

Access provided by:
Universidad Nacional de Colombia
(UNAL)
Sign Out

BrowseMy SettingsGet Help

Advertisement

Back to Results | Next >

Journals & Magazines > IEEE Transactions on Control ... > Volume: 23 Issue: 2

Advertisement

Optimal Hybrid Perimeter and Switching Plans Control for Urban Traffic Networks

<< Results | Next >

4

Mohammad Hajiahmadi ; Jack Haddad ; Bart De Schutter ; Nikolas Geroliminis

Author(s)

View All Authors

34

Paper Citations

815

Full Text Views

Export to
Collabratec

Alerts

Manage
Content Alerts

Add to Citation
Alerts

Related Articles

Solving optimal control problems with state constraints using nonlinear programming and simulation tools
IEEE Transactions on Education
Published: 2004

A Comparison of the Embedding Method With Multiparametric Programming, Mixed-Integer Programming, Gradient-Descent, and Hybrid Minimum Principle-Based Methods
IEEE Transactions on Control Systems Technology
Published: 2014

View All View More

Related Articles

Solving optimal control problems with state constraints using nonlinear programming and simulation tools
IEEE Transactions on Education
Published: 2004

A Comparison of the Embedding Method With Multiparametric Programming, Mixed-Integer Programming, Gradient-Descent, and Hybrid Minimum Principle-Based Methods
IEEE Transactions on Control Systems Technology
Published: 2014

View All View More

See the top organizations patenting in technologies mentioned in this article



Click to Expand

Provided by: InnovationQ PLUS
POWERED BY IEEE AND IP.COM
A PATENT SEARCH AND ANALYTICS TOOL

See the top organizations patenting in technologies mentioned in this article



Click to Expand

Provided by: InnovationQ PLUS
POWERED BY IEEE AND IP.COM
A PATENT SEARCH AND ANALYTICS TOOL

Advertisement

Abstract

Document Sections

Downl
PDF

I. Introduction

II. MFD-Based Modeling of Urban Regions

III. Optimal Hybrid Control for an R -Region MFDs System

IV. Approximation of the R -Region MFDs System

V. Reformulation as MILP

Show Full Outline ▾

Authors

Figures

References

Citations

Keywords

Metrics

Related Articles

Abstract: Since centralized control of urban networks with detailed modeling approaches is computationally complex, developing efficient hierarchical control strategies based on ag... [View more](#)

Metadata

Abstract: Since centralized control of urban networks with detailed modeling approaches is computationally complex, developing efficient hierarchical control strategies based on aggregate modeling is of great importance. The dynamics of a heterogeneous large-scale urban network is modeled as R homogeneous regions with the macroscopic fundamental diagrams (MFDs) representation. The MFD provides for homogeneous network regions a unimodal, low-scatter relationship between network vehicle density and network space-mean flow. In this paper, the optimal hybrid control problem for an R -region MFD network is formulated as a mixed-integer nonlinear optimization problem, where two types of controllers are introduced: 1) perimeter controllers and 2) switching signal timing plans controllers. The perimeter controllers are located on the border between the regions, as they manipulate the transfer flows between them, while the switching controllers influence the dynamics of the urban regions, as they define the shape of the MFDs and as a result affect the internal flows within each region. Moreover, to decrease the computational complexity due to the nonlinear and nonconvex nature of the optimization problem, we reformulate the problem as a mixed-integer linear programming (MILP) problem utilizing piecewise affine approximation techniques. Two different approaches for transformation of the original model and building up MILP problems are presented, and the performances of the approximated methods along with the original problem formulation are evaluated and compared for different traffic scenarios of a two-region urban case study.

Published in: IEEE Transactions on Control Systems Technology (Volume: 23 , Issue: 2 , March 2015)

Page(s): 464 - 478

INSPEC Accession Number: 14930714

Date of Publication: 02 July 2014

DOI: 10.1109/TCST.2014.2330997

ISSN Information:

Publisher: IEEE

Sponsored by: IEEE Control Systems Society

Funding Agency:

Advertisement

 Contents

SECTION I.

Introduction

Large-scale urban networks need efficient traffic management and control schemes. Modeling a large urban network would be a complex task if one wants to study and model the traffic dynamics of each element (i.e., each link and each intersection, including route choice of travelers). On the other hand, centralized control of an urban network with such detailed modeling approach would be computationally complex and makes its implementation in real-time infeasible. Hence, instead of adopting a detailed modeling approach, researchers are investigating alternative possibilities of deriving an aggregate model for the whole traffic network.

The idea of macroscopic fundamental diagrams (MFDs) with optimum accumulations was first proposed in [1] and similar approaches were introduced

later in [2]–[4]. Existence of the MFDs with dynamic features was recently

revealed in [5]. The MFD captures macroscopically (at a network level) the traffic flow characteristics and dynamics of an urban region. It relates the number of vehicles (accumulation) in the region and its production, defined as the trip completion flow of vehicles reaching their destination. The underlying assumption in these previous works has been that the network is homogeneously congested, which is not always the case. Homogeneous networks with a small variance on link densities have a well-defined MFD, i.e., there is a low scatter of flows for the same densities (or accumulations) [6], [7]. A well-defined MFD is schematically shown in Fig. 1. The shape of the MFD can be approximated by a nonsymmetric unimodal curve skewed to the right, i.e., the critical accumulation, n_{cr} (veh), that maximizes network flow is smaller than half the jammed accumulation n_{max} . Note that the network topology, the signal timing plans of the signalized intersections, and the infrastructure characteristics affect the shape of the MFD [8]–[9][10]. Other investigations of the MFD using empirical or simulated data can be found in [11]–[12][13], while routing strategies based on the MFD can be found in [14] and [15].

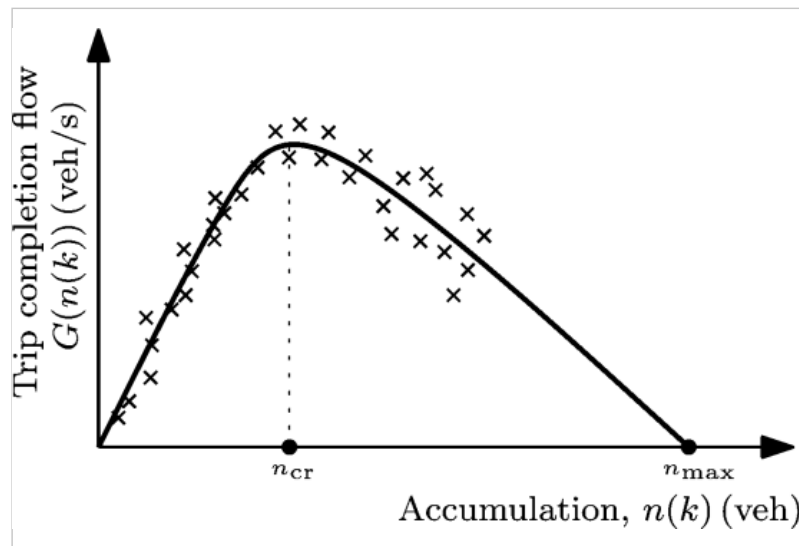


Fig. 1.
Well-defined MFD.

Heterogeneous networks might not have a well-defined MFD, mainly in the congestion (decreasing) part of the MFD, and the scatter becomes higher as accumulation increases [6], [7], [12]. A possible solution is to partition heterogeneous networks (in a static or dynamic way) into more homogeneous regions with small variances of link densities such that each region has a well-defined MFD [16].

The MFD can be utilized to establish efficient and elegant strategies to control network flows. While most of the existing control strategies are locally oriented or distributed at only a small scale, coordinated strategies can decrease delays and increase mobility in large urban networks. Meanwhile, the idea of perimeter control on the borders of urban regions has attracted many researchers. Recently in [17], optimal perimeter control for a two-region urban city is formulated by exploiting the notion of MFD. For stability analysis of perimeter control, the reader can refer to [18], while optimal control for mixed urban freeway networks utilizing MFDs is found in [15]. Perimeter control for single- or multiple-region homogeneous networks has been analyzed with linear multivariable feedback regulators in [19] and [20].

In this paper, we introduce an extra level of control that can manipulate the flow dynamics of each urban region by switching between signal timing plans. Changing timing plans for signalized intersections within regions might alter the shape of the MFD, which will affect the network flow dynamics. Therefore, instead of assuming

one MFD for each region, we introduce a set of MFDs, where each MFD

one MFD for each region, the model can be extended to handle each MFD corresponds to a certain timing plan for intersections inside the region.

Combining switching timing plans and perimeter controllers might significantly increase the network performance, as it gives the ability to control inside and on the border of urban regions, and to adjust to a vast variety of demands and traffic conditions. However, combining these two control inputs is not straightforward, as a mixture of discrete and continuous control inputs is introduced that might have different effects on the flow dynamics. The model of an urban region will be a nonlinear state space model based on the MFD and it has both continuous perimeter control inputs and binary variables for switching the timing plans. Moreover, model predictive control (MPC) [21] is used to solve the optimal control problem. Since we deal with a hybrid system, the resulting open-loop optimization problem is a mixed-integer nonlinear problem. Solving nonlinear and nonconvex optimization problems can be time consuming and showing that a global solution is found is not guaranteed. If the problem is solved multiple times for different initial points, chances are high that a reasonably optimal solution is found. While multistart optimization algorithms or global optimization techniques can be used to overcome this problem, one can try to approximate and transform the model into a mixed linear form and formulate the optimization problem as a mixed-integer linear programming (MILP) problem. The computation time will decrease significantly and one global optimum solution for the MILP problem will be obtained.

To summarize, the paper contributes in three ways. First, a novel hybrid MFD-based model is proposed that is capable of modeling the effect of switching between timing plans on the MFD of an urban network. Second, a MPC scheme is constructed based on the proposed hybrid model and further simplifying mathematical techniques are presented to decrease the computational complexity of the associated optimization problem. Among the techniques are avoiding 2-D piecewise affine (PWA) approximation and using two simpler approaches instead, and quantization of the perimeter control input to solve the problem with input/states multiplications. Finally, we consider practical issues with measuring the traffic variables, trip demands, and also the scattered MFDs observed in real networks; and therefore, we add three types of uncertainties into our hybrid model to make simulation and control of a multiregion urban network more realistic.

The rest of this paper is organized as follows. In Section II, a hybrid MFD-based model of an R -region urban network is presented, while in Section III, the optimal hybrid control problem is formulated. Two mixed linear models based on the PWA approximation of the original model are proposed in Section IV, and the corresponding mixed-integer linear optimization problem is formulated in Section V. The performance of the predictive hybrid controllers (linear and nonlinear) is tested for several case study examples with different scenarios in Section VI. This paper concludes with a discussion about the results and ideas for further research.

SECTION II.

MFD-Based Modeling of Urban Regions

Let us assume that a heterogeneous urban traffic network can be partitioned into R homogeneous regions, each having a well-defined MFD (later we will assume that each homogeneous region can have a set of different MFDs corresponding to the activated signal timing plans) (Fig. 2). In this paper, the model time step counter and the sampling period are denoted by k (—) and T (s), respectively, where $t = k \cdot T$ and $k = 0, 1, 2, \dots, K - 1$. Let $q_{ij}(k)$ (veh/s) be the traffic flow demand generated in region i at time step k with final destination in region j , $i = 1, 2, \dots, R$, and $j \in \mathcal{N}_i$, where $\mathcal{N}_i$ is the set of regions that are directly reachable from region i . Corresponding to the traffic demands, accumulation states are defined to model the dynamic equations: $n_{i,j}(k)$ (veh) denotes the total number

we defined to model the dynamic equation: $n_{ij}(k)$ (veh) denotes the total number of vehicles in region i with direct destination to region j at time step k . Let us denote $n_i(k)$ (veh) as the accumulation or the total number of vehicles in region i at time step k , i.e., $n_i(k) = n_{ii}(k) + \sum_{j \in \mathcal{N}_i} n_{ij}(k)$. The MFD is defined by $G_i(\cdot)$ (veh/s) that is the trip completion flow for region i at $n_i(k)$. The trip completion flow for region i is the sum of transfer flows, i.e., trips from i with destination j , $j \in \mathcal{N}_i$, plus the internal flow, i.e., trips from i with destination i . The transfer flow from i with destination to j , denoted by $M_{ij}(k)$ (veh/s), is calculated corresponding to the ratio between accumulations, i.e., $M_{ij}(k) = n_{ij}(k)/n_i(k) \cdot G_i(n_i(k))$, $j \in \mathcal{N}_i$, while $M_{ii}(k)$ is the internal flow from i with destination to i and calculated by $M_{ii}(k) = n_{ii}(k)/n_i(k) \cdot G_i(n_i(k))$. These relationships assume that the trip lengths for all trips within a region (internal or external) are similar, i.e., the distance traveled per vehicle inside a region is independent of the origin and destination of the trip. For further description, the interested reader is referred to [22], which will not alter the methodology. We utilize a third-order function of $n_i(k)$ to describe the MFD, e.g., $G_i(n_i(k)) = a_i \cdot n_i^3(k) + b_i \cdot n_i^2(k) + c_i \cdot n_i(k)$, where a_i , b_i , and c_i are estimated parameters.

