A Smoothing Method for Ramp Metering

Chuanye Gu, Changzhi Wu, Kok Lay Teo, *Life Senior Member, IEEE*, Yonghong Wu, and Song Wang

***Abstract*— Ramp metering offers great potential to mitigate traffic congestion and improve freeway management efficiency under traffic congestion conditions. This paper proposes an optimization program for freeway dynamic ramp metering based on Cell Transmission Model (CTM). This problem has been formulated as a discrete time optimal control problem with smooth state equations and constraints to meter traffic inflow from on-ramps. In the proposed model, the 'min' operators in the primal CTM are non-differentiable and thus, the corresponding optimal control problem cannot be solved directly using conventional gradient based methods. In this paper, we introduce a smooth approximation to approximate the 'min' operators and then a unified computational approach is developed to solve the problem. Theoretical analysis is carried out, showing that the optimal solution obtained from the approximated problem converges to the optimal solution of the primal CTM. Compared to the classical inequality relaxation method, our method can resolve the flow holding-back problem and reduce under fundamental diagram phenomenon. Compared with the Big-M method, our method has better efficiency. To achieve the desired traffic response control in real application, a series of online optimal control problems are solved using Model Predictive Control (MPC). Simulation studies show that our method can significantly improve freeway traffic management efficiency.**

***Index Terms*— Cell transmission model, ramp metering, optimization control, smooth approximation, Model Predictive Control.**

I. Introduction

TRAFFIC metering has great potential to improve urban network managements by avoiding the loss of capacity at the cost of delaying vehicles entering the congested areas of urban street networks [1], [2]. It can also be used to improve freeway operation by appropriately regulating/limiting the inflow from on-ramps to the freeway so as to keep the traffic state below the saturated flow [3], [4].

Queue-spill overs and gridlock decrease the capacity of freeway handling vehicles, thus wasting commuting time in congested traffic conditions, especially during peak hours. Ramp metering is an efficient way to protect congested areas of the freeway from oversaturated flow conditions. ALINEA, which is one of the first ramp metering controllers, is a feedback control of integral type [5]. Based on the ALINEA, a proportional integral regulator known as Pi-ALINEA was then proposed. Tracking a reference point for the traffic density (or occupancy) is the main purpose of these regulators. There are also some other rule-based control systems, for example coordinated control HERO [6], and others (see [7], [8]).

In addition to these simple regulators, many complex controllers have been designed based on optimization approaches. To alleviate freeway traffic congestion, different cost functions, such as maximising the system throughput or minimizing the total travel time spent by drivers in a traffic network system, are proposed. One of the most widely adopted optimization-based models in freeway control with ramp metering design is the macroscopic model, including the first order models (Cell Transmission Model (CTM) [9], [10]) and the second order models [11]. The second order models have obvious advantages because they can show the phenomenon of capacity drop. However, compared with first-order models, they contain more parameters that are required to be customized. To date, the constructed optimization problems based on these models have only been solved in the sense of local optimality [12], [13] due to the complexity of these optimization problems.

Various formulations and solution techniques have been proposed for the optimal ramp metering problems. The seminal research of optimization-based ramp metering can be traced back to [14], where a static model of traffic behavior was utilized for the formulation of the problem. This model has been extended in various forms in [15]. In particular, a deterministic ramp metering optimization program was proposed in [16] based on the Asymmetric Cell transmission model (ACTM), where it was demonstrated that the solution to the linear relaxation problem is feasible for freeway segments with only on-ramp and off-ramp junctions.

In [17], the Link-Node Cell Transmission Model (LN-CTM) was used to reformulate this problem as a linear program under relaxed piecewise-affine fundamental diagrams. Similarly, the 'min' operators were relaxed in [18] to obtain a feasible solutions by employing traffic demand control in the cell. Particulary, if the fundamental diagrams are assumed to be symmetric triangular, meaning that the backward wave speed equals to the free-flow speed of each cell, then by utilizing the priority control flow to enter the merge junctions, the solution of the relaxed problem is feasible. In [19], a novel first-order multi-lane macroscopic traffic flow model was first proposed

Manuscript received 22 January 2021; revised 19 July 2021 and 21 September 2021; accepted 15 October 2021. Date of publication 5 November 2021; date of current version 9 August 2022. This work was supported in part by the NSFC under Grant 11991020, in part by the Australia Research Council, and in part by the Western Australia (WA) Mainroads and the New South Wales (NSW) Road and Maritime Service, Australia. The Associate Editor for this article was S. Siri. *(Corresponding author: Changzhi Wu.)*

Chuanye Gu is with the School of Management, Guangzhou University, Guangzhou 510006, China, and also with the Department of Mathematics and Statistics, Curtin University, Perth, WA 6845, Australia (e-mail: chuanye.gu@postgrad.curtin.edu.au).

Changzhi Wu is with the School of Management, Guangzhou University, Guangzhou 510006, China (e-mail: changzhiwu@gzhu.edu.cn).

Kok Lay Teo is with the School of Mathematics, Sunway University, Subang Jaya, Selangor Darul Ehsan 47500, Malaysia (e-mail: k.l.teo@curtin.edu.au).

Yonghong Wu and Song Wang are with the Department of Mathematics and Statistics, Curtin University, Perth, WA 6845, Australia (e-mail: y.wu@curtin.edu.au; song.wang@curtin.edu.au).

Digital Object Identifier 10.1109/TITS.2021.3123815

for motorways, which considers the changes of lateral flow and longitudinal flow. In [20], the model proposed in [19] was formulated as a discrete time optimal control problem with linear relaxation through the use of ramp metering and variable speed limits. Recently, a traffic network finite-horizon optimal control model with exact linear relation for ramp metering controls and variable speed limits is proposed in [21], for which the distributed alternating direction method of multipliers (ADMM) is used to optimize the proposed model. A decentralized Model Predictive Control (MPC) approach is proposed in [22] for the freeway system on lossy communication networks under the mainline demand control. The priority parameter is used to ensure the flow of vehicles from both cells.

In [23], a Modified Cell Transmission Model (MCTM) is formulated as a Linear Complementarity System (lCS) which can be efficiently solved, thus successfully avoiding the hard nonlinearity caused by the 'min' operators. In [24], it is shown that if the objective is to minimize the total time spent and the turning rates are invariant, then through the use of ramp metering and partial mainline demand control, convex relaxation can accurately obtain the optimal solution of the original problem by introducing an alternative representation system.

