Delft University of Technology

Delft Center for Systems and Control

Technical report 15-022

Integrated Predictive Control of Freeway Networks Using the Extended Link Transmission Model*

M. Hajiahmadi, G. S. van de Weg, C. M. J. Tampère, R. Corthout, A. Hegyi, B. De Schutter, and H. Hellendoorn

To cite this work, please refer to the published version:

M. Hajiahmadi, G. S. van de Weg, C. M. J. Tampère, R. Corthout, A. Hegyi, B. De Schutter, and H. Hellendoorn, "Integrated predictive control of freeway networks using the extended link transmission model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 1, pp. 65–78, Jan. 2016. doi:10.1109/TITS.2015.2460695

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

* This report can also be downloaded via https://dspub.eu/15-022

Integrated Predictive Control of Freeway Networks Using the Extended Link Transmission Model

Mohammad Hajiahmadi, Goof S. van de Weg, Chris M.J. Tampère, Ruben Corthout, Andreas Hegyi, Bart De Schutter, and Hans Hellendoorn

***Abstract*—In this paper, the recently developed link transmission model (LTM) is utilized in an on-line hybrid model-based predictive control (MPC) framework. The model is extended to include the effects of ramp metering and variable speed limits. Next, an integrated freeway traffic control based on the new model is presented in order to minimize the total time spent in the network. The integrated scheme has the capability of controlling large-scale freeway networks in real-time as the model is computationally efficient and it is yet accurate enough for our control purposes. In addition, the extended model is reformulated as a system of linear inequalities with mixed binary and real variables. The reformulated model along with the linearized total travel time objective function establish a mixed integer linear optimization problem that is more tractable and even faster than the original optimization problem integrated in the MPC scheme. Finally, to investigate the performance of the proposed approaches (nonlinear MPC and the mixed integer linear counterpart), a freeway network layout based on the Leuven Corridor in Belgium is selected. The extended LTM is calibrated for this network using micro-simulation data and next, is used for prediction and control of the large network. Micro-simulation results show that the proposed methods are able to efficiently improve the total travel time.**

I. INTRODUCTION

MODEL-BASED control for traffic networks has been attractive for many researchers. In particular, freeway network modeling and control with the goal of reducing the travel time, reducing the fuel consumption and emissions, increasing the throughput of the network, etc., have been extensively investigated in the literature [1]–[6]. In model-based traffic control, an efficient and accurate model for the evolution of the traffic is required. For traffic networks, a wide range of traffic flow models has been developed, from microscopic models [7] and micro-simulators [8] that represent the behavior of individual vehicles to macroscopic models that describe the traffic dynamics in terms of aggregate variables [7], [9]. Among these models, those that can accurately reproduce the traffic states, the relevant dynamics such as congestion creation and dissolution, blocking back, and the effects of traffic control measures such as ramp metering and variable speed limit (VSL) and meanwhile have reasonable computational complexity are good candidates for being used in the model-based traffic control. Macroscopic models such as the METANET model [10] and its extended versions [3], and the cell transmission model (CTM) [11] are good examples of such models.

Model predictive control (MPC) is an advanced control strategy [12] that has been utilized for traffic network control in recent years [3], [4], [13], [14]. It involves optimization-based calculation of control inputs using a prediction model of the system under control. In [3], [5], the extended METANET is utilized for the prediction of freeway traffic states in the model predictive control framework. The CTM has been also used for predictive freeway and urban network control [14]–[16]. However, the computation effort required for the model-based predictive control schemes increases as the size of the network under control grows and maintaining the real-time control might not be feasible anymore. Hence, research has been focused on finding less complex models that have similar properties as in the more elaborated models.

The link transmission model (LTM) proposed by [17] is a first-order traffic flow model. It was originally developed for dynamic traffic assignment [18]. Results presented in [17], [19], [20] show the capability of the LTM for fast modeling of large-scale networks. This is mainly due to the fact that in the LTM modeling framework, long length links with less number of variables can be used to model freeway roads, while e.g. in the METANET and CTM modeling approaches, road segments of around 500-1000 m are used. Moreover, larger sample times are allowed in the LTM modeling method. Overall, the LTM can provide a good trade-off between accuracy and computational efficiency.

In the previous work [21], we extended the original LTM for incorporating ramp metering and variable speed limits. In this paper, we elaborate more on the theoretical aspects of our previous findings and furthermore, using the extended model, we develop a hybrid model predictive control scheme for freeway network control. We choose the MPC approach mainly because it can incorporate and combine multiple traffic objectives (such as reducing the total time spent, the emission, and the fuel consumption), it can handle constraints on the states and the control inputs, it is effective in response to disturbances affecting the system, and it can integrate traffic control measures such as ramp metering, variable speed limits, and route guidance. Network modeling and predictive control based on the extended LTM is fast. However, further steps towards having *real-time* and efficient control for large-scale

M. Hajiahmadi, B. De Schutter, and H. Hellendoorn are with the Delft Center for Systems and Control (DCSC), Delft University of Technology, Delft, The Netherlands (e-mail: m.hajiahmadi@tudelft.nl).

G.S. van de Weg and A. Hegyi are with the Transport and Planning Department, Delft University of Technology, Delft, The Netherlands (e-mail: g.s.vandeweg@tudelft.nl).

R. Corthout and C.M.J. Tampère are with the Traffic and Infrastructure Department, KU Leuven, Leuven, Belgium.

This work is supported by the European 7th Framework Network of Excellence "Highly-complex and networked control systems (HYCON2)" and by the European COST Action TU1102.

freeway networks is crucial and will be also addressed in this paper.

Apart from the computation time required for repeated model execution and solving the optimization problem in the MPC framework, due to the nonlinear and nonconvex nature of the optimization problem, there might exist multiple local optimal solutions. This adds more complexity to the control problem. In order to overcome this issue, we propose a reformulation of the nonlinear LTM. We use some mathematical techniques for piecewise affine systems [22] along with the simplifying assumption that the number of VSL values is finite (which is consistent with reality), to transform the extended LTM into a linear model composed of linear equations and inequalities with mixed real and integer variables. Having transformed the model, the nonlinear optimization problem in the MPC framework can be replaced by a mixed integer linear optimization problem if the objective function is also linearized. The mixed integer linear optimization problem can be solved faster and in a more tractable way.

The rest of the paper is organized as follows. In Section II, the LTM components are defined and the original mathematical formulations are reviewed. In Section III, the extensions for ramp metering and variable speed limits are presented elaborately. Section IV first discusses the main objectives and structure of the proposed predictive ramp metering and VSL control scheme and next, the approximation and reformulation of the extended LTM towards achieving a more efficient control scheme is presented. In the case study section, the performance of the integrated ramp metering and VSL control scheme is evaluated using a real network layout, the Leuven Corridor. First, the set-up and results of calibration of the LTM are presented. Next, the closed-loop control results are compared and discussed. Finally, the paper ends with concluding remarks and possible future directions.

II. LINK TRANSMISSION MODEL

In this section, the original LTM is introduced using [17] and [19]. The LTM is capable of determining time-dependent link volumes, link travel times, and route travel times in traffic networks. To this aim, the LTM uses the so-called cumulative number of vehicles to represent the traffic evolution. The cumulative numbers of vehicles are updated using flow functions of links and nodes defined in the following subsections.

A. Link model

In the LTM framework, the traffic network is characterized by links that are connected via different types of nodes, as depicted in Fig. 1. A link i starts at an upstream boundary denoted by x_i^0 and ends at a downstream boundary denoted by x_i^{L}. The length of the link is denoted by L_i.

The cumulative number of vehicles $N(x,k)$ is defined only for the upstream and downstream boundaries of each link at the time step k, with sample time T_{s}. In order to obtain the update equations for the cumulative number of vehicles we need to define two quantities for each link; the sending and the receiving number of vehicles. The sending number of vehicles for link i is the maximum number of vehicles that can potentially leave the downstream end of this link during the time interval $\big[k\cdot T_{\mathrm{s}},(k+1)\cdot T_{\mathrm{s}}\big)$ and is defined as:

$$S_i(k)=\min\left[N\left(x_i^0,k+1-\frac{L_i}{\nu_{\mathrm{free},i}\cdot T_{\mathrm{s}}}\right)-N(x_i^{\mathrm{L}},k),\ q_{\mathrm{M},i}\cdot T_{\mathrm{s}}\right], \quad (1)$$

where $\nu_{\mathrm{free},i}$ and $q_{\mathrm{M},i}$ are the free-flow speed and the capacity of link i, respectively. Note that we have assumed a triangular Fundamental Diagram [11] (which considers that all vehicles have the same free-flow speed regardless of the flow). Moreover, in this paper, we assume that the fraction $\frac{L_i}{\nu_{\mathrm{free}}\cdot T_{\mathrm{s}}}$ has integer values and if not, we round it off towards the closest integer value (in the original LTM formulation, interpolation between grid points is used instead). The sending number is constrained by the boundary conditions at the upstream end of the link. If the downstream link boundary at time step $k+1$ is in the free-flow traffic condition, then this state must have been originated from the upstream boundary $\frac{L_i}{\nu_{\mathrm{free},i}\cdot T_{\mathrm{s}}}$ time steps earlier. Note that the sample time must be selected as $T_{\mathrm{s}}\leq\frac{L_i}{\nu_{\mathrm{free}}}$ in order to prevent vehicles from traversing a link within one sampling period.

Similarly, the receiving number of vehicles $R_i(k)$ is the maximum number of vehicles that can enter the upstream end of link i during the time interval $\big[k\cdot T_{\mathrm{s}},(k+1)\cdot T_{\mathrm{s}}\big)$ and it is formulated as:

$$R_i(k)=\min\left[N\left(x_i^{\mathrm{L}},k+1-\frac{L_i}{w_i\cdot T_{\mathrm{s}}}\right)+\rho_{\max,i}L_i-N(x_i^0,k),\ q_{\mathrm{M},i}\cdot T_{\mathrm{s}}\right], \quad (2)$$

where w_i and $\rho_{\max,i}$ are the maximum speed of the congestion wave propagating backward, and the jam density of link i, respectively. Similar to the previous case, the fraction $\frac{L_i}{w_i\cdot T_{\mathrm{s}}}$ is also rounded towards the nearest integer value.

B. Node models

In the LTM framework, links are connected to each other through different types of nodes. For each node, the transition number of vehicles is defined and determined using the sending and receiving numbers of vehicles of its connected links. Basically, the transition number of vehicles represents the maximum number of vehicles that can travel from incoming links to outgoing links of a node during the time interval $\big[k\cdot T_{\mathrm{s}},(k+1)\cdot T_{\mathrm{s}}\big)$. Moreover, we denote the set of the incoming and the outgoing links of each type of node with ℓ_{in} and ℓ_{out}, respectively. In the following, we define the transition number of vehicles for various types of nodes, starting with the simplest case.

