Delft University of Technology

Delft Center for Systems and Control

Technical report 04-002

Anticipative ramp metering control for freeway traffic networks*

T. Bellemans, B. De Schutter, and B. De Moor

If you want to cite this report, please use the following reference instead:
T. Bellemans, B. De Schutter, and B. De Moor, "Anticipative ramp metering control for freeway traffic networks," *Proceedings of the 16th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2004)*, Leuven, Belgium, 12 pp., July 2004. Paper 320.

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.24.73 (secretary)
URL: https://www.dcsc.tudelft.nl

Anticipative Ramp Metering Control for Freeway Traffic Networks

T. Bellemans[†], B. De Schutter[*], and B. De Moor[†]

[†] SISTA (SCD), ESAT, K.U.Leuven, Kasteelpark Arenberg 10, B-3001 Heverlee (Leuven), Belgium, email: {tom.bellemans,bart.demoor}@esat.kuleuven.ac.be

[*] Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands, email: b.deschutter@dcsc.tudelft.nl

Abstract— **We develop an anticipative model-based predictive traffic control approach for ramp metering in freeway traffic networks. If ramp metering is implemented in a freeway network with alternative routes, traffic can spontaneously re-route due to the response of the drivers to the applied control actions. Although spontaneous re-routing can have a significant influence on the resulting traffic situation in the traffic network and on the performance of the traffic network as a whole, re-routing is usually not automatically included in current freeway traffic control frameworks. In this paper, we develop a new method to efficiently calculate and incorporate re-routing effects into a model-based predictive traffic control framework. In this way, anticipative model predictive control for ramp metering in freeway networks is realized. Note that although the focus in this paper is on ramp metering, the approach presented here can also be used for other traffic control measures.**

I. Introduction

A. Overview

Due to the ever increasing need for transportation and mobility more and more cities and regions all around the world face considerable traffic congestion problems. Traffic jams do not only cause considerable costs due to unproductive time losses; they also augment the possibility of accidents, and they have a negative impact on the environment and on the quality of life. On the short term one of the most effective measures against traffic congestion seems to be a better control of the traffic flows by dynamic traffic management. In this paper we consider a model-based approach to determine the optimal settings for the dynamic traffic management measures in freeway networks, where we also take the reaction of the drivers to the traffic measures into account.

In a freeway network with multiple routes from the origins to the destinations, drivers have to choose their route in the network. Given the traffic demands for each origin-destination pair and the topology of the network, the traffic needs to be assigned to the routes before a simulation can be run. We present anticipative ramp metering control, which takes the re-routing effects due to ramp metering into account. Note that although the focus in this paper is on ramp metering, the presented technique can also be applied to other traffic control measures.

This paper is organized as follows. First, a static equilibrium traffic assignment algorithm that assigns the traffic to the different routes based on the collective behavior of the drivers is presented. Next, a new dynamic traffic assignment method based on the static traffic assignment method is proposed. This dynamic traffic assignment method is incorporated in a model predictive control (MPC) framework to realize anticipative MPC-based control that takes the spontaneous re-routing of drivers due to control measures into account.

Before presenting the anticipative MPC traffic control, we first briefly discuss the MPC control approach and the traffic model we will use as prediction model, viz. the METANET model. We also explain how MPC can be used for (regular) ramp metering control.

B. Model predictive control

Model predictive control (MPC) [3], [7], [14] is an on-line model-based predictive control design strategy that has its roots in the process industry but that can also be applied for traffic control. A main advantage of MPC is that constraints such as maximal on-ramp queue lengths, minimal and maximal metering rates, etc. can easily be included in the control design.

Briefly, MPC works as follows. At a given time $t = kT_{\mathrm{ctrl}}$, where k is the control step and T_{ctrl} is the control sample time (typical values for T_{ctrl} are 1 or 5 min), the MPC controller uses a prediction model and (numerical) optimization to determine the optimal control sequence $u^*(k), \ldots, u^*(k+N_{\mathrm{p}}-1)$ that minimizes a given performance indicator over the time horizon $[kT_{\mathrm{ctrl}}, (k+N_{\mathrm{p}})T_{\mathrm{ctrl}}]$ based on the current state of the traffic network and on the expected demands over this period, where N_{p} is called the prediction horizon. This is combined with a receding horizon approach in which at each control step only the first control input $u^*(k)$ is applied to the system; next, the horizon is shifted, new measurements are made, and the process is repeated all over again. In this way the MPC control is able to deal with disturbances, modeling errors, and (slow) changes in the system parameters.

C. The METANET model

The METANET model [16], [23] is a macroscopic traffic flow model developed by Papageorgiou and Messmer. In METANET the freeway network is represented as a graph with nodes and links, where the links correspond to freeway stretches with uniform characteristics; the nodes are placed at on-ramps and off-ramps, and where two or more freeways connect, or where there is a change in the characteristics. Links are divided into one of more segments with a length of about 500 m. The evolution of the traffic system is characterized by the following macroscopic variables for each segment i of each link m at time $t = lT_{\text{sim}}$ where l is the simulation step counter, and T_{sim} is the simulation time step (a typical value for T_{sim} is 10 s):

- traffic density $\rho_{m,i}(k)$ (veh/km/lane) in the segment,
- mean speed $v_{m,i}(k)$ (km/h) of the vehicles in the segment i,
- traffic flow $q_{m,i}(k)$ (veh/h) leaving the segment in the time interval $[lT_{\text{sim}}, (l+1)T_{\text{sim}}]$.

The METANET model describes how these state variables evolve over time using the values of the variables at simulation step l and the inputs (traffic demands. traffic control signals) at simulation step l to compute the new values of the state variables at simulation step $l+1$. For the sake of brevity, we will not repeat these equations here as they are not really necessary for the exposition below. For more information on the METANET model we refer the interested reader to [16], [23].

D. MPC for ramp metering

In this paper we consider the use of MPC for ramp metering in freeway networks. Field studies of traffic measurements usually show an increasing traffic throughput (number of vehicles per hour) with increasing traffic density (number of vehicles per kilometer) until the critical density ρ_{crit} is reached, after which the traffic throughput starts decreasing with increasing traffic density. This phenomenon is known as the fundamental diagram [15] and is presented is Figure 1. The maximal throughput is called the capacity of the freeway and it is denoted by q_{cap}. Free-flow traffic occurs when the traffic density is lower than or equal to the critical density ρ_{crit}. Traffic operation at densities larger than ρ_{crit} corresponds to a congested traffic flow.

