Transportation Research Part B 54 (2013) 17–36

Contents lists available at SciVerse ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Cooperative traffic control of a mixed network with two urban regions and a freeway

Jack Haddad, Mohsen Ramezani, Nikolas Geroliminis *

École Polytechnique Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Urban Transport Systems Laboratory (LUTS), Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 28 September 2012
Received in revised form 21 March 2013
Accepted 22 March 2013

Keywords:
Cooperative decentralized control
Traffic modeling
Macroscopic Fundamental Diagram
Model Predictive Control
Integrated corridor management

ABSTRACT

Currently most optimization methods for urban transport networks (i) are suited for networks with simplified dynamics that are far from real-sized networks or (ii) apply decentralized control, which is not appropriate for heterogeneously loaded networks or (iii) investigate good-quality solutions through micro-simulation models and scenario analysis, which make the problem intractable in real time. In principle, traffic management decisions for different sub-systems of a transport network (urban, freeway) are controlled by operational rules that are network specific and independent from one traffic authority to another. In this paper, the macroscopic traffic modeling and control of a large-scale mixed transportation network consisting of a freeway and an urban network is tackled. The urban network is partitioned into two regions, each one with a well-defined Macroscopic Fundamental Diagram (MFD), i.e. a unimodal and low-scatter relationship between region density and outflow. The freeway is regarded as one alternative commuting route which has one on-ramp and one off-ramp within each urban region. The urban and freeway flow dynamics are formulated with the tool of MFD and asymmetric cell transmission model, respectively. Perimeter controllers on the border of the urban regions operating to manipulate the perimeter interflow between the two regions, and controllers at the on-ramps for ramp metering are considered to control the flow distribution in the mixed network. The optimal traffic control problem is solved by a Model Predictive Control (MPC) approach in order to minimize total delay in the entire network. Several control policies with different levels of urban-freeway control coordination are introduced and tested to scrutinize the characteristics of the proposed controllers. Numerical results demonstrate how different levels of coordination improve the performance once compared with independent control for freeway and urban network. The approach presented in this paper can be extended to implement efficient real-world control strategies for large-scale mixed traffic networks.

1. Introduction

Metropolitan transportation networks have a hierarchical structure which essentially consists of freeways and urban roads providing the interrelated infrastructure for mobility and accessibility. The freeway and the urban network are inherently coupled, but they have dissimilar traffic flow dynamics which challenge the traffic control of mixed networks of two interconnected (urban and freeway) traffic control entities. Although integrating the two entities through an effective mixed control policy could enhance the network performances during heavy congestion conditions, lack of coordination among the urban and freeway network jurisdictions and/or limited means of traffic monitoring and data communication might impede

* Corresponding author. Address: GC C2 389, Station 18, 1015 Lausanne, Switzerland. Tel.: +41 21 69 32481; fax: +41 21 69 35060.
E-mail address: nikolas.geroliminis@epfl.ch (N. Geroliminis).

0191-2615/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trb.2013.03.007

such mixed traffic network ideal goal. To overcome such deficiency, cooperative decentralized or if possible centralized control schemes can be developed as potential solutions, which oblige us to inquire into the traffic dynamics and control of the freeway and urban network to model the mixed traffic network.

Currently most optimization methods for urban transport networks (i) are suited for "toy" networks with simplified dynamics that are far from real-sized networks or (ii) apply decentralized control, which is not appropriate for heterogeneously loaded networks with short links and spillbacks or (iii) investigate good-quality solutions through detailed micro-simulation models and scenario analysis, which due to computational complexity make the problem intractable in real time. In principle, traffic management decisions for different sub-systems of a transport network (urban, freeway) are controlled by operational rules that are network specific and independent from one traffic authority to another. In some cases, the operational decisions of two sub-systems turn out to be competitive. For example, a ramp metering strategy to retain high flows of the freeway sub-system can create long queues in the access ramps that propagate and block the center of the city. In this paper, several control structures with different levels of coordination between the freeway and the urban network control entities are introduced and elaborated for traffic control of the mixed urban-freeway network. Our objective is to investigate how restrictions in coordination among different controllers (e.g. lack of communication or data) can affect the mobility levels of a city. Nevertheless, optimizing in real time all controllers of a city (traffic lights, variable message signs, on-ramps, etc.) in a coordinated way is an infeasible solution due to the computational burden of a very complex model, needed to represent traffic dynamics. Our objective is to integrate realistic aggregated models of urban and freeway traffic with efficient control approaches that will allow for coordinated traffic management.

Recently, a large effort for the development of Integrated Corridor Management (ICM) has been promoted by Federal Highway Administration, with many case studies around US metropolitan areas. Most of the implementations and case studies mainly perform scenario analysis and considers alternative routes under extreme events, e.g. accidents, while it is expected that a more formal optimization approach could lead to a better system performance.

In freeways, ramp metering is the most commonly used controller to manipulate the flow entering the freeway from its urban surrounding roads. Local and coordinated control strategies were proposed and implemented for ramp metering. In local control strategies, the control law for an on-ramp is determined according to the traffic conditions downstream and upstream of the on-ramp (e.g. ALINEA controller in Papageorgiou et al. (1991)). In coordinated strategies, the control law for multiple on-ramps are determined based on the traffic conditions in multiple areas including several on-ramps and sections in the freeway. The coordinated ramp metering is in fact a multi-regulator controller as all ramp meterings attempt to regulate the freeway traffic conditions near the desired densities. Overviews of local and coordinated ramp metering controllers are presented in Papageorgiou and Kotsialos (2002), Papageorgiou et al. (2003), and Geroliminis et al. (2011). The ramp metering approach (even in the coordinated case) might not efficiently operate in case of downstream bottleneck restrictions, e.g. a high demand off-ramp queue spillbacks in the freeway which blocks mainline lanes. Also, in case a freeway ends inside a congested city center, ramp metering might not be able to increase the outflow. In these cases, the freeway and urban network should be controlled in an integrated manner.

For urban networks, the Macroscopic Fundamental Diagram (MFD) aims to simplify the urban traffic micro-modeling, where the collective traffic flow behaviors of subnetworks capture the main characteristics of traffic dynamics, such as the evolution of space-mean flows and densities in different regions of the network. The MFD provides a unimodal, low-scatter relationship between network vehicle density (veh/km) and network space-mean flow or outflow (veh/h) for different network regions, if congestion is roughly homogeneous in the region. Alternatively, the MFD links *accumulation*, defined as the number of vehicles in the region, and *trip completion flow*, defined as the output flow of the region. Urban region flow or trip completion flow increases with accumulation up to a critical point, while additional vehicles in the network cause strong reductions in the flow. The physical model of MFD was initially proposed by Godfrey (1969) and observed with dynamic features in congested urban network in Yokohama by Geroliminis and Daganzo (2008), and investigated using empirical or simulated data by Buisson and Ladier (2009), Ji et al. (2010), Mazloumian et al. (2010), Daganzo et al. (2011), Gayah and Daganzo (2011), Zhang et al. (2013) and others. Earlier works had looked for MFD patterns in data from lightly congested real-world networks or in data from simulations with artificial routing rules and static demands (e.g. Mahmassani et al. (1987); Olszewski et al. (1995) and others), but did not demonstrate that an invariant MFD with dynamic features can arise. Control strategies utilizing the concept of the MFD have been introduced for single-region cities in Daganzo (2007) and later a linear control approach applied for a micro-simulation environment by Keyvan-Ekbatani et al. (2012). These strategies provide some useful insights towards system coordination, but might not operate in an efficient manner and might be far from optimal if congestion is heterogeneously distributed or if many trips have destinations outside the area of analysis, which is the case in many congested cities. Moreover, route guidance strategies with the utilization of MFD have been studied in Knoop et al. (2012) for grid networks.

In case of link density heterogeneity in an urban network, a possible solution to have a well-defined MFD is to partition the heterogenous network into a number of homogeneous smaller regions with small variance of link densities, see Ji and Geroliminis (2012). Recently, Geroliminis et al. (2013) introduced an elegant perimeter control approach to improve traffic conditions in an urban network which has been partitioned into two regions with well-defined MFDs (for stability analysis of the perimeter control see Haddad and Geroliminis (2012)). These results encourage us to utilize the MFD and the perimeter control approach for the mixed urban and freeway network.

All of the above approaches provide a first proof of concept that coordinated real-time control strategies with parsimonious models can create a new generation of smarter cities and improve their mobility. But, still congestion governance in

large-scale systems is currently fragmented and uncoordinated with respect to optimising the goals of travel efficiency and equity for multiple entities. Understanding these interactions for complex and congested cities is a big challenge, which will allow revisiting, redesigning and integrating smarter traffic management approaches to generate cities more livable and sustainable. In this paper, we follow the approach of "system of systems" (SoS), see e.g. Samad and Annaswamy (2011), that aims to model and control a system consists of several independent heterogeneous systems. The mixed network is aligned with an SoS as it consists of an urban network and a freeway system. The perimeter controllers are integrated with the ramp metering controllers to respectively manipulate the flow transfer between urban regions and the inflows from the urban network to the freeway.

Recent studies with empirical data, Geroliminis and Sun (2011) and Saberi and Mahmassani (2012), have shown that a freeway system might not be well-described with an MFD because of strong hysteresis phenomena. Thus, the traffic dynamics of the freeway are modeled in a more detailed approach according to the Asymmetric Cell Transmission Model (ACTM) in Gomes and Horowitz (2006), which is a simplified version of the cell transmission model in Daganzo (1994). Afterwards, the urban and freeway traffic dynamics are extended and integrated to formulate the traffic control problem of the mixed network. A Model Predictive Control (MPC) approach is proposed to minimize the total delay in the whole network. Our optimization findings highlight the importance of coordination of the freeway and urban control entities of a city when congestion is present.

In this paper, there are two main contributions related to traffic modeling and control. In the modeling part, we present a new model that integrates the traffic dynamics of an urban network and a freeway which is not a trivial task or straightforward.

Table 1
List of variables and parameters.

Notation	Description
$C_{\text{on},i}(k)$ (veh/s)	Available flow storage capacity in the on-ramp queue in region i at time step k
$\bar{f}_l$ (veh/h/lane)	Mainline capacity of the freeway cell l
$f_l(k)$ (veh)	Number of vehicles moving from cell l to $l+1$ during time step k
$\bar{f}_{\text{off},l}$ (veh)	Capacity of the off-ramp in cell l
$f_{\text{off},l}(k)$ (veh)	Exit flow of the off-ramp in cell l at time step k
$f_{\text{on},l}(k)$ (veh)	Unmetered on-ramp flow, i.e. the number of vehicles that can enter cell l from its on-ramp during time step k
$G_i(n_i(k))$	Trip completion flow of region i for $n_i(k)$
k	Time step $k = 0,1,\ldots,K-1$
L (–)	Number of the freeway's cells
$\bar{l}_i$ (m)	Average trip length in region i
$M_{ii}(k)$ (veh/s)	Internal trip completion flow from i to i at time step k, $i = 1,2$
$M_{ij}(k)$ (veh/s)	External trip completion flow from i to j at time step k, $i,j = 1,2$
$M_{i3}(k)$ (veh/s)	Maximum output that can flow from region i of the urban network with destination to the freeway at time step k
$n_i(k)$ (veh)	Total number of vehicles in region i at time step k
$n_{i,\text{jam}}$ (veh)	Jam accumulation of region i
$n_{ij}(k)$ (veh)	Total number of vehicles in region i with next destination j (region or freeway) at time step k
$n_{ij,0}$	Initial accumulation of the urban state n_{ij} at $k = 0$
$n_{\text{on},i}(k)$ (veh)	Queue length of the on-ramp in region i at time step k
$n_{\text{on},i,0}$	Initial accumulation of the on-ramp in region i at $k = 0$
$n_{\text{on},i,\max}$ (veh)	Maximum queue length of the on-ramp in region i (storage capacity)
$Q_{ij}(k)$ (veh/s)	A priori known demand generated in origin i to destination j at time step k, $i,j = 1,2,3$, $k = 0,1,\ldots,K-1$
$q_{ij}(k)$ (veh/s)	A generated demand in origin i with direct destination j at time step k that belongs to the trip route $i \to j$
$q_{imj}(k)$ (veh/s)	$m \neq i,j$, a generated demand in origin i with destination j at time step k that belongs to the trip route $i \to m \to j$
$s_{\text{on},i}$ (veh/s)	Maximum number of vehicles that can enter the freeway from the on-ramp belonging to region i
T_k (s)	Time duration of step k
$TT^{\text{f}}_{l_s,l_e}(k)$ (s)	Travel time along the freeway at time step k from cell l_s to cell l_e at time step k
$TT^{\text{r}}_{\text{off}_i}(k)$ (s)	Travel time in off-ramp i at time step k
$TT^{\text{r}}_{\text{on}_i}(k)$ (s)	Travel time in on-ramp i at time step k
$TTD^{\text{u}}_{x,y}(k)$	Travel Time Distribution (TTD) in the urban region for pairs (x,y) at time step k
$u_{\text{on},i}(k)$ (–)	On-ramp metering control input from region i to the freeway at k
$u_{12}(k), u_{21}(k)$ (–)	Perimeter control inputs between the two urban regions at k
$u_{\min}, u_{\max}$	Lower and upper bounds for perimeter and ramp metering controllers
$V_i(k)$ (m/s)	Average speed in urban region i at k
$v_l \in [0,1]$	Normalized free-flow speed of the triangular fundamental diagram
$w_l \in [0,1]$	Normalized congestion wave speed of the triangular fundamental diagram
$x_l(k)$ (veh)	Number of vehicles in freeway cell l at time step k
$x_{l,0}$	Initial accumulation of the freeway cell l at $k = 0$
$x_{l,\max}$ (veh/lane)	Jam accumulation of the triangular fundamental diagram in cell l
$\beta_l(k)$ (–)	Split ratio for the off-ramp (if exists) in cell l at k
$\theta_{ij}(k)$ (–) $\in [0,1]$	Route choice proportion for direct trip between origin i and destination j at k
κ (–)	A preset regulation parameter
ξ_l (–) $\in [0,1]$	On-ramp flow allocation parameter in cell l
γ (–) $\in [0,1]$	On-ramp (if exists) flow blending coefficient
$\eta \in [0,1]$	A predefined ratio

The integration of MFD model with the ACTM arises several unique challenges related to the mixed network as queue spillbacks from the on-ramps to the urban network, the allocation of off-ramp flows, and the route choice behavior as drivers can choose between the urban network and the freeway to travel. Note that by utilizing the new macroscopic modeling, we ease the integrated modeling of urban and freeway systems, which is novel. In addition this study contributes to the control area by addressing the traffic control problem for the mixed freeway and urban network following the control concept of SoS. This work is considered as one step forward and a rigorous contribution towards achieving SoS for transportation infrastructure and networks. In this paper, it is also shown through a numerical example the importance of cooperation between different traffic control jurisdictions.

