Control of Network Systems, 6(1), pp. 415-428.

Munigety, C.R. and Mathew, T.V., 2016. Towards behavioral modeling of drivers in mixed traffic conditions. Transportation in Developing Economies, 2, no.6.

Naderi, M., Papageorgiou, M., Karafyllis, I. and Papamichail, I., 2022. Automated vehicle driving on large lane-free roundabouts. In 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC) (pp. 1528-1535). IEEE.

Papadopoulou, S., Roncoli, C., Bekiaris-Liberis, N., Papamichail, I. and Papageorgiou, M., 2018. Microscopic simulation-based validation of a per-lane traffic state estimation scheme for highways with connected vehicles. Transportation Research Part C: Emerging Technologies, 86, pp.441-452.

Papageorgiou, M., 1984. Multilayer control system design applied to freeway traffic. IEEE Transactions on Automatic Control, 29(6), pp. 482-490.

Papageorgiou, M., Blosseville, J.-M., Hadj-Salem, H., 1990. Modelling and real-time control of traffic flow on the southern part of Boulevard Périphérique in Paris - Part II: Coordinated on-ramp metering. Transportation Research Part A: Policy and Practice, 24(5), pp. 361-370.

Papageorgiou, M., 1995. An integrated control approach for traffic corridors. Transportation Research Part C: Emerging Technologies, 3(1), pp. 19-30.

Papageorgiou, M., 1997, Optimal control as a source of intelligent behaviour. In Proceedings of 12th IEEE International Symposium on Intelligent Control (pp.383-389). IEEE.

Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A. and Wang, Y., 2003. Review of road traffic control strategies. Proceedings of the IEEE, 91(12), pp. 2043-2067.

Papageorgiou, M., Ben-Akiva, M., Bottom, J., Bovy, P.H., Hoogendoorn, S.P., Hounsell, N.B., Kotsialos, A. and McDonald, M., 2007. ITS and traffic management. Handbooks in operations research and management science, 14, pp. 715-774.

Papageorgiou, M., Papamichail, I., Spiliopoulou, A.D. and Lentzakis, A.F., 2008. Real-time merging traffic control with applications to toll plaza and work zone management. Transportation Research Part C: Emerging Technologies, 16, pp. 535-553.

Papageorgiou, M., Leibold, M., Buss, M., 2015. Optimierung: statische, dynamische, stochastische Verfahren für die Anwendung (Optimisation: FFApplied static, dynamic, stochastic methods), 4th Edition. Springer, Berlin.

Papageorgiou, M., Mountakis, K.S., Karafyllis, I., Papamichail, I. and Wang, Y., 2021. Lane-free artificial-fluid concept for vehicular traffic. Proceedings of the IEEE, 109(2), pp. 114-121. Papageorgiou, M., Mountakis, K.S., Karafyllis, I., Papamichail, I. and Wang, Y., 2021. Lane-free artificial-fluid concept for vehicular traffic. Proceedings of the IEEE, 109(2), pp. 114-121.

Papamichail, I., Bekiaris-Liberis, N., Delis, A.I., Manolis, D., Mountakis, K.S., Nikolos, I.K., Roncoli, C. and Papageorgiou, M., (2019). Motorway traffic flow modelling, estimation and control with vehicle automation and communication systems. Annual Reviews in Control, 48, pp.325-346.

Papamichail, I., Schoenn-Anchling, N., Malekzadeh, M., Markantonakis, V., and Papageorgiou, M., 2023. Macroscopic Traffic Flow Model Calibration for Lane-free Automated Vehicle Traffic. In 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC). IEEE.

Patil, B.V., Sampath, L.P.M.I., Krishnan, A. and Eddy, F.Y., (2019). Decentralized nonlinear model predictive control of a multimachine power system. International Journal of Electrical

Power & Energy Systems, 106, pp.358-372.

Pasquale, C., Sacone, S., Siri, S. and Ferrara, A., 2020. Hierarchical centralized/decentralized event-triggered control of multiclass traffic networks. IEEE Transactions on Control Systems Technology, 29(4), pp.1549-1564.

Rajamani, R., 2011. Vehicle dynamics and control. Springer Science & Business Media.

Richards, P.I., 1956. Shock waves on the highway. Oper. Res. 4, pp. 42-51.

Roncoli, C., Papageorgiou, M. and Papamichail, I., 2015. Traffic flow optimisation in presence of vehicle automation and communication systems–Part II: Optimal control for multi-lane motorways. *Transportation* Research Part C: Emerging Technologies, 57, pp. 260-275.

Rudloff, C., Schönauer, R. and Fellendorf, M., 2013. Comparing calibrated shared space simulation model with real-life data. Transportation Research Record, 2390, pp. 44-52.

Siljak, D.D., (2011). Decentralized control of complex systems. Courier Corporation.

Toh, K.C., Todd, M.J. & Tütüncü, R.H. (1999). SDPT3 – A MATLAB software package for semidefinite programming, version 1.3. Optimization Methods and Software, 11(1-4), 545–581.

Troullinos, D., Chalkiadakis, G., Manolis, D., Papamichail, I. and Papageorgiou, M., 2021. Lane-free microscopic simulation for connected and automated vehicles. In 2021 IEEE International Intelligent Transportation Systems Conference (ITSC) (pp. 3292-3299). IEEE.

Troullinos, D., Chalkiadakis, G., Manolis, D., Papamichail, I. and Papageorgiou, M., 2022, September. Extending SUMO for lane-free microscopic simulation of connected and automated vehicles. In SUMO Conference Proceedings (Vol. 3, pp. 95-103).

Trudnowski, D.J. and Pierre, D.A., (1992). Decentralized indirect adaptive control scheme applicable to overlapping interconnected systems. International Journal of Control, 55(2), pp.343-360.

VanAntwerp, J.G. & Braatz, R.D. (2000). A tutorial on linear and bilinear matrix inequalities. Journal of Process Control, 10(4), 363–385.

Wolshon, P.B., Lambert, L, 2004. Convertible roadways and lanes: a synthesis of highway practice. Washington, DC, USA: Transportation Research Board National Research Council (Vol. 340).

Xue, D. and Dong, Z., 2000. An intelligent contraflow control method for real-time optimal traffic scheduling using artificial neural network, fuzzy pattern recognition, and optimization. IEEE Transactions on Control Systems Technology, 8(1), pp. 183-191.

Yanumula, V.K., Typaldos, P., Troullinos, D., Malekzadeh, M., Papamichail, I. and Papageorgiou, M., 2023. Optimal trajectory planning for connected and automated vehicles in lane-free traffic with vehicle nudging. IEEE Transactions on Intelligent Vehicles.

Yuan, K., Knoop, V.L. and Hoogendoorn, S.P., 2015. Capacity drop: Relationship between speed in congestion and the queue discharge rate. Transportation Research Record, 2491, pp. 72-80.

Yuan, K., 2016. *Capacity drop on freeways: Traffic dynamics, theory and modeling*. PhD Thesis, Delft University of Technology, The Netherlands.

Yuan, K., Knoop, V.L. and Hoogendoorn, S.P., 2015. Capacity drop: Relationship between speed in congestion and the queue discharge rate. Transportation Research Record, 2491(1), pp. 72-80.

Zečević, A.I. & Šiljak, D.D. (2005). A new approach to control design with overlapping

information structure constraints. Automatica, 41(2), 265–272.

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.