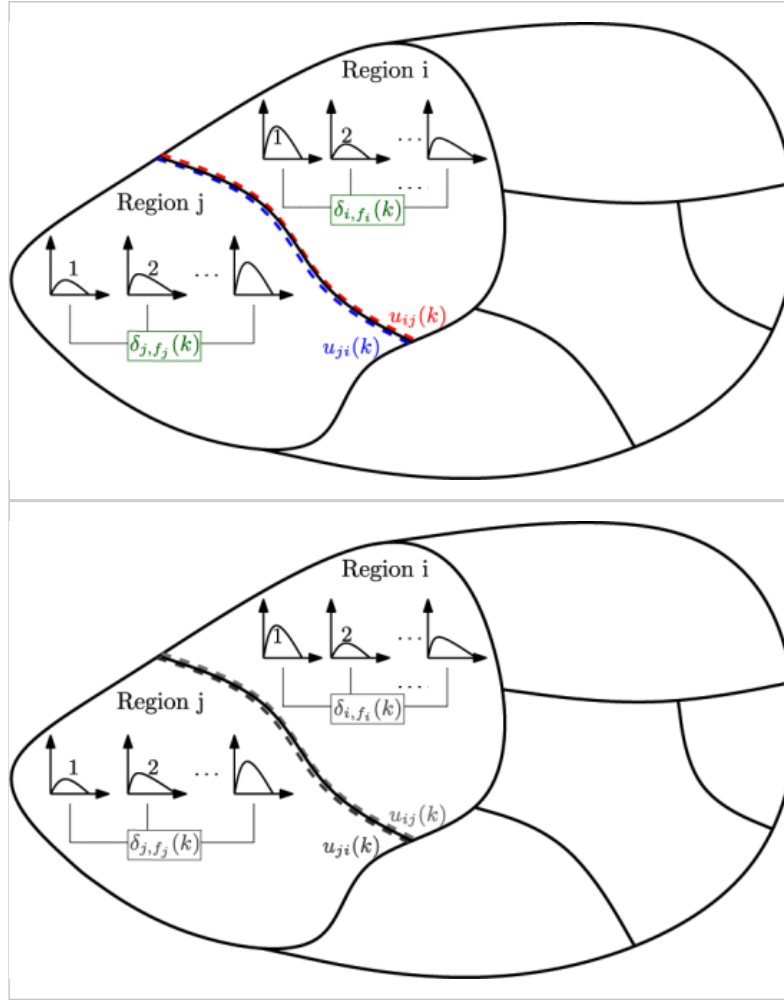


Fig. 2.

Hybrid R -region system with perimeter and switching timing plans control inputs $u_{ij}(k)$ and $\delta_{i,f_i}(k)$ for region i , and $u_{ji}(k)$ and $\delta_{j,f_j}(k)$ for region j .

The vehicle conservation equations (without integrating control measures) of the R -region MFDs system are

$$n_{ii}(k+1) = n_{ii}(k) + T \cdot \left(q_{ii}(k) + \sum_{j \in \mathcal{N}_i} M_{ji}(k) - M_{ii}(k) \right) \quad (1)$$

$$n_{ij}(k+1) = n_{ij}(k) + T \cdot (q_{ij}(k) - M_{ij}(k)) \quad (2)$$

[View Source](#) 

for $i = 1, 2, \dots, R$ and $\forall j \in \mathcal{N}_i$. These equations are a generalized (R regions instead of two) and discretized form of the equations presented in [17]. Note that route choice modeling is not integrated in the dynamic equations.

SECTION III.

Optimal Hybrid Control for an R - Region MFDs System

In the previous section, the MFD-based model (1)–(2) was introduced without any control measure. In the following, two types of controllers are introduced in Section III-A and integrated into the dynamic equations (1) and (2) in Section III-B, while in Section III-C, the optimal hybrid control problem for the R -region MFDs system is formulated.

A. Hybrid Control: Perimeter and Switching Controllers

Two types of controllers are introduced in the hybrid control problem: 1) perimeter controllers, and 2) switching signal timing plans controllers. The perimeter controllers are located on the border between regions, as they manipulate the *transfer flows* between them, while the switching controllers influence the dynamics of the urban regions, as they define the shape of the MFDs. Note that the switching controllers and the perimeter controllers might affect each other but we assume that these effects are negligible.

The signal timing plans alter the shape of the MFD [8]. In this paper, it is assumed that each urban region has a predefined library of signal fixed-timing plans for the signalized intersections inside the region, e.g., fixed-timing plans for the morning and evening peak hours and a typical uncongested hour, where each plan in the library has different green, red, cycle, and offset settings for the intersections. It is also assumed that for each activated signal plan, the region will have a different MFD, i.e., a nonsymmetric unimodal curve skewed to the right, but with different values of the maximum output, and critical accumulations, e.g., three different MFDs for regions i and j in Fig. 2. Therefore, the timing plan library employs a library of MFDs for each region. The switching controller of the region activates one MFD from the library by switching from one signal plan to another.

The optimal perimeter and switching plans decisions are obtained by minimizing the total time spent in the R urban regions. The total time spent (veh · s) is defined as follows:

$$J = T \cdot \sum_{k=1}^{K-1} \sum_{i=1}^R n_i(k). \quad (3)$$

[View Source](#) 

B. Hybrid R -Region MFDs System

Let us denote the perimeter control inputs by $u_{ij}(k)$ ($-$), $i = 1, 2, \dots, R$, $j \in \mathcal{N}_i$, and the switching timing plans control inputs by $\delta_{i,f_i}(k) \in \{0, 1\}$, where $f_i \in \mathcal{F}_i$ and $\mathcal{F}_i$ is the set of MFDs in the library for region i . The control inputs $u_{ij}(k)$, $\delta_{i,f_i}(k)$, and $u_{ji}(k)$, $\delta_{j,f_j}(k)$ are associated with regions i and j , respectively.

The perimeter control inputs $u_{:i}(k)$ and $u_{:j}(k)$ are introduced on the border

the perimeter controller inputs $u_{ij}(k)$ and $u_{ji}(k)$ are introduced on the border between the regions i and j , as shown in Fig. 2, where the purpose is to control the transfer flows between the two regions. The transfer flow $M_{ij}(k)$, $i = 1, 2, \dots, R$, $j \in \mathcal{N}_i$, is controlled such that only a fraction of the flow actually transfers from region i to region j , i.e., $u_{ij}(k) \cdot M_{ij}(k)$, where $0 \leq u_{ij}(k) \leq 1$. Hence, the MFD-based model (1) and (2) is altered by replacing $M_{ij}(k)$ and $M_{ji}(k)$ by $u_{ij}(k) \cdot M_{ij}(k)$ and $u_{ji}(k) \cdot M_{ji}(k)$, respectively. It is also assumed that these controllers will not change the shape of the MFDs.

Since the perimeter controllers exist only on the border between the regions, the internal flows cannot be controlled or restricted. However, the internal flows are determined by the MFDs of the regions. The switching controllers can manipulate indirectly the internal flows by switching the MFDs (or more precisely by switching the signal timing plans of the signalized intersections). Recall that the vehicle conservation equations (1) and (2) assume that each region has only one MFD. Let us now assume that each region i has a predefined MFD library (or set of MFDs denoted by $\mathcal{F}_i$) that corresponds to a signal timing plans library for the signalized intersections. The switching control signal $\delta_{i,f_i}(k)$ activates the f_i th MFD in the set $\mathcal{F}_i$, i.e., $G_{i,f_i}(\cdot)$, if $\delta_{i,f_i}(k) = 1$ and $\delta_{i,r_i}(k) = 0$, $\forall r_i \in \mathcal{F}_i \setminus \{f_i\}$ (so only one $\delta_{i,f_i}(k) = 1$ at any time step, i.e., $\sum_{f_i \in \mathcal{F}_i} \delta_{i,f_i}(k) = 1$). Therefore, the R -region MFDs system (1) and (2) is modified to integrate the switching controllers, as the term $G_i(n_i(k))$ is changed to $\sum_{f_i \in \mathcal{F}_i} \delta_{i,f_i}(k) \cdot G_{i,f_i}(n_i(k))$. The novel hybrid R -region MFDs system is formulated as

$$n_{ii}(k+1) = n_{ii}(k) + T \cdot (q_{ii}(k) + \sum_{j \in \mathcal{N}_i} u_{ji}(k) \cdot M_{ji}(k) - M_{ii}(k)) \quad (4)$$

$$n_{ij}(k+1) = n_{ij}(k) + T \cdot (q_{ij}(k) - u_{ij}(k) \cdot M_{ij}(k)) \quad (5)$$

$$M_{ii}(k) = \frac{n_{ii}(k)}{n_i(k)} \cdot \left[\sum_{f_i \in \mathcal{F}_i} \delta_{i,f_i}(k) \cdot G_{i,f_i}(n_i(k)) \right] \quad (6)$$

$$M_{ij}(k) = \frac{n_{ij}(k)}{n_i(k)} \cdot \left[\sum_{f_i \in \mathcal{F}_i} \delta_{i,f_i}(k) \cdot G_{i,f_i}(n_i(k)) \right] \quad (7)$$

$$n_i(k) = n_{ii}(k) + \sum_{j \in \mathcal{N}_i} n_{ij}(k). \quad (8)$$

[View Source](#) 

C. Optimal Control Problem Formulation

After introducing and integrating the controllers into the hybrid R -region MFDs system, we formulate the optimal hybrid control problem. The scheme of the optimal control problem is shown in Fig. 3. The aim is to minimize the total time spent (3) by manipulating the perimeter controller and by switching between the timing plans of the libraries.

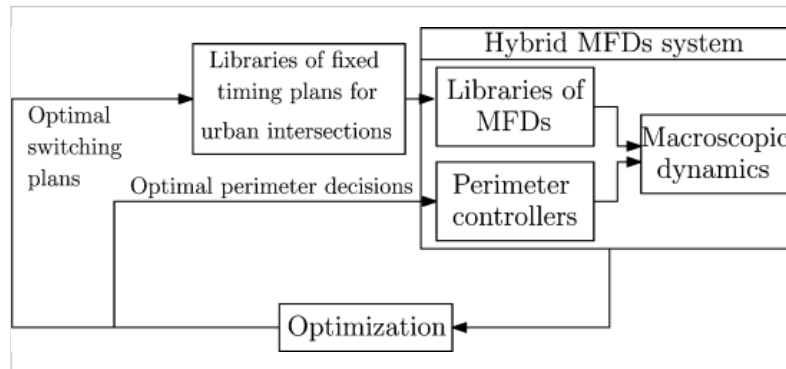


Fig. 3. Optimal hybrid perimeter and switching plans control scheme.

in reality, homogeneous regions are rare and some states, particularly in the congested regime, as shown schematically in Fig. 1. Therefore, errors are expected between the hybrid R -region MFD model (assuming well-defined MFDs) and the real network. Therefore, a closed-loop optimal control scheme is needed to consider the errors between the plant and the model, and also the disturbances, e.g., variations in the expected demands that might affect the system (the differences between the model and the plant will be discussed in details later). Among these schemes is the MPC framework, which has been widely used for different traffic control purposes [23]–[24][25][26][27][28]. The MPC controller determines the optimal control inputs in a receding horizon manner, meaning that at each time step, an objective function is optimized over a prediction horizon of N_p steps and a sequence of optimal control inputs are derived. Then, the first sample of the control inputs is applied to the system and the procedure is repeated with a shifted horizon.