In order to represent a difference in the characteristics of a road such as capacity, speed limits, lane change, etc., an inhomogeneous node n_{nh} can be defined. For the simplest case with one input and one output link, the transition number $G_{ij}(k)$ is formulated as

$$G_{ij}(k)=\min\big[S_i(k),R_j(k)\big],\quad i\in\ell_{\mathrm{in}}(n_{\mathrm{nh}}),\ j\in\ell_{\mathrm{out}}(n_{\mathrm{nh}}), \quad (3)$$

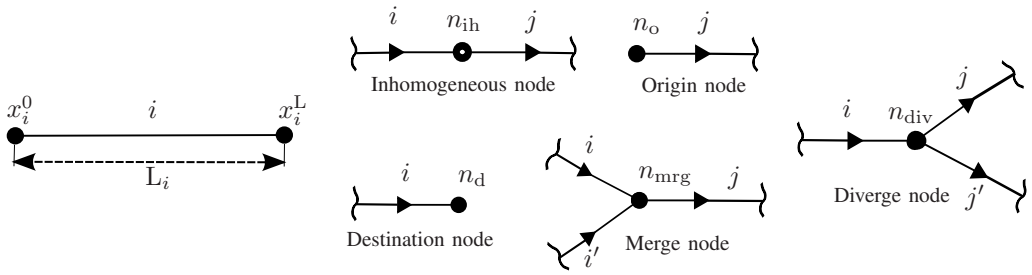


Fig. 1. (a) Link model, (b) Different node types.

where i is the unique incoming link and j is the unique outgoing link of the given node.

For each origin in the network, the corresponding origin node n_{o} is defined and the transition number of vehicles is determined as follows:

$$G_{oj}(k) = \min\left[N_o(k+1) - N(x_j^0, k),\ R_j(k)\right], \quad j \in \ell_{\mathrm{out}}(n_{\mathrm{o}}) \tag{4}$$

where j is the index of the link connected to the origin (we assume that there is only one link connected to an origin or a destination) and N_o denotes the traffic demand in origin o in terms of the cumulative number of vehicles. A simple queue model for origin o is defined as

$$\omega_o(k) = N_o(k) - N(x_j^0, k), \tag{5}$$

where $\omega_o(k)$ and $N(x_j^0, k)$ denote the number of vehicles standing in the queue at origin o and the cumulative number of vehicles that already entered the network at time step k, respectively.

For a destination in the network, the corresponding destination node n_{d} is defined and the transition number of vehicles is determined based on the sending number of vehicles of the incoming link i and possible restrictions of the destination. If the destination accepts all the flows, the transition number of vehicles will be defined as

$$G_{id}(k) = S_i(k), \quad i \in \ell_{\mathrm{in}}(n_{\mathrm{d}}), \tag{6}$$

with i the index of the unique incoming link of destination d. Otherwise, if the destination is treated as a bottleneck with a user-defined constrained outflow R_d, the $G_{id}(k)$ would be the minimum of the sending number of vehicles of the incoming link i and the maximum receiving number of vehicles of the destination d:

$$G_{id}(k) = \min\left[S_i(k), R_d(k)\right], \quad i \in \ell_{\mathrm{in}}(n_{\mathrm{d}}), \tag{7}$$

For merging of links and/or on-ramps in traffic networks, a merge node is defined. Multiple models have been proposed for the merge of links, [23]–[27]. We choose one of the priority-based merge models for two incoming links proposed in [25]. To this aim, the transition number of vehicles from an incoming link i of a merge node to the unique outgoing link j is formulated as follows:

$$\begin{cases} G_{ij}(k) = S_i(k) & \text{if} \quad R_j(k) \geq \Big(S_i(k) + S_{i'}(k)\Big), \\ G_{ij}(k) = \mathrm{median}\Big[S_i(k), R_j(k) - S_{i'}(k), \alpha_{ij} R_j(k)\Big] & \text{otherwise}, \end{cases}$$

$$\alpha_{ij} = \frac{q_{\mathrm{M},i}}{q_{\mathrm{M},i} + q_{\mathrm{M},i'}} \tag{8}$$

where $j \in \ell_{\mathrm{out}}(n_{\mathrm{mrg}})$, $i, i' \in \ell_{\mathrm{in}}(n_{\mathrm{mrg}})$, $i \neq i'$. The distribution fractions α_{ij} reflect priorities that are proportional to the capacities of the incoming links $q_{\mathrm{M},i}$. Note that the sum $\sum_{i \in \ell_{\mathrm{in}}(n_{\mathrm{mrg}})} \alpha_{ij}$ is equal to 1. The reader is referred to [27] for a general merge model with more than two incoming links.

A diverge node connects one incoming link i to its outgoing links $j \in \ell_{\mathrm{out}}(n_{\mathrm{div}})$. For the two outgoing links case, the following model for transition numbers of vehicles has been proposed in [28]:

$$G_{ij}(k) = \min\left[\beta_{ij} S_i(k), R_j(k), \frac{\beta_{ij}}{\beta_{ij'}} R_j(k)\right], \tag{9}$$

where $i \in \ell_{\mathrm{in}}(n_{\mathrm{div}})$, $j, j' \in \ell_{\mathrm{out}}(n_{\mathrm{div}})$, $j \neq j'$. The outflow of the incoming link is divided over the outgoing links according to the turning fractions β_{ij} ($\sum \beta_{ij} = 1$). The turning fractions can be fixed or variable to the route choice [27]. For the general case with more than two outgoing and/or incoming links and other types of nodes (e.g. intersection nodes) the interested reader is referred to [27].

C. Update equations

Having determined the transition number of vehicles of all nodes, the cumulative number of vehicles for the upstream and downstream boundaries of links can be updated using the following equations:

$$N(x_i^{\mathrm{L}}, k+1) = N(x_i^{\mathrm{L}}, k) + \sum_{j \in \ell_{\mathrm{out}}(n)} G_{ij}(k), \quad \text{for all } i \in \ell_{\mathrm{in}}(n), \tag{10}$$

$$N(x_j^0, k+1) = N(x_j^0, k) + \sum_{i \in \ell_{\mathrm{in}}(n)} G_{ij}(k), \quad \text{for all } j \in \ell_{\mathrm{out}}(n), \tag{11}$$

for each node n.

III. EXTENSION OF THE LTM

In this section, the LTM model is extended to include traffic control signals. First we investigate the possibility of extending the LTM for metering of on-ramps. Next, variable speed control using the LTM is discussed and required modifications of the model are explained.

A. Ramp metering

An on-ramp can be treated as a combination of an origin node and a merge node connected by a virtual link with a link length that is equal to 0. We place an origin node for the metered ramp with a constraint on its outflow to a virtual link. Thus, the transition number of vehicles of the on-ramp o to the virtual link i' can be determined as follows (based on (4)):

$$G_{oi'}(k) = \min\Big[N_o(k+1) - N(x^0_{i'}, k), r_o(k)\cdot q_{\mathrm{M},i'}\cdot T_\mathrm{s}\Big], \quad (12)$$

where $q_{\mathrm{M},i'}$ is the capacity of the virtual link i' (veh/h), T_s is the sample time, and $r_o(k) \in [0,1]$ is the metering rate. Moreover, $N_o(k+1)$ denotes the traffic demand in the on-ramp o and $N(x^0_{i'}, k)$ is the cumulative number of vehicles that already entered the virtual link i'.

Next, the transition numbers of vehicles from the virtual link to the outgoing link of the merge node can be determined by (8) using $G_{oi'}(k)$, as the sending number of vehicles of the on-ramp, and the sending number of vehicles of the mainstream incoming link.

Note that using the metering rate $r_o(k)$, one can limit the outflow of an on-ramp in order to prevent traffic congestion on the mainstream road.

B. Variable speed limit control

In this section, we elaborate on the LTM modifications required in order to emulate the effects of variable speed limit signs. Basically, a VSL can be used to modify the time that vehicles spend to reach the downstream boundary of a link. By looking at the LTM model, it can be inferred that manipulating the travel time can be realized using a time-varying speed ν_free in the model. From now on, we denote the time-varying speed ν_free with $\bar{v}(k)$, the speed that holds for all vehicles entering the link during $\big[k\cdot T_\mathrm{s}, (k+1)\cdot T_\mathrm{s}\big)$. In addition to this modification, different traffic conditions that can occur in reality should be investigated and the resulting cases should be integrated in the extended model. In Fig. 2(a) and 2(b), the results of changing the value of the VSL in the free-flow condition are shown for two cases. Before proceeding, note that without loss of generality, the VSL is assumed to be implemented at the upstream boundary of a link.

We start with the case that the speed limit increases at time step k^* to a higher value as depicted in Fig. 2(a). In this case, the vehicles are supposed to reach the downstream boundary faster. However, after the value of the speed limit is changed, there may exist some vehicles still traveling in the link that did not see the new speed limit. These vehicles reach the downstream boundary of the link without following the new speed limit. Therefore, in order to obtain a better update for the cumulative number of vehicles, these vehicles should also be taken into account.

On the other hand, when the speed limit is lowered at time step k^*, the evolution of the cumulative number of vehicles may look similar to Fig. 2(b) (if a free-flow condition is applied, otherwise in congested situations the influence of VSL may not be as apparent as what is depicted here). Vehicles that enter the link after the time instant at which the VSL value is altered, are affected by the new speed limit and will follow the new speed restriction. However, for the vehicles that are already in the link, the new speed limit is not applicable. They reach the upstream boundary with their previously assigned speed limit or the free-flow speed of the freeway. Moreover, since the new speed limit is lower, there will be a time interval in which the cumulative number of vehicles remains constant (this means that no vehicle departs from the downstream end). With this information, we now mathematically formulate these conditions.

1) Increase in the value of VSL: As shown in Fig. 2(a), it takes some time for the vehicles that did not experience the new speed limit to leave the link. Before this time, the sending number of vehicles should be determined using the old value of $\bar{v}$. Moreover, the capacity of the road $q_{\mathrm{M},i}$ is calculated using a triangular fundamental diagram constructed on the old $\bar{v}$ (we assume that the speed of the backward propagating congestion wave remains unchanged). From Fig. 2(c), the capacity $q_{\mathrm{M},i}$ can be determined as follows:

$$q_{\mathrm{M},i} = \rho_{\max,i}\cdot\frac{\bar{v}_i\cdot w_i}{\bar{v}_i + w_i} \quad (13)$$

If the value of the VSL increases at time step k^*, the speed $\bar{v}$ and the capacity of link i will be changed according to the conditions (14a)–(14b), where $\mathrm{VSL}_i(k^*)$ is the value at the time step k^* of the VSL installed at the upstream boundary of link i. In the case (14a), the vehicles that are faced by the new VSL value has not yet reached the end of the link. Therefore, the value of $N(x^\mathrm{L}_i, k+1)$ must be related to the value of $N(x^0_i)$ with a travel delay that is calculated based on the old value of the VSL, i.e. $\frac{L_i}{\bar{v}_i(k^*-1)\cdot T_\mathrm{s}}$. Once the first vehicle reaches the downstream end of the link, the condition (14b) holds and the value of $N(x^\mathrm{L}_i, k+1)$ must be calculated based on the new travel delay $\frac{L_i}{\mathrm{VSL}_i(k^*)\cdot T_\mathrm{s}}$.