Literature review shows that many ramp metering studies used CTM for freeway network loading. According to their approaches of reducing the complexity of the problem, the papers mentioned above can be divided into three categories: (i) The 'min' operators are directly replaced by a set of linear inequality constraints, where the effect of the flow holding-back problem is ignored; and (ii) The 'min' operators are relaxed through the use of inequality constraints under specific assumptions, such as symmetric triangular fundamental diagrams, or combined with traffic demand control; and (iii) The original 'min' operators are transformed equivalently to a series of linear inequality constraints using the Big-M method [1] through the introduction of binary variables. Although this transformation is equivalent, the transformed problem is computationally demanding because many auxiliary binary variables are being introduced, and hence it is not possible to rely on online computation. However, even for methods based on the equivalent transformations, they are specific methods and hard to calculate, and are hard to be extended to solve general problems. Based on the above discussion, these methods either cannot ensure the optimality of the solution obtained or lack generality.

In this paper, we will propose a new solution strategy based on a smooth approximation of the 'min' operators. This approximate model can resolve the holding back problem effectively through appropriately adjusting certain parameters and can be extended to solve general problems without special assumptions. Compared with Big-M method, our model is only required to solve a nonlinear optimal control problem without binary variables and thus it makes possible online computation for traffic response applications. The contributions of our study are as follows:

(i) Propose a smooth CTM-based ramp metering optimization model, where fundamental diagrams are clearly displayed to avoid flow holding-back problem.
(ii) Design a customized MPC solution method based on the co-state system such that the problem under consideration can be solved effectively.
(iii) The proposed solution approach can ensure optimality, convergence and feasibility.
(iv) Nonlinear cases can be solvable by our method because the development of our method does not depend on the linearity assumption on the cost function and the constraints. However, linear inequality relaxation based methods are not applicable.

The rest of the paper is organised as follows. Section II formulates and models the optimization problem of freeway ramp metering. In Section III, our approximate model is proposed and the main convergent results are established. Solution techniques are developed in Section IV. Sections V shows the numerical studies and Section VI gives the conclusion.

II. Optimization Problem

In this section, we will formulate the traffic flow dynamic as an optimal control problem to optimize the inflow from on-ramps to the freeway. Dynamic inflow from on-ramps are the main control variables over the study period. The aim is to improve freeway network performance by regulating the number of vehicles to enter the freeway. The problem is formulated and modeled based on the proposed CTM [9], [10]. The CTM is a numerical method developed based on the space-time discretization of the hydrodynamic traffic flow model [25]. See, for example, [26], [27]. In the formulation of CTM, each freeway is discretized into several homogenous segments, called subsections or cells, and the traffic flow is analysed in each cell through discretize time steps.

We summary in Table I the definitions of all the parameters, sets and decision variables used in this paper.

A. Primal Cell Transmission Model

The proposed cost function of the optimization problem is to optimize the total delay, including the mainline delay and the ramp delay. As shown in [28], this cost function is suitable for the oversaturation of the freeway system, and it tries to allow as many vehicles as possible to reach the destinations. The problem can be formally stated as follows:

$$\text{(PCTM)}\ \min_{r} D = \sum_{t=1}^{T}\sum_{i=1}^{I}(\rho_{i,t}\,\Delta x_i\,\Delta t - \frac{f_{i,t}\,\Delta x_i\,\Delta t}{v_i}) + \sum_{t=1}^{T}\sum_{j=1}^{J} q_{j,t}\,\Delta t, \tag{1}$$

$$\text{s.t. } \rho_{i,t+1} = \rho_{i,t} + \frac{\Delta t}{\Delta x_i} \times (f_{i-1,t} - f_{i,t} + r_{i,t} - s_{i,t}), \quad \forall i = 1,\ldots,I, \forall t, \tag{2}$$

$$f_{i,t} = \min\{v_i\rho_{i,t}, C_i, C_{i+1}, w_{i+1}(\rho_{max,i+1} - \rho_{i+1,t})\}, \quad \forall i, t, \tag{3}$$

$$q_{j,t+1} = q_{j,t} + \Delta t(d_{j,t} - r_{j,t}), \quad \forall j, t, \tag{4}$$

TABLE I
MODELS' VARIABLES AND PARAMETERS

Parameters	
Δx_i	length of cell i
Δt	size of the simulation time step t
C_i	capacity of cell i
v_i	free-speed at cell i
$\rho_{max,i}$	jam density at cell i
w_i	backward wave speed at cell i
$q_{max,j}$	maximum queue length at on-ramp j
$r_{max,j}$	maximum ramp metering at on-ramp j
ϵ	parameter for adjusting approximate error
Sets	
T	set of all time steps
I	set of all cells
J	set of all on-ramps
F	feasible set of Problem PCTM
F^ϵ	feasible set of Problem APCTM
V	convex and compact subset of $\mathbb{R}^J$
$\mathcal{U}$	set of admission controls
Control variable	
$r_{j,t}$	ramp metering at on-ramp j and time step t
Variables	
$\rho_{i,t}$	density at cell i and time step t
$f_{i,t}$	flow at cell i and time step t
$d_{j,t}$	demand flow to enter freeway from on-ramp j at time step t
$q_{j,t}$	queue length waiting for on-ramp j at time step t

$$0 \le \rho_{i,t} \le \rho_{max,i}, \quad \forall i,t, \tag{5}$$
$$0 \le q_{j,t} \le q_{max,j}, \quad \forall j,t, \tag{6}$$
$$0 \le r_{j,t} \le r_{max,j}, \quad \forall j,t. \tag{7}$$

The dynamic of the density in each cell i at time step t is governed by (2). The outflow from the cell i during the time step t is controlled by a piecewise linear fundamental diagram described in (3). The dynamic of the queue length for each on-ramp j at time step t is given by (4). The upper and lower bounds for the density $\rho_{i,t}$, queue length $q_{j,t}$ and ramp-metering $r_{j,t}$ are given by (5)-(7), respectively.

Let $V = \{r = [r_1, r_2, \ldots, r_J] \in \mathbb{R}^J : 0 \le r_j \le r_{max,j}, \forall j\}$, where $r_{max,j}$ is a given constant. It is noted that V is a compact and convex subset of $\mathbb{R}^J$. Let r be a control sequence $\{r_t : t = 1, \ldots, T-1\}$ in V. Then, r is called an admissible control. We use $\mathcal{U}$ to denote the class of all such admissible controls. For a control r in $\mathcal{U}$, if it satisfies the constraints (5)-(7), then it is called a feasible control sequence. Let F be the class of all such feasible controls.

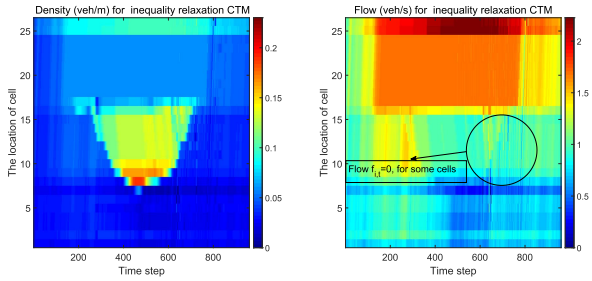
Note that the constraint (3) is equivalent to the following constraint:

$$f_{i,t} = \min\{\min\{v_i\rho_{i,t}, C_i\}, \min\{C_{i+1}, w_{i+1}(\rho_{max,i+1} - \rho_{i+1,t})\}\}, \forall i,t, \tag{8}$$

where $f^D_{i,t} = \min\{v_i\rho_{i,t}, C_i\}$ and $f^S_{i+1,t} = \min\{C_{i+1}, w_{i+1}(\rho_{max,i+1} - \rho_{i+1,t})\}$ represent the demand function and supply function, respectively.