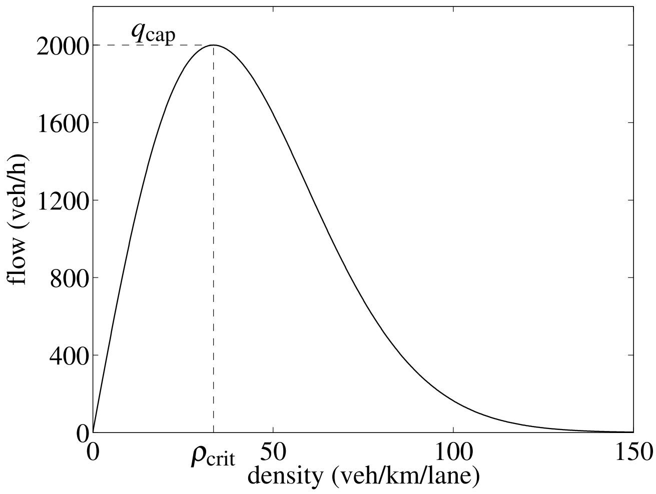

Fig. 1. The flow-density relation of the traffic in a freeway section, also known as the fundamental diagram.

Advanced traffic management systems are implemented to control the traffic flows in such a way that the traffic is in the free-flow state as much as possible. One technique that is frequently used for these purposes is ramp metering. Ramp metering tries to avoid that the traffic density on the freeway becomes larger than the critical traffic density, thus avoiding congestion [4], [30]. The way ramp metering limits the traffic density on the freeway is by restricting the number of vehicles that are allowed to enter the freeway through the on-ramp. This can be implemented by installing a traffic signal at the on-ramp as presented in Figure 2. The green period is selected such that only one car is allowed to enter the freeway per red-green cycle.

We use a discrete-time controller, with as control parameter the metering rate, which is defined as follows [13]:

$$r_i(k) = \frac{q_{\max,i}(k)}{Q_{\text{cap},i}} \quad , \tag{1}$$

where k is the sample step, i is the on-ramp index, $q_{\max,i}(k)$ is the maximal number of cars allowed to enter the freeway via on-ramp i, and $Q_{\text{cap},i}$ is the capacity of the ith on-ramp. Based on the metering rate, a sequence of red and green phases for the traffic signal can be calculated.

If we want to apply MPC to compute optimal ramp metering rates then we first have to select a performance indicator. In this paper we will consider the total time spent (TTS) by all vehicles in network (but note that the proposed approach also works for other performance indicators). Now we can apply MPC using TTS as performance indicator and the METANET model as

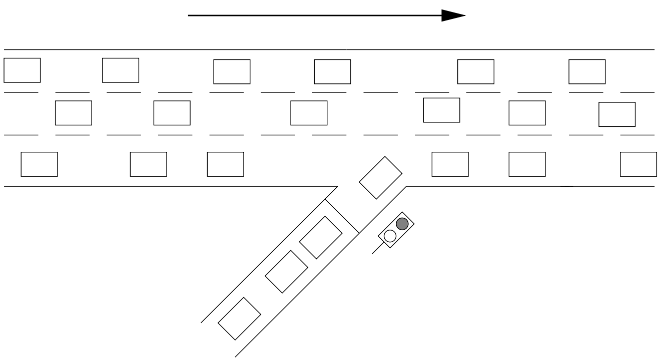

Fig. 2. Schematic representation of ramp metering.

prediction model in order to compute optimal ramp metering rates. For more information on the use of MPC for freeway traffic control (including other control measures such as variable speed limits and mainstream metering) we refer the interested reader to [1], [2], [9]–[11]. Related work is described in [8], [12], [17], [18] and the references therein.

When drivers are confronted with traffic control measures such as ramp metering, variable speed limits, and so on, after a while some of them tend to adapt their route choice so as to limit the effect of the control measures on their own travel time. In this paper we propose a anticipative traffic control framework based on MPC that also takes the reactions and re-routing of the drivers to the traffic control strategy into account. Determining how the traffic flows distribute themselves over different links of the traffic network given the traffic demand, the link capacity, and a route cost model, is called *traffic assignment*. In the context of the METANET model, the traffic assignment is used to compute the splitting rates at the network nodes. In Section II we first discuss static equilibrium traffic assignment. In Section III we will then discuss how the current traffic assignment and the equilibrium traffic assignment can be combined into a dynamic, time-varying traffic assignment that can be used to predict the reactions and rerouting of the drivers to the traffic control signals.

II. STATIC EQUILIBRIUM TRAFFIC ASSIGNMENT

A. Route costs

When traveling in a freeway network, drivers try to find the route that is optimal for themselves. In fact, it appears as if every driver assigns a cost to every alternative route leading to his destination and chooses the one with the smallest cost. Two important factors in the cost assigned to a route are [5], [20]:

- the travel time along the route, and
- the length of the route.

Hence, the cost of a route can be computed as follows:

$$\text{route cost} = \alpha_{\text{tim}}\,\text{travel time} + \alpha_{\text{len}}\,\text{route length} \quad , \qquad (2)$$

where α_{tim} and α_{len} are weighting parameters that express the relative importance of each term. In this paper we assume that the values of α_{tim} and α_{len} used in (2) are the average weight factors, i.e., the average values over all drivers in the network. Since a route consists of consecutive links connecting the origin with the destination, the cost assigned to a route according to (2) can be calculated by adding the link costs.

In order to illustrate the implementation of the static equilibrium traffic assignment method, the reasoning in the remainder of this section will be conducted based on the METANET model. Note however that this choice is free, and that another model can be selected if desired.

In the METANET model, each link is subdivided into segments. Hence, the link cost of link m can be computed as the sum of the segment costs:

$$c_m(q_m(l)) = \alpha_{\text{tim}} \sum_{i \in S_m} \frac{l_m}{v_{m,i}(l)} + \alpha_{\text{len}} \sum_{i \in S_m} l_m \quad , \qquad (3)$$

where $S_m = \{i_1, \ldots, i_{N_m}\}$ is the set of segments in link m, l_m is the length of the segments in link m, and $v_{m,i}(l)$ is the mean speed in segment i of link m at simulation step l, which depends on the link flow vector

$$q_m(l) \stackrel{\text{def}}{=} \left[q_{m,i_1}(l) \ \ldots \ q_{m,i_{N_m}}(l) \right]^{\text{T}}$$

through the model equations (in our case, the METANET model). According to (3), the link cost depends on the traffic demand $q_m(l)$ through the speed in the sections. Indeed, if more drivers use a link, the densities in the sections of that link increase and the desired average speeds in the corresponding sections decrease, resulting in an increase in the travel time (cost) of the link.

The travel time computed by (3) is an instantaneous travel time and can differ from the experienced travel time [25]. However, since we consider static equilibrium assignment in this section, the traffic states in the network are considered invariant in time, and the instantaneous and the experienced travel times are equal. Given the fact that the computation of the experienced travel time is computationally more involved than the computation of the instantaneous travel time, the instantaneous travel time is used in (3). In order to be able to compute the link costs for a whole freeway network using (3), the average speeds in all the segments given a set of link flows have to be computed. Since static equilibrium traffic assignment implies that the traffic flows in the network are invariant in time, it is assumed that the traffic flows in the links are in equilibrium. As a consequence, the equilibrium average segment speeds can be computed using the METANET model equations with $\rho_{m,i}(l+1) = \rho_{m,i}(l)$ and $v_{m,i}(l+1) = v_{m,i}(l)$ for all links m and segments $i \in S_m$.