Now the sending number of vehicles for link i can be determined using $N(x^\mathrm{L}_i, k+1)$ and $q_{\mathrm{M},i}(k)$ obtained from (14a)–(14b):

$$S_i(k) = \min\Big[N(x^\mathrm{L}_i, k+1) - N(x^\mathrm{L}_i, k),\ q_{\mathrm{M},i}(k)\cdot T_\mathrm{s}\Big] \quad (15)$$

On the other hand, in order to determine the receiving number of vehicles $R_i(k)$, the capacity $q_{\mathrm{M},i}$ should be altered *immediately* after the VSL value is changed. This means that the capacity should be always calculated using $\mathrm{VSL}_i(k)$, as follows:

$$R_i(k) = \min\Big[N\Big(x^\mathrm{L}_i, k+1-\frac{L_i}{w_i\cdot T_\mathrm{s}}\Big) + \rho_{\max,i}L_i - N(x^0_i, k),\ \rho_{\max,i}\cdot\frac{\mathrm{VSL}_i(k)\cdot w_i}{\mathrm{VSL}_i(k)+w_i}\cdot T_\mathrm{s}\Big], \quad (16)$$

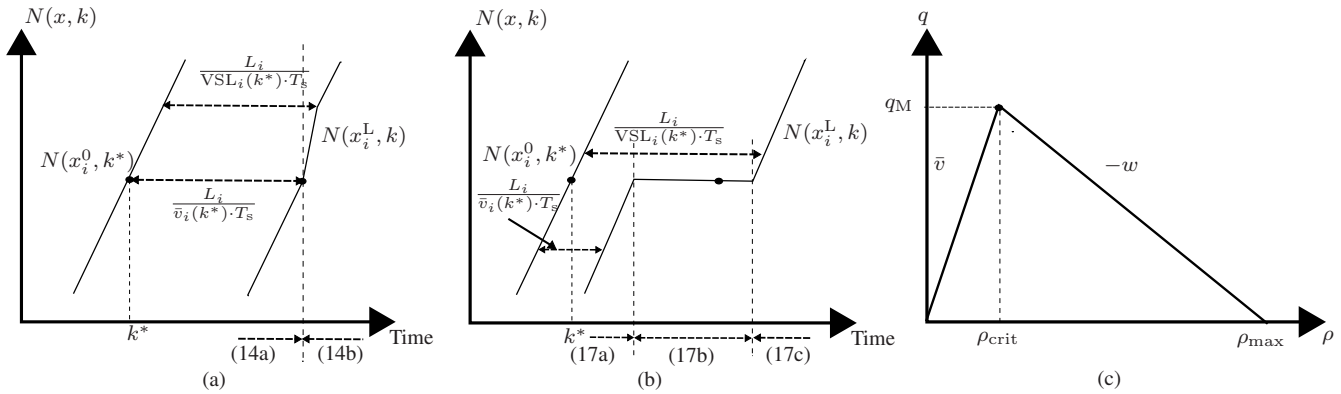


Fig. 2. (a) Increase in the value of VSL, (b) Decrease in the value of VSL, (c) Triangular Fundamental Diagram.

$$\begin{cases} \text{if} \quad N(x_i^\mathrm{L},k) < N(x_i^0,k^*): \begin{cases} N(x_i^\mathrm{L},k+1) = N\Big(x_i^0,k+1-\frac{L_i}{\bar{v}_i(k^*-1)\cdot T_\mathrm{s}}\Big), \\ \bar{v}_i(k) = \bar{v}_i(k^*-1), \\ q_{\mathrm{M},i}(k) = \rho_{\max,i}\cdot\frac{\bar{v}_i(k^*-1)\cdot w_i}{\bar{v}_i(k^*-1)+w_i}, \end{cases} & \text{(14a)} \\ \text{if} \quad N(x_i^\mathrm{L},k) \geq N(x_i^0,k^*): \begin{cases} N(x_i^\mathrm{L},k+1) = N\Big(x_i^0,k+1-\frac{L_i}{\mathrm{VSL}_i(k^*)\cdot T_\mathrm{s}}\Big), \\ \bar{v}_i(k) = \mathrm{VSL}_i(k^*), \\ q_{\mathrm{M},i}(k) = \rho_{\max,i}\cdot\frac{\mathrm{VSL}_i(k^*)\cdot w_i}{\mathrm{VSL}_i(k^*)+w_i}, \end{cases} & \text{(14b)} \end{cases}$$

This is due to the fact that for predecessor links of link i, the capacity of link i is changed when a new speed limit is introduced. But for the sending number of vehicles at the downstream boundary of link i, the capacity remains unchanged until all the vehicles that did not experience the new speed limit pass the end of link i, which is exactly the criterion distinguishing (14a) from (14b).

2) Decrease in the value of VSL: In order to formulate the problem in this case, we note that every vehicle that reaches the downstream end of a link must have entered the link either $\frac{L_i}{\bar{v}_i(k^*-1)}$ or $\frac{L_i}{\mathrm{VSL}_i(k^*)}$ time steps earlier (note that $\mathrm{VSL}_i(k^*) < \bar{v}_i(k^*-1)$). Hence for $k \geq k^*$, $N(x_i^\mathrm{L},k+1)$ can be equal to $N\big(x_i^0,k+1-\frac{L_i}{\bar{v}_i(k^*-1)}\big)$, $N\big(x_i^0,k+1-\frac{L_i}{\mathrm{VSL}_i(k^*)}\big)$ or $N(x_i^0,k^*)$. In (17a)–(17c), different conditions that may occur and the corresponding changes in the model are presented.

The first case (17a) corresponds to vehicles that had entered the link before the new speed limit is imposed. Therefore, they have traveled through the link with taking into account the old speed limit. Hence, the update for the cumulative number of vehicles $N(x_i^\mathrm{L},k+1)$ must be calculated based on the old speed $\bar{v}_i(k^*-1)$. In the second case (17b), due to the lower speed limit for the link, there is no vehicle passing the downstream end for a short period. Thus, $N(x_i^\mathrm{L},k+1)$ should be equal to the cumulative number of vehicles at the upstream boundary by the time that the VSL sign is changed (k^*). The last case (17c) describes the situation that vehicles reach the downstream end while they did encounter the new speed limit. Therefore, the cumulative number $N(x_i^\mathrm{L},k+1)$ should be calculated based on the new speed limit $\bar{v}_i(k) = \mathrm{VSL}_i(k^*)$.

The sending number of vehicles is calculated using $N(x_i^\mathrm{L},k+1)$ and $q_{\mathrm{M},i}(k)$ obtained from the conditions (17a)–(17c) and (15). However, as mentioned in the previous section, in order to determine the receiving number of vehicles $R_i(k)$, the capacity $q_{\mathrm{M},i}$ should be altered right after the new speed limit is announced (this means that for $k \geq k^*$, $q_{\mathrm{M},i}(k) = \rho_{\max,i}\cdot\frac{\mathrm{VSL}_i(k^*)\cdot w_i}{\mathrm{VSL}_i(k^*)+w_i}$ for use in the receiving number of vehicles equation (2)).

Furthermore, all the aforementioned equations in the current section and in the previous section are valid until a new speed limit is introduced. Whenever a new limit is announced, based on its new value (which could be lower or higher than the old one), the evolution equations should be updated as prescribed in this section. However, it should be noted that in our framework, the VSL values should not be updated rapidly. In fact the VSL updating interval must be bigger than the largest free-flow travel time in the link.

In the next section, the extended LTM will be utilized in the model predictive control framework for ramp metering and VSL control of freeway traffic networks.

IV. PREDICTIVE FREEWAY TRAFFIC CONTROL

Model Predictive Control (MPC) [12] is an advanced control method originally developed for control of industrial processes and recently for traffic networks [3], [4], [14], [29]. The main concept is to use a prediction model of the system and an objective function that assesses the desired performance over a given prediction horizon, and next, to find the optimal control inputs using an optimization algorithm. The optimization algorithm finds a sequence of optimal control inputs for the whole prediction horizon, but only the first control input sample is

$$\begin{cases} \text{if} \quad N(x_i^0, k^*) > N\Big(x_i^0, k+1-\dfrac{L_i}{\bar{v}_i(k^*-1)\cdot T_\mathrm{s}}\Big): \begin{cases} N(x_i^\mathrm{L}, k+1) = N\Big(x_i^0, k+1-\frac{L_i}{\bar{v}_i(k^*-1)\cdot T_\mathrm{s}}\Big), \\ \bar{v}_i(k) = \bar{v}_i(k^*-1), \\ q_{\mathrm{M},i}(k) = \rho_{\max,i}\cdot\frac{\bar{v}_i(k^*-1)w_i}{\bar{v}_i(k^*-1)+w_i}, \end{cases} & \text{(17a)} \\ \text{if} \quad N\Big(x_i^0, k+1-\dfrac{L_i}{\bar{v}_i(k^*-1)\cdot T_\mathrm{s}}\Big) \geq N(x_i^0, k^*) \geq N\Big(x_i^0, k+1-\dfrac{L_i}{\mathrm{VSL}_i(k^*)\cdot T_\mathrm{s}}\Big): \begin{cases} N(x_i^\mathrm{L}, k+1) = N(x_i^0, k^*), \\ \bar{v}_i(k) = \bar{v}_i(k^*-1), \\ q_{\mathrm{M},i}(k) = \rho_{\max,i}\cdot\frac{\bar{v}_i(k^*-1)\cdot w_i}{\bar{v}_i(k^*-1)+w_i}, \end{cases} & \text{(17b)} \\ \text{if} \quad N\Big(x_i^0, k+1-\dfrac{L_i}{\mathrm{VSL}_i(k^*)\cdot T_\mathrm{s}}\Big) > N(x_i^0, k^*): \begin{cases} N(x_i^\mathrm{L}, k+1) = N\Big(x_i^0, k+1-\frac{L_i}{\mathrm{VSL}_i(k^*)\cdot T_\mathrm{s}}\Big), \\ \bar{v}_i(k) = \mathrm{VSL}_i(k^*), \\ q_{\mathrm{M},i}(k) = \rho_{\max,i}\cdot\frac{\mathrm{VSL}_i(k^*)w_i}{\mathrm{VSL}_i(k^*)+w_i}, \end{cases} & \text{(17c)} \end{cases}$$

applied to the system and the procedure is repeated for the next control step but with a shift in the prediction period.