B. Existing Solution Methods

1) Linear Inequality Relaxation Based Method: Due to the 'min' operators in constraint (3), the optimization problem PCTM is difficult to be solved directly. To overcome this difficulty, the 'min' operators are relaxed to the following inequality constraints in [16]:



(a) Density (b) Flow

Fig. 1. Density and flow for inequality relaxation.

$$f_{i,t} \le v_i\rho_{i,t}, \quad \forall i,t, \tag{9}$$
$$f_{i,t} \le C_i, \quad \forall i,t, \tag{10}$$
$$f_{i,t} \le C_{i+1}, \quad \forall i,t, \tag{11}$$
$$f_{i,t} \le w_{i+1}(\rho_{max,i+1} - \rho_{i+1,t}), \quad \forall i,t. \tag{12}$$

In general, the optimal value of $f_{i,t}$ obtained for solving Optimal Control Problem PCTM with constraint (3) being relaxed to constraints (9)-(12) is strictly less than the right-hand sides of all the constraints (9)-(12). This problem is known as the flow holding-back problem. In [16], it was shown that for ramp metering control problems, the traffic flow states caused by the inequality relaxation constraints are lower than the fundamental diagram of the CTM. In fact, it is observed that for some cases in our experiments, the traffic flow values of some cells are even zero due to the flow holding-back problem as shown in Fig. 1. In this paper, we will introduce smoothing approximation of the 'min' operator to resolve the holding-back problem (or under the fundamental diagram problem) approximately.

Under some strong assumptions, such as symmetric triangular fundamental diagrams, or using the traffic demand control in each cell, the 'min' operators can be relaxed equivalently to the inequality constraints (9)-(12), see [17], [18], [21], [22]. However, it is doubtful whether the assumption of demand control in every cell is realistic. Thus, this method is not a general problem method.

2) Big-M Method: Another approach is to convert the 'min' operators in constraint (3) into equivalent standard inequality constraints through the introduction of binary variables. This method is known as Big-M method [1]. In fact, constraint (3) can be written equivalently as the following constraints:

$$f_{i,t} \le v_i\rho_{i,t}, \quad \forall i,t, \tag{13}$$
$$f_{i,t} \le C_i, \quad \forall i,t, \tag{14}$$
$$f_{i,t} \le C_{i+1}, \quad \forall i,t, \tag{15}$$
$$f_{i,t} \le w_{i+1}(\rho_{max,i+1} - \rho_{i+1,t}), \quad \forall i,t, \tag{16}$$
$$f_{i,t} \ge v_i\rho_{i,t} - M(1-\alpha_{i,t}), \quad \forall i,t, \tag{17}$$
$$f_{i,t} \ge C_i - M(1-\beta_{i,t}), \quad \forall i,t, \tag{18}$$
$$f_{i,t} \ge C_{i+1} - M(1-\gamma_{i,t}), \quad \forall i,t, \tag{19}$$
$$f_{i,t} \ge w_{i+1}(\rho_{max,i+1} - \rho_{i+1,t}) - M(1-\chi_{i,t}), \quad \forall i,t, \tag{20}$$

$$\alpha_{i,t} + \beta_{i,t} + \gamma_{i,t} + \chi_{i,t} = 1, \quad \forall i, t, \tag{21}$$

$$\alpha_{i,t} \in \{0, 1\}, \beta_{i,t} \in \{0, 1\}, \gamma_{i,t} \in \{0, 1\}, \chi_{i,t} \in \{0, 1\}. \tag{22}$$

Note that after the reformation of the constraints, the resulting optimization problem can be solved only for small-size networks using traditional optimal algorithms. Due to the introduction of a large number of binary variables, it is not feasible to be solved even for medium-sized networks. In this paper, using the structure of the model, an effective solution approach is designed to avoid the flow-holding back problem and it is applicable for large-size networks. Details are given in the next section.

III. Approximate Model

A. Min Approximation

It is difficult to solve Problem PCTM directly because the flow $f_{i,t}$ between cells is required to equal to the minimum value in constraint (3). However, as discussed before, constraint (3) is the key to eliminating internal metering and removing inflow holding-back. To resolve this problem, we propose a novel smoothing approach to approximate the 'min' operators in constraint (3) and show that the solution obtained from the approximate problem converges to the solution of Problem PCTM.

The main technical challenge of constraint (3) is how to deal with its non-smoothness. Based on (8), we can equivalently represent the demand function as $f^D_{i,t} = \min\{v_i\rho_{i,t}, C_i\} = \frac{1}{2}(v_i\rho_{i,t} + C_i - |v_i\rho_{i,t} - C_i|)$, and the supply function as $f^S_{i+1,t} = \min\{C_{i+1}, w_{i+1}(\rho_{max,i+1} - \rho_{i+1,t})\} = \frac{1}{2}(w_{i+1}(\rho_{max,i+1} - \rho_{i+1,t}) + C_{i+1} - |w_{i+1}(\rho_{max,i+1} - \rho_{i+1,t}) - C_{i+1}|)$. Let $\epsilon > 0$ be a small number, we use $f^{D\epsilon}_{i,t}$ and $f^{S\epsilon}_{i+1,t}$ to approximate $f^D_{i,t}$ and $f^S_{i+1,t}$, respectively, i.e.,

$$f^{D\epsilon}_{i,t} = \frac{1}{2}[v_i\rho_{i,t} + C_i - \sqrt{(v_i\rho_{i,t} - C_i)^2 + \frac{\epsilon^2}{4}}], \tag{23}$$

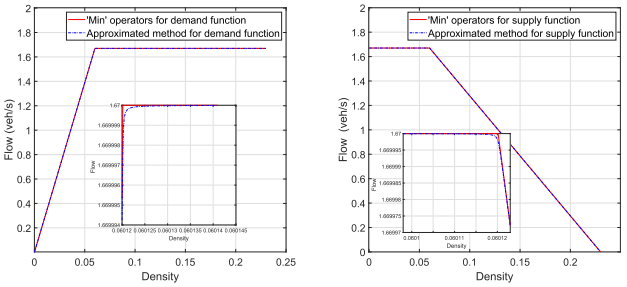
and

$$f^{S\epsilon}_{i+1,t} = \frac{1}{2}[w_{i+1}(\rho_{max,i+1} - \rho_{i+1,t}) + C_{i+1} - \sqrt{(w_{i+1}(\rho_{max,i+1} - \rho_{i+1,t}) - C_{i+1})^2 + \frac{\epsilon^2}{4}}]. \tag{24}$$

It is obvious that the functions $f^{D\epsilon}_{i,t}$ and $f^{S\epsilon}_{i+1,t}$ are smooth and differentiable. Fig. 2 shows the demand function and supply function for the 'min' and the approximate operators with $\epsilon = 10^{-5}$. From Fig. 2, we can observe that the approximate functions not only retain the main characteristics of the 'min' operators, such as the trend, but also ensure differentiability at the inflection point.