B. Wardrop's principle

Wardrop stated in 1952 [27] that the traffic in a network distributes over the links in such a way that an equilibrium occurs where no individual driver can lower his travel time by changing routes. In equilibrium all used routes between an origin-destination pair have the same travel cost and non-used routes have a higher travel cost. The resulting equilibrium is called the user optimal equilibrium since it occurs if every driver individually optimizes his route.

C. Equilibrium traffic assignment algorithms

Now we present how a static equilibrium traffic assignment can be computed using the expression of the route cost presented in (3).

There exist several methods to compute the user optimum defined by Wardrop's principle such as, e.g., the Frank-Wolfe algorithm [24], the method of the successive averages [19], feedback strategies [23], iterative strategies [29], predictive feedback strategies [26], ... (see [20] for an overview and detailed discussion of these methods). Since it is guaranteed to converge to a solution, we discuss the *Method of the Successive Averages* (MSA) [19], [21] here. The MSA is an iterative static equilibrium traffic assignment method that takes the impact of vehicle flows on the link costs into account through the cost function (3). The algorithm uses network flow vectors $\mathbf{q}^i$, which contain the link flows for all the links in the network at iteration i. The algorithm contains the following steps:

Initialization: Set $i = 1$, $\mathbf{q}^i = 0$, $\phi = 1$
repeat
Step 1: Calculate costs $\mathbf{c}^i(\mathbf{q}^i)$ according to (3)
Step 2: Determine $\mathbf{q}^i_{\text{aon}}$ by all-or-nothing assignment
Step 3: Set $\mathbf{q}^{i+1} = (1-\phi)\mathbf{q}^i + \phi\mathbf{q}^i_{\text{aon}}$
Step 4: Set $i = i+1$, $\phi = \frac{1}{i}$
until stopping criterion is satisfied

In Step 1, the vector $\mathbf{c}^i(\mathbf{q}^i)$ containing the link costs associated with the traffic assignment of iteration i is calculated. In Step 2, all the traffic between an origin-destination pair is assigned to the cheapest route. This is called an all-or-nothing assignment. During the ith iteration for every origin-destination pair the cheapest route is searched based on the link cost vector $\mathbf{c}^i(\mathbf{q}^i)$. For smaller networks, the cheapest route can be searched exhaustively, but for larger networks a more advanced method like Dijkstra's shortest path algorithm [6], [28] is required. The flows in the links caused by the all-or-nothing assignments for all origin-destination pairs lead to a flow vector $\mathbf{q}^i_{\text{aon}}$. In Step 3, a new flow vector $\mathbf{q}^{i+1}$ is computed as a convex combination of the previous flow vector and the flow vector $\mathbf{q}^i_{\text{aon}}$ resulting from the all-or-nothing assignments in Step 2. The meaning of ϕ is the following: In iteration i, the value of ϕ is such that the new flow vector that is found is the average of all i flow vectors. Hence the name: the method of successive averages. The stopping criterion for the MSA can be, e.g., a maximal number of iterations, or a maximal difference between two successive flow vectors or cost vectors.

In the next section we will discuss an anticipative traffic assignment strategy that is based on the principles of the static equilibrium traffic assignment presented in this section. One of the main reasons for developing a dynamic traffic assignment based on a static equilibrium traffic assignment, is that in MPC the computational complexity is an important factor. State-of-the-art dynamic traffic assignment methods require too much computation time so that they cannot be used in the on-line MPC control approach. On the other hand, the semi-dynamic anticipative traffic assignment we propose next offers a balanced trade-off between accuracy and computational complexity, so that it is well-suited for use in on-line MPC control.

III. Anticipative traffic assignment

Anticipative traffic assignment takes the response of the drivers to the changes of the traffic state over time into account. The method presented here only uses the experienced traffic state and does not require the traffic demands to be constant. In Section IV the anticipative traffic assignment will be incorporated in an MPC framework for ramp metering in order to allow for the computation of control signals that anticipate the response of the drivers to the control actions.

During their trip in the freeway network, drivers experience the traffic state of the network. Based on the information they gather on the global state of the traffic network, drivers will determine their optimal route. The current traffic state assessment, which is used to determine the optimal route, is typically based on information of the traffic state in the near past.

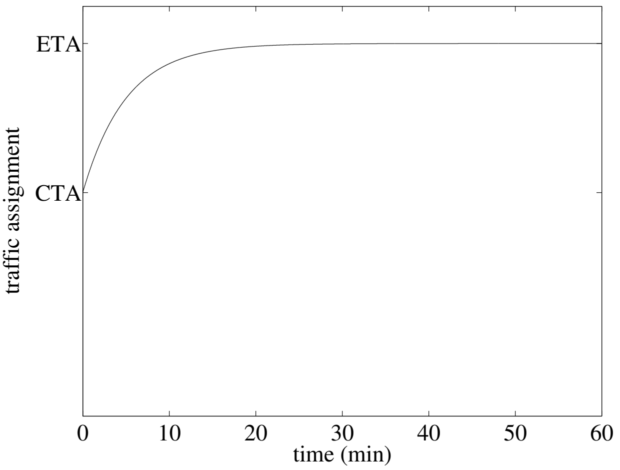

Fig. 3. Exponential evolution from the current traffic assignment (CTA) to the equilibrium traffic assignment (ETA) for $\tau_{\mathrm{evol}} = 5$ min.

This process of gathering traffic state information can be modeled as follows: The traffic state in the near past results from the traffic demands and the metering rates (or other traffic control signals) in the near past. The fact that the gathering of traffic state information takes some time is modeled by averaging the traffic demands and the metering rates over a time horizon τ_{info}. By averaging the traffic demands and the metering rates over the time interval $[t - \tau_{\mathrm{info}}, t]$, we obtain the information about the traffic state and past/expected traffic conditions as they are perceived by the drivers and as they are used at the current time to determine their optimal route. The parameter τ_{info} needs to be tuned and is influenced by the network dimensions and topology, the availability of information (e.g., radio bulletins), and so on.

Drivers use the gathered traffic state information to optimize their route. Using the static equilibrium assignment method presented in Section II, combined with the average values of the traffic demands and the metering rates, the equilibrium traffic assignment (ETA) associated with the traffic situation experienced by the drivers can be computed. Since the average traffic demand is used, the static equilibrium traffic assignment method yields acceptable results even in the case of time varying traffic demands. Based on the assignment of the flows to the routes, the splitting rates at the bifurcation nodes in the freeway network are computed.