A. Traffic objective and constraints

For a traffic network, one can define different objective functions based on travel time, fuel consumption of vehicles, emissions, etc. The objective function we choose here is the total time spent (TTS) in the traffic network, consisting of the time vehicles spend in queues at mainstream origins and on-ramps and the travel time on the freeway. The TTS objective function in the MPC framework is formulated as follows:

$$J_{\mathrm{TTS}}(k_\mathrm{c}) = T_\mathrm{s}\cdot\sum_{k=Mk_\mathrm{c}}^{M(k_\mathrm{c}+N_\mathrm{p})-1}\Bigg[\sum_{o\in O_\mathrm{all}}\omega_o(k)+\sum_{i\in\ell_\mathrm{all}}\big(N(x_i^0,k)-N(x_i^\mathrm{L},k)\big)\Bigg], \quad (18)$$

where T_s is the simulation sample time, k_c is the controller time step counter, and k is the model time step counter. In fact, we assume that the controller sample time T_c is an integer multiple of the simulation sample time: $T_\mathrm{c} = MT_\mathrm{s}$. In the time intervals between consecutive control time steps, the control inputs are not altered. Moreover, N_p is the prediction horizon, ω_o is the queue length at origin o, and ℓ_all and O_all are the set of all links and the set of all origins, respectively. Moreover, the optimal control inputs obtained from the MPC controller may in general have undesired fluctuations over time. Note that the control inputs are in fact the metering rate and the values shown on the VSL signs. Therefore, to avoid large fluctuations a penalty term on the control input deviations is introduced and added to the objective function. The penalty term on the ramp metering input is formulated as

$$\zeta_\mathrm{r}\sum_{l=k_\mathrm{c}}^{k_\mathrm{c}+N_\mathrm{p}-1}\sum_{o\in O_\mathrm{ramp}}\big|r_o(l)-r_o(l-1)\big|, \quad (19)$$

where r_o is the metering signal, O_ramp is the set of indices of metered ramps[1], and ζ_r is a weighting factor. Similarly, penalizing the VSL input can be formulated as

$$\zeta_\mathrm{v}\sum_{l=k_\mathrm{c}}^{k_\mathrm{c}+N_\mathrm{p}-1}\sum_{i\in\ell_\mathrm{vsl}}\big|\mathrm{VSL}_i(l)-\mathrm{VSL}_i(l-1)\big|, \quad (20)$$

where VSL_i is the VSL input, ℓ_vsl is the set of indices of links equipped with VSL signs, and ζ_v is a weighting factor.

Moreover, to reduce the complexity, control variables are sometimes taken constant after passing a predefined control horizon N_c. Taking this into account, N_p in (19) should be replaced by N_c. Moreover, to take into account the physical limitation of queues at on-ramps, we use a hard constraint on the queue lengths. The total objective function along with the queue length constraint and the LTM as prediction model constitute a nonlinear nonconvex optimization problem that has to be solved at every control step in the MPC framework to find the optimal control signals. There is no guarantee to find a unique global solution for the optimization problem and furthermore, solving the nonlinear optimization may take considerable time. In the next section, a solution to this problem is proposed. More specifically, we will transform the nonlinear nonconvex optimization problem into a mixed integer linear programming (MILP) problem.

Using the methods proposed in [30], [31], one can transform the model and the objective function into a system of linear equations and inequalities involving real and integer variables and formulate an MILP problem. The MILP problem can be efficiently solved using existing MILP solvers like CPLEX or GLPK (see [32]). Note that MILP solvers can find the *global* optimum of the MILP problem.

B. Reformulation of the LTM

In order to obtain an MILP problem, we first transform the extended LTM into a system of linear inequalities with mixed real and binary variables, inspired by the method proposed in [31]. Compared to the mixed logical dynamical form proposed in [33], this system of linear inequalities is less complex, as it needs less number of variables and less number of inequalities to model the system.

In the following, we present the transformation approach for different parts of the LTM. First, we consider the parts of the model that are not affected by the VSL extension, and next,

[1]It should be noted that outflows of mainstream origins can also be controlled in some cases (mainstream metering), so in that case they can also be included in the set O_ramp.

we present the reformulation of the proposed VSL extensions into the linear form.

For links that are not equipped with VSL signs, the delays $\frac{L_i}{\nu_{\text{free},i}\cdot T_\text{s}}$ and $\frac{L_i}{w_i\cdot T_\text{s}}$ would be constant over time. Moreover, the transition number of vehicles (3) for a homogeneous node is the minimum of the following three affine functions:

$$f_{G_{ij},1}(k) = N\left(x_i^0, k+1-\frac{L_i}{\nu_{\text{free},i}\cdot T_\text{s}}\right) - N(x_i^\text{L}, k), \quad (21)$$

$$f_{G_{ij},2}(k) = q_{\text{M},i}(k)\cdot T_\text{s}, \quad (22)$$

$$f_{G_{ij},3}(k) = N\left(x_i^\text{L}, k+1-\frac{L_i}{w_i\cdot T_\text{s}}\right) + \rho_{\max,i}L_i - N(x_i^0, k). \quad (23)$$

We introduce two binary variables $\delta_{G_{ij},1}$ and $\delta_{G_{ij},2}$, and we also define the constraint $\delta_{G_{ij},1}(k)+\delta_{G_{ij},2}(k)\leq 1$ so that we can have only three combinations for $(\delta_{G_{ij},1}, \delta_{G_{ij},2})$. The transformation of (3) to a set of linear inequalities is:

$$\delta_{G_{ij},1}(k)+\delta_{G_{ij},2}(k)\leq 1, \quad (24)$$

$$\big(\delta_{G_{ij},1}(k)+\delta_{G_{ij},2}(k)\big)\cdot M^- \leq G_{ij}(k) - f_{G_{ij},1}(k) \leq 0, \quad (25)$$

$$\big(1-\delta_{G_{ij},1}(k)+\delta_{G_{ij},2}(k)\big)\cdot M^- \leq G_{ij}(k) - f_{G_{ij},2}(k) \leq 0, \quad (26)$$

$$\big(1+\delta_{G_{ij},1}(k)-\delta_{G_{ij},2}(k)\big)\cdot M^- \leq G_{ij}(k) - f_{G_{ij},3}(k) \leq 0, \quad (27)$$

where M^- is a negative number with a large absolute value that ensures:

$$|M^-| > \Big|\min_{n\in\{1,2,3\},k}[G_{ij}(k)-f_{G_{ij},n}(k)]\Big| \quad (28)$$

The equivalence of (3) and (24)–(27) is validated in (29a)–(29c). The same procedure can be applied to origin and destination nodes.

The transition number of vehicles for merging nodes can also be transformed into linear inequalities. Consider the merging of two incoming links i and i':

$$\begin{cases} G_{ij}^\text{mrg}(k)=S_i(k),\ G_{i'j}^\text{mrg}(k)=S_{i'}(k) \\ \qquad\qquad \text{if} \quad \text{R}_\text{j}(\text{k}) \geq \Big(\text{S}_\text{i}(\text{k})+\text{S}_\text{i'}(\text{k})\Big), \\ \begin{cases} G_{ij}^\text{mrg}(k)=\text{median}\Big[S_i(k), R_j(k)-S_{i'}(k), \alpha_{ij}R_j(k)\Big], \\ G_{i'j}^\text{mrg}(k)=\text{median}\Big[S_{i'}(k), R_j(k)-S_i(k), \alpha_{i'j}R_j(k)\Big] \end{cases} \\ \qquad\qquad \text{otherwise}, \end{cases} \quad (30)$$

with $\alpha_{ij}=\dfrac{q_{\text{M},i}}{q_{\text{M},i}+q_{\text{M},i'}}$, $\qquad \alpha_{i'j}=\dfrac{q_{\text{M},i'}}{q_{\text{M},i}+q_{\text{M},i'}}$.

We consider the transformation of the transition number of vehicles from link i to the outgoing link j, $G_{ij}^\text{mrg}(k)$ (a similar approach can be applied to $G_{i'j}^\text{mrg}(k)$). First of all, for condition $R_j(k)\geq S_i(k)+S_{i'}(k)$, we can define a binary variable $\delta_\text{mrg}(k)$. We assume that the binary variable is set to 1 whenever the condition holds. Now the condition can be transformed to a linear form using the following basic rule [33]:

$$[f(x)\leq 0]\Leftrightarrow[\delta=1], \quad \text{iff} \quad \begin{cases} f(x)\leq M\cdot(1-\delta), \\ f(x)\geq \epsilon+(m-\epsilon)\cdot\delta, \end{cases} \quad (31)$$

with f an affine function defined over a bounded set X of the input variable x, m and M the lower and upper bounds of f over X, $\epsilon>0$ a small tolerance namely the machine precision[2]. The transformed condition is formulated as follows:

$$S_i(k)+S_{i'}(k)-R_j(k)\leq M^\text{mrg}\cdot\big(1-\delta_\text{mrg}(k)\big), \quad (32)$$

$$S_i(k)+S_{i'}(k)-R_j(k)\geq \epsilon+\big(m^\text{mrg}-\epsilon\big)\cdot\delta_\text{mrg}(k), \quad (33)$$

where M^mrg and m^mrg denote the upper and the lower bounds of $S_i(k)+S_{i'}(k)-R_j(k)$. Note that an estimation of the bounds $m^\text{mrg}, M^\text{mrg}$ can be obtained based on the trajectories of $S_i, R_i, S_{i'}$ obtained from simulation and historical data from the traffic network. Note that a tight upper/lower bound estimation is not crucial, although it is better from a computational point of view.

Moreover, we include the following constraint:

$$\big(1-\delta_\text{mrg}(k)\big)\cdot m_{ij} \leq G_{ij}^\text{mrg}(k)-S_i(k) \leq \big(1-\delta_\text{mrg}(k)\big)\cdot M_{ij}, \quad (34)$$

with m_{ij} and M_{ij} the lower and the upper bounds of $G_{ij}^\text{mrg}(k)-S_i(k)$, respectively. Now if $\delta_\text{mrg}=1$, then $G_{ij}^\text{mrg}(k)=S_i(k)$.