Before carrying out further analysis, we estimate the bounds of $f^D_{i,t} - f^{D\epsilon}_{i,t}$ and $f^S_{i+1,t} - f^{S\epsilon}_{i+1,t}$. By the definitions of $f^{D\epsilon}_{i,t}$ and $f^{S\epsilon}_{i+1,t}$, it holds that

$$0 \le f^D_{i,t} - f^{D\epsilon}_{i,t} \le \frac{\epsilon}{4}, \tag{25}$$



(a) Demand function (b) Supply function

Fig. 2. Demand and supply function for 'min' and approximate operators.

Similarly, we can show that

$$0 \le f^S_{i+1,t} - f^{S\epsilon}_{i+1,t} \le \frac{\epsilon}{4}. \tag{26}$$

By virtue of relation (8), we have

$$f_{i,t} = \frac{1}{2}(f^D_{i,t} + f^S_{i+1,t} - |f^D_{i,t} - f^S_{i+1,t}|). \tag{27}$$

Similarly, let

$$f^{\epsilon\epsilon}_{i,t} = \frac{1}{2}(f^{D\epsilon}_{i,t} + f^{S\epsilon}_{i+1,t} - |f^{D\epsilon}_{i,t} - f^{S\epsilon}_{i+1,t}|). \tag{28}$$

Using the same approximation method, we obtain

$$f^{\epsilon}_{i,t} = \frac{1}{2}(f^{D\epsilon}_{i,t} + f^{S\epsilon}_{i+1,t} - \sqrt{(f^{D\epsilon}_{i,t} - f^{S\epsilon}_{i+1,t})^2 + \frac{\epsilon^2}{4}}). \tag{29}$$

Based on inequalities (28) and (29), we have

$$0 \le f^{\epsilon\epsilon}_{i,t} - f^{\epsilon}_{i,t} \le \frac{\epsilon}{4}. \tag{30}$$

B. Approximate Model

Now, the approximate problem of Problem PCTM may be stated as given below:

$$\text{(APCTM)} \min_r D^\epsilon = \sum_{t=1}^{T}\sum_{i=1}^{I}(\rho^\epsilon_{i,t}\Delta x_i\Delta t - \frac{f^\epsilon_{i,t}(\rho^\epsilon_{i,t})\Delta x_i\Delta t}{v_i}) + \sum_{t=1}^{T}\sum_{j=1}^{J} q_{j,t}\Delta t, \tag{31}$$

$$\text{s.t. } \rho^\epsilon_{i,t+1} = \rho^\epsilon_{i,t} + \frac{\Delta t}{\Delta x_i} \times (f^\epsilon_{i-1,t}(\rho^\epsilon_{i-1,t}) - f^\epsilon_{i,t}(\rho^\epsilon_{i,t}) + r_{i,t} - s_{i,t}), \quad \forall i, t, \tag{32}$$

$$0 \le \rho^\epsilon_{i,t} \le \rho_{max,i}, \quad \forall i, t, \tag{33}$$

and constraints (4), (6), (7), (23), (24) and (29).

For a control r in $\mathcal{U}$, if it satisfies the constraints (6), (7) and (33), then it is called a feasible control sequence. Let F^ϵ be the class of all such feasible controls. Clearly, Problem APCTM is a smooth discrete time optimal control problem which can be efficiently solved using traditional optimization approaches, such as sequential quadratic programming algorithm (SQP) [29].

Theorem 1 below shows that the solution of our approximate model APCTM will converge to the solution of Problem PCTM through appropriately controlling the parameter ϵ.

Theorem 1: Under the conditions of Lemma 2, $\overline{r}$ is an optimal control vector of Problem PCTM.

Proof: By induction, it follows from Lemma 2 (see Appendix) and the differentiability of $q_{j,t}$ that, for each $j = 1, \ldots, J, t = 1, \ldots, T$,

$$\lim_{\epsilon \to 0} q_{j,t}^{*\epsilon} = \overline{q}_{j,t}, \tag{34}$$

and

$$\lim_{\epsilon \to 0} \overline{q}_{j,t}^{\epsilon} = q_{j,t}^{*}. \tag{35}$$

Based on Lemma 1 (see Appendix), we have, for each i, t,

$$\lim_{\epsilon \to 0} f_{i,t}^{\epsilon}(\overline{\rho}_{i,t}) = \overline{f}_{i,t}, \tag{36}$$

and

$$\lim_{\epsilon \to 0} f_{i,t}^{\epsilon}(\rho_{i,t}^{*}) = f_{i,t}^{*}. \tag{37}$$

By Lemma 2 (see Appendix), $\lim_{\epsilon \to 0} \rho_{i,t}^{*\epsilon} = \overline{\rho}_{i,t}$ and $\lim_{\epsilon \to 0} \overline{\rho}_{i,t}^{\epsilon} = \rho_{i,t}^{*}$ for each i and t. Since $f_{i,t}^{\epsilon}$ is continuously differentiable with respect to each of the components of ρ and r, we have

$$\lim_{\epsilon \to 0} f_{i,t}^{\epsilon}(\rho_{i,t}^{*\epsilon}) = \lim_{\epsilon \to 0} f_{i,t}^{\epsilon}(\overline{\rho}_{i,t}). \tag{38}$$

and

$$\lim_{\epsilon \to 0} f_{i,t}^{\epsilon}(\overline{\rho}_{i,t}^{\epsilon}) = \lim_{\epsilon \to 0} f_{i,t}^{\epsilon}(\rho_{i,t}^{*}). \tag{39}$$

Combining (36) and (38) yields

$$\lim_{\epsilon \to 0} f_{i,t}^{\epsilon}(\rho_{i,t}^{*\epsilon}) = \overline{f}_{i,t}. \tag{40}$$

By (37) and (39), we have

$$\lim_{\epsilon \to 0} f_{i,t}^{\epsilon}(\overline{\rho}_{i,t}^{\epsilon}) = f_{i,t}^{*}. \tag{41}$$

By virtue of (73) of Lemma 2 (see Appendix), (34) and (40), we obtain

$$\lim_{\epsilon \to 0} D^{*\epsilon}(\rho^{*\epsilon}, r^{*\epsilon}, q^{*\epsilon}) = D(\overline{\rho}, \overline{r}, \overline{q}), \tag{42}$$

where $D^{*\epsilon}(\rho^{*\epsilon}, r^{*\epsilon}, q^{*\epsilon})$ is the optimal function value of Problem APCTM and $D(\overline{\rho}, \overline{r}, \overline{q})$ is the associated objective function value of Problem PCTM. Similarly, by (75) of Lemma 2 (see Appendix), (35) and (41), we obtain

$$\lim_{\epsilon \to 0} D^{\epsilon}(\overline{\rho}^{\epsilon}, \overline{r}^{\epsilon}, \overline{q}^{\epsilon}) = D^{*}(\rho^{*}, r^{*}, q^{*}), \tag{43}$$

where $D^{\epsilon}(\overline{\rho}^{\epsilon}, \overline{r}^{\epsilon}, \overline{q}^{\epsilon})$ is the associated objective function value of Problem APCTM and $D^{*}(\rho^{*}, r^{*}, q^{*})$ is the optimal function value of Problem PCTM.