This paper tackles an SoS with two interconnected urban regions and a freeway crossing with a couple of ramps in each region. More complex systems can be addressed with similar modeling and optimization principles if combined with efficient algorithms for partitioning the urban part in homogeneous regions with low scatter MFDs. Traffic control policies that minimize delays of the mixed network can affect the route choice of vehicles, which is the flow distribution between the urban network and the freeway, since for example vehicles can travel from the periphery to the city center by traveling either through the freeway or the urban network. This effect should not be ignored as restricting access to one region, might transfer demand to another region of a city and change the distribution of congestion (e.g. very aggressive ramp metering might force people to utilize the urban network and vice versa). The route choice models can be classified into two types according to the behavioral supposition of drivers (Wardrop, 1952): the User Equilibrium (UE) when users choose the best route with minimum cost and no user has an incentive to change his route; and the System Optimum (SO) when users cooperate with one another to minimize the total cost for the entire network. In uncongested networks, the UE route choice flow tends to the SO (Prashker and Bekhor, 2000), though to reach the SO in congested networks, it is essential to manage the users' traveling behaviors, such as departure time, mode or route choice. In the mixed traffic network model, we first propose a dynamic simple route choice, which considers that travelers make choices with respect to the sequence of regions (and not a detailed sequence of links) that they will travel to reach their destination. Further we integrate the proposed MPC controller with an SO route choice model to explore the effect of route choice on the overall performance of the mixed network.

The outline of the remainder is as follows. In Sections 2 and 3, the mixed urban and freeway network modeling and the control problem formulation are introduced, respectively. Afterwards, various control policies for traffic control of the mixed network are elaborated in Section 4 and results of the control policies are presented in Section 5. The comparison between the dynamic simple route choice and SO route choice models is investigated in Section 6. Finally in the last section, conclusions are drawn and future work is summarized. Table 1 provides a nomenclature for the different variables and parameters utilized in the paper.

2. Traffic modeling of a mixed urban and freeway network

Consider a mixed urban and freeway network as schematically shown in Fig. 1, where it is assumed that the urban network is partitioned into two homogeneous urban regions, denoted by 1 and 2, having their own MFDs. In addition, there is a freeway, denoted by 3, as an alternative commuting route that passes through both urban regions having one on-ramp and one off-ramp within each region. The freeway also carries demand which is generated and finishes outside the two urban regions. Thus, a 3×3 time-dependent origin destination matrix is associated with the network demand.

To keep the traffic dynamic formulations elegant, the following assumptions are made regarding the trip routes in the mixed network: (A1) the freeway can be used at most once during the trip (exit and re-enter is not allowed), (A2) there is at most one urban region transfer during the trip, e.g. traveling from 1 to 2 then to 1 is not allowed. Under these assumptions, an origin–destination trip might have at most two trip routes since the traveler can choose between using the urban

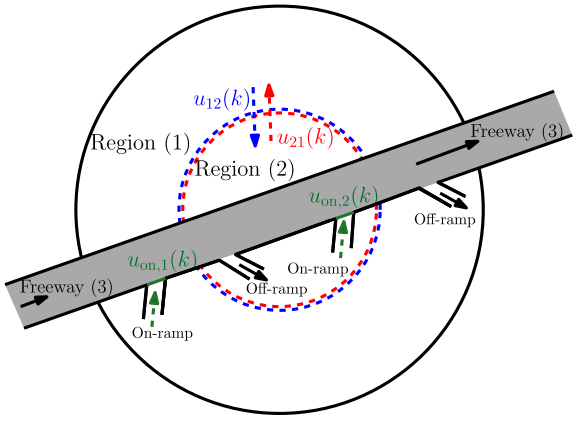

Fig. 1. A mixed urban and freeway network: two regions 1 and 2 with two perimeter control inputs $u_{12}(k)$ and $u_{21}(k)$, a freeway with two on-ramps and two off-ramps, and two on-ramp metering control inputs $u_{\text{on},1}(k)$ and $u_{\text{on},2}(k)$.

Table 2
Trip routes and demands in the mixed network.

Origin	Destination		
	1	2	3
1	q_{11}: 1 → 1 q_{131}: 1 → 3 → 1	q_{12}: 1 → 2 q_{132}: 1 → 3 → 2	q_{13}: 1 → 3 q_{123}: 1 → 2 → 3
2	q_{21}: 2 → 1 q_{231}: 2 → 3 → 1	q_{22}: 2 → 2	q_{23}: 2 → 3 q_{213}: 2 → 1 → 3
3	q_{31}: 3 → 1 q_{321}: 3 → 2 → 1	q_{32}: 3 → 2 q_{312} 3 → 1 → 2	q_{33}: 3 → 3

network or the freeway, or using both the urban network and the freeway (if this option exists), e.g. there are two trip routes from 1 to 1: 1 → 1 traveling from 1 to 1 using the urban region 1, and 1 → 3 → 1 traveling within region 1 to the on-ramp 1, entering and travelling through the freeway, and then exiting from the off-ramp 1 and complete the trip in region 1. All origin–destination trip routes in the mixed network are summarized in Table 2. Note that from region 2 to 2 there is only one route through the urban region, as the off-ramp in region 2 is assumed to be prior to its on-ramp, as shown in Fig. 1. Alternations of the aforementioned assumptions can be integrated in the methodology, but the formulation might be more tedious. Corresponding to the aforementioned O–D matrix, let $Q_{ij}(k)$ (veh/s) be a priori known demand generated in origin i to destination j at time step k, $i,j = 1,2,3$, $k = 0,1,\ldots,K-1$, which is distributed between two choices for the same origin–destination: $q_{ij}(k)$ (veh/s) denotes a generated demand in origin i with direct destination j at time step k that belongs to the trip route $i \to j$, while $q_{imj}(k)$ (veh/s), $m \neq i,j$, denotes a generated demand in origin i with destination j at time step k that belongs to the trip route $i \to m \to j$. All origin–destination demands are also summarized in Table 2. These choices depend on the development and propagation of congestion at different parts of the network. Errors in $Q_{ij}(k)$ are discussed later.

In the following subsections, the traffic dynamics of the mixed freeway and urban network are elaborated. First, the traffic dynamics of the urban regions and the freeway are respectively introduced and modified according to the ACTM and MFD. A challenge is to integrate the two models (meso and macro), as there are transfer flows at the boundaries of freeway and urban network that depend on traffic conditions in both regions. Details of the dynamic route choice modeling between the freeway and the urban network are also discussed. Later, the entire mixed network traffic control problem will be presented in Section 3.

2.1. The urban traffic modeling

Corresponding to the aforementioned traffic demands, see Table 2, six accumulation states are introduced to model the dynamic equations of the urban network: $n_{ij}(k)$ (veh), $i = 1,2$; $j = 1,2,3$, where $n_{ij}(k)$ is the total number of vehicles in region i with next destination j at time step k. Let us denote $n_i(k)$ (veh) as the accumulation or the total number of vehicles in region i at time step k, i.e. $n_i(k) = \sum_{j=1}^{3} n_{ij}(k)$. The output of MFD, $G_i(n_i(k))$ (veh/s), represents the trip completion flow of region i for $n_i(k)$ which is the sum of transfer flows, i.e. external flows as trips from i with destination j, $i \neq j$, plus the internal flow as trips from i with destination i. The transfer flow from i to j is calculated corresponding to the proportion between accumulations, i.e. the external flow is $M_{ij}(k) = n_{ij}(k)/n_i(k) \cdot G_i(n_i(k))$, $i \neq j$, while the internal flow from i to i is calculated by $M_{ii}(k) = n_{ii}(k)/n_i(k) \cdot G_i(n_i(k))$. These relationships assume that trip length of *regional trips* (internal or external) are similar. For a description of different cases (which will not alter the methodology) the reader can refer to Geroliminis (2009).

Simulation and empirical results in Geroliminis and Daganzo (2008) show that the shape of MFD can be approximated by a non-symmetric unimodal curve skewed to the right, i.e. critical density that maximizes network flow is smaller than half of the jam density. Thus, we utilize a 3rd-order function of $n_i(k)$, e.g. $G_i(n_i(k)) = a_i \cdot n_i^3 + b_i \cdot n_i^2 + c_i \cdot n_i$, where a_i, b_i, c_i can be estimated from real data. Analytical ways to estimate MFDs as a function of topology and signal setting have been presented in Daganzo and Geroliminis (2008), Helbing (2009), and Geroliminis and Boyacı (2012). The laws of mass conservation for the six urban state variables of the mixed network are derived in Section 3, see (18)–(23).

2.2. The freeway modeling

Given that an MFD cannot consistently describe the dynamics in a freeway, we consider the traffic dynamics of the freeway in the mixed network based on the ACTM. The mass conservation equations of the on-ramps are adjusted to fit the mixed network problem, as the input demands of the on-ramps are the output of the MFDs. Traditionally, the off-ramp flows are considered independent of the applied control and defined a priori. Given the interactions between the freeway and urban system, the freeway off-ramp flows are the input demand for the urban network, therefore, new equations are derived to split the off-ramp flows to different urban accumulation states. In the sequel, a brief description of the ACTM is presented, followed by mixed network modeling adjustments. The reader can refer to Gomes and Horowitz (2006) and Daganzo (1994) for a full description of ACTM. Note that the CTM (and consequently ACTM) does not reflect the capacity drop phenomenon during congestion. This issue can be addressed by integrating inverse λ fundamental diagrams in the CTM or by employing a higher order freeway model that can capture the capacity drop. Moreover, ACTM simplifies the CTM modeling of merging

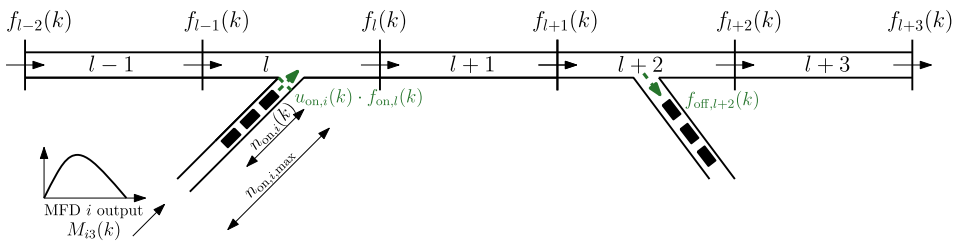

Fig. 2. Representative freeway cells in the ACTM. Cell l has an on-ramp belonging to region i, while cell $l+2$ has an off-ramp. The on-ramp is fed from a demand generated in urban region i and calculated from the MFD, and flows exit at the off-ramp continuing their trip to reach their final destinations.

behavior by assuming independence between the on-ramp and mainline flows during merging, which is not fully consistent with the physics of traffic. While this simplification is applied to derive linear on-ramp control in Gomes and Horowitz (2006), still with proper parameter calibration of ACTM (parameter ξ_l) an accurate enough behavior of merging in both the mainline and the on-ramps can be observed, especially under congested conditions.

In the ACTM, the freeway is divided to L cells, where each cell l of the freeway contains at most one on- or one off-ramp. Five cells of the freeway are schematically shown in Fig. 2 as cell l has an on-ramp belonging to urban region i, and cell $l+2$ has an off-ramp. The number of vehicles in cell l at time step k, $k = 0,1,\ldots,K-1$, is denoted by $x_l(k)$ (veh), while $f_l(k)$ (veh) is the number of vehicles moving from cell l to $l+1$ during time step k. The on-ramp is fed from $M_{i3}(k)$ (veh/s), the maximum output that can flow from region i of the urban network with destination to the freeway at time step k and calculated using the MFD. Let $n_{\mathrm{on},i}(k)$ (veh) be the queue length of the on-ramp in region i at time step k, and $n_{\mathrm{on},i,\max}$ (veh) be the maximum queue length of the on-ramp in region i. It is assumed that each cell l has a triangular fundamental diagram with the following parameters: $w_l \in [0,1]$ is the normalized congestion wave speed, $v_l \in [0,1]$ is the normalized free-flow speed, $x_{l,\max}$ (veh/lane) is the jam accumulation, and $\bar{f}_l$ (veh/h/lane) is the mainline capacity.