It takes some time before the traffic flows in the freeway network will reorganize according to the ETA. This is due to the fact that not all drivers will decide to use the new route immediately. The transition from the current traffic assignment (CTA) to the ETA is modeled as an exponential evolution (with time constant τ_{evol}) of the splitting rates over time as shown in Figure 3. The time constant τ_{evol} needs to be tuned such that the evolution from CTA to ETA occurs in a realistic time frame.

IV. Anticipative MPC-based traffic control strategy

Now we propose the anticipative MPC-based traffic control strategy combines the MPC-based control strategy and the anticipative traffic assignment technique. A schematic representation of the overall control strategy is presented in Figure 4.

The MPC-based control module produces control signals in the form of metering rates, which are applied to the traffic system. The state of the traffic situation is measured (e.g., every minute) and fed back to the MPC module. These measurements are used to update the state and to subsequently start a new optimization over the shifted prediction horizon (i.e., a receding horizon approach).

During the traffic assignment, a prediction of the ETA is made using information of the traffic situation experienced by the drivers during the interval $[t - \tau_{\mathrm{info}}, t]$, where t is the time at which the ETA is computed. This experienced traffic situation is computed based on the traffic demands and the metering rates as described in Section III. The ETA and the CTA are combined to compute a description of the evolution of the splitting rates in time. This evolution is fed to the MPC control module in order to be used for the MPC prediction during the prediction horizon. By supplying the MPC with the evolution of the splitting rates from CTA to ETA, the controller is able to take the re-routing behavior of the drivers into account. Since the dynamics of the re-routing are slower than the dynamics of the traffic system near the on-ramps, it suffices to update the traffic assignment at a slower pace than the traffic states (e.g., every T_{anticip} = 15 min).

The assumption of the dynamics of traffic re-routing being slower than the dynamics of the traffic behavior on the freeway near the on-ramps is an important assumption. The slower the dynamics of the re-routing process are, the larger the parameter T_{anticip} can be chosen. This is important since the computation of the ETA is rather computationally intensive. Also from the point of view of stability of the anticipative MPC-based ramp metering controller it is desirable that the dynamics of the re-routing are sufficiently small compared to the dynamics of the traffic operation [22].

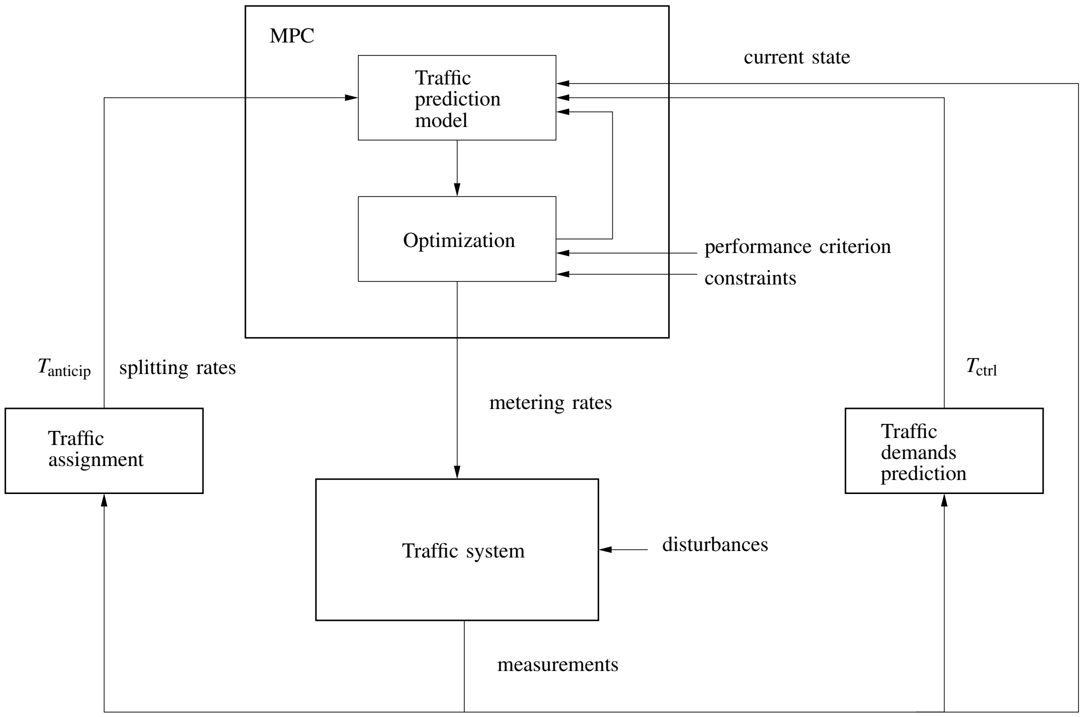

Fig. 4. Schematic representation of the anticipative MPC-based control strategy for ramp metering.

Remark 4.1 In the METANET traffic simulation software package [23], several dynamic traffic assignment algorithms are incorporated ranging from feedback to iterative strategies. However, since the dynamic traffic assignment is updated at every simulation step, as is also described in [25], this results in a high computational complexity compared to the simulation of the METANET model without dynamic traffic assignment. The anticipative MPC-based control strategy presented in this paper reduces the computational complexity by updating the dynamic traffic assignment at a (slower) pace corresponding to the re-routing dynamics.

V. SIMULATION EXAMPLE

Now we illustrate anticipative MPC-based ramp metering control using a simulation example.

A. Network layout

The traffic network for the simulation example is presented in Figure 5. It consists of a freeway with four lanes that bifurcates into two branches of two lanes each. Downstream both branches join in a four-lane freeway. Both four-lane freeway links are 3 km long. The lower two-lane branch is called the primary branch. The primary branch is 6 km long, and an on-ramp is present in the middle of the branch. The secondary branch is longer than the primary branch and is 8 km long. During simulation, the traffic originating from the mainstream origin distributes over the two branches. The distribution of the traffic over both branches depends on the state of both branches and is modeled using the anticipative traffic assignment discussed in Section III.

B. Traffic scenario

We simulate a traffic scenario with road maintenance works on the primary branch as shown in Figure 5. The maintenance works result in a reduction of the number of lanes from 2 to 1 in the last segment (i.e., the last 500 m) of the primary branch. The maintenance works start at 5.30 a.m. and persist during the remainder of the simulation.

For illustrative purposes the traffic demand on the mainstream is considered constant and equal to 4500 veh/h in this simulation. The traffic demand on the on-ramp is equal to 200 veh/h with a peak traffic demand of 500 veh/h around 6 a.m. (see the full line in the top plot of Figure 7).