We directly formulate the problem in the MPC framework. Let k_c (–) and T_c (s) be the control time step and the control sample time, respectively. It is assumed that the controller time step length is an integer multiple of the model time step length, i.e., $T_c = M \cdot T$. Then, the overall optimization problem is formulated as follows:

$$\min_{\tilde{u}_{ij}(k_c), \tilde{\delta}_{i,f_i}(k_c), \tilde{n}_{ii}(k_c), \tilde{n}_{ij}(k_c)} T \cdot \sum_{k=M \cdot k_c}^{M \cdot (k_c + N_p) - 1} \sum_{i=1}^R n_i(k) \quad (9)$$

[View Source](#) ?

subject to

$$\text{Model equations (4)–(8)} \quad (10)$$

$$0 \leq n_i(k) \leq n_{i,\text{jam}} \quad (11)$$

$$u_{ij,\min} \leq u_{ij}(k) \leq u_{ij,\max} \quad (12)$$

$$u_{ij}(k) = u_{ij}^c(k_c) \text{ if } k \in \{M \cdot k_c, \dots, M \cdot (k_c + 1) - 1\} \quad (13)$$

$$\delta_{i,f_i}(k) = \delta_{i,f_i}^c(k_c) \text{ if } k \in \{M \cdot k_c, \dots, M \cdot (k_c + 1) - 1\} \quad (14)$$

$$\delta_{i,f_i}(k) \in \{0, 1\}, \forall f_i \in \mathcal{F}_i \quad (15)$$

[View Source](#) ?

for $i = 1, 2, \dots, R$ and $\forall j \in \mathcal{N}_i$, where $n_{i,\text{jam}}$ (veh) is the jam accumulation for region i , and $u_{ij,\min}$ and $u_{ij,\max}$ (–) are, respectively, the lower and upper bounds for the perimeter control signals for regions i and j . The optimization variables defined over the prediction horizon N_p are

$$\tilde{n}_{ij}(k_c) = [n_{ij}(M \cdot k_c), \dots, n_{ij}(M \cdot (k_c + N_p) - 1)]^T,$$

$$\tilde{n}_{ii}(k_c) = [n_{ii}(M \cdot k_c), \dots, n_{ii}(M \cdot (k_c + N_p) - 1)]^T,$$

$$\tilde{u}_{ij}(k_c) = [u_{ij}^c(k_c), \dots, u_{ij}^c(k_c + N_p - 1)]^T \text{ and}$$

$$\tilde{\delta}_{i,f_i}(k_c) = [\delta_{i,f_i}^c(k_c), \dots, \delta_{i,f_i}^c(k_c + N_p - 1)]^T, \text{ where } u_{ij}^c(k_c + l) \text{ and } \delta_{i,f_i}^c(k_c + l)$$

for $l = 0, \dots, N_p - 1$ are the perimeter and switching control inputs at every control time step k_c , respectively. The current model equations do not directly consider downstream restrictions, e.g., the boundary capacity. One more term can be added, the boundary capacity, which is a function of the receiving region accumulation and restricts the transfer flow if the receiving region is highly congested. This constraint is ignored during the optimization process. The physical reasoning behind this assumption is that: 1) the boundary capacity decreases for accumulations that are much larger than the critical accumulation [29] and 2) the control inputs will not allow the system to get close to gridlock.

The problem (9)–(15) is a mixed-integer nonlinear programming (MINLP) optimization problem, and it can be solved using mixed-integer nonlinear optimization algorithms [30]. However, since here we deal with both real and binary decision variables, and also since the model equations have nonlinear terms, the optimization problem could have multiple (local) optimal points. Moreover, as it will be demonstrated in Section VI, the optimization algorithm takes considerable time. This is mainly because the MINLP algorithm is executed for several random initial points, to find the lowest possible value of the objective function. Thus, in the next section we simplify and reformulate the problem to eventually establish a

nonconvex, its simplifying and reformulating the problem to eventually solution a mixed-integer *linear* optimization problem.

SECTION IV.

Approximation of the R -Region MFDs System

Solving the nonlinear and nonconvex (the nonconvexity is because of having a hybrid nonlinear model with a mixture of continuous perimeter control inputs and binary decision variables to switch between MFDs) optimization problem (9)–(15) can be time-consuming and not tractable for real-time implementation. In the following two sections, we will recast the problem into a mixed-integer *linear* optimization problem. The nonlinear model in the MPC framework (9)–(15) is replaced by an approximated model following PWA approximation techniques and some mathematical simplifications. The idea of PWA approximation of MFDs was presented in a hierarchical control framework for intelligent vehicle highway systems in [31].

Basically, the nonlinearity in the dynamic equations is present in: 1) the internal and transfer trip completion flows, see $M_{ii}(k)$ in (6) and $M_{ij}(k)$ in (7), respectively, and 2) the product between the perimeter controllers and the transfer trip completion flows [(4), (5)]. In the following, we address these nonlinearities and obtain two different approximated models. The first model is less computationally complex but less accurate than the second one. In the case study section, a performance evaluation of the two methods along with the original nonlinear approach will be presented.

A. First Approach (PWA Approximation Along With Forward Simulation)

The multiplication of $n_{ii}(k)$ [or $n_{ij}(k)$] with the other variables in the square brackets in (6) [or (7)] results in multiple products of real variables. In principle, each product needs to be approximated by a PWA function [32], [33]. A function $f : \Omega \rightarrow \mathbb{R}^m$ is PWA if there exists a polyhedral partition $\{\Omega_i\}_{i \in \mathcal{I}}$ ($\cup_{i \in \mathcal{I}} \Omega_i = \Omega$, $\Omega_i \cap \Omega_j = \emptyset$, $\forall i \neq j$) of $\Omega \subseteq \mathbb{R}^n$ such that f is affine on each polyhedron Ω_i . One can approximate a nonlinear function by PWA functions with arbitrary accuracy and by considering a sufficiently large number of regions. However, for our particular case (bivariate function of accumulations), the PWA approximation is a tedious task as more parameters have to be introduced [34]. In other words, we have to deal with a 2-D PWA approximation [34], [35] and to get enough accuracy in the modeling, the resulting PWA function would need a large number of affine pieces. This may add more complexity to the associated optimization problem. Therefore, as a main consideration in the PWA approximation, the number of affine functions should be kept small while providing a close match to the original nonlinear function.

Hence, to simplify the approximation, we estimate the variables $n_{ii}(k)$ and $n_{ij}(k)$ in the transfer flows by forward simulation as follows: we first simulate the R -region MFDs system according to the model presented in (4) and (5) over a prediction horizon with control inputs and initial accumulations obtained from the previous time step, and subsequently the variables $n_{ii}(k)$ and $n_{ij}(k)$ in $M_{ii}(k)$ and $M_{ij}(k)$ are replaced with the values obtained from the simulation. Hence, we no longer deal with multiplication of variables but only with multiplication with time-varying but known parameters. Nevertheless, this creates errors in the optimization algorithm and might affect the overall performance.

1) PWA Approximation of the Trip Completion Flows:

The nonlinearity in the internal trip completion flows $M_{ii}(k)$ is approximated as follows (a similar procedure is applied to the transfer flows $M_{ij}(k)$). Substituting

the third-order function $G_{i,i}(n_{ii}(k)) = a_{i,i} \cdot n_{ii}^3(k) + b_{i,i} \cdot n_{ii}^2(k) + c_{i,i} \cdot n_{ii}(k)$

into (6), one can rewrite the internal flows $M_{ii}(k)$ for $i = 1, 2, \dots, R$ as follows:

$$M_{ii}(k) = n_{ii}(k) \cdot \left[\sum_{f_i \in \mathcal{F}_i} \delta_{i,f_i}(k) \cdot (a_{i,f_i} \cdot n_i^2(k) + b_{i,f_i} \cdot n_i(k) + c_{i,f_i}) \right]. \quad (16)$$

[View Source](#) 

The function $P_{i,f_i}(n_i(k)) = a_{i,f_i} \cdot n_i^2(k) + b_{i,f_i} \cdot n_i(k) + c_{i,f_i}$ [inside the parentheses in (16)] defined on the interval $[n_{i,\min}, n_{i,\max}]$ can be approximated by a continuous PWA function $\hat{P}_{i,f_i}(n_i(k))$ with three intervals as follows:

$$\hat{P}_{i,f_i}(n_i(k)) = \begin{cases} \gamma_{i,f_i} + \frac{n_i(k) - n_{i,\min}}{\alpha_{i,f_i} - n_{i,\min}} \cdot (\xi_{i,f_i} - \gamma_{i,f_i}) & \text{for } n_{i,\min} \leq n_i(k) < \alpha_{i,f_i} \\ \xi_{i,f_i} + \frac{n_i(k) - \alpha_{i,f_i}}{\beta_{i,f_i} - \alpha_{i,f_i}} \cdot (\epsilon_{i,f_i} - \xi_{i,f_i}) & \text{for } \alpha_{i,f_i} \leq n_i(k) < \beta_{i,f_i} \\ \epsilon_{i,f_i} + \frac{n_i(k) - \beta_{i,f_i}}{n_{i,\max} - \beta_{i,f_i}} \cdot (\zeta_{i,f_i} - \epsilon_{i,f_i}) & \text{for } \beta_{i,f_i} \leq n_i(k) < n_{i,\max} \end{cases} \quad (17)$$

[View Source](#) 

where the set of parameters $\theta_{i,f_i} = \{\gamma_{i,f_i}, \alpha_{i,f_i}, \beta_{i,f_i}, \xi_{i,f_i}, \epsilon_{i,f_i}, \zeta_{i,f_i}\}$ can be estimated by solving the following nonlinear least-squares optimization problem:

$$\min_{\theta_{i,f_i}} \int_{n_{i,\min}}^{n_{i,\max}} (P_{i,f_i}(n_i(k)) - \hat{P}_{i,f_i}(n_i(k)))^2 dn_i. \quad (18)$$

[View Source](#) 

This optimization problem can be solved by multistart nonlinear optimization algorithms [36].

2) Approximation of the Product Between the Perimeter Controllers and the Transfer Flows:

The transfer flows are multiplied with the perimeter controller inputs in (4) and (5). These products cannot be replaced with values obtained from simulation as the optimal perimeter inputs should be determined from the optimization algorithm. As discussed before, the perimeter control inputs determine the percentage of flows that are allowed to transfer between regions, and thus they take values in the interval $[0, 1]$. Considering the practical case in which the perimeter control is going to be implemented, the perimeter control is realized by changing the signal settings of intersections. Consequently, the perimeter signals take values from a finite set in the interval $[0, 1]$. This means that we can make the control inputs $u_{ij}(k)$ quantized as follows [37]:

$$u_{ij}(k) = u_{ij,0} \cdot \left(\sum_{l=0}^r 2^l \cdot \omega_{ij,l}(k) \right) \quad (19)$$

[View Source](#) 

where $u_{ij,0}$ are *a priori* given constants and $\omega_{ij,l}(k) \in \{0, 1\}$ are the optimization variables. The set of possible input values is then finite and its cardinality is 2^{r+1} , while the difference between two consecutive values is determined by $2^r u_{ij,0}$. Having a sum of weighted binary variables for each perimeter control input, the problem with multiplication of control inputs with transfer flow functions will be

simplified since multiplication with binary variables can be easily handled with the

complexity, since multiplication with binary variables can be easily handled with the techniques presented in Section V.

Remark 1:

Another way to tackle the problem with the multiplication of perimeter control input, and the transfer flow function is to introduce a new variable $\tilde{M}_{ji}$ and rewrite (4) as

$$n_{ii}(k+1) = n_{ii}(k) + T \cdot \left(q_{ii}(k) + \sum_{j \in S_i} \tilde{M}_{ji}(k) - M_{ii}(k) \right) \quad (20)$$

[View Source](#) ?

with an additional constraint

$$0 \leq \tilde{M}_{ji}(k) \leq M_{ji}(k). \quad (21)$$

[View Source](#) ?

However, if the control input u_{ji} explicitly appears in the objective function (e.g., in the penalty term), this method cannot be applied.

B. Second Approach (Recasting 2-D PWA Approximation)

In the previous section, one way to tackle the problem with multiplication of real variables was presented: using forward simulation. This method can deliver satisfactory results for some cases. However, to achieve more accuracy, one can iterate on forward simulation and optimization inside each MPC control step that would introduce additional computation time. We mention that instead of using forward simulation to estimate the multiplication terms $[n_{ii}(k)$ and $n_{ij}(k)$ with the square brackets in (6) and (7)], one can directly approximate the bilinear functions following 2-D PWA approximation methods in the literature, e.g., the one in [34]. However, there are methods to reduce this tedious 2-D PWA approximation with a 1-D problem. In the sequel, we treat the nonlinear terms in the model using these reducing methods proposed in [37] and [38]. We give detailed descriptions only for nonlinear terms in $M_{ii}(k)$ (16), but a similar explanation holds for $M_{ij}(k)$. According to (16), we have to deal with two nonlinear terms: $n_{ii}(k) \cdot n_i(k)$ and $n_{ii}(k) \cdot n_i^2(k)$ (note that the two nonlinear terms in $M_{ij}(k)$ would be $n_{ij}(k) \cdot n_i(k)$ and $n_{ij}(k) \cdot n_i^2(k)$).

1) PWA Approximation of $n_{ii}(k) \cdot n_i(k)$:

The term $n_{ii}(k) \cdot n_i(k)$ can be rewritten as [38]

$$n_{ii}(k) \cdot n_i(k) = \frac{1}{4} \left[(n_i(k) + n_{ii}(k))^2 - (n_i(k) - n_{ii}(k))^2 \right]. \quad (22)$$

[View Source](#) ?