The freeway is integrated with the urban network through its on- and off-ramps. In the following, we first determine the on-ramp freeway flow taking into account the MFD output, and afterwards the off-ramp flows are split as inputs to different urban accumulations.

The unmetered on-ramp flow $f_{\mathrm{on},l}(k)$ (veh) is the number of vehicles that can enter cell l from its on-ramp during time step k. It is calculated as follows:

$$f_{\mathrm{on},l}(k) = \min[n_{\mathrm{on},i}(k) + M_{i3}(k)\cdot T_k,\ \xi_l\cdot(x_{l,\max}-x_l(k)),\ s_{\mathrm{on},i}\cdot T_k], \tag{1}$$

where i is the region that the on-ramp belongs to, T_k (s) is the time step size, $s_{\mathrm{on},i}$ (veh/s) is the maximum number of vehicles that can enter the freeway from the on-ramp belonging to region i, and ξ_l [−] $\in [0,1]$ is the on-ramp flow allocation parameter, see Gomes and Horowitz (2006). The on-ramp metering control inputs, denoted by $u_{\mathrm{on},i}(k)$ (−), $i = 1,2$, are introduced on the entrance of the freeway to control the flow entering from the region i to the freeway, see Fig. 1. The queue dynamic for the on-ramp belonging to region i with ramp metering control input, $u_{\mathrm{on},i}(k)$, taking into the account the on-ramp maximum queue length, is as follows:

$$n_{\mathrm{on},i}(k+1) = \min(n_{\mathrm{on},i}(k) + M_{i3}(k)\cdot T_k - u_{\mathrm{on},i}(k)\cdot f_{\mathrm{on},l}(k), n_{\mathrm{on},i,max}). \tag{2}$$

The mainline flow in the freeway is calculated as follows:

$$f_l(k) = \min[(1-\beta_l(k))\cdot v_l\cdot(x_l(k)+\gamma\cdot u_{\mathrm{on},i}(k)f_{\mathrm{on},l}(k)), F_l(k), w_{l+1}\cdot(\bar{x}_{l+1}-x_{l+1}(k)-\gamma\cdot u_{\mathrm{on},i}(k)\cdot f_{\mathrm{on},l+1}(k))], \tag{3}$$

where $\beta_l(k)$ (−) is the split ratio for the off-ramp (if exists) in cell l, γ (−) $\in [0,1]$ is the on-ramp (if exists) flow blending coefficient, and $F_l(k) \triangleq \min\{\bar{f}_l,\ (1-\beta_l(k))/\beta_l(k)\cdot\bar{f}_{\mathrm{off},l}\}$, where $\bar{f}_{\mathrm{off},l}$ (veh) is the off-ramp capacity. The exit flow of the off-ramp in cell l, $f_{\mathrm{off},l}(k)$ (veh), is determined as follows:

$$f_{\mathrm{off},l}(k) = \frac{\beta_l(k)}{1-\beta_l(k)}\cdot f_l(k). \tag{4}$$

Finally, the mainline mass conservation is

$$x_l(k+1) = x_l(k) + f_{l-1}(k) + u_{\mathrm{on},i}(k)\cdot f_{\mathrm{on},l}(k) - f_l(k) - f_{\mathrm{off},l}(k), \tag{5}$$

for $l = 1,2,\ldots,L$, and $k = 0,1,\ldots,K-1$, where $f_{\mathrm{on},l}(k) = 0$ and/or $f_{\mathrm{off},l}(k) = 0$ if cell l does not contain an on-ramp and/or an off-ramp, respectively.

The ACTM does not keep track of the origin and the destination of vehicles as off-ramps are usually the final destinations of the trip. However, in the mixed network vehicles exiting from the off-ramps can have various destinations, e.g. finishing their trip in the urban region connected to the off-ramp or continuing their trip to the other urban region. To calculate the off-ramp flow distribution, we assume that the split ratios are similar to the instantaneous corresponding O–D demands, i.e. the off-ramp exit flows denoted by "hat" variables are calculated as follows

$$[\hat{q}_{31}(k),\hat{q}_{312}(k),\hat{q}_{231}(k)] = f_{\text{off},1}(k)\cdot\frac{[q_{31}(k)+q_{131}(k),q_{312}(k),q_{231}(k)]}{q_{31}(k)+q_{131}(k)+q_{312}(k)+q_{231}(k)}, \tag{6}$$

$$[\hat{q}_{32}(k),\hat{q}_{321}(k),\hat{q}_{132}(k)] = f_{\text{off},2}(k)\cdot\frac{[q_{32}(k),q_{321}(k),q_{132}(k)]}{q_{32}(k)+q_{321}(k)+q_{132}(k)}. \tag{7}$$

The above "hat" demands will be further integrated in the dynamic equations of the mixed network in Section 3.1. A more accurate estimation would require that in each cell we keep track of the final destination of each vehicle. Moreover, (6) and (7) are precise when O–D patterns are not varying rapidly. If O–D patterns change significantly the off-ramp flow distribution can be derived according to the experienced corresponding O–D demands, i.e. the current split ratios at time step k can be associated with the O–D demands of time step k minus mean of O–D experienced travel time. However, both enhancements would make ACTM dynamics more tedious. These can be a future direction of research.

2.3. The route choice modeling

Real-time control can affect the route choices of users. To be able to integrate the modeling framework in a real-time control environment, a parsimonious model of route choice has to be developed. Instead of considering a detailed route choice, a path choice is modeled as a sequence of regions during a trip, as expressed in Table 2. Nevertheless, given that MFD (or ACTM) modeling assumes that all vehicles traveling at a region (or cell) at time step k have equal speeds, we decide to add stochasticity in the urban region's trip length to make the model more realistic (otherwise all or nothing choices will be made). By assuming that all origins and destinations are homogeneously distributed in each of the urban regions and that the space is continuous, the Trip Length Distributions (TLDs) within each urban region for different trips can be estimated either in an analytical or experimental way. The TLD between x and y is the distribution of the distance between a random point in x and a random point in y. The possible pairs of (x,y) according to the trip routes of Table 2 are: (I) (i,i) two random points inside urban region i, (II) $(i,\partial(ij))$ a random point inside urban region i and the nearest point on the border between regions i and j, (III) (i,on_i) a random point inside urban region i and the on-ramp of urban region i, (IV) $(\partial(ij),\text{on}_j)$ the border between urban regions i and j and the on-ramp of region j, and (V) $(\text{off}_i,\partial(ij))$ the off-ramp of urban region i and the nearest point on the border between regions i and j. If GPS data are available for a fraction of the vehicles these TLDs can be estimated more accurately. Let us denote $TTD^{\text{u}}_{x,y}(k)$ the Travel Time Distribution (TTD) in the urban region for pairs (x,y) at step k (the continuous time t is divided to K time steps, as each time step k has the time duration of T_k (s)). Then, the TTDs are estimated as follows:

$$TTD^{\text{u}}_{x,y}(k) = \frac{TLD^{\text{u}}_{x,y}}{V_i(k)}, \tag{8}$$

where $V_i(k)$ (m/s) is the average speed in urban region i at k, which is calculated according to the MFD of region i as follows:

$$V_i(k) = \frac{G_i(n_i(k))\cdot\bar{l}_i}{n_i(k)}, \tag{9}$$

where $G_i(n_i(k))$ is the MFD value of region i at time step k, $n_i(k)$ is the accumulation of region i at time step k, and $\bar{l}_i$ (m) is the average region trip length (which can be estimated from real data, see Geroliminis and Daganzo (2008)). Note that $\bar{l}_i$, is considered time-independent, but as a regional route choice is embedded in the model, trip length between origin i and destination j varies with the traffic conditions.

Let $TT^{\text{f}}_{l_s,l_e}(k)$ (s) be the instantaneous travel time along the freeway at time step k from cell l_s to cell l_e (note that both cells take values between 1 and L, and that the mainline freeway cell associated with on- or off-ramp i is denoted as on_i or off_i for simplicity), and $TT^{\text{r}}_{\text{on}_i}(k)$ (or $TT^{\text{r}}_{\text{off}_i}(k)$) (s) be the travel time in on-ramp i (or off-ramp i) at time step k. Then, $TT^{\text{r}}_{\text{on}_i}(k)$ is approximated by

$$TT^{\text{r}}_{\text{on}_i}(k) = \frac{n_{\text{on},i}(k)\cdot T_k}{u_{\text{on},i}(k)\cdot f_{\text{on},l}(k)}, \tag{10}$$

where $n_{\text{on},i}(k)$ (veh) is the queue length of on-ramp i at time step k, $u_{\text{on},i}(k)$ (–) is the ramp metering control value $(0\leqslant u_{\text{on},i}(k)\leqslant 1)$ at k that constraints $f_{\text{on},l}(k)$ (veh), i.e. the unmetered outflow of on-ramp i at k that enters to its connected freeway cell l, see Fig. 2. Further, the travel time of off-ramp i at k, $TT^{\text{r}}_{\text{off}_i}(k)$ (s), is assumed to be equal to the free flow travel time (note that if this assumption is relaxed then one needs to keep track of the destination of vehicles in order to model the traffic state of the off-ramps, which is not the case in ACTM. Keeping track of the origin and the destination of vehicles hinders the macroscopic modeling of freeways).

To have a dynamic route choice modeling, one should split the a priori known $Q_{ij}(k)$ O–D demand between possible alternative trip routes. Let $\theta_{ij}(k)$ (–) be the route choice proportion for direct trip between origin i and destination j, $\theta_{ij}(k)\in[0,1]$, where $q_{ij}(k)=\theta_{ij}(k)\cdot Q_{ij}(k)$ and $q_{imj}(k)=(1-\theta_{ij}(k))\cdot Q_{ij}(k)$ if $q_{imj}(k)$ exists, otherwise $q_{ij}(k)=Q_{ij}(k)$. e.g. $q_{11}(k)=\theta_{11}(k)\cdot Q_{11}(k)$, and $q_{131}(k)=(1-\theta_{11}(k))\cdot Q_{11}(k)$, while $q_{22}(k)=Q_{22}(k)$. The value of $\theta_{ij}(k)$ is calculated dynamically according to the traffic conditions in the urban regions and the freeway, based on the estimation of the trip route travel times at time step k. Note

that there is no need for route choice calculations of $2 \to 2$ and $3 \to 3$ O–D pairs, since they only have a single trip route choice.

Now let us consider an O–D from i to j with trip routes $i \to j$ and $i \to m \to j$. Recall that $TTD^{u}_{x,y}$ denotes the travel time distributions in the urban regions between x and y, $TT^{f}_{l_s,l_e}$ denotes the total travel time along the freeway from cell l_s to cell l_e, and $TT^{r}_{on_i}$ (or $TT^{r}_{off_i}$) denotes the travel time in on-ramp i (or off-ramp i). Note that the travel time variability is only considered in the urban regions. The two trip route Travel Time Distributions TTD_{ij} and TTD_{imj} at time step k are calculated as follows

$$TTD_{ii}(k) = TTD^{u}_{i,i}(k) \qquad i = 1 \tag{11}$$

$$TTD_{ij}(k) = \begin{cases} TTD^{u}_{i,\partial(ij)}(k) + TTD^{u}_{\partial(ij),j}(k) & \text{if } ij \in \{12,21\}, \\ TTD^{u}_{i,on_i}(k) + TT^{r}_{on_i}(k) + TT^{f}_{on_i,L}(k) & \text{if } ij \in \{13,23\}, \\ TT^{f}_{1,off_j}(k) + TT^{r}_{off_j}(k) & \text{if } ij \in \{31,32\}, \end{cases} \tag{12}$$

$$TTD_{imj}(k) = \begin{cases} TTD^{u}_{i,on_i}(k) + TT^{r}_{on_i}(k) + TT^{f}_{on_i,off_j}(k) + TT^{r}_{off_j}(k) & \text{if } imj \in \{131,132,231\}, \\ TTD^{u}_{i,\partial(im)}(k) + TTD^{u}_{\partial(im),on_m}(k) + TT^{r}_{on_m}(k) + TT^{f}_{on_m,L}(k) & \text{if } imj \in \{123,213\}, \\ TT^{f}_{1,off_m}(k) + TT^{r}_{off_m}(k) + TTD^{u}_{off_m,\partial(mj)}(k) + TTD^{u}_{\partial(mj),j}(k) & \text{if } imj \in \{321,312\}. \end{cases} \tag{13}$$

Ultimately, given the distributions of travel time of both alternative trip routes, one can calculate $\theta_{ij}(k)$ as follows:

$$\theta_{ij}(k) = \Pr(TTD_{ij}(k) - TTD_{imj}(k) < 0). \tag{14}$$

Eq. (14) determines θ_{ij} as the proportion of vehicles of O–D from i to j, whose travel time of trip route $i \to j$ is less than trip route $i \to m \to j$. Eq. (14) requires first, to reflect the $TTD_{imj}(k)$ about vertical axis with zero travel time because of the negative sign, then convolve it with $TTD_{ij}(k)$. This procedure reads

$$(TTD_{ij}(k) - TTD_{imj}(k))(\tau) = \int_{-\infty}^{\infty} (TTD_{ij}(k))(\omega) \cdot (TTD_{imj}(k))(-\tau + \omega) \mathrm{d}\omega \tag{15}$$

$$\theta_{ij}(k) = \int_{-\infty}^{0} (TTD_{ij}(k) - TTD_{imj}(k))(\tau) \mathrm{d}\tau. \tag{16}$$

Finally, the probability of $(TTD_{ij}(k) - TTD_{imj}(k)) < 0$ is the θ_{ij}. The above equations assume independence between the values of the two distributions, which can be relaxed if real data is available for TLD estimation. An example of such procedure is shown in Fig. 3b for O–D from region 1 to 1 at $k = 0$. This dynamic simple route choice model assumes that drivers (users) have the real-time information to select their trip route, though this real-time information is based on the current condition of the network, see (8)–(16). It is worth to mention that more detailed dynamic traffic assignment strategies can also be proposed in a way to keep the control problem tractable from an optimization point of view. Following up (16) the average trip length between origin i and destination j (given that there are two possible routes for every pair) can be estimated as $\bar{l}_{ij} = \theta_{ij}(k) \cdot E\{TLD_{ij}(k)\} + (1 - \theta_{ij}(k)) \cdot E\{TLD_{imj}(k)\}$.