C. Model and controller parameters

The parameter τ_{info} of the anticipative traffic assignment determines the size of the horizon over which the drivers gather information to assess the global state of the network. The larger τ_{info}, the slower the response of the route choice behavior of the drivers on varying

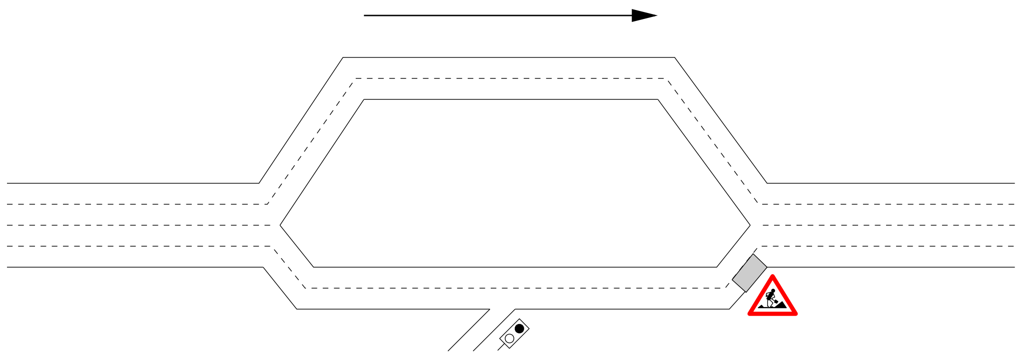

Fig. 5. Schematic representation of the freeway network studied in the simulation example.

traffic demands and metering rates will be. We take τ_{info} = 30 min.

The parameter τ_{evol} models the swiftness of the response of the drivers to a difference between the current traffic assignment and the equilibrium traffic assignment. Since the maintenance works are located near the downstream end of the 6-km main two-lane branch, it takes some time for the disturbance caused by this lane reduction to propagate upstream and to reach the drivers that still need to make a route choice. The time it takes for the drivers to adapt to the ETA depends, e.g., on the situation and the familiarity of the drivers with the freeway network. We choose τ_{evol} = 45 min.

The rate at which the traffic assignment needs to be updated depends on the re-routing dynamics in the network, which are typically much slower than the dynamics of the traffic system near the on-ramps and which depend on the topology of the network. A trade-off can be made between the accuracy of the traffic assignment and the computational complexity of the anticipative traffic assignment by tuning the time T_{anticip} between two ETA updates. In this simulation example we choose T_{anticip} = 15 min.

For the link cost weights (cf. equation (3)) we take $\alpha_{\text{tim}} = 1$ and $\alpha_{\text{len}} = 0$, i.e., we choose the average instantaneous travel time through a link to be the cost of that link.

We also impose a maximal on-ramp queue length of 100 vehicles for the controlled case. Furthermore, we set the minimal metering rate of the on-ramp r_{min} to 0.1.

Before discussing anticipative MPC-based ramp metering for the set-up and the traffic scenario described above, we first consider a simulation without ramp metering but with anticipative traffic assignment included.

D. No ramp metering

At the start of the simulation, the traffic network is in steady state and the instantaneous travel times along both alternative routes are equal (cf. the Wardrop principle discussed in Section II-B).

At 5.30 a.m., the traffic demand on the on-ramp increases while at the same time the maintenance works start at the downstream end of the primary branch. Due to the increasing traffic demand on the on-ramp the traffic density in the first segment downstream of the on-ramp starts to increase immediately, which results in a decrease of the average speed in the segment (see Figure 6).

Due to the maintenance works, the number of lanes of the primary branch is reduced from two to one. Since the traffic flow carried by the primary branch is too high for one single lane, congestion sets in. In the upper plot of Figure 6 the traffic density of the first segment after the on-ramp is shown. Since the reduction of the number of lanes is located downstream of this segment, it takes some time for the congestion to propagate to this segment. This is observed as the delay between the start of the construction works at 5.30 a.m. and the peak in the traffic density in the segment. The traffic density in the segment exceeds the critical density of the segment and congestion sets in in the segment. The increased traffic density in the primary branch results in a decrease of the average speed in the branch (see Figure 6), and thus in an increased travel time. This phenomenon can be observed in the lower plot of Figure 6, which presents the travel time from the mainstream origin to the destination for both alternative routes. The travel time of the route including the primary branch (solid line) increases immediately with the start of the maintenance works in the primary branch and with the increase in traffic demand at the on-ramp.

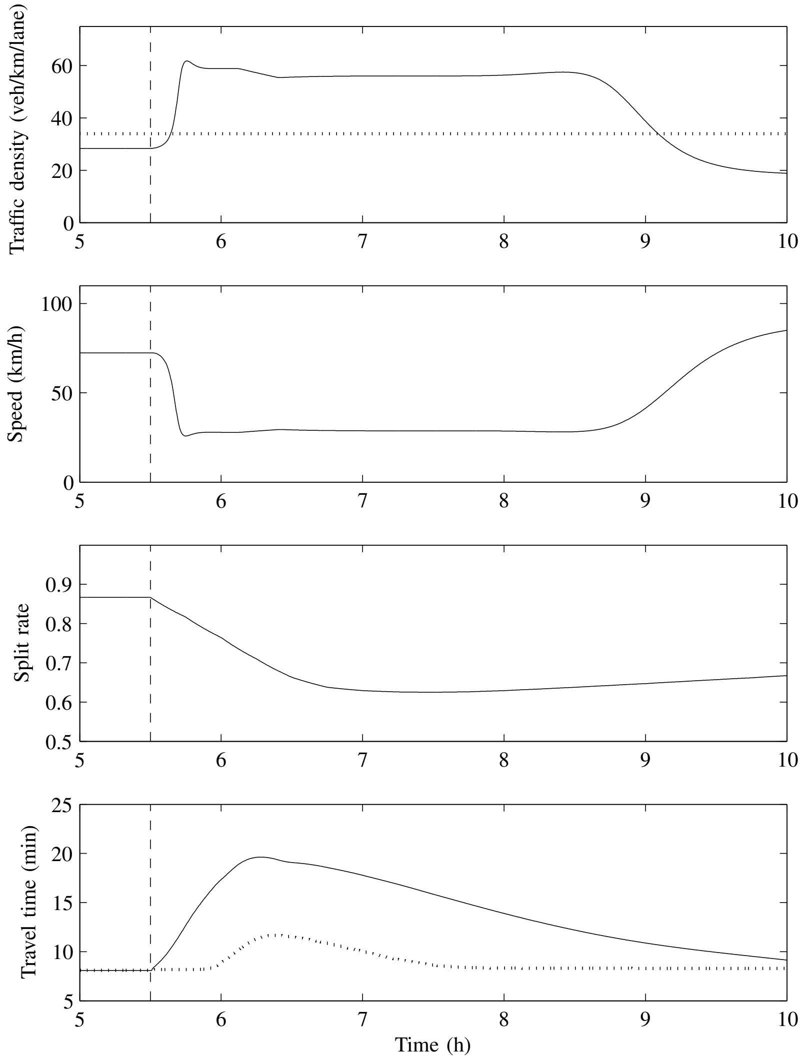

Fig. 6. Simulation results of anticipative traffic assignment without control. The vertical dashed line represents the start of the construction works on the primary branch. In the lower plot the instantaneous travel time is shown for the primary (solid line) and the secondary (dotted line) route from the mainstream origin to the destination.