Now for simplicity, we assign new names for the affine functions in the argument of the median operator:

$$g_1(k)=S_{i'}(k), \quad (35)$$

$$g_2(k)=R_j(k)-S_{i'}(k), \quad (36)$$

$$g_3(k)=\alpha_{ij}R_j(k). \quad (37)$$

Moreover, three binary variables $\delta_{\text{med},1}(k)$, $\delta_{\text{med},2}(k)$, $\delta_{\text{med},3}(k)$ are defined. Six combinations may occur and therefore we add the following two constraints:

$$\delta_{\text{med},1}(k)+\delta_{\text{med},2}(k)+\delta_{\text{med},3}(k)\geq 1, \quad (38)$$

$$\delta_{\text{med},1}(k)+\delta_{\text{med},2}(k)+\delta_{\text{med},3}(k)\leq 2. \quad (39)$$

to cover all the possible conditions. It can be verified that the constraints[3] (40)–(49) along with (32)–(34) and (38)–(39) are an equivalent representation of (30). Note that $M^-\ll 0$ and $M^+\gg 0$ in (40)–(49) should be chosen in a similar way as in (28).

Now we consider the links that have speed limit signs installed and activated at their upstream boundary. According to Section III-B, the delay term in the sending number of vehicles is time-varying (note that the congestion wave speed w is assumed not to be altered. Hence, the delay term $\frac{L_i}{w_i\cdot T_\text{s}}$ is constant). Therefore, the function (21) is no longer affine. Moreover, in the presence of speed limits, the conditions introduced in Section III-B need to be taken into account in

[2]It is mainly used to change a strict inequality into a non-strict inequality.

[3]the time index k is dropped for the ease of readability.

$$\begin{cases} f_{G_{ij},1}(k) \le f_{G_{ij},2}(k) \text{ and } f_{G_{ij},1}(k) \le f_{G_{ij},3}(k) \iff G_{ij}(k) = f_{G_{ij},1}(k),\ \big(\delta_{G_{ij},1}(k), \delta_{G_{ij},2}(k)\big) = (0,0) & \text{(29a)} \\ f_{G_{ij},2}(k) \le f_{G_{ij},1}(k) \text{ and } f_{G_{ij},2}(k) \le f_{G_{ij},3}(k) \iff G_{ij}(k) = f_{G_{ij},2}(k),\ \big(\delta_{G_{ij},1}(k), \delta_{G_{ij},2}(k)\big) = (1,0) & \text{(29b)} \\ f_{G_{ij},3}(k) \le f_{G_{ij},1}(k) \text{ and } f_{G_{ij},3}(k) \le f_{G_{ij},2}(k) \iff G_{ij}(k) = f_{G_{ij},3}(k),\ \big(\delta_{G_{ij},1}(k), \delta_{G_{ij},2}(k)\big) = (0,1) & \text{(29c)} \end{cases}$$

$$g_1 - g_3 \le (1 - \delta_{\text{med},1} + \delta_{\text{med},3}) \cdot M^+, \tag{40}$$
$$g_3 - g_1 \le (1 - \delta_{\text{med},3} + \delta_{\text{med},1}) \cdot M^+, \tag{41}$$
$$g_2 - g_1 \le (1 - \delta_{\text{med},1} + \delta_{\text{med},2}) \cdot M^+, \tag{42}$$
$$g_1 - g_2 \le (2\delta_{\text{med},1} + \delta_{\text{med},2} + \delta_{\text{med},3} - 1) \cdot M^+, \tag{43}$$
$$g_3 - g_2 \le (3 - 2\delta_{\text{med},1} - \delta_{\text{med},2} - \delta_{\text{med},3}) \cdot M^+, \tag{44}$$
$$g_2 - g_3 \le (1 + \delta_{\text{med},1} - \delta_{\text{med},2}) \cdot M^+, \tag{45}$$
$$(\delta_{\text{med},1} + 2\delta_{\text{med},2} + \delta_{\text{med},3} - 1) \cdot M^- + \delta_{\text{mrg}} \cdot M^- \le G_{ij}^{\text{mrg}} - g_1 \le (\delta_{\text{med},1} + 2\delta_{\text{med},2} + \delta_{\text{med},3} - 1) \cdot M^+ + \delta_{\text{mrg}} \cdot M^+, \tag{46}$$
$$(1 + \delta_{\text{med},1} - \delta_{\text{med},2} + \delta_{\text{med},3}) \cdot M^- + \delta_{\text{mrg}} \cdot M^- \le G_{ij}^{\text{mrg}} - g_2 \le (1 + \delta_{\text{med},1} - \delta_{\text{med},2} + \delta_{\text{med},3}) \cdot M^+ + \delta_{\text{mrg}} \cdot M^+, \tag{47}$$
$$(2 - \delta_{\text{med},1} + \delta_{\text{med},2} - \delta_{\text{med},3}) \cdot M^- + \delta_{\text{mrg}} \cdot M^- \le G_{ij}^{\text{mrg}} - g_2 \le (2 - \delta_{\text{med},1} + \delta_{\text{med},2} - \delta_{\text{med},3}) \cdot M^+ + \delta_{\text{mrg}} \cdot M^+, \tag{48}$$
$$(3 - \delta_{\text{med},1} - 2\delta_{\text{med},2} - \delta_{\text{med},3}) \cdot M^- + \delta_{\text{mrg}} \cdot M^- \le G_{ij}^{\text{mrg}} - g_3 \le (3 - \delta_{\text{med},1} - 2\delta_{\text{med},2} - \delta_{\text{med},3}) \cdot M^+ + \delta_{\text{mrg}} \cdot M^+, \tag{49}$$

order to determine the correct delay in (1) and also the capacity $q_{\text{M},i}(k)$.

In order to simplify the transformation, we assume that the VSL can take values only from a finite set. This is a realistic assumption since the VSL signs on roads typically show only 3-5 discrete numbers for the speed limit (e.g. 50, 70, 100, 120 km/h). Therefore, the sending number of vehicles can be reformulated as sum of cumulative number of vehicles with different discrete delays:

$$S_i(k) = \min\Big[q_{\text{M},i},\ \sum_{n=1}^{N_{\text{speed}}} \delta_n(k) \cdot N\Big(x_i^0, k + 1 - \frac{L_i}{\text{VSL}_{i,n} \cdot T_{\text{s}}}\Big) - N(x_i^{\text{L}}, k)\Big], \tag{50}$$

with N_{speed} the total number of discrete VSL values. On the other hand, we can define two binary variables $\delta_{i,\text{inc}}$ and $\delta_{i,\text{dec}}$, as follows:

$$\text{VSL}_i(k) - \text{VSL}_i(k-1) - \epsilon \ge 0 \Longleftrightarrow \delta_{i,\text{inc}}(k) = 1, \tag{51}$$
$$\text{VSL}_i(k) - \text{VSL}_i(k-1) + \epsilon \le 0 \Longleftrightarrow \delta_{i,\text{dec}}(k) = 1. \tag{52}$$

Hence, we can capture and store the value $N(x_i^0, k^*)$ in $Z_{N,i}$, formulated as

$$Z_{N,i}(k) = \big[\delta_{i,\text{inc}}(k) + \delta_{i,\text{dec}}(k)\big] \cdot N(x_i^0, k). \tag{53}$$

Similarly, we can store the VSL value in an auxiliary variable $Z_{\text{VSL},i}$:

$$Z_{\text{VSL},i}(k) = \big[\delta_{i,\text{inc}}(k) + \delta_{i,\text{dec}}(k)\big] \cdot \text{VSL}_i(k). \tag{54}$$

If $k = k^*$, then $Z_{\text{VSL},i}(k^*) = \text{VSL}_i(k^*)$, otherwise $Z_{\text{VSL},i}(k) = 0$. Now we review two basic rules adopted from [33]. The product of two binary variables δ_1 and δ_2 can be replaced by an auxiliary binary variable $\delta_3 \triangleq \delta_1 \cdot \delta_2$. It can be verified that [30], [33]

$$\delta_3 = \delta_1 \cdot \delta_2 \ \text{ iff } \begin{cases} -\delta_1 + \delta_3 \le 0, \\ -\delta_2 + \delta_3 \le 0, \\ \delta_1 + \delta_2 - \delta_3 \le 1. \end{cases} \tag{55}$$

Moreover, multiplication of a binary variable δ with an affine function $f(\cdot)$ defined over a bounded set X of the variable x can be replaced by an auxiliary variable $z \triangleq \delta \cdot f(x)$, meaning that $z = 0$ when $\delta = 0$ and $z = f(x)$ in case $\delta = 1$. It can be proved that [30], [33]

$$z = \delta \cdot f(x) \ \text{ iff } \begin{cases} z \le M \cdot \delta, \\ z \ge m \cdot \delta, \\ z \le f(x) - m \cdot (1 - \delta), \\ z \ge f(x) - M \cdot (1 - \delta), \end{cases} \tag{56}$$

with m and M the lower and upper bounds of $f(\cdot)$ over the set X, respectively. Using the equivalent forms (31), (55), (56), we can transform (50)–(54) and subsequently, the VSL conditions (14a)–(14b) and (17a)–(17c) into a system of linear equations and inequalities.

C. Final mixed integer linear optimization problem

After transforming the LTM into a linear form, the total objective function should also be reformulated. The TTS objective function is already linear. But the penalty terms (19) and (20) are piecewise affine. It can be transformed into a mixed-integer linear form by defining additional binary and auxiliary variables. However, there exists a more efficient way to recast the penalty terms as linear problems without introducing binary variables. It can be easily proved that the following optimization problems have the same optimal

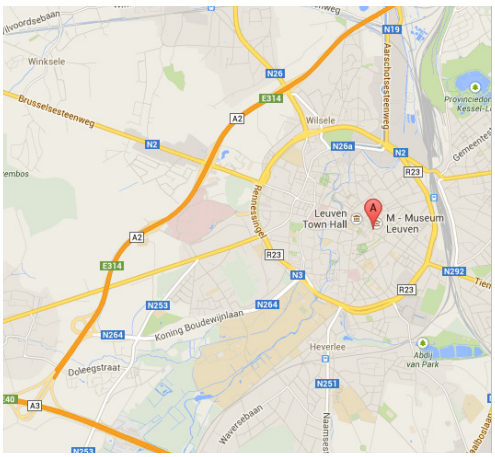


Fig. 3. Google maps illustration of the A2 freeway in Leuven, Belgium. The South-North direction is used for the evaluation.

Fig. 4. Schematic representation of the network.

solution:

$$\min_{\theta} \sum |\theta_i| \Longleftrightarrow \begin{cases} \min_{\theta,\beta} \sum \beta_i \\ \beta_i \geq \theta_i \\ \beta_i \geq -\theta_i \end{cases} \tag{57}$$

Using this technique along with the reformulated model, the final MILP problem can be constructed.