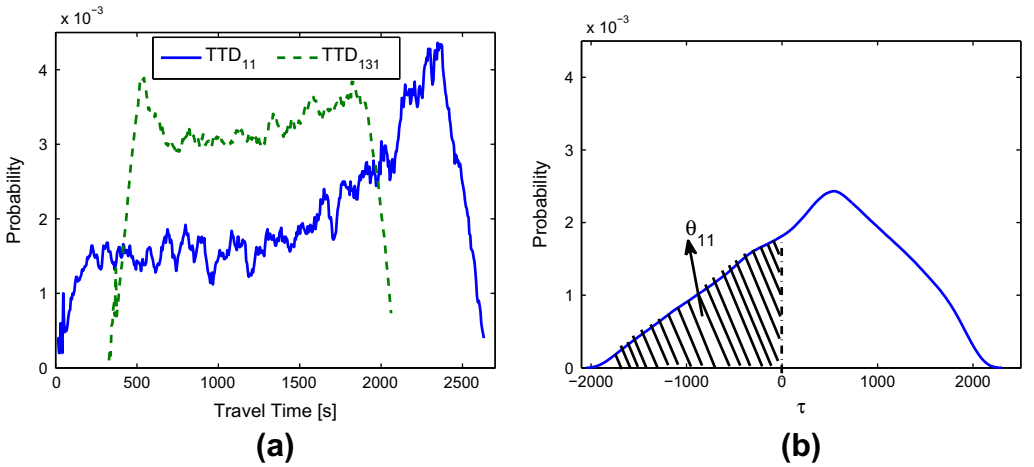

Fig. 3. An example of route choice procedure for origin 1 to destination 1: (a) the TTD of both trip routes at $k = 0$, (b) the probability distribution function of $(TTD_{11}(0) - TTD_{131}(0))$. The hatched area is $\theta_{11}(0) = 0.35$.

3. The mixed network control problem

3.1. Problem formulation

Given the traffic dynamics of the previous section, we can now formulate an optimization problem. In the mixed network control problem, there are two types of controllers to minimize the network total delay: the perimeter controllers for the urban regions and the on-ramp meterings for the freeway. The perimeter control inputs denoted by $u_{12}(k)$ and $u_{21}(k)$ (–) are on the border between the two urban regions as shown in Fig. 1, to control the urban external transfer flows. Since the perimeter controllers exist only on the border of the two regions, the internal flows cannot be controlled or restricted, while the external transfer flows are controlled such that, only a proportion of inter-region demand flows can pass the perimeter at time step k. i.e. $u_{12}(k)$ and $u_{21}(k)$, where $0 \leqslant u_{12}(k), u_{21}(k) \leqslant 1$, are the proportion of the transfer flows that goes from region 1 to 2 and region 2 to 1 at time step k, respectively. Likewise, the on-ramp metering control inputs, $u_{\mathrm{on},i}(k)$, $i = 1,2$, are on the entrance of the freeway to control the flow entering the freeway, such that a proportion of demand input from the urban region can enter into the freeway, i.e. $0 \leqslant u_{\mathrm{on},i}(k) \leqslant 1$, $i = 1,2$.

In the traffic dynamics of the mixed network, there are several types of state variables enumerated as: six state variables describing the dynamics of the urban regions, two state variables describing the queue dynamics of the freeway on-ramps, and L state variables describing accumulation of the freeway cells. Therefore, the mixed traffic network control problem is formulated as follows:

$$J = \min_{\substack{u_{12}(k),u_{21}(k),\\ u_{\mathrm{on},1}(k),u_{\mathrm{on},2}(k);\\ \text{for } k=0,\ldots,K-1}} T_k \cdot \left[\sum_{k=0}^{K-1}(n_1(k)+n_2(k)) + \left[\sum_{k=0}^{K-1}\sum_{l=1}^{L}x_l(k) + \sum_{k=0}^{K-1}\sum_{i=1}^{2}n_{\mathrm{on},i}(k)\right]\right] \tag{17}$$

subject to

$$n_{11}(k+1) = n_{11}(k) + T_k \cdot \left[\frac{\hat{q}_{321}(k)+q_{21}(k)}{\hat{q}_{321}(k)+q_{213}(k)+q_{21}(k)} \cdot u_{21}(k) \cdot M_{21}(k) + q_{11}(k) + \hat{q}_{231}(k) + \hat{q}_{31}(k) - M_{11}(k)\right] \tag{18}$$

$$n_{12}(k+1) = n_{12}(k) + T_k \cdot [q_{12}(k) + q_{123}(k) + \hat{q}_{312}(k) - u_{12}(k) \cdot M_{12}(k)] \tag{19}$$

$$n_{13}(k+1) = n_{13}(k) + T_k \cdot \left[\frac{q_{213}(k)}{\hat{q}_{321}(k)+q_{213}(k)+q_{21}(k)} \cdot u_{21}(k) \cdot M_{21}(k) + q_{13}(k) + q_{131}(k) + q_{132}(k) - \min(M_{13}(k), C_{\mathrm{on},1}(k))\right] \tag{20}$$

$$n_{21}(k+1) = n_{21}(k) + T_k \cdot [q_{21}(k) + q_{213}(k) + \hat{q}_{321}(k) - u_{21}(k) \cdot M_{21}(k)] \tag{21}$$

$$n_{22}(k+1) = n_{22}(k) + T_k \cdot \left[\frac{q_{12}(k)+\hat{q}_{312}(k)}{q_{12}(k)+q_{123}(k)+\hat{q}_{312}(k)} \cdot u_{12}(k) \cdot M_{12}(k) + q_{22}(k) + \hat{q}_{132}(k) + \hat{q}_{32}(k) - M_{22}(k)\right] \tag{22}$$

$$n_{23}(k+1) = n_{23}(k) + T_k \cdot \left[\frac{q_{123}(k)}{q_{12}(k)+q_{123}(k)+\hat{q}_{321}(k)} \cdot u_{12}(k) \cdot M_{12}(k) + q_{23}(k) + q_{231}(k) - \min(M_{23}(k), C_{\mathrm{on},2}(k))\right] \tag{23}$$

$$0 \leqslant \sum_{j=1}^{3} n_{ij}(k) \leqslant n_{i,\mathrm{jam}} \quad i = 1,2 \tag{24}$$

$$u_{\min} \leqslant u_{ij}(k) \leqslant u_{\max} \quad i = 1,2;\ j = 3-i \tag{25}$$

$$u_{\min} \leqslant u_{\mathrm{on},i}(k) \leqslant u_{\max} \quad i = 1,2 \tag{26}$$

$$n_{ij}(0) = n_{ij,0} \quad i = 1,2;\ j = 1,2,3 \tag{27}$$

$$x_l(0) = x_{l,0} \quad l = 1,2,\ldots,L \tag{28}$$

$$n_{\mathrm{on},i}(0) = n_{\mathrm{on},i,0} \quad i = 1,2 \tag{29}$$

and (1)–(16)

for $k = 0,1,2,\ldots,K-1$, where $n_{ij,0}$, $n_{\mathrm{on},i,0}$, and $x_{l,0}$ are the initial accumulations of the urban states, on-ramps, and freeway cells at $k = 0$; $n_{i,\mathrm{jam}}$ (veh) is the jam accumulation of region i; $u_{\min}$ and $u_{\max}$ are the lower and upper bounds for perimeter and

ramp metering controllers; and $C_{on,i}(k)$ (veh/s) is the available flow storage capacity in the on-ramp queue, i.e. $C_{on,i}(k) = (n_{on,i,max} - n_{on,i}(k))/T_k$.

The term min $(M_{i3}(k), C_{on,i}(k))$, $i = 1,2$, in (20) and (23) represents the flow of vehicles that leaves the urban region to enter the freeway, which is the minimum of outgoing flow and the available flow capacity in on-ramp i. Moreover, the term $u_{ij}(k) \cdot M_{ij}(k)$, $i = 1,2$; $j = 3 - i$, expresses the outflow from region i to region j, which comprises of two parts; (i) the part that is from origin i to destination j, and (ii) the other part that is from origin i to destination 3 through region j. Thus to distinguish between these two parts, the urban interflow, $u_{ij}(k) \cdot M_{ij}(k)$, are divided accordingly to the ratio of demand flow. E.g. $(q_{12}(k) + \hat{q}_{312}(k))/(q_{12}(k) + q_{123}(k) + \hat{q}_{312}(k)) \cdot u_{12}(k) \cdot M_{12}(k)$ and $q_{123}(k)/(q_{12}(k) + q_{123}(k) + \hat{q}_{321}(k)) \cdot u_{12}(k) \cdot M_{12}(k)$ respectively represent the outflow from origin 1 to destination 2 and outflow from origin 1 to destination 3 through region 2. Note that, Eqs. (18)–(23) are the conservation of mass laws for urban accumulations, $n_{ij}(k)$, while (24) is the lower and upper bound constraints on region i accumulation. Recall that $M_{ij}(k) = n_{ij}(k)/n_i(k) \cdot G_i(n_i(k))$, $i = 1,2$; $j = 1,2,3$, the Eqs. (8)–(16) derive the $q_{ij}(k)$ and $q_{imj}(k)$ based on the proposed dynamic route choice modeling, and (1)–(7) describe the freeway dynamics. If one compares (18)–(23) for MFD dynamics with (1)–(3) for ACTM dynamics it might observe some inconsistency in the mass conservation, as the MFD equations do not directly consider downstream restrictions (while ACTM does), but apply an exit function. One could add one more term, the boundary capacity, which is a function of the receiving region accumulation. Nevertheless, this constraint can be ignored during the optimization process. The physical reasoning besides this assumption is that (i) boundary capacity decreases for accumulations much larger than the critical accumulation (see Geroliminis and Daganzo (2007)), and (ii) the control inputs will not allow the system to get close to gridlock.

3.2. Solution approach – an MPC controller

The optimal control problem for the mixed network is solved by the MPC approach. The MPC obtains optimal solutions with feedback control for dynamic systems. It can tackle errors between the model and the plant (reality) by utilizing a feedback monitored-information. In our problem, this is a crucial issue due to the scatter in the MFD, mainly in the congested regime, as errors are expected between the MFD model and the plant. In addition, noise in the traffic demand is expected between predicted and actual demands.

The MPC has been used for optimization in different traffic control problems, e.g. ramp metering of freeway networks in Bellemans et al. (2006), variable speed limits and route guidance for freeway networks in Kotsialos et al. (2002) and Hegyi et al. (2005a,b), signal control for large-scale urban networks in Gartner et al. (2002), Aboudolas et al. (2010), and Lin et al. (2011), and mixed urban and freeway networks in van den Berg et al. (2007). A historical survey for industrial applications (other than traffic control) of MPC can be found in Qin and Badgwell (2003), while theoretical issues of MPC can be found in Garcia et al. (1989), Camacho and Bordons (1999), and Mayne et al. (2000).

For the sake of brevity, the full description of the solution method is not presented here, however, the reader can refer to Geroliminis et al. (2013) for further information. In the following, we present the outlines of the mixed network MPC through the block diagrams presented in Fig. 4.

On the top of Fig. 4, the "mixed urban-freeway plant" block presents the dynamic evolution of the mixed network in reality. This block can be implemented in several ways: (I) real field implementations, e.g. one can apply the proposed scheme on Yokohama network that experiences a well-defined MFD (as was found in Geroliminis and Daganzo (2008)) or Buisson and Ladier (2009), (II) simulation-based plant, e.g. using micro-simulations of the San-Francisco business district center presented in Geroliminis and Daganzo (2007), Ji et al. (2010) or Keyvan-Ekbatani et al. (2012), and (III) model-based plant, e.g. Geroliminis et al. (2013) and Daganzo (2007). In this paper, we follow the later to present the reality of mixed networks as the model (8)–(29) is utilized but with integrating noise in demand and errors in MFDs, similar to Eqs. (23)–(27) in Geroliminis et al. (2013), such that the dynamics evolutions of the plant and the model are not the same. Demand noise is considered fully random, while errors in the MFD have some heteroscedastic behavior, i.e. the error is expected to be larger when accumulation grows.

One of the main features of the MPC framework is the "rolling time horizon" (or receding horizon), see right of Fig. 4. The optimization technique is applied to finite time duration, called prediction horizon, which is much smaller than the total control process time. The control process time is covered by overlapping several time horizons, where at each time step the open-loop problem of the mixed urban-freeway model is optimized. Then, only the first control sequence is applied to the plant and the same procedure is carried out again till the final time of the problem.