As a result of the increase of the travel time on the primary route, the secondary route becomes the fastest route. Hence, the drivers will be inclined to start using the secondary route. This can be observed in the third plot of Figure 6 where the evolution of the split rate at the bifurcation node over time is plotted. The split rate starts to decrease from the moment the travel time on the primary route becomes larger than the travel time on the secondary route. In the lower plot of Figure 6 we observe an increase in the travel time on the secondary route, which results from the increased traffic volume on the secondary route but also from the spill-back of congestion of the primary branch into the upstream four-lane freeway link. The congestion in the four-lane freeway link starts resolving as soon as a sufficiently large number of drivers starts using the longer secondary route. As the four-lane mainstream link is common to both the primary and the secondary route, this can be observed as the decrease in travel time for both routes around 7 a.m.

Once the congestion on the mainstream link is resolved, the congestion in the primary branch starts to resolve since more drivers are now using the secondary route. Since more and more segments of the primary branch are becoming uncongested, the travel time of the primary route decreases. Around 9 a.m. the traffic density in the segment right after the on-ramp drops below the critical density and the congestion in this segment is resolved.

Eventually, the system reaches a new equilibrium state with equal travel times for both routes. In this new equilibrium state there still is congestion in the

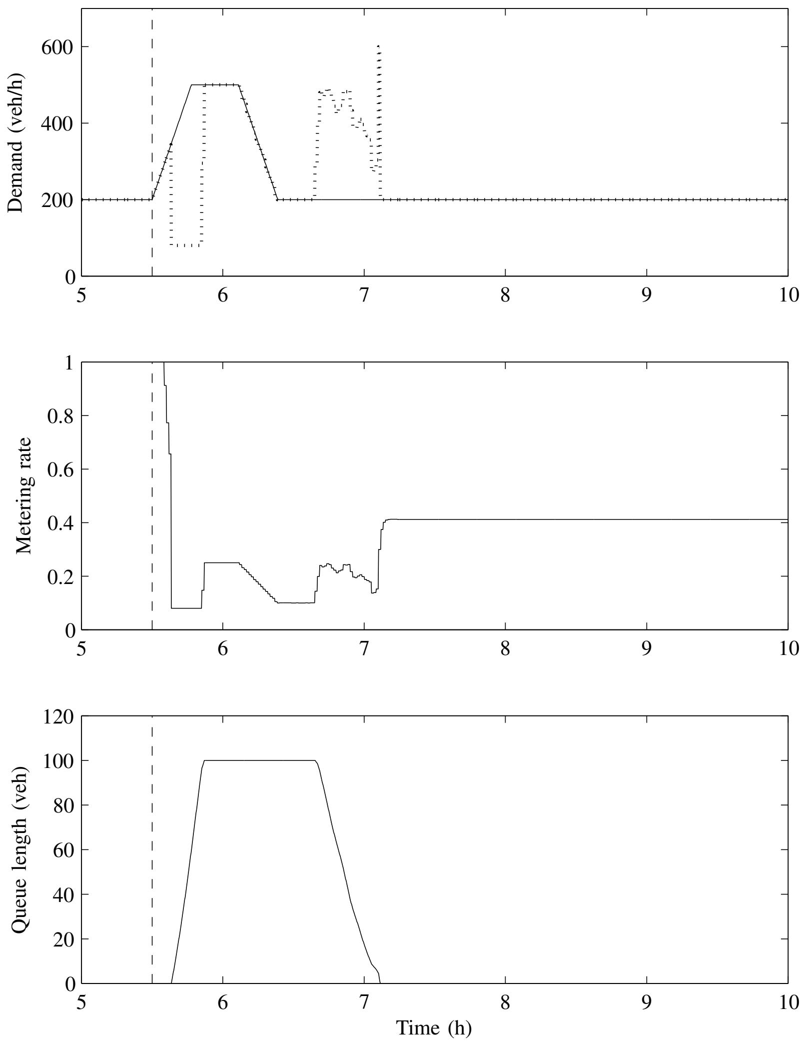

Fig. 7. Simulation results of anticipative MPC-based ramp metering. In the top plot the solid line represents the traffic demand, and the dashed line gives traffic that is actually allowed to enter the primary branch.

segments directly upstream of the maintenance works. However, in the upstream segments of the primary branch, the traffic density during the new equilibrium is lower than the traffic density in the equilibrium state at the beginning of the simulation due to the lower traffic volume using the primary branch. As an illustration of this we refer to the traffic density in the segment behind the on-ramp at the start and at the end of the simulation period (see Figure 6).

For the scenario presented above total time spent in the network (i.e., the performance indicator) is 4886 veh.h for the no-control case.

E. Anticipative MPC-based ramp metering control

The simulation example with anticipative MPC-based ramp metering control starts from the same equilibrium state as the simulation without ramp metering control. At 5.30 a.m., the traffic demand at the on-ramp starts to increase as presented in Figure 7, and the maintenance works start, which results in a reduction of the number of lanes from 2 to 1 near the downstream end of the primary branch.

As a result of the increased traffic demand at the on-ramp and the reduced number of lanes in the last segment of the primary branch, the traffic density in the first segment downstream of the on-ramp starts to increase immediately as can be observed in the upper plot of Figure 8. In order to avoid the traffic density in the primary branch from becoming larger than the critical density ρ_{crit}, ramp metering becomes active as can be seen in the middle plot of Figure 7 where the evolution of the metering rate at the on-ramp over time is presented. In the upper plot of Figure 7 we see the traffic demand at the on-ramp plotted in a solid line