Defining two new auxiliary variables

$$y_{1,i}(k) = n_i(k) + n_{ii}(k), \quad y_{2,i}(k) = n_i(k) - n_{ii}(k) \quad (23)$$

[View Source](#) ?

one gets

$$n_{ii}(k) \cdot n_i(k) = \frac{1}{4} (y_{1,i}(k)^2 - y_{2,i}(k)^2).$$

[View Source](#) ?

Instead of performing a 2-D PWA approximation, we now have to deal with the PWA approximation of two separated single-variable functions $y_{1,i}(k)^2$ and $y_{2,i}(k)^2$. The function $f(y_i) = y_i^2$ can be approximated by PWA functions using a nonlinear least-squares optimization formulation as in (17). However, the domain of the

functions should be defined properly and according to the domains of the original variables. For instance, the domain of $f(y_{1,i}) = y_{1,i}^2$ is $[y_{1,i,\min}, y_{1,i,\max}]$ with $y_{1,i,\min} = \min\{n_i + n_{ii} | n_{i,\min} \leq n_i \leq n_{i,\max}, n_{ii,\min} \leq n_{ii} \leq n_{ii,\max}\}$ and $y_{1,i,\max} = \max\{n_i + n_{ii} | n_{i,\min} \leq n_i \leq n_{i,\max}, n_{ii,\min} \leq n_{ii} \leq n_{ii,\max}\}$.

2) PWA Approximation of $n_{ii}(k) \cdot n_i^2(k)$:

We follow the same procedure as above. Defining two variables $y_{3,i}(k)$ and $y_{4,i}(k)$, $n_{ii}(k) \cdot n_i^2(k)$ can be rewritten as $\frac{1}{4}(y_{3,i}(k)^2 - y_{4,i}(k)^2)$ with

$$y_{3,i}(k) = n_i^2(k) + n_{ii}(k), \quad y_{4,i}(k) = n_i^2(k) - n_{ii}(k). \quad (24)$$

[View Source](#) 

However, there is still a nonlinear term in $y_{3,i}(k)$ and $y_{4,i}(k)$. The simple solution for that is to approximate the term $n_i^2(k)$ with a set of affine functions determined from an identification procedure like in (17) with an appropriate domain for $n_i(k)$ and next replace $n_i^2(k)$ with its PWA approximation in $y_{3,i}(k)$ and $y_{4,i}(k)$.

3) Multiplication With Control Inputs $u_{ij}(k)$:

As discussed before, the transfer flows are multiplied by the perimeter control inputs. One can use the same procedure explained in Section IV-B1 for approximation of the multiplication of $u_{ij}(k)$ with $n_{ij}(k) \cdot n_i^2(k)$ and $n_{ij}(k) \cdot n_i(k)$. However, this would introduce more variables and make the model more complicated for optimization use. In this case, we assume that the perimeter control inputs are quantized and defined as in (19). Hence, instead of having multiplication of real variables, we deal with multiplication of binary decision variables and PWA approximated transfer flow functions. In the next section, a way for dealing with this type of multiplications is presented.

As a comparison of the two approximation methods, the second approach is expected to give results closer to those of the original nonlinear approach. This is because in the first method, we replace some variables with simulated data and the values will remain unchanged during the optimization. However, the computation time required in the second approach is expected to be higher than the first one since in the second approach more auxiliary variables are defined. These expectations are confirmed in the case studies section.

SECTION V. Reformulation as MILP

The approximate models cannot be directly used in a linear or PWA MPC framework (9)–(15). This is due to the fact that in the approximated models two sets of binary variables are introduced; one set is associated with switching between the intervals of the PWA functions, and the other set contains the switching signals for both the timing plans and also the perimeter control inputs (as we made them quantized). On the other hand, due to the large number of regions that the combination of different affine pieces in the model introduces, the evaluation of the approximated models for several times as a part of the optimization algorithm inside the MPC scheme, is not efficient. Therefore, we make a conversion of the approximated models to a system of the following form:

$$\begin{aligned} x(k+1) &= A \cdot x(k) + B_1 \cdot u(k) + B_2 \cdot \delta(k) + B_3 \cdot z(k) + b \\ y(k) &= C \cdot x(k) + D_1 \cdot u(k) + D_2 \cdot \delta(k) + D_3 \cdot z(k) \\ d &\geq E_1 \cdot x(k) + E_2 \cdot u(k) + E_3 \cdot \delta(k) + E_4 \cdot z(k) \end{aligned}$$

[View Source](#) 

where $x(k) \in \mathbb{R}^{n_x}$ and $y(k) \in \mathbb{R}^{n_y}$ respectively represent the state and output

where $x(k) \in \mathbb{R}^{n_x}$ and $y(k) \in \mathbb{R}^{n_y}$, respectively, represent the state and output vectors, while $\delta(k) \in \{0, 1\}^{n_\delta}$ and $z(k) \in \mathbb{R}^{n_z}$ are auxiliary binary and real-valued variables, resulting from a procedure explained subsequently. Further, b and e are constant vectors that along with the system matrices $\{A, B_i, C, D_i, E_i\}$ specify a mixed logical dynamic model [37]. In this model representation, the binary (defined for PWA function, switching between MFDs, and quantization of perimeter control signals) and auxiliary variables required to define the regions are directly included in model through additional constraints. Compared with the models derived in the previous section, one large but tractable model applies that is composed by stacking the individual linear and affine equations along with auxiliary linear inequalities, and thus resulting in a model size that grows linearly with increasing the number of regions.

In this section, we transform the approximated models presented in the previous sections to the form of (25). Consider an affine function $f(\cdot)$ defined over a bounded set $\mathcal{X}$ of the input variable x , with upper and lower bounds M and m over $\mathcal{X}$. Having a binary decision variable $\delta \in \{0, 1\}$, it can be proved that the following statement holds [37], [38]:

$$[f(x) \leq 0] \Leftrightarrow [\delta = 1], \text{ iff } \begin{cases} f(x) \leq M \cdot (1 - \delta) \\ f(x) \geq \epsilon + (m - \epsilon) \cdot \delta \end{cases} \quad (25)$$

[View Source](#) 

with ϵ being a small tolerance used to change a strict inequality into a nonstrict inequality. Moreover, the product of two binary variables δ_1 and δ_2 can be replaced by an auxiliary binary variable $\delta_3 \triangleq \delta_1 \cdot \delta_2$. Next, it can be verified that

$$\delta_3 = \delta_1 \cdot \delta_2 \text{ is equivalent to } \begin{cases} -\delta_1 + \delta_3 \leq 0 \\ -\delta_2 + \delta_3 \leq 0 \\ \delta_1 + \delta_2 - \delta_3 \leq 1. \end{cases} \quad (26)$$

[View Source](#) 

Finally, multiplication of a binary variable δ with an affine function $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ 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 is easy to verify that

$$z = \delta \cdot f(x) \text{ is equivalent to } \begin{cases} z \leq M \cdot \delta \\ z \geq m \cdot \delta \\ z \leq f(x) - m \cdot (1 - \delta) \\ z \geq f(x) - M \cdot (1 - \delta) \end{cases} \quad (27)$$

[View Source](#) 

with m and M the minimum and maximum of $f(\cdot)$ over the set $\mathcal{X}$, respectively.

Using the above mentioned rules, one can rewrite the approximated models presented in the previous section into the form of (25). For instance, the PWA function (17) can be rewritten as

$$\hat{P}_{i,f_i}(n_i(k)) = \sum_{j=1}^3 (\mathcal{A}_{i,f_i}^j \cdot n_i(k) + \mathcal{B}_{i,f_i}^j) \cdot \delta_{i,f_i}^j \quad (28)$$

[View Source](#) 

where δ_{i,f_i}^j correspond to the intervals defined in (17) ($\delta_{i,f_i}^j = 1$ when $n_i(k)$ is in the interval j) and $\mathcal{A}_{i,f_i}^j$ and $\mathcal{B}_{i,f_i}^j$ can be calculated from the formulation presented in (17). Then, it is straightforward to rewrite (29) into the form of (25) with the help of (26) and (28).

After reformulation of the approximated models presented in Section IV we get a

into a system of linear equations and linear inequality constraints including real and integer variables. Getting back to the optimization problem, the selected performance index (total time spent) is already in the linear form. However, one can add a penalty term to prevent undesired fluctuations in the perimeter control inputs and the decision switching variables. The penalty on the perimeter control inputs can be defined as follows:

$$\sum_{l=1}^{N_p-1} \underbrace{|u_{ij}^c(k_c + l) - u_{ij}^c(k_c + l - 1)|}_{p_l}. \quad (29)$$

[View Source](#) 

The above term can be transformed into a linear form by defining auxiliary variables as follows:

$$\sum_{l=1}^{N_p-1} q_l \text{ subject to } -p_l \leq q_l \leq p_l. \quad (30)$$

[View Source](#) 

It can be easily proved that minimizing (31) over p_l and q_l would result in the same optimal solution as in case of minimizing (30) over p_l .

All in all, the problem of minimizing the total time spent in the network subject to the obtained mixed linear model of the system and other linear constraints on the inputs and states is formulated as a mixed-integer linear optimization problem (MILP) that is solved in the MPC framework. This problem is tractable and can be solved using advanced solvers [39].

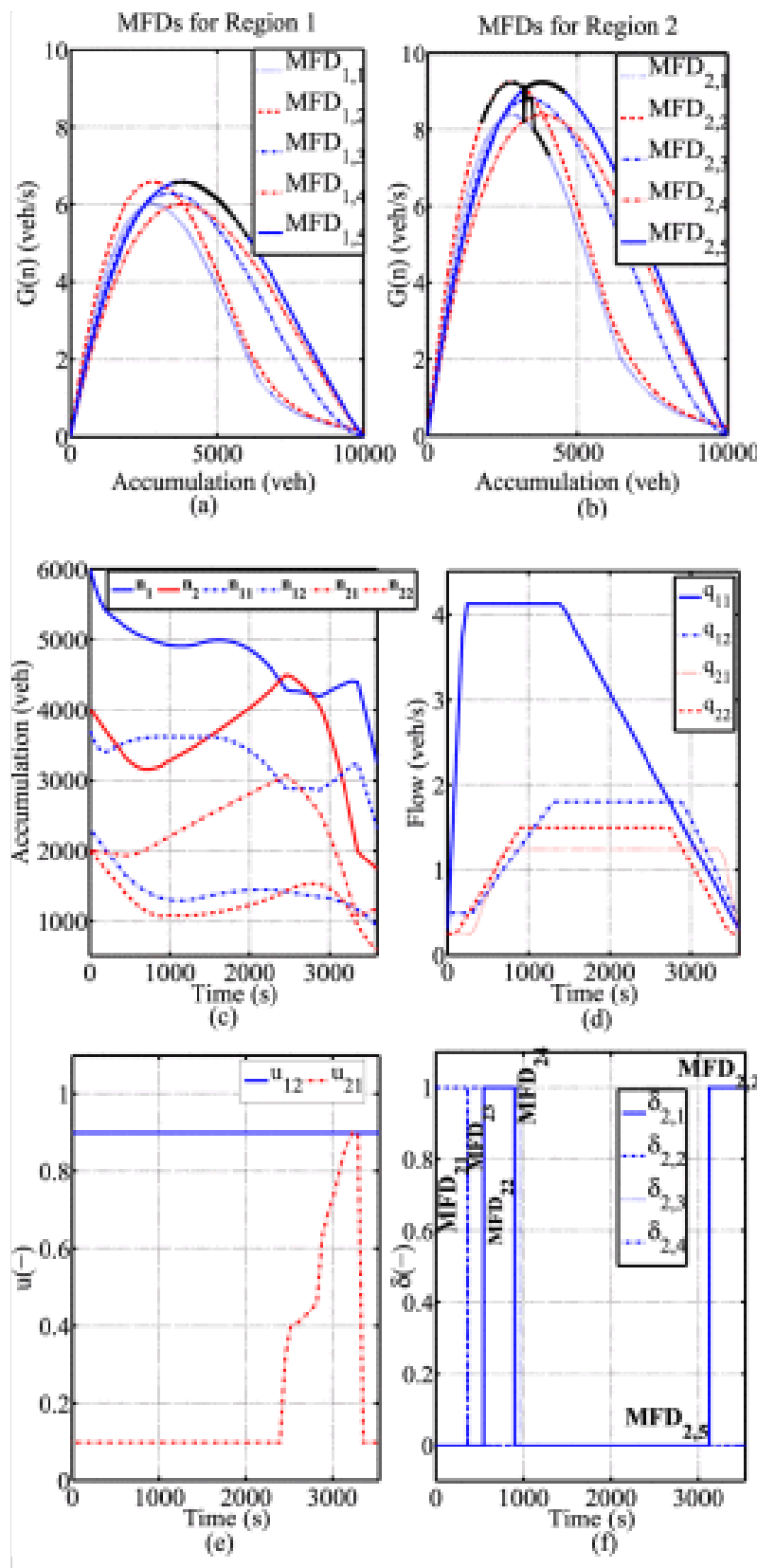
SECTION VI. Case Studies

In this section, we implement and evaluate the performances of the proposed hybrid schemes using simulation. We stick to the macroscopic level to investigate and highlight the performance of our proposed control methods. We use a simulation model to represent the urban traffic network and a prediction model to estimate the traffic states inside the MPC framework. We start with low mismatch between the simulation model and the prediction model to evaluate how the proposed control algorithms deal with the general traffic congestion control problem. Next, we perform several extensive tests with the introduction of different types of uncertainties in the simulation model, to better represent the reality and to evaluate our control approaches under more realistic scenarios. In Example 1, we investigate the performance of the hybrid perimeter and switching timing plans control (the original MINLP approach) and show that additional improvements are obtained, compared with perimeter control or switching timing plans only, if both control entities are coordinated and considered in the mixed-integer nonlinear optimization. The performance of the proposed hybrid scheme is further compared with a greedy feedback perimeter controller. In Example 2, the two approximation methods are implemented and their performances are compared with the MINLP approach in terms of computation time and total cost. Two different demand profiles are selected in the examples to show that the proposed hybrid scheme is able to handle different traffic scenarios. Finally, we present the results of evaluating our proposed methods for different types of uncertainties (in state measurements, MFDs, trip demands) introduced in the simulation model. In the following, we describe in full details the urban network under study and different types of uncertainties that might exist in reality and needed to be modeled in the network simulation model.