The "MPC controller" (bottom of the figure) contains: (I) the mixed urban-freeway model (8)–(29) which is used to predict the dynamics of the mixed network, and (II) the optimization technique that minimizes the total delays of the network according to (17). The mixed urban-freeway model integrates the ACTM of the freeway and the two-region MFDs model of the urban network through the traffic flows (green arrows in the figure) according to (2), (6), and (7).

The MPC controller obtains the optimal control sequences for the current horizon by solving an optimization problem using the direct sequential method, also referred to single-shooting or control vector parametrization (CVP) in the literature, see Geroliminis et al. (2013). Note that several "control policies", see left-bottom of the figure, which will be introduced in Section 4, determine the level of coordination between the two control entities, and the control structure, e.g. centralized or decentralized.

The predicted O–D is given in advance, however, route choice demands are determined corresponding to the chosen strategy of route choice, i.e. a simple route choice (8)–(16) or system optimum as described in Section 6.

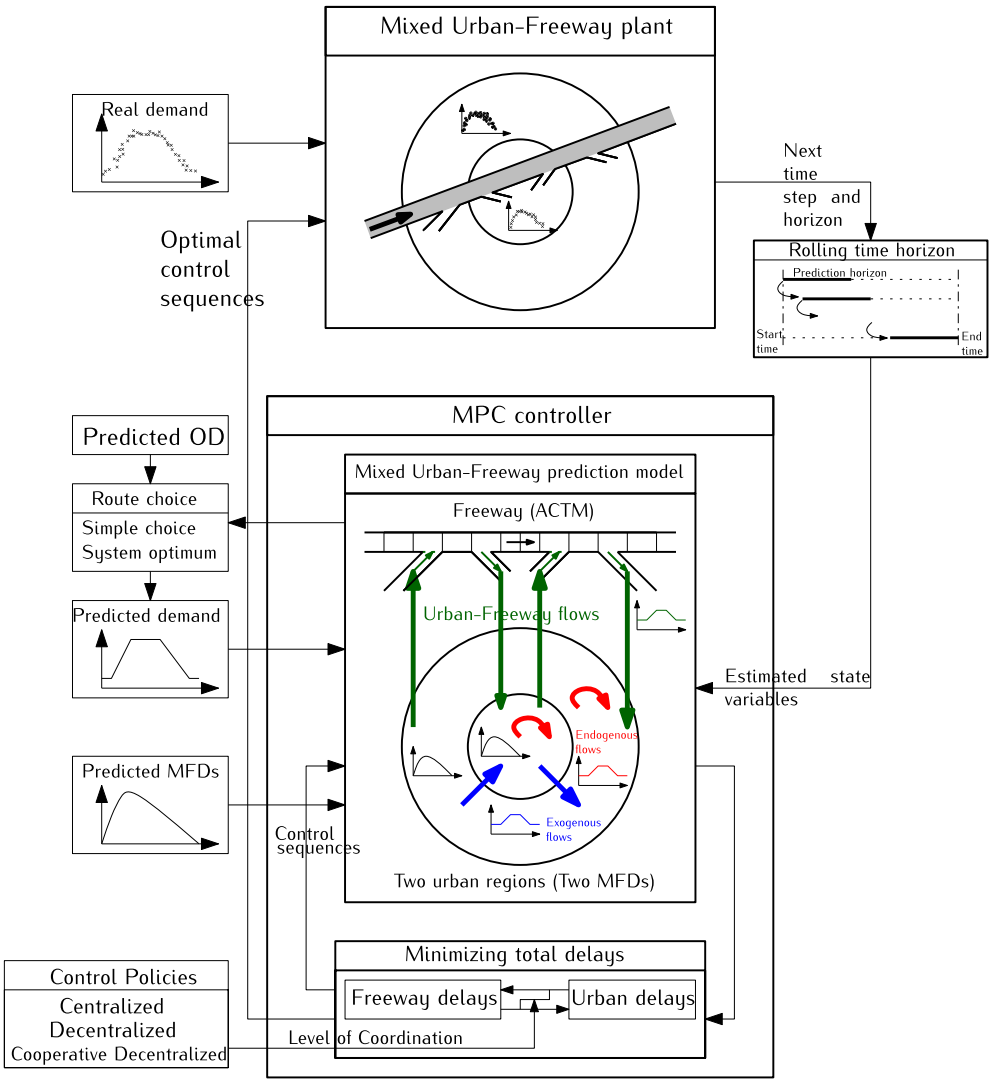

Fig. 4. Model Predictive Control scheme for mixed urban-freeway networks.

3.3. Test study example

In this subsection, we present a case study example to investigate the characteristics of the proposed MPC controller. This MPC controller is fully centralized and the comparison with other control policies such as decentralized and cooperative decentralized will be discussed later in Section 5. Note that the simple route choice model is integrated in the mixed network dynamics, while an SO route choice model will be integrated later in Section 6. In this example, the time varying demand is simulating one hour of morning peak followed by half an hour of low demand. Furthermore, the O–D demands, which are assumed to have trapezoidal shapes, are perturbed by a normal random component (see perturbed demand in Fig. 5a), in order to investigate the robustness of the MPC controller to the accuracy of demand prediction. The MPC controller tries to find the optimal control sequences assuming that the demands are without errors (trapezoidal shapes) in the optimization model, while the demands are applied to the plant, i.e. the mixed traffic network in reality, with errors (perturbed lines). Without loss of generality, we assume both regions have the same MFD consistent with the MFD observed in Yokohama, see Fig. 2 in Geroliminis and Daganzo (2008). While the MPC model assumes the MFD as aforementioned 3rd-order function (see the two coincide lines in Fig. 5b), the MFDs of the MPC plant are perturbed by a random component (see red and blue points in Fig. 5b), in order to both capture the scatter behavior of MFD and test the robustness of the proposed MPC controller to the modeling error of MFD.

The selected MPC controller parameters are: the prediction horizon $N_p = 20$, the control horizon $N_c = 2$, the control lower bound $u_{min} = 0.1$, and the upper bound $u_{max} = 0.9$. Furthermore, the freeway consists of 17 cells each has length of 0.5 (km) except the first cell which is long enough to accommodate all the vehicles in the entrance queue of the network.

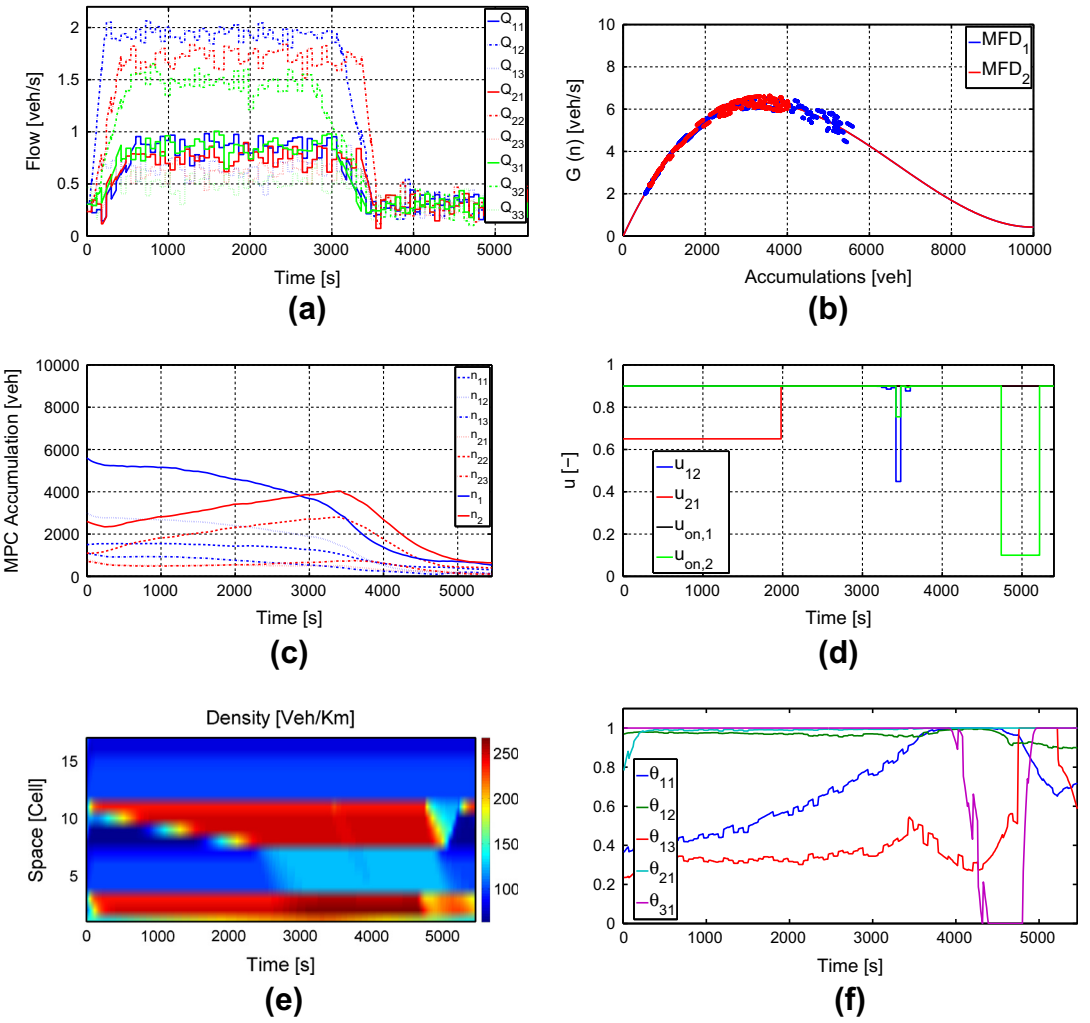

Fig. 5. The results of case study example for fully coordinated: (a) O–D demands, (b) urban region MFDs, (c) urban accumulations, (d) MPC control sequences, (e) freeway density contour, and (f) evolution of route choice parameters.

The parameters of the triangular fundamental diagram of the freeway cells are: the jam accumulation $x_{l,\max} = 125$ (veh/km/lane), the mainline capacity $\bar{f}_l = 2200$ (veh/h/lane), and the free flow speed equals to 88.5 (km/h). Other setups of the simulation are: the split ratios of off-ramps, $\beta_1 = 0.15$ and $\beta_2 = 0.25$ (–), the maximum queue length of on-ramps, $n_{\text{on},i,\max} = 300$ (veh), and the on-ramps capacity flow, $s_{\text{on},i} = 6000$ (veh/h). The values of the last two parameters might be considered high for a single on-ramp. Indeed, they represent a hyper-ramp per urban region (i.e. all ramps are grouped in one, by considering similar characteristics and queue lengths). Coordinated ramp metering strategies, e.g. Papamichail and Papageorgiou (2011) and Geroliminis et al. (2011), can be utilized to equally distribute the queues among the ramps. The reasoning behind this is that if multiple ramps are considered within each region, then a ramp choice model should be added to the developed formulation. More complex city structures with more urban regions and ramps is a future research priority.

In this numerical example, regions 1 and 2 are initially congested and uncongested, respectively, i.e. the initial accumulations are $n_1(0) = 5600$ (veh) and $n_2(0) = 2600$ (veh), and region 2 as the central business district attracts most of the trips. The evolution of accumulations $n_{ij}(k)$ over 1.5 (h) of simulation are presented in Fig. 5c. Note that at the beginning of the control process, the MPC controller decreases the total accumulation in region 1, $n_1(k)$, by choosing $u_{21}(k) = 0.65$ to restrict the flow from region 2 to region 1. Afterwards, the MPC controller tries to keep the both accumulations uncongested by changing u_{21} to $u_{\max}$ at $t = 2000$ (s) (see Fig. 5d) to let more vehicles enter region 1.

The effect of on-ramp controllers on the condition of freeway is more comprehensible with the help of Fig. 5e, where the density contour of freeway is illustrated. The main points of interest are cells with on-ramps (3 and 11), where actuating $u_{\max}$ for the MPC on-ramp controllers, $u_{\text{on},1}(k)$ and $u_{\text{on},2}(k)$, let more vehicles enter the freeway to avoid on-ramp spillbacks to the urban regions. Consequently, there are fewer vehicles in the urban regions and more vehicles queuing in the freeway, which seems sensible for mixed traffic network controllers to keep the vehicles in the freeway instead of urban network during the

rush hour. At $t = 4800$ (s), $u_{on,2}(k)$ switches to u_{min} for 400 (s) leading to a sudden decrement in the density of cell 11 (on-ramp) and its former cells. Note that the critical density of the cells is 99.4 (veh/km) and cells 7 and 15 have off-ramps.

In Fig. 5f, the time series of route choice parameters θ_{11}, θ_{12}, θ_{13}, θ_{21}, and θ_{31} are depicted. The enduring low value of urban travel time between regions 1 and 2 leads θ_{12} and θ_{21} to have almost constant values near to the maximum value. In contrast, θ_{11}, θ_{13}, and θ_{31} follow different trends. At the beginning of simulation as region 1 is congested, its speed is very low, thus about 65% of vehicles of the O–D demand from 1 to 1 prefer to choose the alternative trip route through the freeway ($1 \to 3 \to 1$). As the time passes, region 1 accumulation decreases and hence, the speed increases and more vehicles are in favor of choosing trip route $1 \to 1$ instead of $1 \to 3 \to 1$. The same logic is apparent for the case of θ_{13} and θ_{31}. The θ_{23} and θ_{32} are not depicted since they are constantly equal to 1 for all the examples, which is clear as e.g. traveling through trip route $2 \to 3$ is always faster than trip route $2 \to 1 \to 3$. Overall, the outcome of the MPC controller demonstrates that the traffic can be efficiently managed and accommodated in the mixed traffic network with the centralized MPC control scheme.