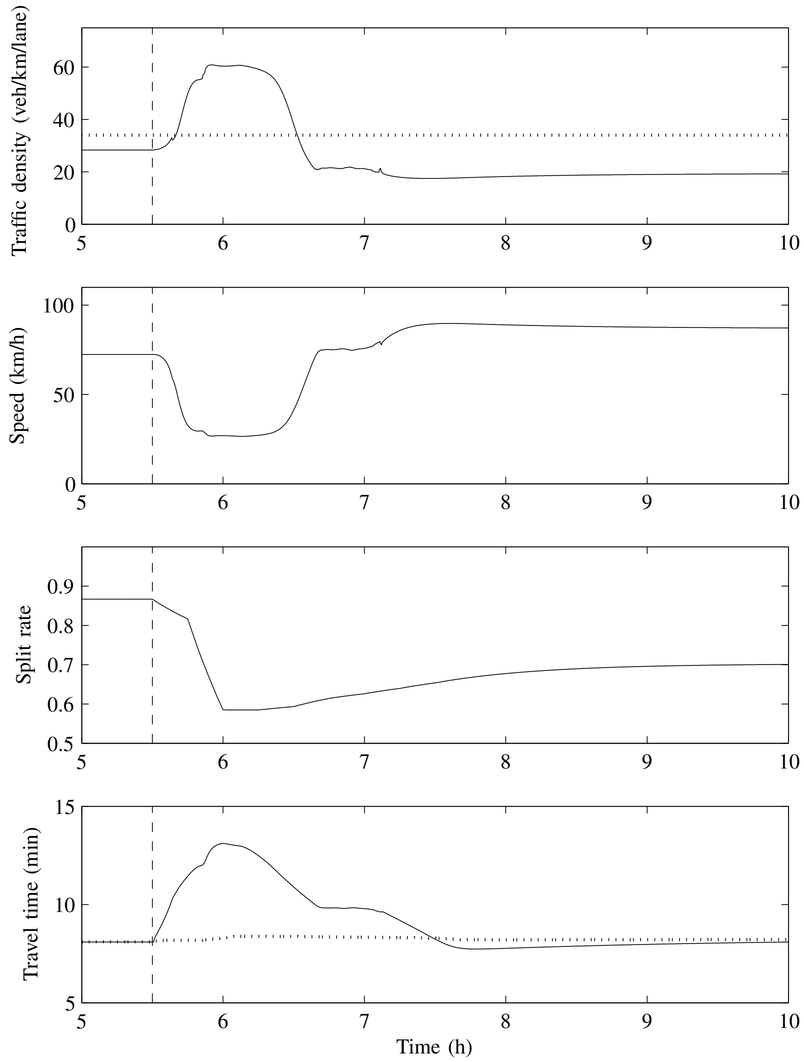

Fig. 8. Simulation results for anticipative MPC-based ramp metering. The plots at the top give the density and average speed in the first segment downstream of the on-ramp. The dotted line in the upper plot represents the critical traffic density ρ_{crit}. In the lower plot the instantaneous travel time is shown for the primary (solid line) and the secondary (dotted line) route from the mainstream origin to the destination.

and the traffic flow that is allowed to enter the freeway plotted as a dotted line. In the lower plot, the evolution of the queue length at the on-ramp over time is shown.

Shortly after the activation of the ramp metering, the metering rate drops to its minimal value ($r_{\min} = 0.1$) and after some time the queue length becomes equal to the maximal number of 100 vehicles that is allowed at the on-ramp. The MPC controller is able to take the constraint on the queue length into account by increasing the metering rate. The metering rate must be large enough to prevent the queue from growing and the metering rate must be as small as possible to keep the traffic density on the freeway below the critical density ρ_{crit}. This trade-off results in a metering rate that is such that the number of vehicles that is allowed to enter the freeway is equal to the traffic demand (see Figure 7).

In the upper plot of Figure 8 we observe that despite ramp metering the traffic density in the first segment downstream of the on-ramp grows larger than the critical density ρ_{crit}. This is due to the congestion from the maintenance works spilling back in the upstream direction. However, if we compare the traffic density in the control case (Figure 8) with the traffic density in the no-control case (Figure 6), we observe that the traffic density in the ramp metering control case remains lower than in the no-control case. This results in a lower instantaneous travel time for the primary route in the controlled case.

Some minutes before 6 a.m. the MPC-based ramp metering controller is forced by the constraint on the queue length to allow more traffic to enter the freeway in order to avoid the queue becoming too long. This

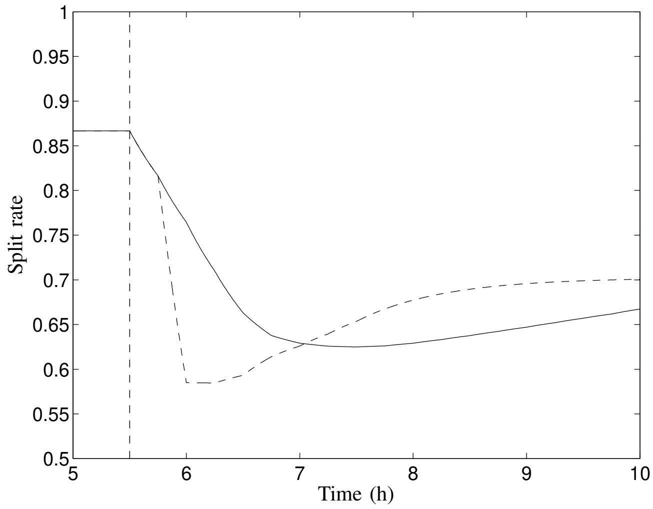

Fig. 9. Split rates for the no-control case (full line) and the controlled case (dashed line).

results in an immediate increase of the traffic density in the segment fed by the on-ramp as can be seen in Figure 8. The increased traffic density leads to a decrease of the average speed in the segment and thus the travel time of the primary route increases due to the increased metering rate (see the lower plot of Figure 8).

The third plot in Figure 8 shows the evolution of the split rate over time. Due to the larger travel time on the primary branch, a fraction of the traffic shifts from the primary to the secondary route. The impact of the shift of traffic from the primary to the secondary freeway branch on the instantaneous travel time of the secondary route is rather small.

As the traffic demand at the on-ramp decreases again after 6 a.m., the traffic density on the freeway starts to decrease as well. At a certain point, the traffic density on the primary freeway branch is low enough for the controller to start dissolving the queue. After 7 a.m., all the traffic that wants to enter the primary freeway branch through the on-ramp is allowed to do so, despite the metering rate that differs from 1. Indeed, we observe in Figure 7 that the queue length is 0 and that the traffic flow entering the freeway through the on-ramp is equal to the traffic demand at the on-ramp after 7 a.m. The metering rate differs from 1, but it does not restrict the on-ramp traffic.

As was the case in the no-control case, the traffic evolves to a new equilibrium state with equal travel times for both alternative routes in the controlled case. In this new equilibrium state, there is some congestion in the segments directly upstream of the maintenance works. Due to the delays and the increased travel time caused by this congestion in the primary branch, there are more drivers using the longer secondary route.

Figure 9 shows the split rates (i.e., the fraction of the vehicles opting for the primary branch) for the no-control case and the controlled case in the same plot. We see that the congestion on the primary branch, which is higher in the no-control case, causes more drivers to select the alternative route than in the controlled case.

For the anticipative MPC-based ramp metering control, the total time spent in the network was 3781 veh.h, which is an improvement of about 22 % compared to the no-control case.

VI. Conclusions

Since drivers tend to use the cheapest route available, a traffic control action can cause a re-routing of traffic. The anticipative MPC-based ramp metering control strategy that we have developed in this paper, takes into account that due to the control actions in the freeway network, the drivers may choose an alternative route. The proposed anticipative MPC-based ramp metering strategy optimizes the metering rates in such a way that the desired (lowest cost) traffic state is obtained despite the re-routing behavior of the drivers. It was assumed that the re-routing takes place at a slower time-scale since drivers do not re-route instantaneously. The re-routing of the drivers was modeled as an exponential transition from the current traffic assignment (CTA) to the equilibrium traffic assignment (ETA). The ETA is determined based on the information that is available to the drivers, i.e., the traffic situation in the near past. In case there is some kind of information provided to the drivers, this can be modeled by shortening the transition time from CTA to ETA. We have also illustrated the new anticipative MPC-based ramp metering control strategy with a simulation example.