For any $\overline{r} \in F$ and $\overline{r}^{\epsilon} \in F^{\epsilon}$, we have

$$D^{*}(\rho^{*}, r^{*}, q^{*}) \leq D(\overline{\rho}, \overline{r}, \overline{q}), \tag{44}$$

and

$$\lim_{\epsilon \to 0} D^{*\epsilon}(\rho^{*\epsilon}, r^{*\epsilon}, q^{*\epsilon}) \leq \lim_{\epsilon \to 0} D^{\epsilon}(\overline{\rho}^{\epsilon}, \overline{r}^{\epsilon}, \overline{q}^{\epsilon}). \tag{45}$$

Combining (42), (43), (44) and (45), we obtain

$$\begin{aligned} D^{*}(\rho^{*}, r^{*}, q^{*}) &\leq D(\overline{\rho}, \overline{r}, \overline{q}) = \lim_{\epsilon \to 0} D^{*\epsilon}(\rho^{*\epsilon}, r^{*\epsilon}, q^{*\epsilon}) \\ &\leq \lim_{\epsilon \to 0} D^{\epsilon}(\overline{\rho}^{\epsilon}, \overline{r}^{\epsilon}, \overline{q}^{\epsilon}) = D^{*}(\rho^{*}, r^{*}, q^{*}). \end{aligned}$$

Thus,

$$D(\overline{\rho}, \overline{r}, \overline{q}) = D^{*}(\rho^{*}, r^{*}, q^{*}).$$

This completes the proof. ∎

Problem APCTM is a smooth discrete time optimal control problem which can be solved efficiently using gradient-based optimization methods. The required gradient formulas for the objective and constraint functions will be derived in the next section.

IV. Gradient Formulas

To solve Problem APCTM, we need the gradients of the objective and the constraint functions with respect to the variables r.

A. Gradient Formulas

Let $y_t = \{\rho_{1,t}^{\epsilon}, \rho_{2,t}^{\epsilon}, \ldots, \rho_{I,t}^{\epsilon}, q_{1,t}, q_{2,t}, \ldots, q_{J,t}\}^{\top} \in \mathbb{R}^{I+J}$ and $r_t = \{r_{1,t}, r_{2,t} \ldots, r_{J,t}\}^{\top} \in \mathbb{R}^{J}$ be the state and control vectors, respectively. Then, for any $t = 1, 2, \ldots, T-1$, the state equations of (4) and (32) are written in compact form as given below:

$$\begin{aligned} y_{t+1} = &\begin{bmatrix} \rho_{1,t}^{\epsilon} \\ \rho_{2,t}^{\epsilon} \\ \vdots \\ \rho_{I,t}^{\epsilon} \\ q_{1,t} \\ q_{2,t} \\ \vdots \\ q_{J,t} \end{bmatrix} + \begin{bmatrix} \frac{\Delta t}{\Delta x_1} \cdot r_{0,t} \\ \frac{\Delta t}{\Delta x_2} \cdot f_{1,t}^{\epsilon} \\ \vdots \\ \frac{\Delta t}{\Delta x_I} \cdot f_{I-1,t}^{\epsilon} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \begin{bmatrix} \frac{\Delta t}{\Delta x_1} \cdot f_{1,t}^{\epsilon} \\ \frac{\Delta t}{\Delta x_2} \cdot f_{2,t}^{\epsilon} \\ \vdots \\ \frac{\Delta t}{\Delta x_I} \cdot f_{I,t}^{\epsilon} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ &+ \begin{bmatrix} \frac{\Delta t}{\Delta x_1} \cdot r_{1,t} \\ \frac{\Delta t}{\Delta x_2} \cdot r_{2,t} \\ \vdots \\ \frac{\Delta t}{\Delta x_I} \cdot r_{I,t} \\ -\Delta t \cdot r_{1,t} \\ -\Delta t \cdot r_{2,t} \\ \vdots \\ -\Delta t \cdot r_{J,t} \end{bmatrix} + \begin{bmatrix} -\frac{\Delta t}{\Delta x_1} \cdot s_{1,t} \\ -\frac{\Delta t}{\Delta x_2} \cdot s_{2,t} \\ \vdots \\ -\frac{\Delta t}{\Delta x_I} \cdot s_{I,t} \\ \Delta t \cdot d_{1,t} \\ \Delta t \cdot d_{2,t} \\ \vdots \\ \Delta t \cdot d_{J,t} \end{bmatrix}. \end{aligned} \tag{46}$$

Let $F(t, y_t, r_t)$ denote the right hand side of the difference equation (46). It contains state variables, control variables and time. The initial condition for the system of difference equations is

$$y_1 = [\rho_{1,1}, \ldots, \rho_{I,1}, q_{1,1}, \ldots, q_{J,1}]^{\top} \in \mathbb{R}^{I+J}. \tag{47}$$

We now consider the following class of discrete time optimal control problems in canonical formulation. Let

$$\Phi_0(y_T(r)) = \sum_{i=1}^{I}(\rho_{i,T}^{\epsilon}\Delta t\Delta x_i - \frac{f_{i,T}^{\epsilon}\Delta t\Delta x_i}{v_i}) + \sum_{j=1}^{J} q_{j,T}\Delta t,$$

$$L_0(y_t(r)) = \sum_{i=1}^{I}(\rho_{i,t}^{\epsilon}\Delta t\Delta x_i - \frac{f_{i,t}^{\epsilon}\Delta t\Delta x_i}{v_i} + \sum_{j=1}^{J} q_{j,t}\Delta t.$$