4. Control policies for the mixed network

In this section, several control policies are introduced and elaborated for the decentralized traffic control of the mixed urban-freeway network. The policies have different control structures and different levels of coordination between the freeway and the urban network control entities. Our objective is to investigate how coordination among different controllers can improve the mobility patterns of a city. In the following, seven control policies (CPs) are introduced from fully decentralized with no communication between urban and freeway entities to fully coordinated, with a comprehensive description for each policy: (CP1) ALINEA control for freeway and u_{max} for urban network (ALINEA + urban u_{max}), (CP2) ALINEA control with queue constraint for freeway and u_{max} for urban network (ALINEA Q + urban u_{max}), (CP3) ALINEA control for freeway and MPC for urban network (ALINEA + urban MPC), (CP4) ALINEA control with queue constraint for freeway and MPC for urban network (ALINEA Q + urban MPC), (CP5) Decentralized MPC (D-MPC), (CP6) Cooperative Decentralized MPC (CD-MPC), and (CP7) Centralized mixed network MPC (MPC).

In CP1, the ALINEA control law (Papageorgiou et al., 1991) is applied for the two on-ramps of the freeway, while the perimeter control inputs are set to their upper bounds in order to minimize the restriction for vehicle flows, i.e. $u_{12}(k) = u_{21}(k) = u_{max}$. The ALINEA policy for the on-ramp metering controllers determines the number of vehicles entering the freeway, $f_{on,l}(k+1)$, corresponding to the difference between the current accumulation of cell l and its desired accumulation, i.e. $x_l(k) - x_{ref}$, and a preset regulation parameter κ (–),

$$f_{on,l}(k+1) = f_{on,l}(k) + \kappa \cdot (x_l(k) - x_{ref}). \tag{30}$$

CP2 is similar to CP1, except that there are queue constraints on the on-ramps because of the limitation of queue space at the on-ramps. This policy can be considered the simplest step towards coordination as it tries to avoid spillbacks in the urban regions, but without analyzing the direct effect in system delays and outflow. Therefore, the ALINEA control law is modified to tackle the queue constraint as follows (Bellemans et al., 2006):

$$f_{on,l}(k+1) = \begin{cases} f_{on,l}(k) + \kappa \cdot (x_l(k) - x_{ref}) & \text{if } n_{on,i}(k) \leqslant \eta \cdot n_{on,i,max}, \\ s_{on,i} \cdot T_k & \text{otherwise}, \end{cases} \tag{31}$$

where $\eta \in [0,1]$. In this manner, the ALINEA controller with queue constraint tries to keep the length of on-ramp queue below a predefined ratio η of the maximum queue length $n_{on,i,max}$, which can effectively hinder the freeway spillbacks to the urban regions.

In CP3, the ramp metering controllers follow the ALINEA control law in (30), while the optimal perimeter control values are obtained from solving the mixed control problem using the MPC approach. This means that the MPC solver optimizes the total delay function in (17) *only* for $u_{12}(k)$, $u_{21}(k)$ while values of $u_{on,1}(k)$ and $u_{on,2}(k)$ for $k = 0, \ldots, K-1$ are determined separately according to (30).

CP4 is similar to CP3, except that the ramp metering controllers follow the ALINEA with queue constraints (control law in (31)). In this control policy, the optimal perimeter control is obtained from solving the mixed control problem using the MPC approach, as in CP3 but values for $u_{on,1}(k)$, $u_{on,2}(k)$ are determined separately according to (31).

CP5 is the Decentralized MPC (D-MPC) policy that considers the freeway and urban network controllers as two separate entities without coordination between them, i.e. the D-MPC policy applies two parallel and separate MPC problems for the freeway and the urban network. The D-MPC splits the optimal control problem of the mixed network to two smaller-sized optimization problems: (i) freeway total delay minimization by manipulating *only* the ramp metering controllers, and (ii) urban total delay minimization by manipulating *only* the perimeter controllers. In this manner, the objective function (17) is divided to two terms; the urban total delay term, i.e. $T_k \cdot \left[\sum_{k=0}^{K-1}[n_1(k) + n_2(k)]\right]$, which is optimized solely for the perimeter controllers, $u_{12}(k)$ and $u_{21}(k)$, and the freeway total delay term, i.e. $T_k \cdot \left[\sum_{k=0}^{K-1}\sum_{l=1}^{L} x_l(k) + \sum_{k=0}^{K-1}\sum_{i=1}^{2} n_{on,i}(k)\right]$, which is minimized solely for the on-ramp metering control inputs $u_{on,1}(k)$ and $u_{on,2}(k)$. Under this policy the two different objectives might be conflicting.

In order to improve the performance of the D-MPC policy, we propose a modification to establish a cooperation between the D-MPC control entities in CP6. The modification is such that each of the two freeway and urban controllers tries to minimize the whole network total delay as formulated in (17). Meaning that, the urban controller minimizes (17) for the current urban optimization horizon *only* by manipulating the perimeter controllers, $u_{12}(k)$ and $u_{21}(k)$, as the ramp metering controllers, $u_{on,1}(k)$ and $u_{on,2}(k)$, are assumed to be known and constant from the previous step of freeway optimization. Likewise, the freeway controller minimizes (17) for the current freeway optimization horizon *only* by manipulating the on-ramp meterings, $u_{on,1}(k)$ and $u_{on,2}(k)$, as the perimeter controllers, $u_{12}(k)$ and $u_{21}(k)$, are assumed to be known and constant from the previous step of urban optimization. This policy is named Cooperative Decentralize MPC (CD-MPC) because the two control entities are not centralized and fully coordinated, yet they communicate with partial information in a way to achieve the mixed network optimal performance. The CD-MPC policy is crucial in case of limited traffic monitoring or data communication.

CP7 is a fully centralized MPC controller for the whole mixed network consistent with the formulated problem (1)–(29).

5. Comparison of control policies

We test all the previous control policies for further investigation on the same example setup, as described in Section 3.3. The total delay of all policies for the two urban regions, the freeway, and the on-ramps are summarized in Table 3. The primary control policy to inspect is the integration of ALINEA freeway ramp metering and u_{max} for the urban perimeter controllers. The corresponding evolution of urban accumulations over time is depicted in Fig. 6a, revealing that region 1

Table 3
Total delay (veh·s· 10^5).

	ALINEA + *Urbanu*$_{max}$	*ALINEAQ* + *Urbanu*$_{max}$	ALINEA + Urban MPC	ALINEA Q + Urban MPC	D-MPC	CD-MPC	MPC
Region 1	524	438	504	262	537	254	196
Region 2	121	97.2	162	160	279	133	143
Freeway	129	172	129	189	124	217	218
On-Ramp 1	13.2	12.8	12.6	14.8	16.0	6.55	7.84
On-Ramp 2	16.0	16.0	15.8	16.2	16.3	16.1	15.7
Network	803	736	823	642	972	626	581

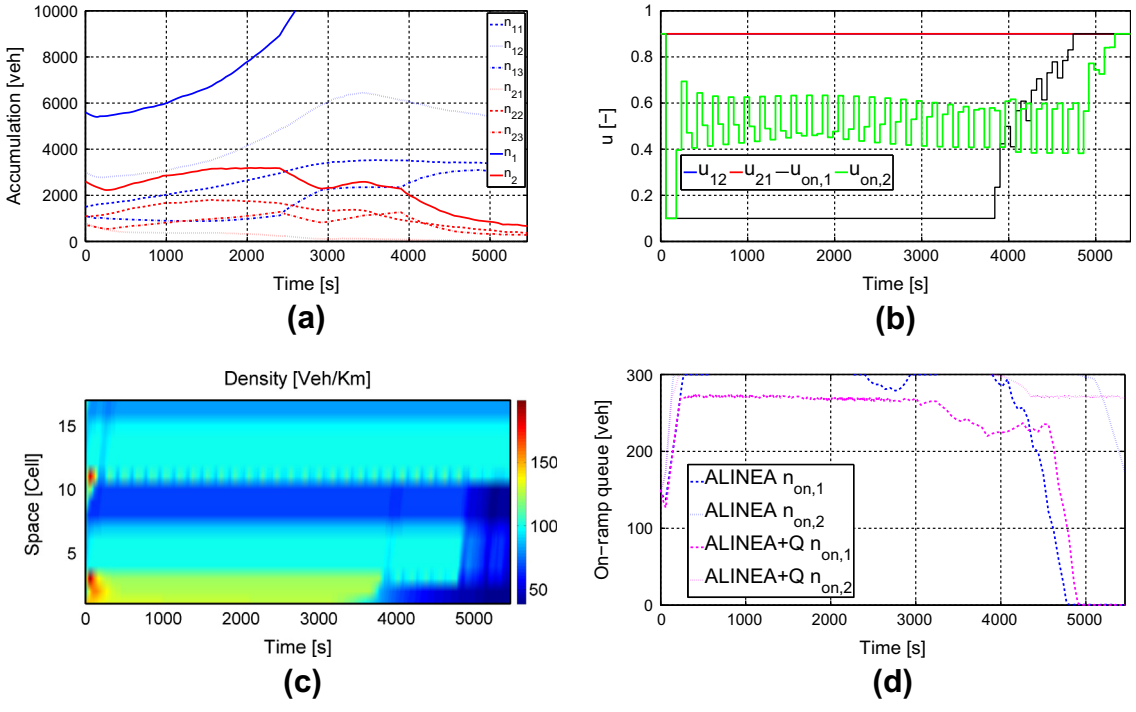

Fig. 6. The results of case study example for CP1 (ALINEA ramp metering and u_{max} urban perimeter control): (a) urban accumulations, (b) control sequences, (c) freeway density contour, and (d) comparison of on-ramp queues for CP1 and CP2 (ALINEA with queue constraint ramp metering).

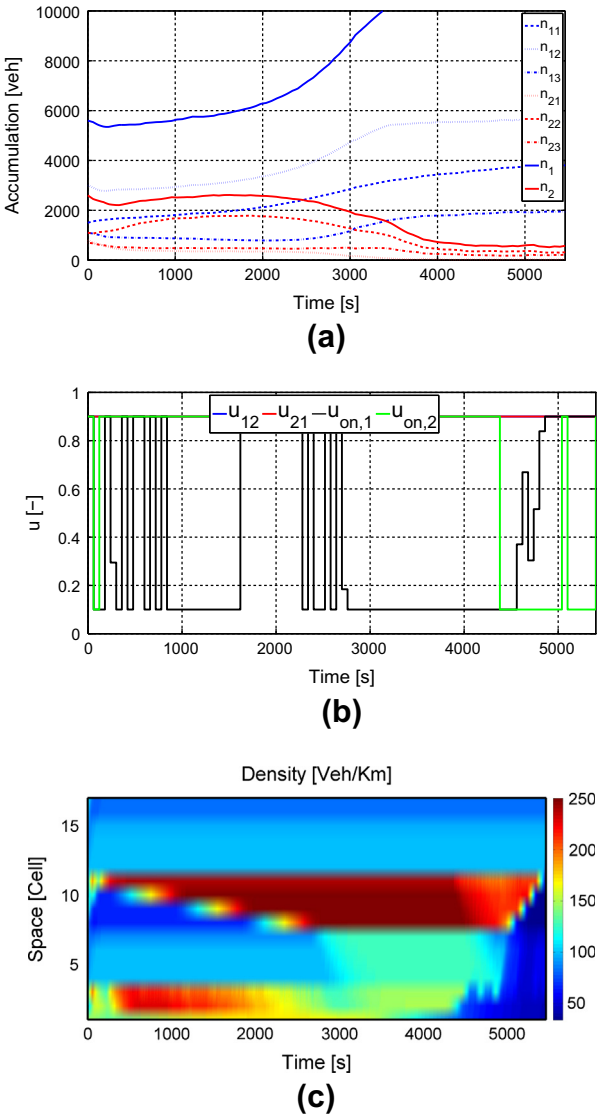

Fig. 7. The results of case study example for CP2 (ALINEA with queue constraint ramp metering and u_{max} urban perimeter control): (a) urban accumulations, (b) control sequences, and (c) freeway density contour.

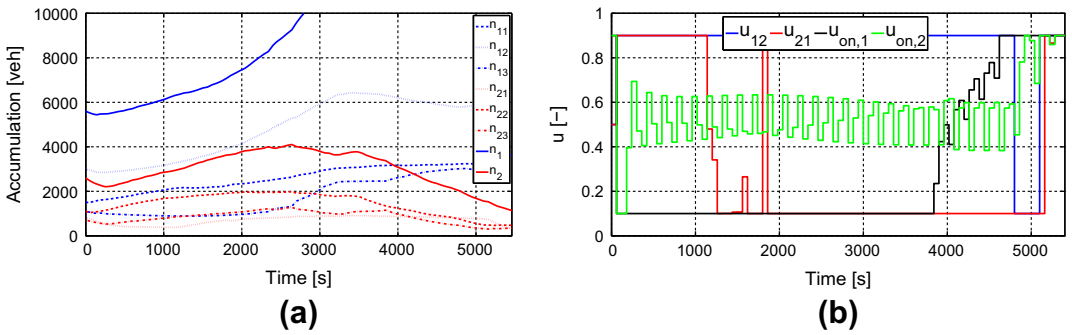

Fig. 8. The results of case study example for CP3 (ALINEA ramp metering and MPC urban perimeter control): (a) urban accumulations and (b) control sequences.