Topics for future research include: inclusion of other static and dynamic traffic assignment methods, and investigation of their effect on the trade-off between accuracy or performance and computational complexity; investigation of other scenarios and larger networks; and inclusion of other traffic control measures.

Acknowledgments

Research supported the Flemish Government (Research Council K.U.Leuven: Concerted Research Action Mefisto-666; FWO projects: G.0240.99 and G.0256.97; FWO-Research Communities: ICCoS and ANMMM; IWT projects: EUREKA 1562-SINOPSYS, EUREKA 2063-IMPACT, STWW), the Belgian Federal Government: (IUAP V-22, PODO-II); the NWO-CONNEKT project 014-34-523

"Advanced multi-agent control and information for integrated multi-class traffic networks (AMICI)"; the STW project "Multi-agent control of large-scale hybrid systems" (DWV.6188); and the TU Delft spearhead program "Transport Research Centre Delft: Towards Reliable Mobility".

VII. REFERENCES

[1] T. Bellemans, *Traffic Control on Motorways*. PhD thesis, Faculty of Applied Science, K.U.Leuven, Leuven, Belgium, May 2003.

[2] T. Bellemans, B. De Schutter, and B. De Moor, "Model predictive control with repeated model fitting for ramp metering," in *Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems (ITSC'02)*, Singapore, pp. 236–241, Sept. 2002.

[3] E.F. Camacho and C. Bordons, *Model Predictive Control in the Process Industry*. Berlin, Germany: Springer-Verlag, 1995.

[4] Cambridge Systematics, Inc., Oakland, California, *Twin Cities Ramp Meter Evaluation – Final Report*, Feb. 2001. Prepared for the Minnesota Department of Transportation.

[5] C. F. Daganzo and Y. Sheffi, "On stochastic models of traffic assignment," *Transportation Science*, vol. 11, pp. 253–274, 1977.

[6] E. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, pp. 269–271, 1959.

[7] C.E. García, D.M. Prett, and M. Morari, "Model predictive control: Theory and practice — A survey," *Automatica*, vol. 25, no. 3, pp. 335–348, May 1989.

[8] H. Haj-Salem and M. Papageorgiou, "Ramp metering impact on urban corridor traffic: Field results," *Transportation Research Part A*, vol. 29, no. 4, pp. 303–319, 1995.

[9] A. Hegyi, *Model Predictive Control for Integrating Traffic Control Measures*. PhD thesis, Delft University of Technology, Delft, The Netherlands, Feb. 2004.

[10] A. Hegyi, B. De Schutter, H. Hellendoorn, and T. van den Boom, "Optimal coordination of ramp metering and variable speed control — An MPC approach," in *Proceedings of the 2002 American Control Conference*, Anchorage, Alaska, pp. 3600–3605, May 2002.

[11] A. Hegyi, B. De Schutter, and J. Hellendoorn, "Optimal coordination of variable speed limits to suppress shock waves," *Transportation Research Record*, no. 1852, pp. 167–174, 2004.

[12] A. Kotsialos, M. Papageorgiou, M. Mangeas, and H. Haj-Salem, "Coordinated and integrated control of motorway networks via non-linear optimal control," *Transportation Research Part C*, vol. 10, no. 1, pp. 65–84, Feb. 2002.

[13] A. Kotsialos, M. Papageorgiou, and A. Messmer, "Optimal coordinated and integrated motorway network traffic control," in *Proceedings of the 14th International Symposium of Transportation and Traffic Theory (ISTTT)*, Jerusalem, Israel, pp. 621–644, July 1999.

[14] J.M. Maciejowski, *Predictive Control with Constraints*. Harlow, England: Prentice Hall, 2002.

[15] A.D. May, *Traffic Flow Fundamentals*. Englewood Cliffs, New Jersey: Prentice-Hall, 1990.

[16] A. Messmer and M. Papageorgiou, "METANET: A macroscopic simulation program for motorway networks," *Traffic Engineering and Control*, vol. 31, pp. 466–470, 1990.

[17] M. Papageorgiou, J.M. Blosseville, and H. Haj-Salem, "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*, vol. 24, no. 5, pp. 361–370, Sept. 1990.

[18] M. Papageorgiou, H. Hadj-Salem, and J.M. Blosseville, "ALINEA: A local feedback control law for on-ramp metering," *Transportation Research Record*, no. 1320, pp. 58–64, 1991.

[19] W.B. Powell and Y. Sheffi, "The convergence of equilibrium algorithms with predetermined step sizes," *Transportation Science*, vol. 16, no. 1, pp. 45–55, Feb. 1982.

[20] Y. Sheffi, *Urban Transportation Networks*. Englewood Cliffs, New Jersey: Prentice-Hall, 1985.

[21] Y. Sheffi and W. Powell, "An algorithm for the equilibrium assignment problem with random link times," *Networks*, vol. 12, pp. 191–207, 1982.

[22] H. Taale and H.J. van Zuylen, "Traffic control and route choice: Occurrence of instabilities," in *Proceedings of the TRAIL 5th Annual Congress 1999: Five Years Crossroads of Theory and Practice*, vol. 2, pp. 1–19, 1999.

[23] Technical University of Crete, Dynamic Systems and Simulation Laboratory and A. Messmer, *METANET – A simulation program for motorway networks*, July 2000.

[24] D. Van Vliet, "The Frank-Wolfe algorithm for equilibrium traffic asignment viewed as a variational inequality," *Transportation Research Part B*, vol. 21, no. 1, pp. 87–89, 1987.

[25] Y. Wang and M. Papageorgiou, "A predictive feedback routing control strategy for freeway network traffic," in *Proceedings of the American Control Conference*, Anchorage, Alaska, pp. 3606–3611, IEEE, May 2002.

[26] Y. Wang, M. Papageorgiou, and A. Messmer, "A predictive feedback routing control strategy for freeway network traffic," in *Proceedings of the 82nd Annual Meeting of the Transportation Research Board*, Washington, D.C, Jan. 2003.

[27] J.G. Wardrop, "Some theoretical aspects of road traffic research," *Proceedings of the Institution of Civil Engineers, Part II*, vol. 1, pp. 325–378, 1952.

[28] W.L. Winston, *Operations Research: Applications and Algorithms*. Duxbury Press, 4th ed., 2004.

[29] M.B. Wisten and M.J. Smith, "Distributed computation

of dynamic traffic equilibria," *Transportation Research C*, vol. 5, pp. 77–93, 1997.

[30] H. Zhang, S.G. Ritchie, and W.W. Recker, "Some general results on the optimal ramp control problem," *Transportation Research Part C*, vol. 4, no. 2, pp. 51–69, 1996.