A. Setup

A. Setup

We consider an urban network partitioned into two homogeneous regions, i.e., $R = 2$, the periphery (region 1) and the city center (region 2). The libraries of the signal timing plans and MFDs are given *a priori* for each region. In the first example, a set of five MFDs are defined for the periphery (region 1) and the city center (region 2), as shown in Fig. 6(a) and (b), respectively. In Fig. 6(b), the set consists of MFD_{1,3} adopted from [5] with maximum trip completion flow 6.3 veh/s corresponding to critical accumulation 3400 veh, jam accumulation 10 000 veh (thus the parameters of the third-order polynomial would be $a = 1.4877 \times 10^{-7}/3600$, $b = -2.9815 \times 10^{-3}/3600$, $c = 15.0912/3600$) and four other MFDs that are obtained based on deviation from the critical accumulation and the maximum trip completion flow of MFD_{1,3}. The percentages of the deviations are $\pm 10\%$ and $\pm 5\%$ for the critical accumulation and the maximum trip completion flow, respectively. Moreover, it is assumed that the sizes of the two regions are different, hence the MFDs of the city center (region 2) are the periphery MFDs multiplied by a coefficient (1.4), as shown in Fig. 6(a). In practice, these MFDs can be obtained by changing the signal settings of intersections and can be estimated with [9] and [8]. In Example 2, each region is assumed to have three MFDs [the same MFD_{1,2}, MFD_{1,3}, and MFD_{1,4}, as shown in Fig. 6(b)].



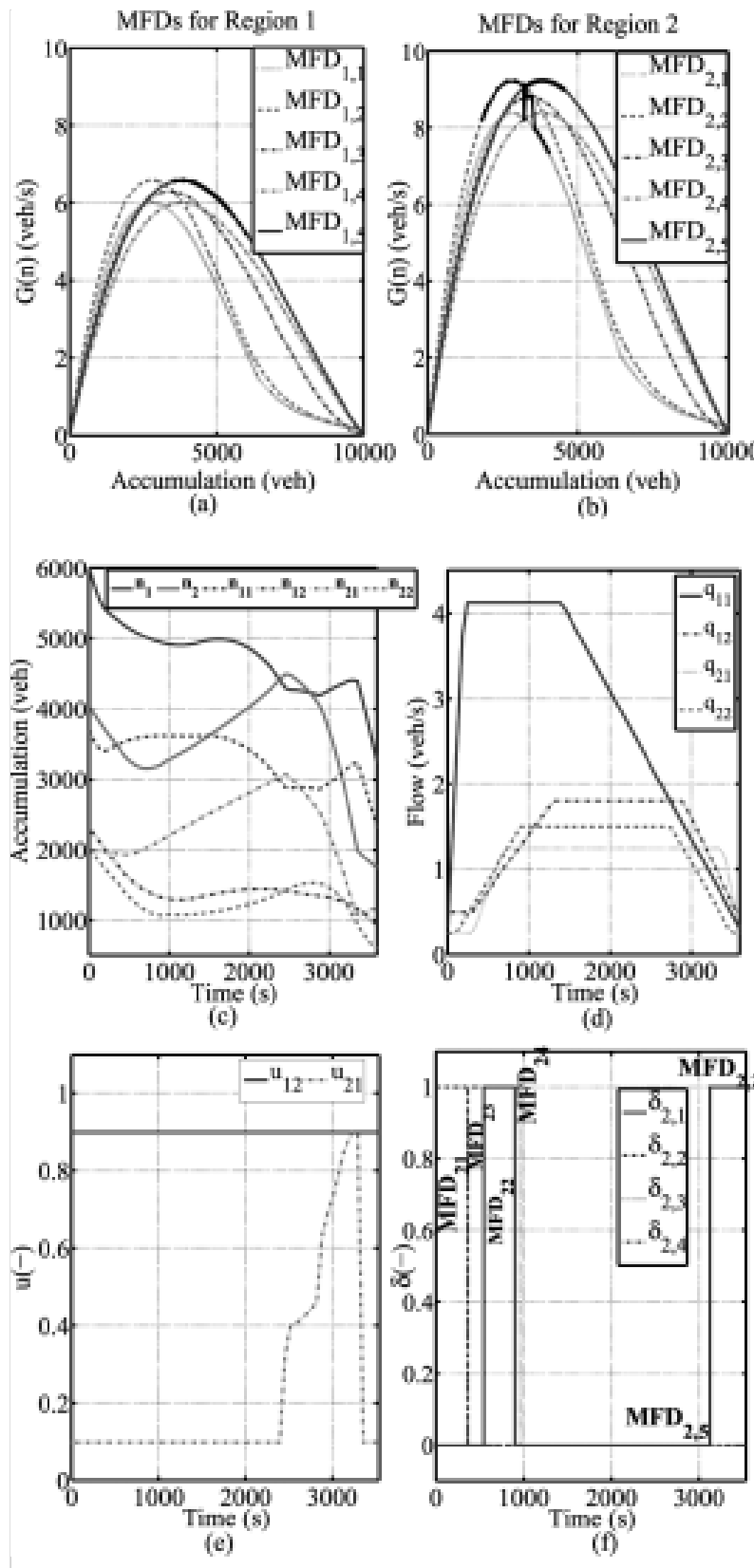


Fig. 6.

Example 1: Performance overview of the nonlinear hybrid scheme. (a) and (b) MFDs used in the prediction model together with schematic optimal switching between MFDs. (c) Accumulations. (d) Average demand profiles used in MPC. (e) Optimal perimeter control input. (f) Optimal switching signals for region 2 (the controller always chooses $MFD_{1,5}$ for region 1 in this case).

1) Uncertainties in Plant:

The dynamic equations of the simulation model (plant/reality) differ from the

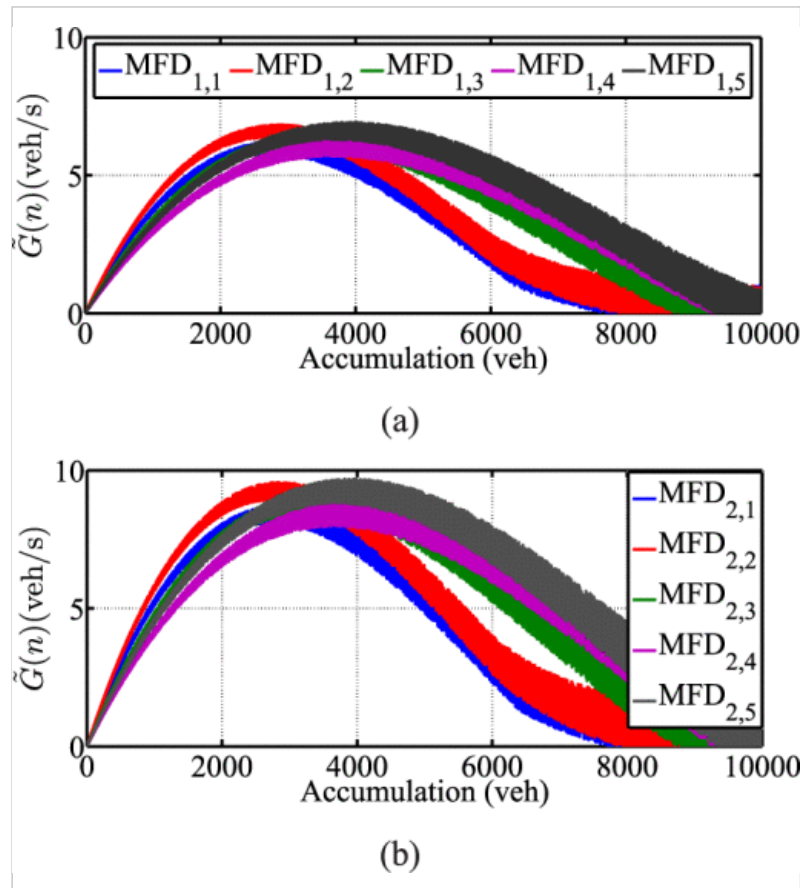
The dynamic equations of the simulation model (primary reality) differ from the prediction model used in the MPC framework as they contain different types of error explained in the following. Please note that the presented MFDs in Fig. 6(a) and (b) are utilized for the MPC prediction model, while the network (reality) is assumed to include errors in the MFDs following the error formulation in [17]. In reality, an MFD is extracted based on several data collection experiments in the network. Based on the level of homogeneity of the network, the MFD will exhibit scattering. By scattering, we mean that in general corresponding to each accumulation, there exist multiple trip production points. The level of scattering increases when the accumulation grows. Therefore, there is no explicit mathematical equation for the MFD. In the following, we approximate the MFD using a third-order polynomial but to consider the scattering, we assume a uniformly distributed additive noise with zero mean and a variance that is proportional to the accumulation level. For all simulation scenarios, we add the error e_i (veh/s) to the MFDs of the simulation model as follows:

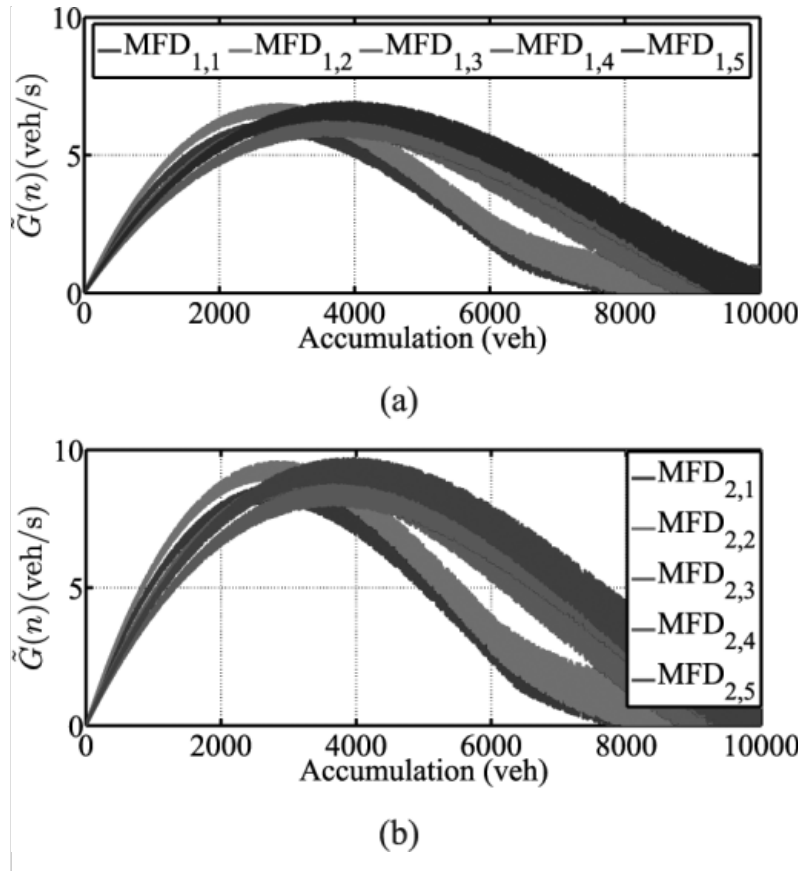
$$e_i(k) \sim U(-C_i \cdot n_i(k), C_i \cdot n_i(k)) \quad (31)$$

$$\tilde{G}_{i,f_i}(n_i(k)) = G_{i,f_i}(n_i(k)) + e_i(k) \quad (32)$$

[View Source](#) 

with $C_i = 0.2/3600$. Hence, we get a model where the scattering increases with the increase in the level of accumulations. The MFDs $\tilde{G}_{i,f_i}$ used to simulate the urban network are shown in Fig. 4.