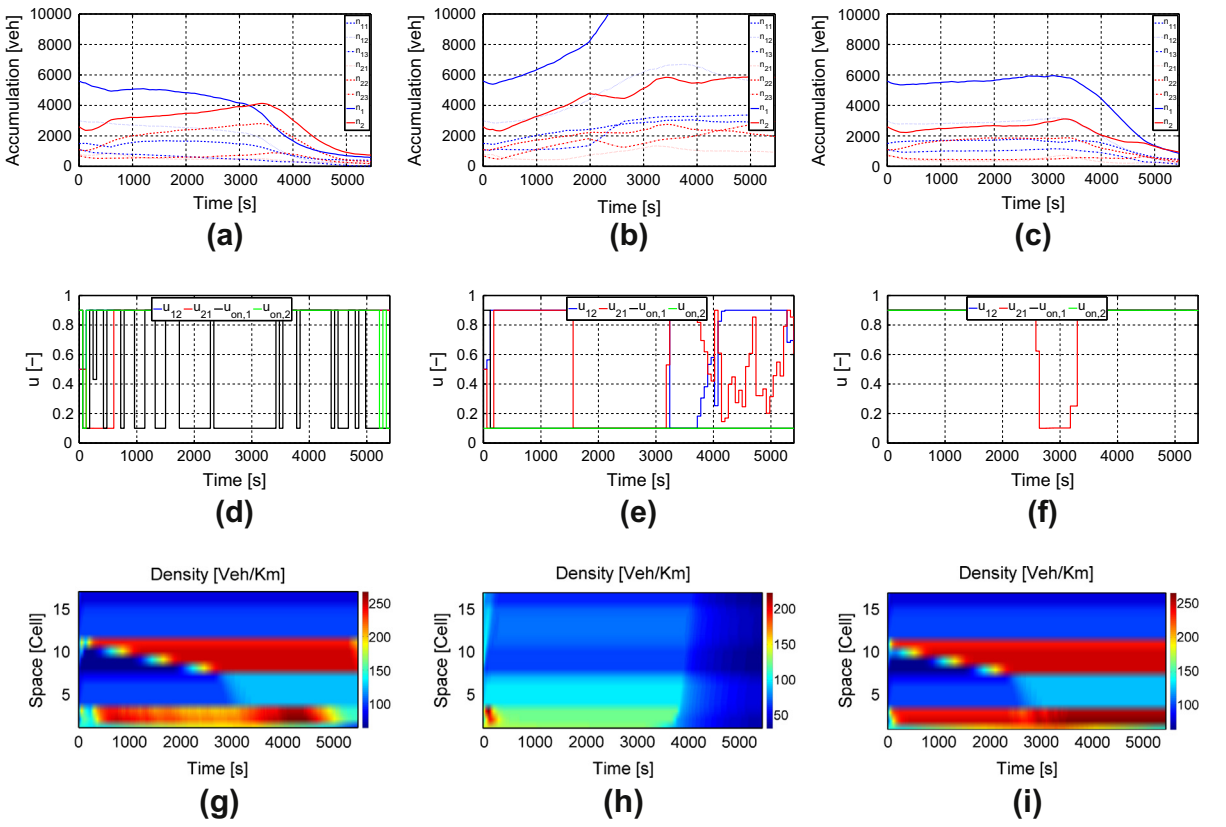

Fig. 9. The results of case study example for (i) CP4 (ALINEA with queue constraint ramp metering and MPC urban perimeter control): (a) urban accumulations, (d) control sequences, (g) freeway density contour; (ii) CP5 (decentralized MPC control (D-MPC)): (b) urban accumulations, (e) control sequences, (h) freeway density contour; and (iii) CP6 (cooperative decentralized MPC control (CD-MPC)): (c) urban accumulations, (f) control sequences, (i) freeway density contour.

encounters the gridlock condition. The ALINEA control sequences are illustrated in Fig. 6b demonstrating a jumpy non-smooth trend as a result of density alternation of corresponding on-ramp cells 3 and 11. Fig. 6c shows the density contour of freeway cells. The cell 3 density is over the desired value from $t = 0$ to $t = 3800$ (s), which forces the ALINEA controller $u_{on,1}$ be equal to u_{min}. As the time passes on, the cell 3 density decreases because of the reduction in the demands, which results in $u_{on,1}$ increasing jumps to regulate the density of cell 3 at the desired value. The same reasoning is valid for on-ramp 2 while the density of cell 11 is always around its preset desired value, which makes the $u_{on,2}$ going up and down to finely regulate the cell 11 density. It is apparent that the freeway is under-utilized because of local control scope of ALINEA, that prevents the vehicle to use the freeway and keeps them in the urban regions, which increases the chance of urban gridlock occurrence.

The on-ramp queues for both ALINEA and ALINEA with queue constraint freeway controllers are depicted in Fig. 6d. It is apparent that using ALINEA ramp metering law, both the on-ramp queues reach the maximum possible queue length, however, integrating queue constraint within ALINEA forces the queues to be less than the maximum value and be equal to the predefined queue length threshold. For on-ramp 1 this phenomenon is obvious, but for on-ramp 2, since its input demand is so high that even with $u_{on,2} = u_{max}$ this constraint cannot be met during the rush hour. After $t = 4300$ (s) with reduction in the demand, the on-ramp 2 queue decreases to the predefined queue length and below. The corresponding urban accumulations and control sequences of ALINEA freeway controller with queue constraint and u_{max} urban perimeter controller are shown in Fig. 7a and b, respectively. The freeway density contour, Fig. 7c, demonstrates that the activation of queue constraint permits more vehicles to enter the freeway, which increases the freeway accumulation and consequently, less chance of urban gridlock and less total delay for the whole network.

The two aforementioned control policies with the urban perimeter fixed control lead the urban network to gridlock. Thus, we propose to replace the fixed controllers with the MPC controllers such that the new control policy consists of ALINEA controller for freeway and MPC for urban perimeter control (CP3). The corresponding evolution of urban accumulations is depicted in Fig. 8a and once compared with Fig. 6a reveals that the time that region 1 reaches gridlock is slightly increased, which is not beneficial. The control sequences are illustrated in Fig. 8b showing the same jumpy non-smooth trend similar to combination of ALINEA with u_{max} urban perimeter control. The same trend of on-ramp queues and freeway density is also seen which indicates that the integration of freeway ALINEA and urban MPC is not fruitful for the mixed network traffic control and higher coordination level is needed between the two entities.

In contrast, the performance of integration of urban MPC approach and ALINEA freeway ramp metering with queue constraint (CP4) demonstrates an advantageous control policy. The urban accumulations are depicted in Fig. 9a in which both regions end at the uncongested regime implying the effectiveness of the traffic control policy. To investigate more, control sequences are shown in Fig. 9d and once compared to Fig. 7b, the MPC urban perimeter controllers are different while ALINEA controllers are to some extent the same, still manifests lack of cooperation between urban and freeway controllers. The u_{21} starts with the value of 0.5 for one control step then switches to $u_{\min}$ during which the accumulation of region 1 decreases and the accumulation of region 2 increases. Note the accumulations between $t = 0$ and $t = 600$ (s) in Figs. 9a and 7a. This not very significant alternation in the urban perimeter control values makes region 1 accumulation decreases from the edge of gridlock to a more stable regime, which yields to stability of both regions during the peak hour. In addition, by inspecting the freeway density contour, Fig. 9g, we observe more vehicles traveling along the freeway which is the reason that the urban regions could operate at the uncongested regime.

To improve the performance of the above control policies, the D-MPC controller which consists of one urban and one freeway control entities is introduced and tested on the same case study example. Each of the urban and freeway controllers tries to minimize the total delay *only* in their subnetwork of interest, hence, the freeway MPC controller forces both the ramp metering controllers, i.e. $u_{\text{on},1}$ and $u_{\text{on},2}$, to be equal to $u_{\min}$ as shown in Fig. 9e. This prevents the vehicles to use the freeway for commuting, causes the freeway to be underutilized (Fig. 9h) and consequently increases the accumulation in the urban regions rapidly which leads to urban gridlock (Fig. 9b). To avoid the competitiveness between the D-MPC control entities, we add an additional level of coordination that adjusts the objective function of MPC entities to modify them into CD-MPC configuration. The results in Fig. 9c, f, and i show that the CD-MPC can control the traffic demand very efficiently compared to D-MPC. Table 3 summarizes the numerical results of all control policies averaged over 10 runs demonstrating that the best policy is the centralized MPC. Yet for cases with larger number of state variables, e.g. cases with more urban regions and/or more freeway cells, the MPC optimization module is not tractable. In this case or when the centralized MPC controller is not applicable because of technical limitations, the proposed CD-MPC controller, which consists of two smaller sized control entities seems promising with insignificant performance loss (less than 7%).

Note that the performances of CPs (percentage of improvement) are fully dependent on the level of congestion, e.g. if urban regions are uncongested the marginal improvement of CP6 (CD-MPC) or CP7 is not significant. This statement was verified as we have run the same example with 10% less demand of O–Ds $1 \to 2$ and $2 \to 2$. The results show that with less demand still the CP1, CP3, and CP5 cannot handle the traffic load and lead the urban regions to high congestion. The improvements of CP6 (CD-MPC) and CP7 (MPC) over the CP2 (ALINEA Q + urban $u_{\max}$) are about 3% and 4%, respectively, meaning that the simple cooperative CP2 is good enough to control this scenario. Though in the base scenario, the corresponding improvements are 15% and 21%, respectively.

6. System optimum route choice in the mixed network

The introduced simple route choice modeling in Section 2.3 is based on the user equilibrium (UE) assumption, indicating that the drivers (users) choose the route with the smallest travel time (cost) among their route choice set. We model the estimation of route travel time based on the current condition of the mixed network, while alternative estimation models (e.g. integrating prediction to capture future evolution of the network) have also great potential for future investigation. Independently of route travel time estimation modeling, the UE assumption still holds as a sound description of non-cooperative route choice behavior of drivers. It is evident that from the stand point of the system, i.e. mixed transport network, the simple route choice cannot provide the optimal solution. Nevertheless, the proposed MPC framework provides the opportunity to integrate the system optimum (SO) route choice in the mixed network traffic control problem. Though, the real-world implementation of SO traffic control policy is not fully operational, because of inherent unpredictability of human behavior (Jahn et al., 2005), theoretically, we assume that it is possible to fully control (guide) the vehicles to a specific route to commute their origin–destination trip. Even if this is not fully implementable with traffic management schemes (variable message signs, pricing, etc.), it provides a comparison of what can be the most efficient way to control a transport network with a given demand and infrastructure.

In the following, we implement an SO route choice model only for the CP7–Centralized mixed network MPC (and not for all control policies presented in Section 4), since it provides the upper bound of total improvement. We consider the same example setup to test the implementation of SO route choice model. In this model, the route choice proportions, i.e. θ_{ij} for various origin destinations, are regarded as the output of the MPC optimization module, contrary to the simple route choice model where route choice proportions are the input to the optimization module. Thus, adjustments of mixed traffic dynamic equations are confined to omission of (8)–(16) and instead addition of all the θ_{ij} as optimization variables in the objective function (17). The results of the mixed network control with SO route choice model, averaged over 10 runs and presented

Table 4
Total delay for MPC fully coordinated controller with system optimum route choice (veh·s·10^5).