Then,

$$g_0(r) = \Phi_0(y_T(r)) + \sum_{t=1}^{T-1} L_0(t, y_t(r)),$$

where $g_0(r)$ is the objective function which is to be optimized subject to $r \in \mathcal{U}$. Similarly, we rewrite the constraints as follows:

$$\begin{aligned} g_l(r) &= \Phi_l(y_T(r)) + \sum_{t=1}^{T-1} L_l(t, y_t(r), r_t) \\ &= -\rho_{i,t}^{\epsilon}, \quad 1 \le l \le IT, \\ &\quad i = 1\ldots, I,\ t = 1, \ldots, T, \\ g_l(r) &= \Phi_l(y_T(r)) + \sum_{t=1}^{T-1} L_l(t, y_t(r), r_t) \\ &= \rho_{i,t}^{\epsilon} - \rho_{max,i},\ IT + 1 \le l \le 2IT, \\ &\quad i = 1\ldots, I,\ t = 1, \ldots, T, \\ g_l(r) &= \Phi_l(y_T(r)) + \sum_{t=1}^{T-1} L_l(t, y_t(r), r_t) \\ &= -q_{j,t},\ 2IT + 1 \le l \le (2I + J)T, \\ &\quad j = 1\ldots, J,\ t = 1, \ldots, T, \\ g_l(r) &= \Phi_l(y_T(r)) + \sum_{t=1}^{T-1} L_l(t, y_t(r), r_t) \\ &= q_{j,t} - q_{max,i},\ (2I + J)T + 1 \le l \le 2(I + J)T, \\ &\quad j = 1\ldots, J,\ t = 1, \ldots, T. \end{aligned}$$

These constraint functions are said to be in canonical form, because they are in the same form as the objective function. Now, we can derive the gradient formulas of the objective and constraint functions in a unified way. Define

$$r = [(r_1)^\top, (r_2)^\top, \ldots, (r_{T-1})^\top]^\top.$$

Let the control vector r be perturbed by $\xi\widehat{r}$, where $\xi > 0$ is a small constant and $\widehat{r}$ is an arbitrary but fixed perturbation of r given by

$$\widehat{r} = [(\widehat{r}_1)^\top, (\widehat{r}_2)^\top, \ldots, (\widehat{r}_{T-1})^\top]^\top.$$

Then, we have

$$r(\xi) = r + \xi\widehat{r} = [(r_1(\xi))^\top, (r_2(\xi))^\top, \ldots, (r_{T-1}(\xi))^\top]^\top,$$

where

$$r_t(\xi) = r_t + \xi\widehat{r}(\xi), t = 1, \ldots, T-1.$$

Consequently, the state of the system will be perturbed, and so are the objective and constraint functions.

Define

$$y_t(\xi) = y_t(r(\xi)), t = 2, \ldots, T. \tag{48}$$

Then,

$$y_{t+1}(\xi) = F(t, y_t(\xi), r_t(\xi)). \tag{49}$$

The variation of the state for $t = 1, 2, \ldots, T-1$ is:

$$\Delta y_{t+1} = \frac{\partial F(t, y_t, r_t)}{\partial y_t}\Delta y_t + \frac{\partial F(t, y_t, r_t)}{\partial r_t}\widehat{r}_t \tag{50}$$

with

$$\Delta y_1 = 0. \tag{51}$$

For the $l-$th function ($l = 0$ is the objective function), we have

$$\frac{\partial g_l(r)}{\partial r}\widehat{r} = \frac{\partial \Phi_l(y_T)}{\partial y_T}\Delta y_T + \sum_{t=1}^{T-1}[\frac{\partial L_l(t, y_t, r_t)}{\partial y_t}\Delta y_t + \frac{\partial L_l(t, y_t, r_t)}{\partial r_t}\widehat{r}_t]. \tag{52}$$

For each $l = 0, 1, \ldots, 2(I + J)T$, define the Hamiltonian

$$H_l(t, y_t, r_t, \lambda_{t+1}^l) = L_l(t, y_t, r_t) + (\lambda_{t+1}^l)^\top F(t, y_t, r_t),$$

where $\lambda_t^l \in \mathbb{R}^{I+J}, t = T, T-1, \ldots, 2$, denotes the co-state sequence for the $l-$th canonical constraint. Then, it follows from (52) that

$$\begin{aligned} \frac{\partial g_l(r)}{\partial r}\widehat{r} = \frac{\partial \Phi_l(y_T)}{\partial y_T}\Delta y_T + \sum_{t=1}^{T-1}\{&\frac{\partial H_l(t, y_t, r_t, \lambda_{t+1}^l)}{\partial y_t}\Delta y_t \\ &- (\lambda_{t+1}^l)^\top\frac{\partial F(t, y_t, r_t)}{\partial y_t}\Delta y_t + \frac{\partial H_l(t, y_t, r_t, \lambda_{t+1}^l)}{\partial r_t}\widehat{r}_t \\ &- (\lambda_{t+1}^l)^\top\frac{\partial F(t, y_t, r_t)}{\partial r_t}\widehat{r}_t\}. \end{aligned} \tag{53}$$

Based on (50) and (51), we have

$$\Delta y_{t+1} = \frac{\partial F(t, y_t, r_t)}{\partial y_t}\Delta y_t + \frac{\partial F(t, y_t, r_t)}{\partial r_t}\widehat{r}_t. \tag{54}$$

Let the co-state λ_t^l be determined by the following system of difference equations:

$$(\lambda_t^l)^\top = \frac{\partial H_l(t, y_t, r_t, \lambda_{t+1}^l)}{\partial y_t}, t = T-1, T-2, \ldots, 2, \tag{55}$$

and

$$(\lambda_T^l)^\top = \frac{\partial \Phi_l(y_T)}{\partial y_T}. \tag{56}$$

By virtue of (53), (54), (55), (56), (48) and (49), we obtain

$$\frac{\partial g_l(r)}{\partial r}\widehat{r} = [\frac{\partial H_l(1, y_1, r_1, \lambda_2^l)}{\partial r_1}, \ldots, \frac{\partial H_l(T-1, y_{T-1}, r_{T-1}, \lambda_T^l)}{\partial r_{T-1}}]\widehat{r}.$$

Because $\widehat{r}$ is arbitrary, we have the following gradient formula:

$$\frac{\partial g_l(r)}{\partial r} = [\frac{\partial H_l(1, y_1, r_1, \lambda_2^l)}{\partial r_1}, \ldots, \frac{\partial H_l(T-1, y_{T-1}, r_{T-1}, \lambda_T^l)}{\partial r_{T-1}}]. \tag{57}$$

Now we summarize the gradient computation in the following theorem.

Theorem 2: Consider Problem APCTM. *Then, for each* $l = 0, 1, \ldots, 2(I + J)T$, *the gradient of* $g_l(r)$ *with respect to control vector* r *is given by (57), where* $r = [(r_1)^\top, (r_2)^\top, \ldots, (r_{T-1})^\top]^\top$.

B. Algorithm

Problem APCTM is essentially a nonlinear mathematical programming problem where the decision vector is the control vector r. Many gradient-based optimization methods, such as sequential quadratic programming (SQP), can be used to solve it. To apply gradient-based optimization methods, for each $r \in V$, we need the values of the objective function $g_0(r)$ and the constraint functions $g_l(r)$, $l = 1, 2, \ldots, 2(I + J)T$, together with their corresponding gradients. Detailed computation is given in Algorithm 1.