V. CASE STUDY

This section describes the evaluation of the proposed integrated ramp metering and variable speed limit control scheme. First, the benchmark network and the selected traffic scenario will be described. Next, identification and calibration of the LTM for the benchmark network is presented. Finally, closed-loop control results of the freeway network using the proposed model predictive control schemes will be presented.

A. Benchmark network and simulation set-up

The South-North direction of the A2 freeway near Leuven, Belgium, is taken as the benchmark network. Fig. 3 shows a Google maps illustration of this freeway. The freeway consists of 4 on-ramps and 4 off-ramps and congestion is triggered at the most downstream on-ramp. Ramp metering installations are placed at every on-ramp and variable speed limits are placed directly downstream of every off-ramp. Fig. 4 shows a schematic representation of the freeway network. The freeway is divided up into 11 links with two lanes which have the following lengths in kilometers, ordered from upstream to downstream: 1, 1.18, 0.42, 1.03, 0.53, 0.52, 0.72, 0.82, 0.4, 1.8, and 1.0 km. Inductive loop detectors measuring the number of vehicles that have passed, and their average speed are located at the upstream and downstream end of every link.

The four off-ramps are located at positions 2.18, 3.63, 4.68, and 6.22 km. The downstream end of the diverging section is considered as the off-ramp location. The on-ramps are located at positions 2.6, 4.16, 5.4, and 6.86 km. The beginning of the merging section is taken as the on-ramp location. At every on-ramp, the stop-line of the ramp metering installation is located 150 meters upstream of the merging area. The on-ramp queue has a storage space of 900 meters. Loop detectors are placed at the stop-line, and at the maximum queue length, 900 meters upstream of the stop-line.

The freeway is simulated using the microscopic simulation software package VISSIM 5.30. The Wiedemann 99 model is used to reproduce the driving behavior. The following parameters of VISSIM have been altered from the default settings: CC0 3.50 m, CC1 1.1 s, CC2 8.00 m, CC3 -8.00 (-), CC4 -0.50 (-), CC5 0.60 (-), CC6 6.00 (-), CC7 -0.25 m/s^2, CC8 1.00 m/s^2, and CC9 1.50 m/s^2. Using these parameters, a capacity flow of 1800 veh/h/lane, and a queue discharge rate of 1800 (veh/h/lane) are obtained, thus, there is no capacity drop present. The sampling time in Vissim is 0.2 seconds. Furthermore, the demand profiles for the mainstream road and the on-ramps are presented in Table I. The (fixed) split fractions are 28.09% of the mainstream flow for off-ramp 1, 6.79% for off-ramp 2, 12.90% for off-ramp 3, and 10.26% for off-ramp 4. Note that the demand of on-ramp 4 has an increase to 1000 (veh/h) for a short period and then it decreases to a lower level. This high on-ramp demand causes congestion on the mainstream road that propagates all the way back to the upstream end of the freeway (as also illustrated in Fig. 6(a)).

TABLE I
MAINSTREAM AND ON-RAMPS DEMANDS (VEH/H)

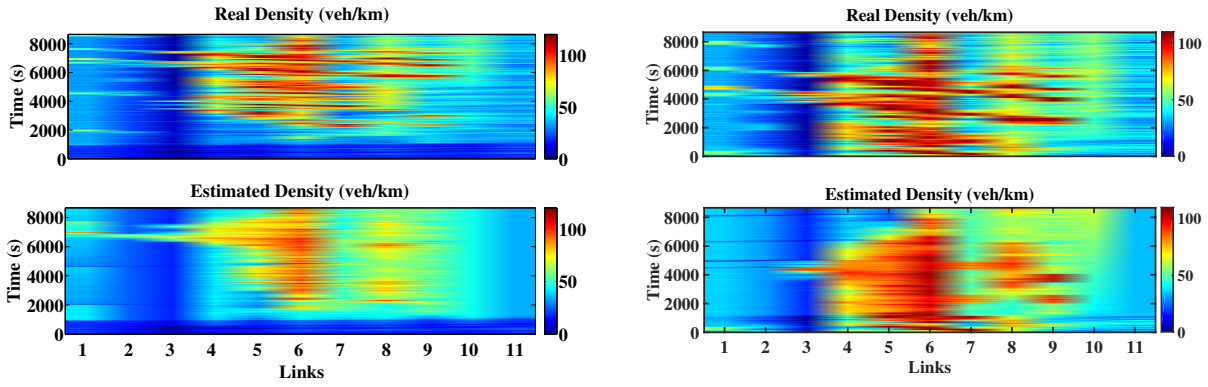
Timing	Main	Ramp 1	Ramp 2	Ramp 3	Ramp 4
0-900 s	2225	240.5	223	265	250
900-1800 s	4450	750	446	530	500
1800-2100 s	4450	750	446	530	1000
2100-2400 s	4450	750	446	530	250
2400-6300 s	4450	750	446	530	500
6300-7200 s	2225	240.5	223	265	250

MATLAB is used to compute the optimal ramp metering and VSL signals. The simulation sample time in MATLAB is 5 seconds. The MPC controller determines the optimal control inputs every 60 seconds using the TOMLAB optimization toolbox (the *patternsearch* solver is used to solve the nonlinear optimization problem, and the *CPLEX* solver inside the TOMLAB toolbox is used to solve the MILP problem) in MATLAB on a computer with a 3.6 GHz processor and 8Gb RAM.

B. Calibration of the LTM

Using simulation data from VISSIM, the LTM is calibrated. The identification procedure for estimating the parameters of the LTM is formulated as a nonlinear optimization problem solved using the global optimization solver *patternsearch*. The objective is to minimize the difference between the real densities and the estimated densities from the LTM, formulated as follows:

$$\mathcal{J} = \frac{1}{n_{\mathrm{d}} n_{\ell}} \sum_{i=1}^{n_{\ell}} \sum_{k=1}^{n_{\mathrm{d}}} \left(\rho_i(k) - \widehat{\rho}_i(k)\right)^2, \tag{58}$$



(a) (b)

Fig. 5. Calibration of the LTM: real and estimated densities of all links for two traffic scenarios with different demand profiles: (a) congestion starts in the middle of the simulation period, (b) network is already congested at the beginning of the simulation period.

TABLE II
ESTIMATED PARAMETERS FOR EACH LINK

Link	ν_{free} (km/h)	w (km/h)	$\rho_{\max}$ (veh/km)	q_{M} (veh/h)
1	119.00	23.00	290.34	5596
2	119.84	24.00	271.24	5423
3	115.00	22.37	219.78	4115
4	114.06	24.00	263.23	5219
5	110.50	23.97	213.65	4208
6	116.96	20.25	289.23	4992
7	114.01	18.87	232.76	3777
8	112.34	22.28	268.53	4992
9	118.25	23.00	218.15	4200
10	120.00	20.00	288.41	4944
11	118.91	22.90	283.33	5440

where n_{d}, n_ℓ, ρ_i, and $\widehat{\rho}_i$ denote the number of data samples, the number of links, the real density and the density predicted by the LTM, respectively. The density of links can be calculated using the cumulative number of vehicles at upstream and downstream boundaries of links, as follows:

$$\rho_i(k) = \frac{N(x_i^0, k) - N(x_i^{\text{L}}, k)}{L_i}. \tag{59}$$

The nonlinear optimization problem is solved using the function *patternsearch* from the *Global Optimization* toolbox of MATLAB. The optimization algorithm is run 10 times for different random initial points in order to prevent reaching a local optimum only. The obtained parameters of the LTM are presented in Table II. The links are numbered from upstream to downstream. Note that in the network's layout, there are extra lanes from 200 m before the off-ramps and also for 200 m after the on-ramps. However, we do not define extra links in the LTM to model these small parts, but we take into account the cumulative number of vehicles leaving (entering) these links to fit 2-lane LTM links to the data. This is consistent with having different maximum densities for different links in Table II.

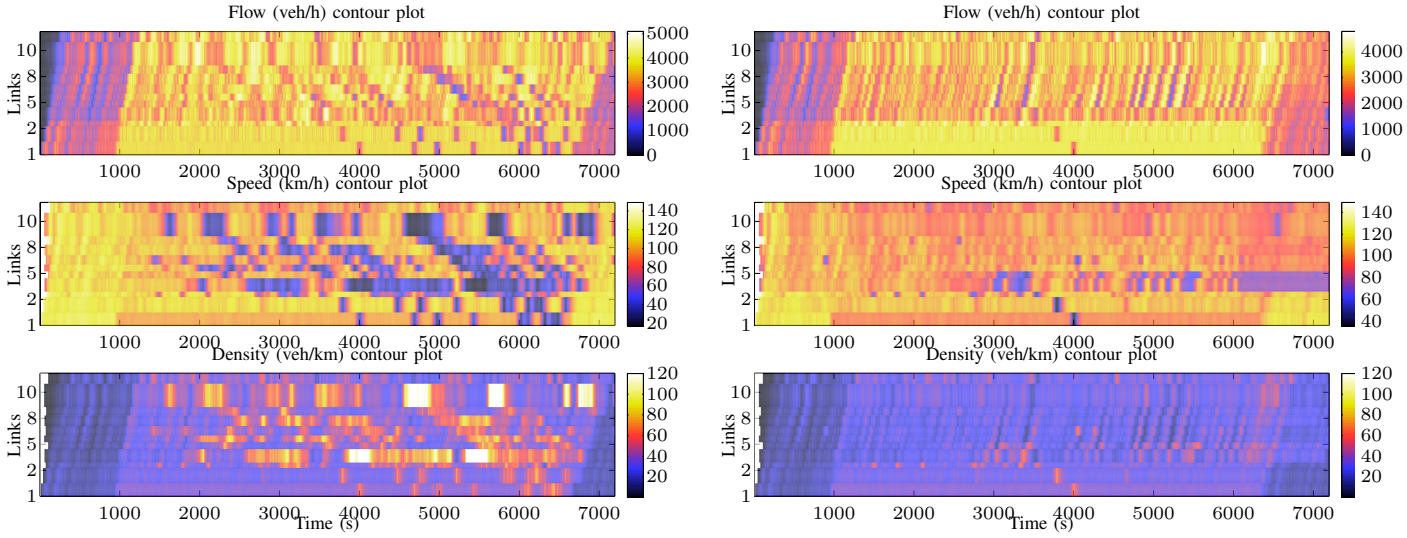
Moreover, results presented in Fig. 5 verify that the calibrated LTM is able to estimate traffic densities close to the ones obtained from the simulation data. In order to numerically evaluate the calibrated model, we choose to compare the total time spent on the freeway for both real data and the estimated data from the calibrated LTM. This is very important as the main performance measure in our simulations is the TTS and the model should be able to provide a sufficiently accurate estimation of the TTS values over a prediction horizon. Using the real and estimated densities, the TTS values are calculated for the whole simulation time horizon and the obtained deviation is around $0.62 - 1.22\%$ with respect to the real TTS of 3370.5 $(\text{veh} \cdot \text{h})$. This amount of deviation is acceptable for our model-based traffic control purpose presented in the next section.