**Fig. 4.**

Uncertain MFDs representing the urban network under control. (a) MFDs of the periphery. (b) MFDs of the center.

In reality, there is uncertainty about the measured states of the network, especially in the estimation of the number of vehicles with destinations inside regions or across the regions (n_{ii} and n_{ij}). Hence, the effects of errors in the measured states should be considered. However, to be consistent with reality, one should expect larger errors in the n_{ij} than in the total number of vehicles inside the region, i.e., n_i . This is due to the fact that the estimation of the total number of vehicles inside a region is easier than obtaining an estimation of the number of vehicles with certain destinations that can be inside a region or in other regions in the neighborhood (for example n_i can be estimated with fixed sensors in certain locations of the network, while n_{ij} would require tracking devices of vehicles and destination will not be fully known until vehicles reach it). Therefore, we model the error in the states as follows:

$$\tilde{n}_{ii}(k) = n_{ii}(k) + \omega_{ii} \cdot n_{ii}(k) \cdot \varepsilon_{ii}(k) \quad (33)$$

$$\tilde{n}_{ij}(k) = n_{ij}(k) + \omega_{ij} \cdot n_{ij}(k) \cdot \varepsilon_{ij}(k) \quad (34)$$

[View Source](#)

where the values for ω_{ii} and ω_{ij} are first set to 0.05 and then to 0.1, to simulate a 5% and a 10% error in the measurements, respectively. Moreover, the error vector $\varepsilon(k) = (\varepsilon_{ii}(k), \varepsilon_{ij}(k))^T$ has a normal distribution with the mean value of zero and the covariance matrix as

$$\text{Cov}(\varepsilon) = \begin{bmatrix} 1 & -0.75 \\ -0.75 & 1 \end{bmatrix}. \quad (35)$$

[View Source](#)

The total of number of vehicles inside the region i ($\tilde{n}_i(k) = \tilde{n}_{ii}(k) + \tilde{n}_{ij}(k)$) will

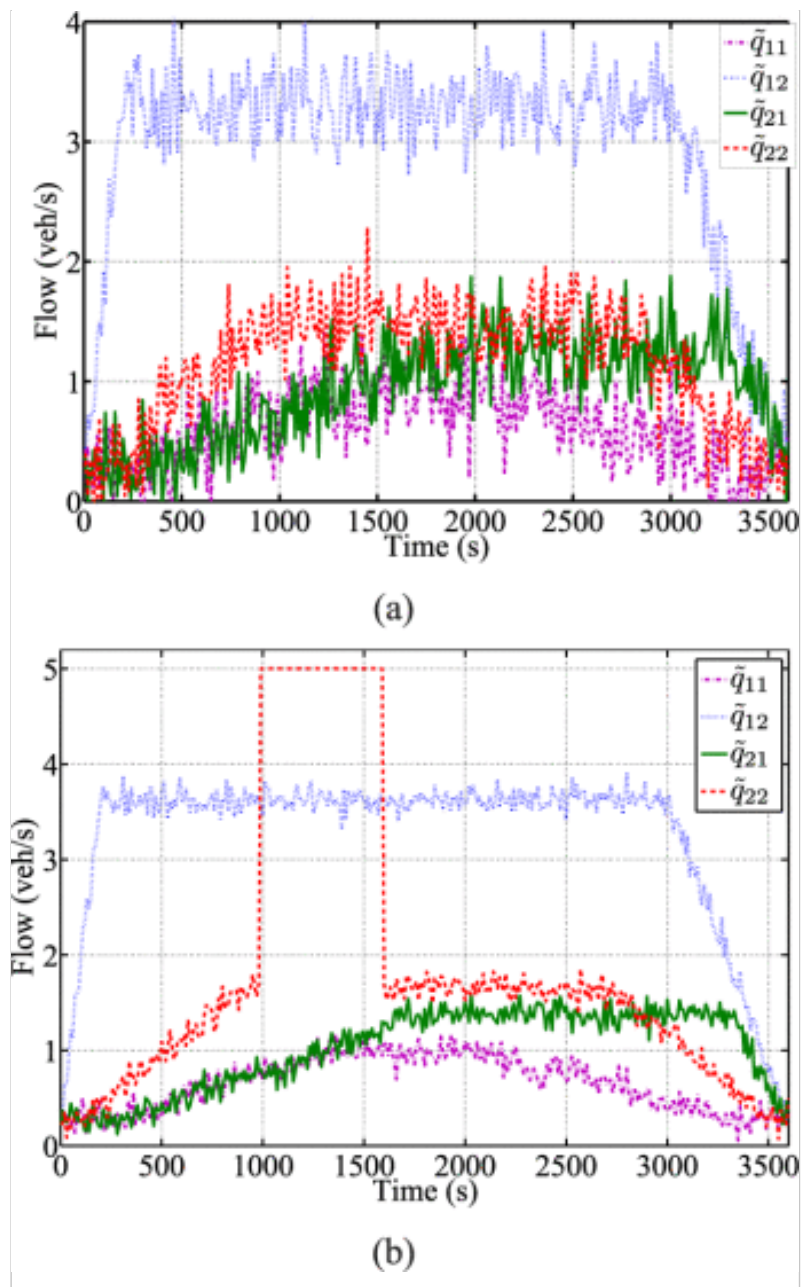
contain the sum of the elements of the error vector $\varepsilon(k)$, thus with the variance $1/2$ (as the variance of the sum of two correlated variables $\text{Var}(X + Y)$ is $\text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$).

Furthermore, we also consider the uncertainty in trip demands. The prediction model in the MPC framework takes the average profile as e.g., shown in Fig. 6(d), while the network simulation model assumes noisy demand profiles to represent uncertain variations of demands from day to day and also to include events that temporarily affect the demand profiles. For the first case, the unbiased demand is assumed to have an additive white Gaussian noise, as follows:

$$\tilde{q}_{ij}(k) = q_{ij}(k) + \mathcal{N}(0, \sigma_{ij}^2) \quad (36)$$

[View Source](#) 

with $i, j = 1, 2$ and σ_{ij}^2 (veh/s)² the variance of the noise. For the simulations, we consider a large noise with $\sigma = 0.5$.³ On the other hand, in the biased case, the demand profile has a sudden jump, as well as an additive nonzero mean Gaussian noise. This jump is not known to the MPC controller and only included in the network simulation model. In Fig. 5(a), the demand profile corrupted with unbiased noise is shown, while the biased demand is shown in Fig. 5(b). These demand profiles are used in the network model of Example 2. Note that since we deal with aggregated regional-based trip demands, they suffer less from the effects of fluctuations that exist in regular origin-destination demand estimations. Furthermore, the level of the noise added to the trip demands is in line with the results of practical experiments presented in [40] and [41]. In Examples 1 and 2, we have only included the errors in the MFDs, while in the last section (robustness evaluation), we investigate the effects of all possible uncertainties on the performance of the hybrid control approaches.



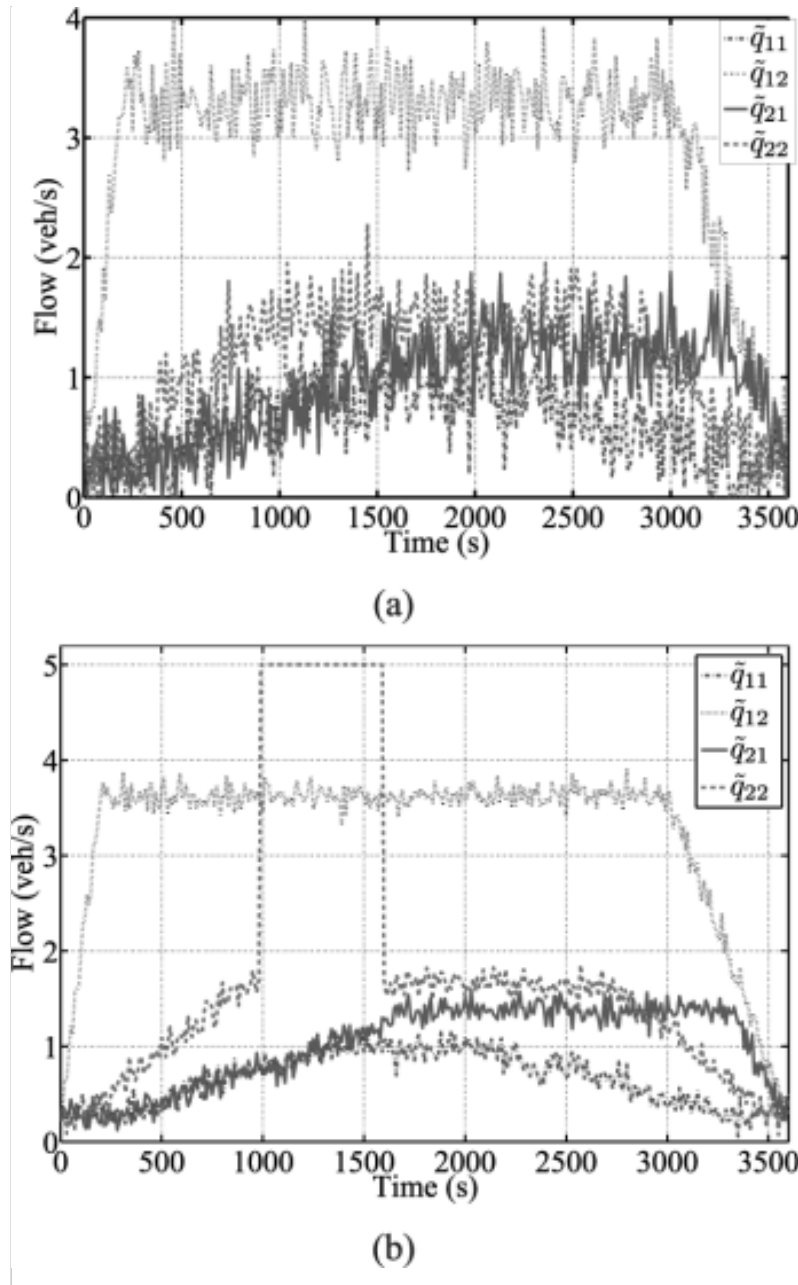


Fig. 5. Noisy demand profiles used in the simulation model (plant/network). (a) High unbiased noise in demand. (b) Biased noise (sudden jump in q_{22}).

2) Simulation Parameters:

For simulation of the system, we choose the sample time as $T = 30$ s. Moreover, the control sample time is selected as $T_c = 60$ s for Examples 1 and 2, while it varies for the last section (robustness to measurement noise). The hybrid model predictive controllers (the MINLP approach and the PWA-MILP methods) use the prediction horizon $N_p = 20$ min and the control horizon $N_c = 2$ min.⁴ Moreover, the penalty term (30) is added to the objective function with a weight of 10 (this choice is obtained based on the nominal values of the total time spent objective function and the penalty term. More discussions about finding proper weights are provided in [17]). Furthermore, the lower and upper bounds of the perimeter input are selected as $u_{ij,\min} = 0.1$ and $u_{ij,\max} = 0.9$. Therefore, the flows between regions are neither completely allowed, nor fully blocked.

B. Example 1

The demand profiles for trins inside each region and between them are shown in

The demand profile for trips inside each region and between them are shown in Fig. 6(d). There is a high demand for trips inside the periphery, see $q_{11}(\cdot)$ in the figure. Further, both regions are initially congested, i.e., the initial accumulations are larger than the critical accumulations [$n_{11}(0) = 3700$ veh, $n_{12}(0) = 2300$ veh, $n_{21}(0) = 2000$ veh, $n_{22}(0) = 2000$ veh].

The hybrid controller finds the optimal perimeter control inputs along with the optimal timing plan, as shown in Fig. 6(e) and (f), for each region using mixed-integer nonlinear optimization. The values of MFDs calculated during the optimization are shown in the black curve, e.g., the calculated values of MFD for region 1 belong to $MFD_{1,5}$, which is the optimal MFD (or signal timing plan) during the whole time period as switching of plans does not occur.

In the absence of control or having only the perimeter control on the borders, one or both regions would get to a gridlock situation. But with optimal switching between timing plans and assisted by perimeter control, both regions will escape from high-level congestion and they will be eventually uncongested by the end of the simulation interval, as shown in Fig. 6(c).