Algorithm 1 : Algorithm to Compute Gradients of Objective Function and Constraints

1: Initialization: For a given $r \in V$ and initial condition (47).
2: Output: Compute the solution $y_{t+1}(r)$, $t = 1, 2, \ldots, T-1$ of system (46) forward in time from $t = 1$ to $t = T - 1$.
3: **for** $l = 0$ to $2(I + J)T$ **do**
4: Compute $\Phi_l(y_T(r))$.
5: **for** $t = 1$ to $T - 1$ **do**
6: Compute $L_l(t, y_t(r), r_t)$.
7: **end for**
8: Compute $\sum_{t=1}^{T-1} L_l(t, y_t(r), r_t)$.
9: Compute $g_l(r) = \Phi_l(y_T(r)) + \sum_{t=1}^{T-1} L_l(t, y_t(r), r_t)$.
10: **for** $t = T$ to 2 **do**
11: Solve the system of the co-state system (55) and (56) backward in time. Let $\lambda_t^l(r)$ be the solution obtained.
12: **end for**
13: Calculate the gradients of g_l using (57).
14: **end for**

Step 2 in Algorithm 1 is to compute the values of $y_t(r)$ corresponding to each given r. Then, $g_l(r)$ is computed in Step 9 based on Step 4 and Step 8. After that, the co-state system (55) and (56) is solved backward in time from $t = T$ to $t = 2$ to acquire $\lambda_t^l(r)$. Finally, the gradients of the objective function and the constraint functions are calculated using the gradient formulas given in Theorem 2.

V. Numerical Studies

In this section, numerical performance of a ramp metering method based on the proposed model is given. A Model Predictive Control (MPC) approach is utilized to achieve traffic response control.

The MPC approach has been widely utilized in freeway traffic control problems, see [30]–[32]. Based on the current state of the traffic system, the MPC approach utilizes a traffic model to predict dynamic of the state, and finds an optimal control signal which gives the optimal value of the objective function. This property guarantees that the controller can take advantage of potentially larger future gains at a current (smaller) cost, thereby avoiding short-sighted control action.

After optimization, the values for the control variables of the first sample of the optimal control action are applied to the process. The remaining control signals are recalculated in a finite rolling horizon scheme. Readers can refer to [33] for a detailed description of the MPC method.

A. MPC Design

In this section, we first redefine the MPC objective function, which minimizes the total delay of the freeway system, including mainline delay and on-ramp delay for $l = t, \ldots, t + N_p - 1$. It is similar to the cost function defined in (31). The optimization problem is reformulated as follows:

$$D_t = \sum_{l=t}^{t+N_p-1} [\sum_{i=1}^{I} (\rho_{i,l} \Delta x_i \Delta t - \frac{f_{i,l} \Delta x_i \Delta t}{v_i}) + \sum_{j=1}^{J} q_{j,l} \Delta t]. \quad (58)$$

Note that the tuning rules used to select the appropriate value of the prediction horizon N_p and control sample time N_c are very important to the performance of MPC [3]. Normally, the value of N_p should be larger than the typical travel time from the controlled segment to the exit of the network. The reason is that if the prediction horizon N_p is shorter than the typical travel time, the vehicles affected by the current control action have no effect on the network operation before exiting. On the other hand, N_p should not be too large due to the computational complexity of the MPC optimization problem. Hence, we choose N_p as the typical time in the network based on this reasoning. For the control sample time N_c, we will choose a value that represents the trade-off between performance and computational effort.

Performance of the following models are compared:

- Primal cell transmission model (PCTM)
- Linear inequality relaxation based method (LICTM)
- Big-M method (BMM)
- Approximate model (APCTM)
- Generalized non-holding back linear programming formulation (GNHBLP) proposed by Zhu [34] to address holding back problems.

All the experiments were ran on a computer with Intel(R) Core(TM) i7-8565U CPU-1.80GHz 1.99 GHz and RAM 16GB. We use the MATLAB implementation of the SQP algorithm (fmincon) to solve the models PCTM, LICTM, APCTM and GNHBLP, and use the GUROBI [35] to solve the model BMM.

B. Example 1

1) Scenario: The example is selected from the Kwinana Freeway in the vicinity of Perth in Australia. This section of freeway is divided into $I = 26$ cells with $J = 8$ on-ramps and 4 off-ramps. Each segment has a longitude of $\Delta x_i = 500$m for any $i = 1, \ldots, I$. Cells $i = 1, 3, 4, 6, \ldots, 24$ have 3 lanes and cells $i = 2, 5, 25, 26$ have 4 lanes. Fig. 3 shows the details of this example.

There are eight control signals: ramp metering in cells 2, 5, 8, 9, 10, 16, 17, 25, respectively. We measured density,

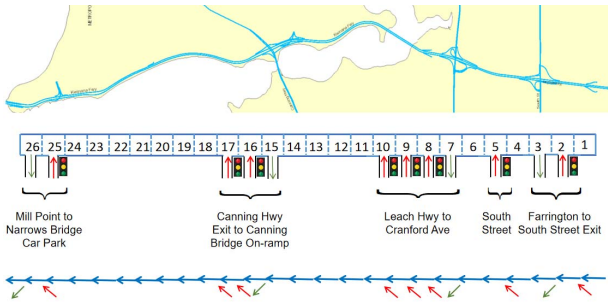
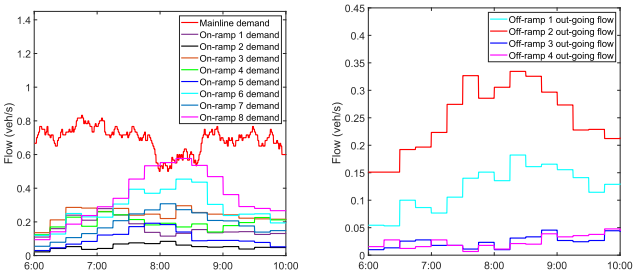


Fig. 3. Road graph.