C. Control results

Fig. 6(a) shows an illustration of the uncontrolled situation. It can be observed that congestion forms near the most downstream on-ramp and propagates upstream. Once it reaches the most upstream on-ramp, the congestion increases close to this on-ramp. The total time spent, which is the sum of the mainstream traveling time and the time spent in queues at on-ramps is 1115.63 $(\text{veh} \cdot \text{h})$. The total time spent on the mainstream road only is 1061.41 $(\text{veh} \cdot \text{h})$.

Now the uncontrolled case is compared with the cases in which predictive ramp metering and VSL control is applied. Two proposed methods are implemented, nonlinear MPC based on the original formulation of the extended LTM, and MILP-MPC. The optimization problems integrated in both methods have a queue length constraint of 100 vehicles for all on-ramps. The improvement in the TTS values for the nonlinear MPC and the MILP case along with the average computation time (required for solving each optimization step) are compared in Table III. As can be inferred, the control approaches are able to provide approximately $10 - 14\%$ reduction in the TTS value. Moreover, the MPC approaches provide a significant reduction in the total time spent on the mainstream road (around 44% less than the uncontrolled case). Although this comes at the price of having longer queues at the on-ramps, the overall TTS is considerably lower than the uncontrolled case.

Results of closed-loop simulation using the nonlinear MPC method with $N_{\text{p}} = 7, N_{\text{c}} = 3$ are illustrated in Fig. 6(b) and Fig. 8. Moreover, the results of the MILP-MPC approach



(a) (b)

Fig. 6. Flow, speed and density plots for all links over time: (a) Uncontrolled case, (b) Controlled using nonlinear MPC.

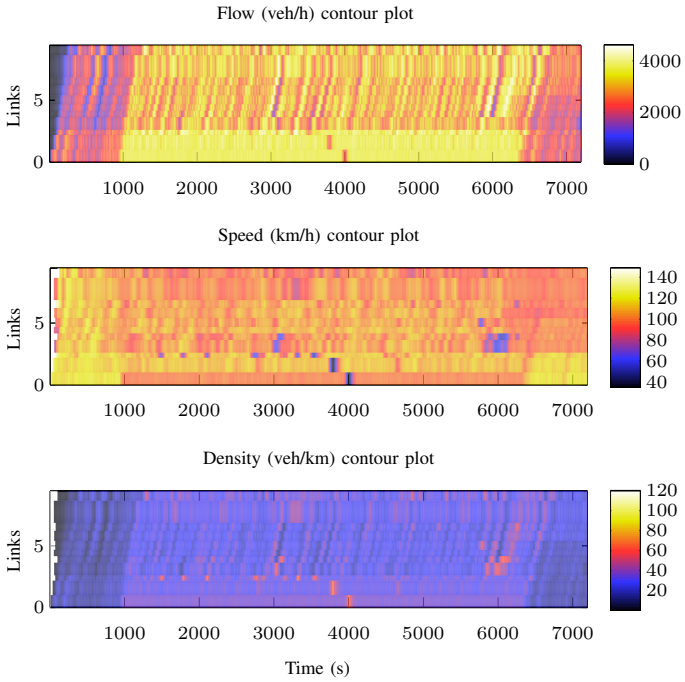


Fig. 7. Flow, speed and density plots for all links over time: MILP-MPC approach.

are presented in Fig. 7 and Fig. 9. Comparing Fig. 6(a) and Fig. 6(b), it can be observed that the MPC controller is able to resolve the congestion caused by the high demand in the most downstream on-ramp. In fact, the MPC controller (in both the nonlinear and the MILP case) achieves this by limiting the outflow of the on-ramps (especially the first 3 on-ramps, as also shown in Fig. 8(a)) and by imposing the speed limit mostly for the first links of the freeway, as depicted in Fig. 8(b).

Moreover, the queue lengths at the on-ramps for both uncontrolled and controlled (ramp metering) cases are shown in Fig. 10. As can be seen, the queue lengths in the controlled case are considerably higher than in the uncontrolled condition. However, they do not exceed the 100 (veh) constraint on the queue length. In addition, using the MPC schemes, the total time spent in the network is improved and moreover, the congestion (with reduced mean speed and high densities of vehicles in several links) is significantly attenuated, as illustrated in Fig. 6(b) and Fig. 7. Note that because the micro-simulation has a stochastic nature, it is possible that sometimes the nonlinear MPC approach performs better than the MILP-MPC method and vice versa. It should be noted that in both methods we use rounding approximations to make the delays integer variables. Moreover, in the MILP-MPC approach, we just reformulate the LTM. Therefore, we expect that the performance of both methods should be close to each other. As can be observed in Fig. 6(b) and Fig. 7, in this run of the closed-loop simulation, the congestion level is less in the MILP-MPC control case. Instead, the queue lengths in the nonlinear MPC case are a bit smaller than in the MILP-MPC case. Furthermore, as discussed in Section IV, we use a penalty term in the total objective function to reduce fluctuations in the control inputs, as can be seen in Fig. 8 and Fig. 9. The penalty term and the TTS objective function are normalized by their nominal values (the uncontrolled case) in the total objective function. In addition, the penalty term is weighted by 0.2.

Regarding the computation time, as can be inferred from Table III, the nonlinear optimization problem is solved for several random initial points in each MPC control step, since there may exist multiple local optimal solutions. On the contrary, the MILP approach is more efficient and it provides the global solution of the reformulated problem. Moreover, due to the stochastic nature of the micro-simulation, the performance of nonlinear MPC is sometimes worse than the MILP approach (note that both N-MPC and MILP-MPC use rounding approxi-

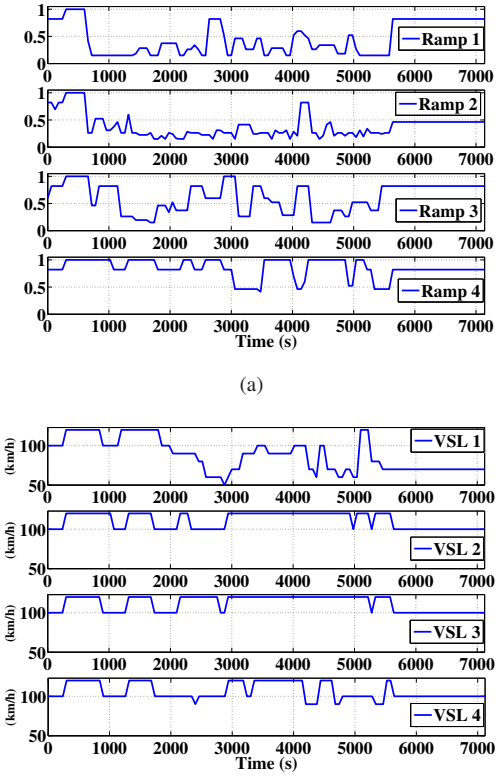


Fig. 8. Nonlinear MPC control inputs: (a) Ramp metering signals,(b) VSL signals.

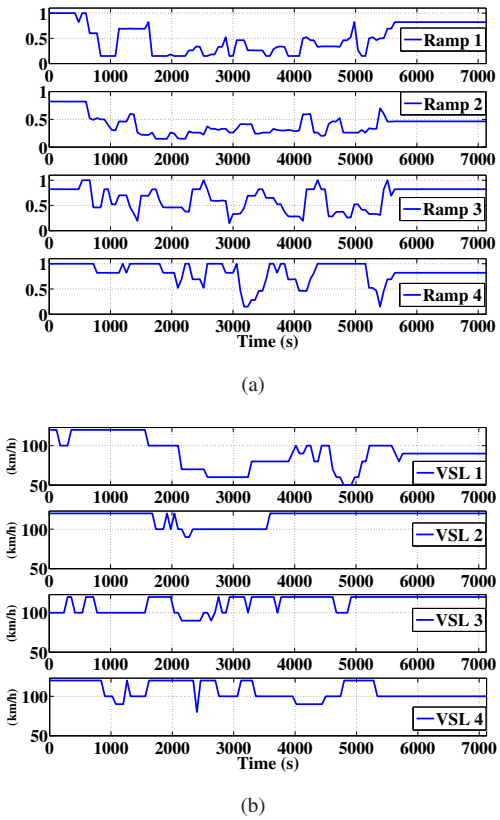


Fig. 9. MILP-MPC control inputs: (a) Ramp metering signals, (b) VSL signals.

mations for the delays in the model). Furthermore, experiments for $N_{\mathrm{p}} > 7$ and $N_{\mathrm{c}} > 3$ show that the computation time (particularly for the nonlinear approach) grows exponentially, while the reduction in the TTS is not significant. Moreover, for small N_{p} the queues are not dissolved until the end of the simulation period. One way to prevent this is to increase N_{p}. However, for large values of N_{p}, it might be the case that MPC focuses more optimizing the future behavior rather than the current conditions. One can also increase N_{c} to prevent this, but this comes at the price of computational complexity and also more fluctuations in the control inputs. Another solution is therefore to add an end-point penalty function to the total objective function (as it is also performed in [34]). The end-point term expresses the time required for vehicles present in the network by the end of the prediction horizon to exit the network. Without this end-point term, it might be the case that the MPC just postpones the formation of queues and if the prediction horizon is not long enough, long queues are still not dissolved by the end of the simulation period (although the demand level may have been lowered). For our case study setup, a prediction horizon between 7 and 9 is enough as by reducing the demands, the queues start to discharge before the end of simulation period (as can be observed from Fig. 10). Moreover, adding the end-point penalty term in this case would have little benefits. This is mainly because the lowest speed on the freeway in the controlled case is around 65 (km/h) on Links 4 and 5 (see Fig. 6(b)) and therefore, the maximum travel time required to exit the freeway for vehicles in Link 4 is less than $(1.56/65 + 5.26/100) \cdot 60 = 4.59$ min (1.56 km is the total length of Links 4 and 5, and 5.26 km is the total distance from the upstream of Link 6 to the downstream end of the freeway, and 100 (km/h) is the average speed on proceeding links, as can be observed from Fig. 6(b)), which is less than $N_{\mathrm{p}} = 7$ min. Hence, the MPC scheme with this prediction horizon is able to regulate the queues with taking into account the constraint on the queue length while significantly improving the travel time on the mainstream road.