To evaluate the MPC hybrid controller results, the total time spent for the whole period of simulation (1 h) is compared for several control schemes, as shown in Table I: 1) only perimeter control with different combinations of MFDs (the MFDs for both regions are fixed during the simulation period. Since there exist 5 MFDs in each of the libraries, 25 combinations would be possible to choose); 2) switching timing plans control only; and 3) a greedy feedback perimeter controller. The greedy perimeter controller is a simple state-feedback *perimeter* controller with the policy of protecting regions with high accumulations and high-trip destinations. The control laws of the greedy controller are as follows: if both regions are uncongested, the perimeter control inputs are maximized and if both regions are congested, the perimeter control inputs $u_{i,j}$ and $u_{j,i}$ are, respectively, set to the maximum and minimum values, if region j is more congested than region i and vice versa. Note that the greedy control has been tested for all 25 combinations of MFDs.

TABLE I Performance Evaluation for Different Control Schemes in Example 1

Control scheme	Total time spent (veh · s)
Uncontrolled	Gridlock
Greedy feedback controller	Gridlock
Perimeter control with $MFD_{1,5}$ and $MFD_{2,1}$	3.47×10^7
Perimeter control with $MFD_{1,5}$ and $MFD_{2,2}$	3.42×10^7
Perimeter control with $MFD_{1,5}$ and $MFD_{2,3}$	3.38×10^7
Perimeter control with $MFD_{1,5}$ and $MFD_{2,4}$	3.37×10^7
Perimeter control with $MFD_{1,5}$ and $MFD_{2,5}$	3.36×10^7
Perimeter control with other MFD combinations	Gridlock
Switching timing plans	Gridlock
Hybrid controller (MINLP approach)	2.78×10^7

The results shown in Table I imply that the MPC hybrid controller is superior for all control schemes in the sense that at least 17% improvement in total time spent is achieved when both controllers are applied instead of only perimeter control. In addition, note that applying only the switching timing plan control or using the greedy feedback control still leads to gridlock situations in the regions.

C. Example 2

In this example, we provide a scenario to evaluate the performance of the proposed approximated methods and the original mixed-integer nonlinear optimization approach. Moreover, to have a better performance evaluation of the approximation approaches, the results are compared with the greedy perimeter controller as well.

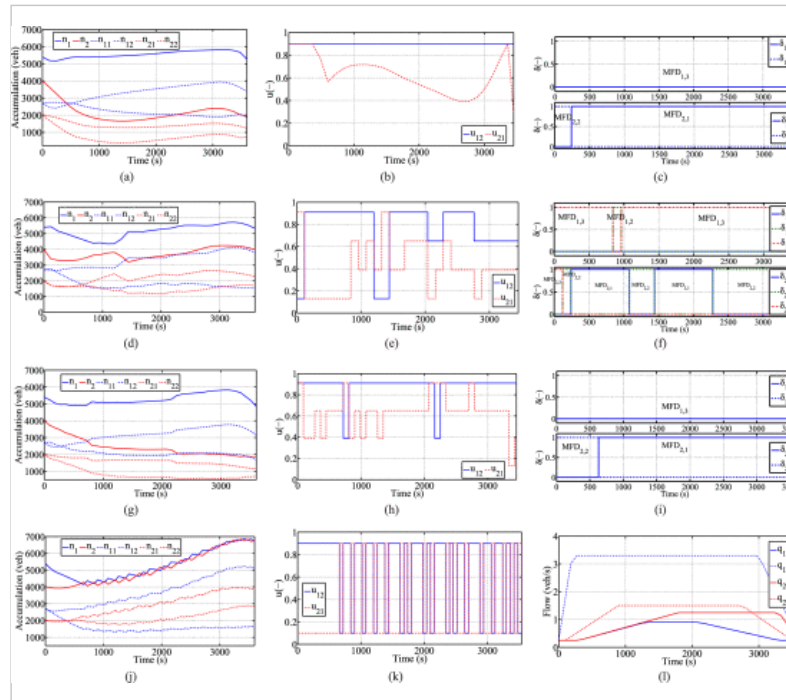
The demand profile simulates a peak morning hour with high demand $q_{11}(\cdot)$ for

The demand profile simulates a peak morning hour with high demand $q_{12} \setminus /$ trips from region 1 (the periphery) to region 2 (the city center), as shown in Fig. 7. The closed-loop system is simulated for a period of 1 h. The initial accumulations are $n_{11}(0) = 2700$ veh, $n_{12}(0) = 2700$ veh, $n_{21}(0) = 2000$ veh, and $n_{22}(0) = 2000$ veh. The accumulations of the regions are measured and fed to the MPC controller. There are three cases of MPC controllers; one with embedded MINLP optimization based on the nonlinear prediction model, one with MILP optimization based on the first approximated model as prediction model (we call it PWA-MILP1), and one with MILP optimization based on the second approximated model as prediction model (we call it PWA-MILP2). The quantized perimeter input is formulated as

$$u_{ij}(k) = u_{ij,0} \cdot (0.5 + 2^0 \cdot \omega_{ij,1}(k) + 2^1 \cdot \omega_{ij,2}(k)) \quad (37)$$

[View Source](#) ?

with $u_{ij,0} = 0.26$. Therefore, in the PWA-MILP cases, the perimeter control input takes values from the set $\{0.13, 0.4, 0.65, 0.9\}$.



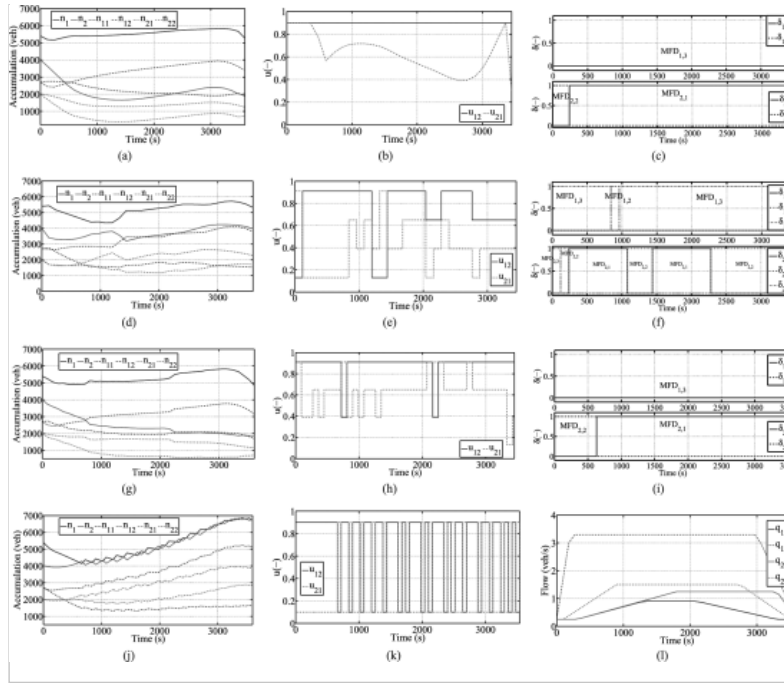


Fig. 7.

Example 2: The results obtained from the MINLP approach, PWA-MILP1, PWA-MILP2 methods, and the greedy perimeter controller for a two-region urban network. (a), (d), (g), and (j) Accumulations. (b), (e), (h), and (k) Perimeter control inputs. (c), (f), and (i) Switching timing plans inputs. (l) Average demand profiles for trips inside each region and between them.

The evolution of accumulations over time corresponding to the MINLP approach, the first approximation method PWA-MILP1, the second approximation method PWA-MILP2, and the greedy controller are shown in Fig. 7(a), (d), (g), and (j), respectively. These figures demonstrate the effectiveness of the control measures as they show that the control inputs prevent the two regions from moving forward toward gridlock (as all accumulations are less than the jam accumulations). In the absence of control, the gridlock circumstance would occur. The MINLP approach results in a better performance compared with both PWA-MILP approaches, in particular for the accumulations of region 2. For the PWA-MILP1 approach, this can be explained by the fact that we have approximated the second-order polynomials with two affine functions (17), and also because of the forward simulation method that has been introduced to overcome the multiplication of variables. Hence, the performance of the PWA-MILP1 method can be improved by approximating the polynomials with a larger number of affine functions and using more iterations in each control time step.

Nevertheless, a more accurate way to tackle the problem with multiplication of variables was proposed in the second approximation method PWA-MILP2. Therefore, the performance of the PWA-MILP2 method is closer to MINLP approach in terms of the sum of accumulations over the whole simulation period. Moreover, to further verify the advantage of the MILP formulation of the problem, the greedy perimeter controller results are compared with the results of the hybrid approaches. Comparing with Fig. 7, the greedy perimeter controller's performance is much worse than all three hybrid approaches. With the greedy controller, the accumulations of both regions will exceed 7000 vehicles at the end of the simulation time, and the total time spent is much higher.

The optimal perimeter control inputs for the MINLP, PWA-MILP1, PWA-MILP2 approaches, and the greedy controller are shown in Fig. 7(b), (e), (h), and (k), respectively. The perimeter inputs $u_{12}(k)$ of the MINLP approach are close to the maximum to allow more vehicles to leave region 1 while u_{21} varies more over time. Moreover, the optimal switching timing plans for the MINLP, PWA-MILP1, and PWA-MILP2 approaches are, respectively, shown in Fig. 7(c), (f), and (i), for both

regions 1 and 2. It can be observed in this scenario that the optimization algorithms

regions and it can be observed in this scenario that the optimization algorithms mostly take the envelope of the three MFDs for each region. This is more clear in the MINLP and PWA-MILP2 approaches while in the PWA-MILP1 case, the switching between MFDs occurs more often.

The computation time and total time spent are compared for different proposed algorithms and for different values of prediction horizon in Table II. The average computation time for the scenario $N_p = 20$ is 51.52 s for one run of the MINLP algorithm,⁵ while it is 1.143 and 5.3934 s for the PWA-MILP1 and PWA-MILP2 approaches. Note that the MINLP algorithm has been executed 10 times in each control time step for different random initial points to prevent reaching local optimal solution. Therefore, the actual computation time of the MINLP method is multiple of the aforementioned number.

TABLE II Performance Evaluation for Example 2

Prediction horizon	Total time spent (veh · s)				Average computation time (s)			
	MINLP	PWA-MILP 1	PWA-MILP 2	PWA-MILP 3	MINLP	PWA-MILP 1	PWA-MILP 2	PWA-MILP 3
10	3.35×10^7	3.91×10^7	3.47×10^7	3.42×10^7	20.33×10	0.527	1.1323	1.0165
20	2.85×10^7	3.31×10^7	2.87×10^7	2.85×10^7	51.52×10	1.143	5.3934	3.5341
30	2.79×10^7	3.18×10^7	2.84×10^7	2.80×10^7	164.12×10	4.103	9.3129	7.8275

It can be inferred from Table II that the PWA-MILP2 method has a better performance in terms of the total time spent (veh · s) compared with the PWA-MILP1 approach but slightly worse than the MINLP case. The computation time of the PWA-MILP2 approach is larger than PWA-MILP1 approach but much smaller than nonlinear case. Furthermore, using each of the two PWA-MILP approaches result in less total time spent than the greedy perimeter controller (3.75×10^7 veh · s). Only in the case $N_p = 10$, the first approximation method gives a slightly larger time spent compared with the one achieved from the greedy controller.

Moreover, as mentioned before in Remark 1, the quantization of the perimeter control input can be prevented by introducing an auxiliary variable $\tilde{M}$ and by adding some extra inequality constraints, provided that the perimeter control input is not penalized. Results of using this technique are presented in Table II, under the name PWA-MILP3. In fact, we follow the same approach as in PWA-MILP2, but without quantizing the perimeter input and without considering the penalty term on the perimeter input. The obtained results show a slight decrease in the computation time and an improvement in the total objective function (compared with the other approximation methods). Therefore, if one prefers to penalize the control input (e.g., to prevent instability or other undesired behaviors due to oscillations in the control input), PWA-MILP1 or PWA-MILP2 is suggested. However, if penalizing the control input is not deemed necessary, clearly PWA-MILP3 is the best choice.

D. Robustness to Measurement Errors and Uncertain Demands

In this section, we first study the effects of measurement errors on the performance of the proposed schemes and propose a solution for the drawbacks caused by these errors. Next, we study the impacts of the unbiased and biased noise in trip demands. The selected scenario is identical to Example 2 with addition of the two new types of uncertainties introduced in the simulation model. The prediction horizon is $N_p = 20$ min, the control horizon is $N_c = 2$ min, the simulation sample time 30 s, and the total simulation time is 1 h.