Region 1	Region 2	Freeway	On-Ramp 1	On-Ramp 2	Network
167	118	209	10.5	7.60	512

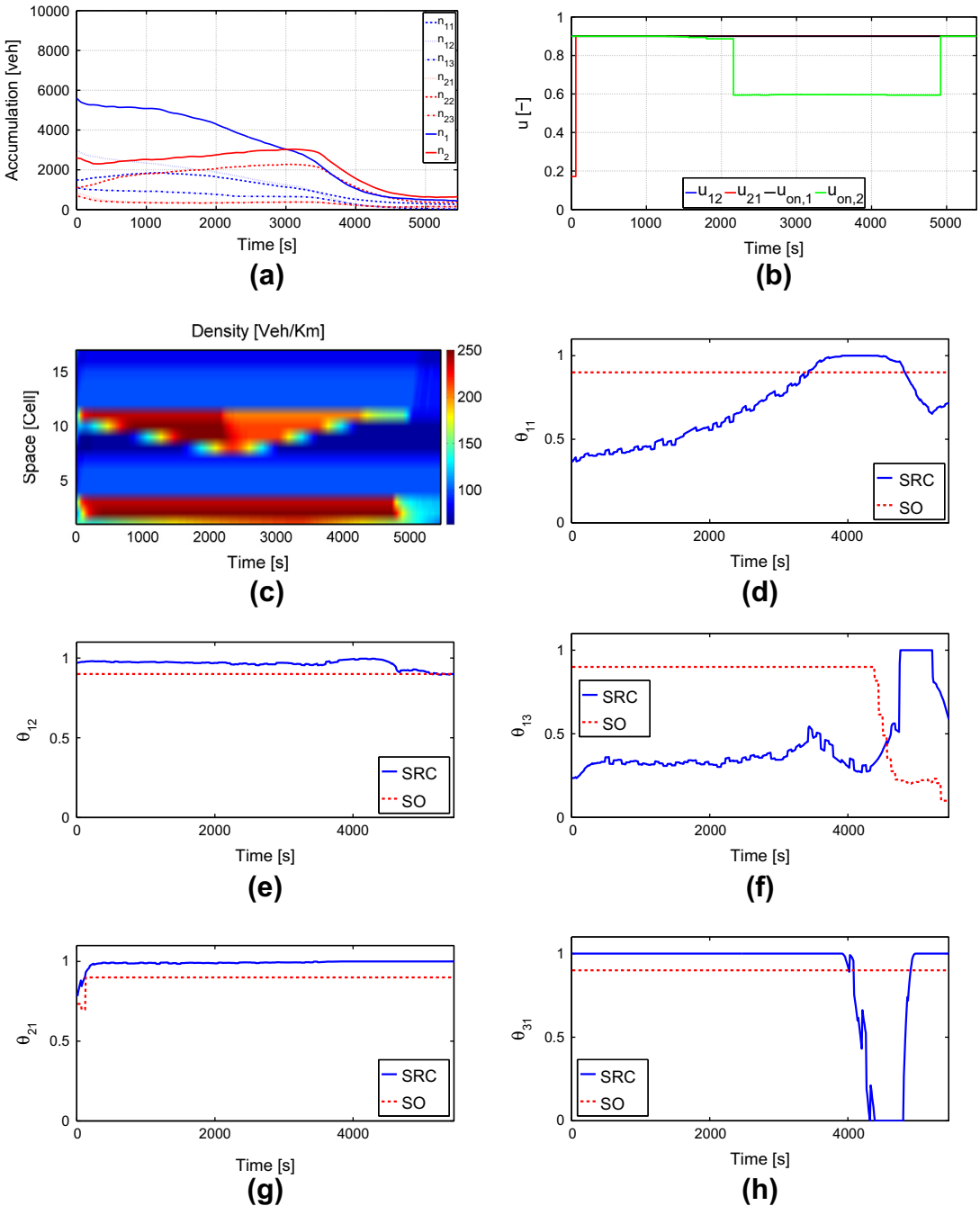

Fig. 10. The results for the augmented MPC controller with system optimum route choice model: (a) urban accumulations, (b) control sequences, (c) freeway density contour, and the route choice results corresponding to the system optimum (SO) and the simple route choice model (SRC) for: (d) θ_{11}, (e) θ_{12}, (f) θ_{13}, (g) θ_{21}, (h) θ_{31}.

in Table 4, indicating 12% improvement over the simple route choice model, in terms of decreasing the network total delay. Fig. 10a–c depict the urban accumulations, the control sequences, and the freeway density contour of the augmented MPC controller with SO route choice model, respectively.

The θ_{11}, θ_{12}, θ_{13}, θ_{21}, and θ_{31} corresponding to both SO and the simple route choice models are illustrated in Fig. 10d, e, f, g, and h, respectively. It is evident that most of the time the SO route choice model forces the vehicles to use their first trip route choice, which is the direct one without any region to region or region to freeway transfer.

7. Conclusions and future work

The control problem of a large-scale mixed traffic network, consisting of two urban regions with MFD representation, and one alternative freeway route modeled with the asymmetric cell transmission model is formulated. For traffic control purposes, we consider two controllers on the perimeter of regions to manipulate the urban inter-transfer flows; in addition, two on-ramp controllers to control the traffic flow from urban regions to the freeway. The optimal traffic control problem is solved by an MPC scheme. Several control policies with different controller structures and levels of urban-freeway coordination are introduced and scrutinized. The results demonstrate the advantage of centralized control over combination of a simple freeway ramp metering with urban MPC controller, and also the great importance of cooperation in decentralized MPC approach in cases with lack of full data communication and coordination between urban and freeway control entities, i.e. when the centralized MPC is not tractable. In addition, the system optimum route choice is integrated within the centralized MPC, which leads to increase the network performance by 12%.

These results can be beneficial for municipal administrators to develop efficient hierarchical control strategies for metropolitan mixed traffic networks. This work contributes one step forward towards achieving SoS approach for transportation infrastructure and networks. The outcome of this optimization does not provide the exact phase settings for traffic signals in the boundary between the two regions. Nevertheless, recent work for single (Keyvan-Ekbatani et al., 2012) and multiple regions (Aboudolas and Geroliminis, 2013) provide the necessary tools to dynamically change the signal settings to meet the controllers' inputs u_{12} and u_{21} from the aggregated optimization. In case that local queues are developed in the proximity of the controllers (ramps and boundaries between urban regions), analysis of Geroliminis and Boyacı (2012) can identify signal parameters in the individual regions of a city in such a way to move traffic smoothly at the desired flows, without concentrating a large number of vehicles at the boundaries of the regions. Traffic control problem of networks with more complex structure and dynamic traffic assignment is ongoing research. Ongoing work also involves the development of efficient control strategies for networks with a larger number of urban regions, on- and off-ramps. In this case, the size of regions might change over time due to congestion propagation, which will require a dynamic partitioning of the city in different parts. This is a challenging problem, both from an optimization (higher computational effort) and modeling (more complex dynamics). A field implementation and/or detailed micro-simulation of proposed control policies are important to identify the effect of the control and the heterogeneity of spatial distribution of congestion on the shape of the MFD and the efficiency of the control method. Another research direction is related to traffic monitoring and the objective is to process real-time data from multiple sensors to estimate the necessary variables (accumulations, travel times etc.) involved in the control problem.

Acknowledgments

This research was financially supported by the Swiss National Science Foundation (SNSF) Grant # 200021-13016. The authors thank Dr. Gabriel Gomez and Dr. Alex Kurzhanskiy from UC Berkeley for their suggestions and help with CTMSim Simulator.

References

Aboudolas, K., Geroliminis, N., 2013. Feedback perimeter control for multi-region and heterogeneous congested cities. In: Transportation Research Board 92nd Annual Meeting, Washington, DC.

Aboudolas, K., Papageorgiou, M., Kouvelas, A., Kosmatopoulos, E., 2010. A rolling-horizon quadratic-programming approach to the signal control problem in large-scale congested urban road networks. Transportation Research Part C 18 (5), 680–694.

Bellemans, T., De Schutter, B., Moor, B.D., 2006. Model predictive control for ramp metering of motorway traffic: a case study. Control Engineering Practice 14 (7), 757–767.

Buisson, C., Ladier, C., 2009. Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundamental diagrams. Transportation Research Record 2124, 127–136.

Camacho, E.F., Bordons, C., 1999. Model Predictive Control. Springer-Verlag, Berlin, Germany.

Daganzo, C.F., 1994. The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. Transportation Research Part B 28 (4), 269–287.

Daganzo, C.F., 2007. Urban gridlock: macroscopic modeling and mitigation approaches. Transportation Research Part B 41 (1), 49–62.

Daganzo, C.F., Gayah, V.V., Gonzales, E.J., 2011. Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability. Transportation Research Part B 45 (1), 278–288.

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

Garcia, C.E., Prett, D.M., Morari, M., 1989. Model predictive control: theory and practice – a survey. Automatica 25 (3), 335–348.

Gartner, N.H., Pooran, F.J., Andrews, C.M., 2002. Optimized policies for adaptive control strategy in real-time traffic adaptive control systems, implementation and field testing. Transportation Research Record 1811, 148–156.

Gayah, V.V., Daganzo, C.F., 2011. Clockwise hysteresis loops in the macroscopic fundamental diagram: an effect of network instability. Transportation Research Part B 45 (4), 643–655.

Geroliminis, N., 2009. Dynamics of peak hour and effect of parking for congested cities. In: Transportation Research Board 88th Annual Meeting. No. 09-1685. Washington, DC.

Geroliminis, N., Boyacı, B., 2012. The effect of variability of urban systems characteristics in the network capacity. Transportation Research Part B 46 (10), 1607–1623.

Geroliminis, N., Daganzo, C.F., 2007. Macroscopic modeling of traffic in cities. In: Transportation Research Board 86th Annual Meeting, Washington, DC, Paper No. 07-0413.

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

Geroliminis, N., Haddad, J., Ramezani, M., 2013. Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: a model predictive approach. IEEE Transactions on Intelligent Transportation Systems 14 (1), 348–359.
Geroliminis, N., Srivastava, A., Michalopoulos, P., 2011. A dynamic-zone-based coordinated ramp-metering algorithm with queue constraints for minnesota's freeways. IEEE Transactions on Intelligent Transportation Systems 12 (4), 1576–1586.
Geroliminis, N., Sun, J., 2011. Hysteresis phenomena of a macroscopic fundamental diagram in freeway networks. Transportation Research Part A 45 (9), 966–979.
Godfrey, J.W., 1969. The mechanism of a road network. Traffic Engineering and Control 11 (7), 323–327.
Gomes, G., Horowitz, R., 2006. Optimal freeway ramp metering using the asymmetric cell transmission model. Transportation Research Part C 14 (4), 244–262.
Haddad, J., Geroliminis, N., 2012. On the stability of traffic perimeter control in two-region urban cities. Transportation Research Part B 46 (1), 1159–1176.
Hegyi, A., De Schutter, B., Hellendoorn, H., 2005a. Model predictive control for optimal coordination of ramp metering and variable speed limits. Transportation Research Part C 13 (3), 185–209.
Hegyi, A., De Schutter, B., Hellendoorn, J., 2005b. Optimal coordination of variable speed limits to suppress shock waves. IEEE Transactions on Intelligent Transportation Systems 6 (1), 102–112.
Helbing, D., 2009. Derivation of a fundamental diagram for urban traffic flow. The European Physical Journal B 70 (2), 229–241.
Jahn, O., Mhring, R.H., Schulz, A.S., Moses, N.E., 2005. System-optimal routing of traffic flows with user constraints in networks with congestion. Operations Research 53 (4), 600–616.
Ji, Y., Daamen, W., Hoogendoorn, S., Hoogendoorn-Lanser, S., Qian, X., 2010. Macroscopic fundamental diagram: Investigating its shape using simulation data. Transportation Research Record 2161, 42–48.
Ji, Y., Geroliminis, N., 2012. On the spatial partitioning of urban transportation networks. Transportation Research Part B 46 (10), 1639–1656.
Keyvan-Ekbatani, M., Kouvelas, A., Papamichail, I., Papageorgiou, M., 2012. Exploiting the fundamental diagram of urban networks for feedback-based gating. Transportation Research Part B 46 (10), 1393–1403.
Knoop, V.L., Hoogendoorn, S.P., Van Lint, J.W.C., 2012. Routing strategies based on the macroscopic fundamental diagram. Transportation Research Record 2315, 1–10.
Kotsialos, A., Papageorgiou, M., Mangeas, M., Haj-Salem, H., 2002. Coordinated and integrated control of motorway networks via non-linear optimal control. Transportation Research Part C 10 (1), 65–84.
Lin, S., De Schutter, B., Xi, Y., Hellendoorn, H., 2011. Fast model predictive control for urban road networks via MILP. IEEE Transactions on Intelligent Transportation Systems 12 (3), 846–856.
Mahmassani, H., Williams, J., Herman, R., 1987. Performance of urban traffic networks. In: Gartner, N., Wilson, N. (Eds.), Proceedings of the 10th International Symposium on Transportation and Traffic Theory. Elsevier, Amsterdam, The Netherlands.
Mayne, D.Q., Rawlings, J.B., Rao, C.V., Scokaert, P.O.M., 2000. Constrained model predictive control: stability and optimality. Automatica 36 (6), 789–814.
Mazloumian, A., Geroliminis, N., Helbing, D., 2010. The spatial variability of vehicle densities as determinant of urban network capacity. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 368 (1928), 4627–4647.
Olszewski, P., Fan, H.S.L., Tan, Y.-W., 1995. Area-wide traffic speed-flow model for the Singapore CBD. Transportation Research Part A 29A (4), 273–281.
Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., Wang, Y., 2003. Review of road traffic control strategies. Proceedings of the IEEE 91 (12), 2043–2067.
Papageorgiou, M., Haj-Salem, H., Blosseville, J., 1991. ALINEA a local feedback control law for on-ramp metering. Transportation Research Record 1320, 58–64.
Papageorgiou, M., Kotsialos, A., 2002. Freeway ramp metering: an overview. IEEE Transactions on Intelligent Transportation Systems 3 (4), 271–281.
Papamichail, I., Papageorgiou, M., 2011. Balancing of queues or waiting times on metered dual-branch on-ramps. IEEE Transactions on Intelligent Transportation Systems 12 (2), 438–452.
Prashker, J.N., Bekhor, S., 2000. Some observations on stochastic user equilibrium and system optimum of traffic assignment. Transportation Research Part B 34 (4), 277–291.
Qin, S.J., Badgwell, T.A., 2003. A survey of industrial model predictive control technology. Control Engineering Practice 11 (7), 733–764.
Saberi, M., Mahmassani, H., 2012. Exploring properties of network-wide flow-density relations in a freeway network. In: Transportation Research Board 91st Annual Meeting, Washington, DC.
Samad, T., Annaswamy, A.M. (Eds.), 2011. The Impact of Control Technology. IEEE Control Systems Society. <www.ieeecss.org>.
van den Berg, M., Hegyi, A., De Schutter, B., Hellendoorn, J., 2007. Integrated traffic control for mixed urban and freeway networks: a model predictive control approach. European Journal of Transport and Infrastructure Research 7 (3), 223–250.
Wardrop, J.G., 1952. Some theoretical aspects of road traffic research. Proceedings of the Institution of Civil Engineers, Part II 1 (2), 352–362.
Zhang, L., Garoni, T., de Gier, J., 2013. A comparative study of macroscopic fundamental diagrams of arterial road networks governed by adaptive traffic signal systems. Transportation Research Part B 49, 1–23.