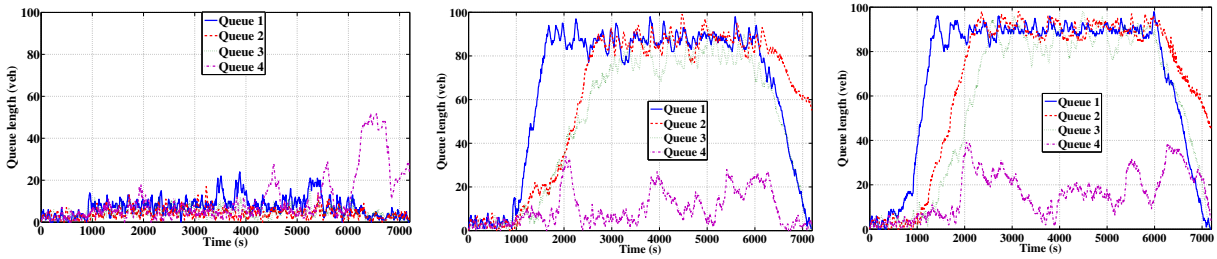
VI. CONCLUSION AND FUTURE RESEARCH

The Link Transmission Model has been extended, reformulated and utilized in a hybrid model predictive control framework. We first modified the model in order to incorporate the effects of traffic control measures, ramp metering and variable speed limits. Next, we established two integrated predictive ramp metering and VSL control schemes, nonlinear MPC based on the extended LTM, and a mixed integer linear programming approach based on the transformed LTM. Finally, the performance of the proposed control schemes was evaluated using micro-simulation for the Leuven Corridor. The obtained results show that the MPC schemes are able to achieve considerable improvement in the total time spent in the network, and moreover, the computation time required for control of such network with several links, on/off-ramps and control inputs is reasonably low (specially for the mixed integer linear programming method).

TABLE III

COMPARISON OF TTS (veh · h) AND CPU TIME (s) FOR TWO APPROACHES, NONLINEAR MPC AND MILP-MPC. THE PERCENTAGES SHOW THE REDUCTION IN THE TTS WITH RESPECT TO THE UNCONTROLLED CASE. THE CPU TIMES FOR N-MPC ARE MULTIPLIED BY 5 TO EXPRESS THAT THE NONLINEAR OPTIMIZATION IS SOLVED 5 TIMES IN EACH CONTROL TIME STEP.

Control parameters	Total time spent (TTS) (veh · h)		Average CPU time (s)		Time spent on mainstream road (veh · h)	
	N-MPC	MILP	N-MPC	MILP	N-MPC	MILP
$N_\mathrm{p} = 3, N_\mathrm{c} = 2$	1036.14 (-7.1%)	1032.19 (-7.5%)	26.88×5	7.32	597.25 (-43.73%)	595.91 (-43.85%)
$N_\mathrm{p} = 5, N_\mathrm{c} = 2$	987.53 (-11.4%)	994.84 (-10.8%)	43.29×5	10.51	590.62 (-44.36%)	592.91 (-44.13%)
$N_\mathrm{p} = 7, N_\mathrm{c} = 3$	972.81 (-12.7%)	963.90 (-13.6%)	124.33×5	21.79	588.99 (-44.51%)	585.12 (-44.87%)
$N_\mathrm{p} = 9, N_\mathrm{c} = 4$	969.94 (-13.0%)	965.38 (-13.4%)	297.87×5	51.39	587.02 (-44.69%)	587.12 (-44.68%)



(a) (b) (c)

Fig. 10. Queues at on-ramps: (a) uncontrolled, (b) controlled by nonlinear MPC, (c) controlled using MILP-MPC.

Extensions to the current work would be, 1) further extending the LTM to include the effects of possible capacity drops at merge nodes, and 2) incorporating robust model predictive techniques in order to better cope with uncertainties in demand profiles and incidents, 3) field implementation of the proposed LTM-based control schemes and 4) extending the LTM-based modeling and control approach for urban and mixed urban/freeway networks.

REFERENCES

[1] A. Alessandri, A. D. Febbraro, A. Ferrara, and E. Punta, "Nonlinear optimization for freeway control using variable-speed signaling," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 6, pp. 2042–2052, 1999.

[2] A. Kostsialos, M. Paageorgiou, M. Mangeas, and H. Haj-Salem, "Coordinated and integrated control of motorway networks via nonlinear optimal control," *Transportation Research Part C*, vol. 10, no. 1, pp. 65–84, 2002.

[3] A. Hegyi, B. De Schutter, and H. Hellendoorn, "Model predictive control for optimal coordination of ramp metering and variable speed limits," *Transportation Research Part C*, vol. 13, no. 3, pp. 185–209, 2005.

[4] I. Papamichail, A. Kotsialos, I. Margonis, and M. Papageorgiou, "Coordinated ramp metering for freeway networks – a model-predictive hierarchical control approach," *Transportation Research Part C*, vol. 18, no. 3, pp. 311–331, 2010.

[5] S. Zegeye, B. De Schutter, J. Hellendoorn, E. Breunesse, and A. Hegyi, "A predictive traffic controller for sustainable mobility using parameterized control policies," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 3, pp. 1420–1429, Sep. 2012.

[6] R. C. Carlson, I. Papamichail, and M. Papageorgiou, "Integrated feedback ramp metering and mainstream traffic flow control on motorways using variable speed limits," *Transportation Research Part C: Emerging Technologies*, vol. 46, pp. 209–221, 2014.

[7] S. Hoogendoorn and P. Bovy, "State-of-the-art of vehicular traffic flow modelling," *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 215, no. 4, pp. 283–303, 2001.

[8] M. Fellendorf and P. Vortisch, "Microscopic traffic flow simulator VISSIM," in *Fundamentals of Traffic Simulation*, ser. International Series in Operations Research and Management Science, J. Barceló, Ed. New York, USA: Springer, 2010, vol. 145, pp. 63–93.

[9] A. May, *Traffic Flow Fundamentals*. Englewood Cliffs, NJ, United States: Prentice Hall, 1990.

[10] A. Messmer and M. Papageorgiou, "Metanet: A macroscopic simulation program for motorway networks," *Traffic Engineering and Control*, vol. 31, no. 9, pp. 466–470, 1990.

[11] C. Daganzo, "The cell transmission model: A simple dynamic representation of freeway traffic," *Transportation Research B*, vol. 28, no. 4, pp. 269–287, 1994.

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

[13] M. Kamal, M. Mukai, J. Murata, and T. Kawabe, "Model predictive control of vehicles on urban roads for improved fuel economy," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 3, pp. 831–841, 2013.

[14] A. Ferrara, S. Sacone, and S. Siri, "Supervisory model predictive control for freeway traffic systems," in *Proceedings of IEEE 52nd Annual Conference on Decision and Control (CDC)*, Florence, Italy, 2013, pp. 905–910.

[15] G. Gomes and R. Horowitz, "Optimal freeway ramp metering using the asymmetric cell transmission model," *Transportation Research Part C: Emerging Technologies*, vol. 14, no. 4, pp. 244–262, 2006.

[16] D. Pisarski and C. Canudas-de-Wit, "Optimal balancing of road traffic density distributions for the Cell Transmission Model," in *Proceedings of IEEE 51st Annual Conference on Decision and Control (CDC)*, Maui, HI, United States, 2012, pp. 6969–6974.

[17] I. Yperman, *The Link Transmission Model for Dynamic Network Loading*. Leuven, Belgium: Ph.D. dissertation, Katholieke Universiteit Leuven, 2007.

[18] F. Viti and C. M. Tampére, "Dynamic traffic assignment: recent advances and new theories towards real time applications and realistic travel behaviour," in *New Developments in Transport Planning*. Edward Elgar Publishing, 2010.

[19] I. Yperman, S. Logghe, C. M. Tampère, and B. Immers, "The multicommodity link transmission model for dynamic network loading," in *Proceedings of the 85th Annual Meeting of the Transportation Research Board*, Washington, DC, United States, 2006.

[20] W. Himpe, R. Corthout, C. M. Tampère, and B. Immers, "An implicit solution scheme for the link transmission model," in *Proceedings of the IEEE Intelligent Transportation Systems Conference*, The Hague, The Netherlands, 2013.

[21] M. Hajiahmadi, R. Corthout, C. Tampère, B. De Schutter, and H. Hellendoorn, "Variable speed limit control based on the extended link transmission model," in *Proceedings of the 92nd Annual Meeting of the Transportation Research Board*, Washington, DC, Jan. 2013.

[22] E. Sontag, "Nonlinear regulations: The piecewise linear approach," *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 346–357, 1981.

[23] J. Lebacque, "The Godunov scheme and what it means for first order traffic flow models," in *Proceedings of the 13th International Symposium*

of Transportation and Traffic Theory (ISTTT), Lyon, France, 1996, pp. 621–644.

[24] W. Jin and H. Zhang, "On the distribution schemes for determining flows through a merge," *Transportation Research B*, vol. 37, no. 6, pp. 521–540, 2003.

[25] C. Daganzo, "The cell transmission model, part II: Network traffic," *Transportation Research B*, vol. 29, no. 1, pp. 79–94, 1995.

[26] D. Ni and J. Leonard, "A simplified kinematic wave model at a merge bottleneck," *Applied Mathematical Modelling*, vol. 29, no. 11, pp. 1045–1072, 2005.

[27] C. Tampère, R. Corthout, D. Cattrysse, and L. H. Immers, "A generic class of first order node models for dynamic macroscopic simulation of traffic flows," *Transportion Research Part B*, vol. 45, no. 1, pp. 289–309, 2011.

[28] G. Newell, "A simplified theory of kinematic waves in highway traffic, part iii: multi-destination flows," *Transportation Research Part B*, vol. 27, no. 4, pp. 281–313, 1993.

[29] T. Bellemans, B. De Schutter, and B. De Moor, "Model predictive control for ramp metering of motorway traffic: A case study," *Control Engineering Practice*, vol. 14, no. 7, pp. 757–767, 2006.

[30] H. Williams, *Model Building in Mathematical Programming*. New York, United States: Wiley, 1993.

[31] Y. Pavlis and W. Recker, "A mathematical logic approach for the transformation of the linear conditional piecewise functions of dispersion-and-store and cell transmission traffic flow models into linear mixed-integer form," *Transportation Science*, vol. 43, no. 1, pp. 98–116, 2009.

[32] A. Atamtürk and M. W. P. Savelsbergh, "Integer-programming software systems," *Annals of Operation Research*, vol. 140, no. 1, pp. 67–124, 2005.

[33] A. Bemporad and M. Morari, "Control of systems integrating logic, dynamics, and constraints," *Automatica*, vol. 35, no. 3, pp. 407–427, 1999.

[34] S. Liu, B. De Schutter, and H. Hellendoorn, "Multi-class traffic flow and emission control for freeway networks," in *Proceedings of the IEEE Intelligent Transportation Systems Conference*, The Hague, Netherlands, 2013, pp. 2223–2228.