By adding the measurement errors to the plant, as in (34) and (35), the performance of the hybrid controller gets affected by introducing fluctuations in the perimeter control inputs and by slightly increasing the number of switching between MFDs. Simulation results in case of having 10% error in the measured n_{ij} are shown in Fig. 8(a). The perimeter control inputs have considerable jumps, and therefore not useful for practical situations. Traffic operators expect more stable control profiles with smaller changes in the pattern. To overcome this problem, we propose, in addition to penalizing the control input variations, to select a control

sample time larger than the simulation sample time while keeping the obtained

sample time might cause the simulation sample time while keeping the obtained control inputs constant between two consecutive control time steps. By performing this, the perimeter control inputs will have a smoother behavior over time, as can be observed in Fig. 8, while the total time spent in the network might not be altered significantly.

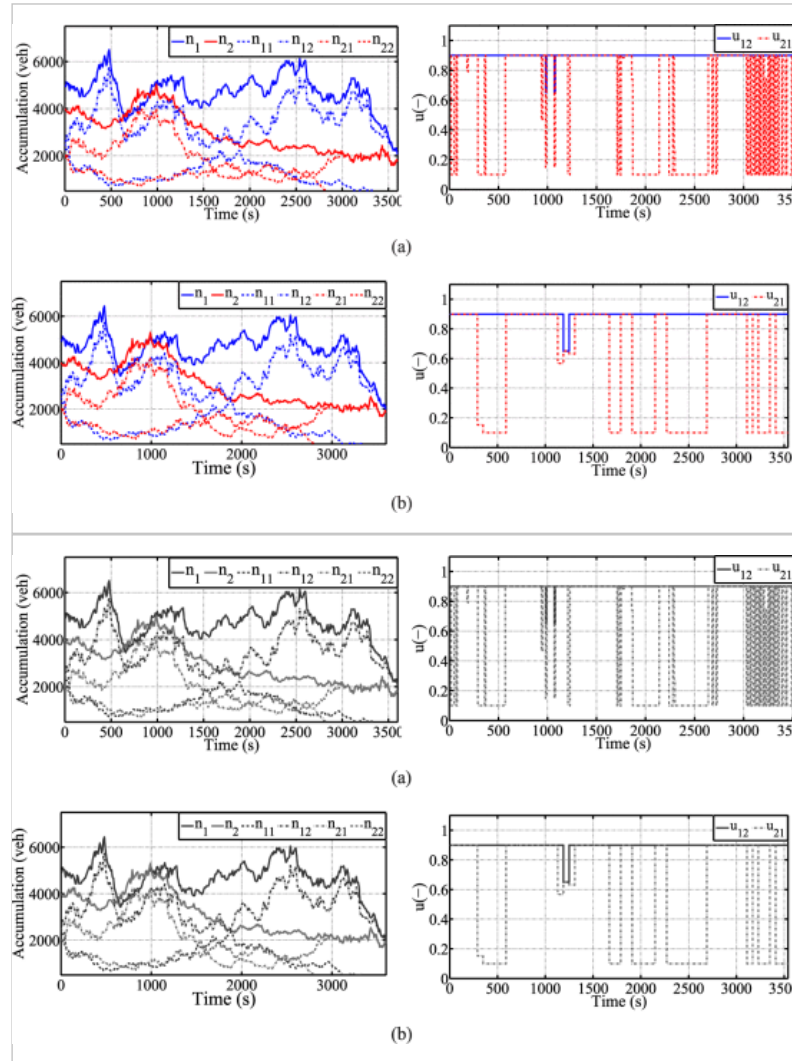


Fig. 8.

(a) Effect of measurement errors on states and perimeter control inputs. (b) Smoothing the perimeter control inputs (MINLP approach).

Results for simulation with different T_c are presented in Table III. Note that due to the addition of errors in the system, we expect different total time spent values for different runs with the same set of control inputs. Thus, the values presented in the table are the average over 10 runs for each case of control sample time. It can be observed that for less frequent calls of the controller, we achieve around the same results but with less oscillations in the control inputs. However, for the ratio $T_c/T = 6$ and higher the performance will get worse. Furthermore, the obtained results show that the PWA-MILP1 approach has worse results compared with the other two approaches due to the forward simulation technique.

TABLE III TTS VALUES ($\times 10^7 \text{veh} \cdot \text{s}$) FOR Example 2, in the PRESENCE of MEASUREMENT ERROR and NOISE in DEMANDS (in Addition to Error in MFDs)

Noise in trip demands	$\frac{\sigma}{\bar{r}}$	No measurement error			Measurement error (5%)			Measurement error (10%)		
		MINLP	PWA-MILP 1	PWA-MILP 2	MINLP	PWA-MILP 1	PWA-MILP 2	MINLP	PWA-MILP 1	PWA-MILP 2
no noise	2	2.85	3.31	2.87	2.87	3.42	2.92	3.26	3.76	3.28
	3	-	-	-	2.90	3.46	2.99	3.28	3.78	3.29
	6	-	-	-	3.43	3.86	3.48	3.62	3.97	3.66
unbiased	2	2.88	3.22	2.86	2.94	3.45	3.01	3.15	3.81	3.18
	3	-	-	-	2.97	3.53	3.13	3.19	4.01	3.21
	6	-	-	-	3.56	3.78	3.64	3.73	gridlock	3.91
biased	2	3.14	3.97	3.20	3.32	4.21	3.30	3.45	gridlock	3.55
	3	-	-	-	3.36	4.35	3.38	3.56	gridlock	3.61
	6	-	-	-	3.74	gridlock	3.86	3.88	gridlock	4.02

Now we study the effects of adding noise in the trip demands in the simulation model (note that the prediction model still takes the average demand profile). In the first case, the demand profile shown in Fig. 5(a) is selected. Simulation results are shown in Fig. 9(a) and (b) for two control strategies MINLP and PWA-MILP1. The numerical results for other cases are presented in Table III. As can be observed, the hybrid control strategies are able to handle the unbiased noise in the demands. Only small size fluctuations in the control inputs occur, which can be prevented by penalizing the control inputs and also by increasing the control sample time. It should be noted that increasing the control sample time more than three times the simulation sample time affects the performance.

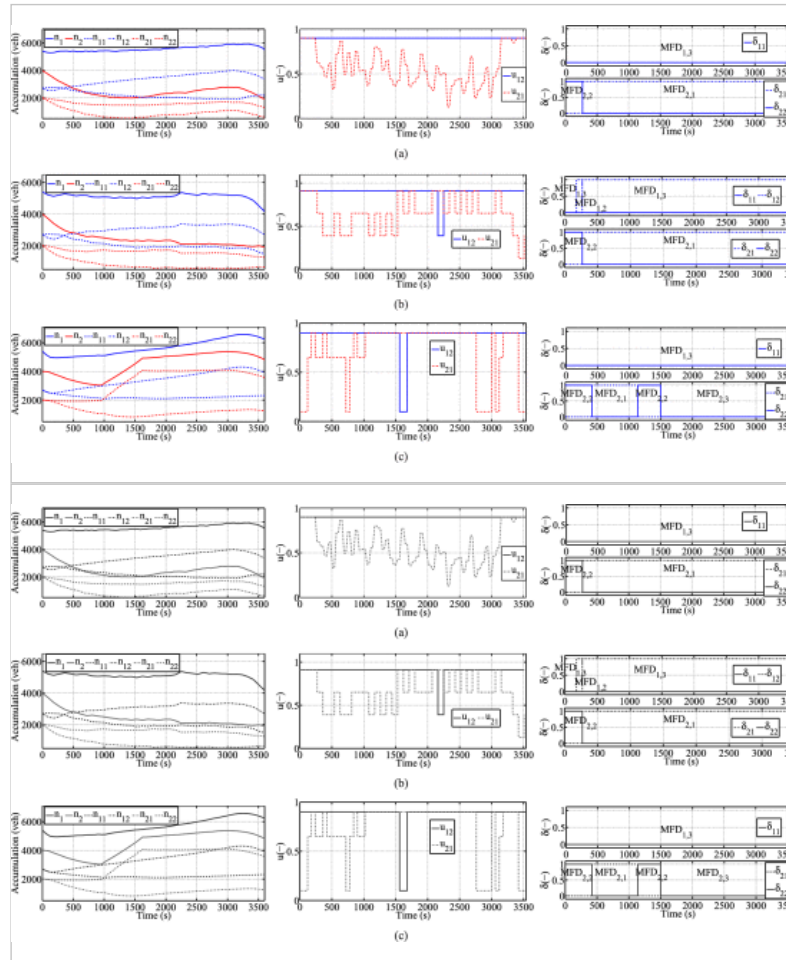


Fig. 9.

Robustness to noise in demand. (a) Unbiased noise, control using MINLP approach. (b) Unbiased noise, control using PWA-MILP2 method. (c) Biased noise (peak in q_{22}), control using PWA-MILP2.

Next, we use the biased noise in the demand profile of the simulation model. We have added Gaussian noise a mean of 10% of the average profile and a variance $\sigma_{ij}^2 = 0.2^2$, and a jump in q_{22} as shown in Fig. 5(b). Simulation result for the PWA-MILP2 approach is shown in Fig. 9(c), while numerical results for all approaches are presented in Table III. Overall, it can be inferred that the proposed hybrid control strategies are robust to different types of uncertainties in the urban network (reality). When comparing all modeling errors, we notice that the approaches are

most sensitive to measurement errors in the states since such an error changes the

more sensitive to measurement errors in the states since such an error changes the initial condition and subsequently, also the predicted state evolution in a significant way. However, note that the 10% error in the measurements, in the order of $0.1 \times 5000 = 500$ vehicles, is consistent with recent findings in the literature on the estimation of accumulations and MFDs [41], [42]. As can be inferred from Table III, the PWA-MILP1 approach has poor performance when combined MFD, measurement and demand noise exist in the simulation model (plant). The MINLP approach has impressive and robust performance under different conditions, only it suffers from high computation time for large-scale problems. Hence, for cases in which the computation time is not crucial, the multistart MINLP approach is suggested, while for larger problems, the three approximation methods can be chosen based on the structure of the network, the type and level of uncertainties exist, and the acceptable level of oscillation in the control inputs.

SECTION VII. Conclusion

Within the hierarchical multilevel approach for control of large-scale urban traffic networks, we introduced a new control scheme, the switching timing plans together with the perimeter controllers to manage and control a large-scale urban network.

The optimal control solutions are obtained in an MPC scheme for two different open-loop optimization problems: MINLP and MILP. The MILP problem is obtained after approximation of the nonlinear model using some techniques along with the PWA approximation. The results of the case studies show the importance of the approximated model regarding the required computation time. The computation times for solving the MILP problems were much lower than the MINLP for the two regions example. This is very crucial for real-time implementation in networks with a large number of regions, as the MINLP might not be tractable. Furthermore, it should be noted that the MILP results have some very small deviations (specially the second approach) compared with the MINLP results.

The effectiveness of hybrid (perimeter and switching) control has been compared quantitatively with perimeter control only, i.e., the results in [17]. It is apparent that the switching timing plans controllers can enhance the network performance when they collaborate with the perimeter controllers, as they can utilize more efficiently the network capacity to decrease the total time spent in the network. However, several research questions are still open in this direction, e.g., investigation of other approximation methods that might enhance the MILP approach. Formulating the control problem using the optimal control theories and numerical methods, such as solving a two point boundary value problem would also be interesting to investigate. Simulation of the closed-loop system using microsimulation software packages and field implementation of the proposed methods would shed more light on how these controllers can change the spatial distribution of congestion. In the hierarchical framework, lower level local controllers must be properly designed to realize the optimal control inputs determined by the high-level schemes. The effect of control decisions in the route choice of users is also a research direction. Monitoring techniques [40]–[41][42] for different types of sensors and penetration rates to decrease the measurement errors in the state variables and in the demands should be studied as well.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewer for suggesting the interesting approach mentioned in Remark 1. M. Hajiahmadi would like to thank the Urban Transport Systems Laboratory, EPFL, Lausanne for their great hospitality, excellent working conditions, and valuable collaboration. A significant

part of this work was performed when M. Hajiahmadi, J. Haddad, and N.

Authors

Figures

References

Citations

Keywords

Metrics

IEEE Account

Profile Information

Purchase Details

Need Help?

Other

A not-for-profit organization, IEEE is the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.
© Copyright 2018 IEEE - All rights reserved. Use of this web site signifies your agreement to the terms and conditions.

US & Canada: +1 800 678 4333
Worldwide: +1 732 981 0060

IEEE Account	Purchase Details	Profile Information	Need Help?
» Change Username/Password	» Payment Options	» Communications Preferences	» US & Canada: +1 800 678 4333
» Update Address	» Order History	» Profession and Education	» Worldwide: +1 732 981 0060
	» View Purchased Documents	» Technical Interests	» Contact & Support

About IEEE *Xplore* | Contact Us | Help | Accessibility | Terms of Use | Nondiscrimination Policy | Sitemap | Privacy & Opting Out of Cookies

A not-for-profit organization, IEEE is the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.
© Copyright 2018 IEEE - All rights reserved. Use of this web site signifies your agreement to the terms and conditions.