TABLE II
MODEL PARAMETERS

Name of parameter	Value
$v_i, i = 1, \ldots, 24$	27.7778m/s
$v_i, i = 25, 26$	22.2222m/s
$w_i, i = 1, 3, 4, 6, \ldots, 24$	9.8029m/s
$w_i, i = 2, 5$	9.7895m/s
$w_i, i = 25, 26$	10.783m/s
$C_i, i = 1, 3, 4, 6, \ldots, 24$	1.6667veh/s
$C_i, i = 2, 5, 25, 26$	2.2222veh/s
$q_{max,j}, j = 7$	120veh
$q_{max,j}, j = 1, \ldots, 6, 8$	60veh
$r_{max,j}, j = 1, \ldots, 8$	0.55veh/s
$\rho_{max,i}, i = 1, 3, 4, 6, \ldots, 24$	0.23veh/m
$\rho_{max,i}, i = 2, 5, 25, 26$	0.3067veh/m
Δt	15s
ϵ	0.00000001

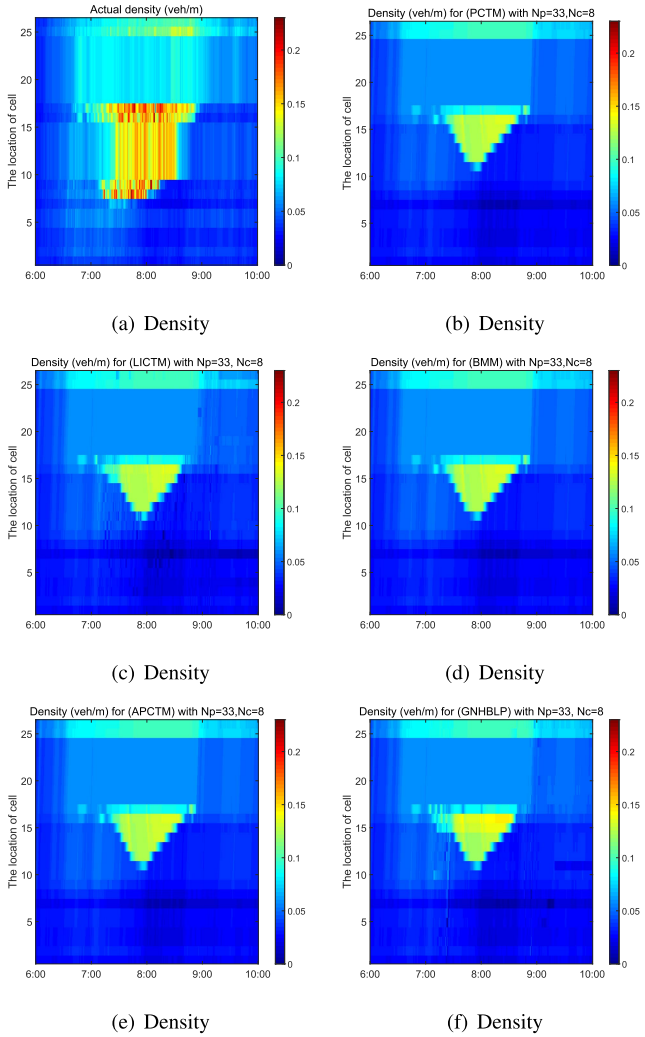


(a) Mainline demand and on-ramp demands (b) Out-going flow

Fig. 4. Mainline demand, on-ramp demands and out-going flows.

flow, speed, on-ramp demands and out-going flows at each sample time t. We can obtain the flow-density fundamental diagrams from the data measured by detectors installed in the Kwinana Freeway. These model parameters are calculated and are given in Table II. The mainline demand, ramp demands and out-going flows are measured from the detectors, see Fig. 4. Traffic control can be used to improve the performance of the freeway system. The time chosen is 4h from 6 : 00 am to 10 : 00 am, which corresponds to 960 steps. In this example, we select $N_p = 33$, i.e., 8 minutes as predictive horizon, and $N_c = 8$, i.e., 2 minutes for control sample time, which meet the above requirements.

2) Experiment Results: The density results for the six models (actually being measured, PCTM, LICTM, BMM, APCTM and GNHBLP) are shown in Fig. 5. The bars on the right side



(a) Density (b) Density

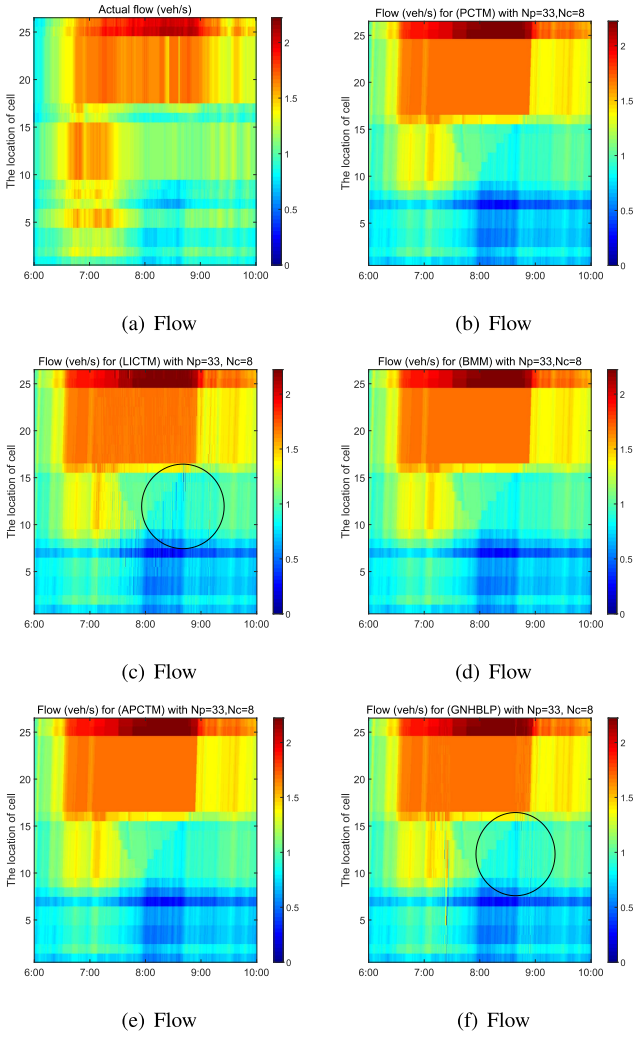
(c) Density (d) Density

(e) Density (f) Density

Fig. 5. Density for actually being measured, PCTM, LICTM, BMM, APCTM and GNHBLP.

of each of Figs. 5(a)-5(f) show the changes in the values of the corresponding state variables that increase from bottom to top. From Figs. 5(a)-5(f), we can observe the evolution for density for each of the models (PCTM, LICTM, BMM, APCTM and GNHBLP). As the demands for the mainline and on-ramps increase, the density at the junctions of the mainline and on-ramps increases. Consequently, the speed decreases, congestion gradually occurs, and the congestion wave propagates upstream from the junction. This situation lasts approximately two hours, which is consistent with the traffic flow during the morning rush hour, i.e., from 7:00 am to 9:00 am. From Figs. 5(b), 5(d) and 5(e), we observe that the changing trend of our model APCTM is almost the same as that of PCTM and BMM. This is consistent with our theoretical analysis.

The traffic flow results for the six models (actually being measured, PCTM, LICTM, BMM, APCTM and GNHBLP) are shown in Fig. 6. From Figs. 6(a)- 6(f), we can observe the evolution of traffic flow of the models (PCTM, LICTM, BMM, APCTM and GNHBLP). Compared the result of LICTM with those of other models, we see that PCTM, BMM, APCTM and

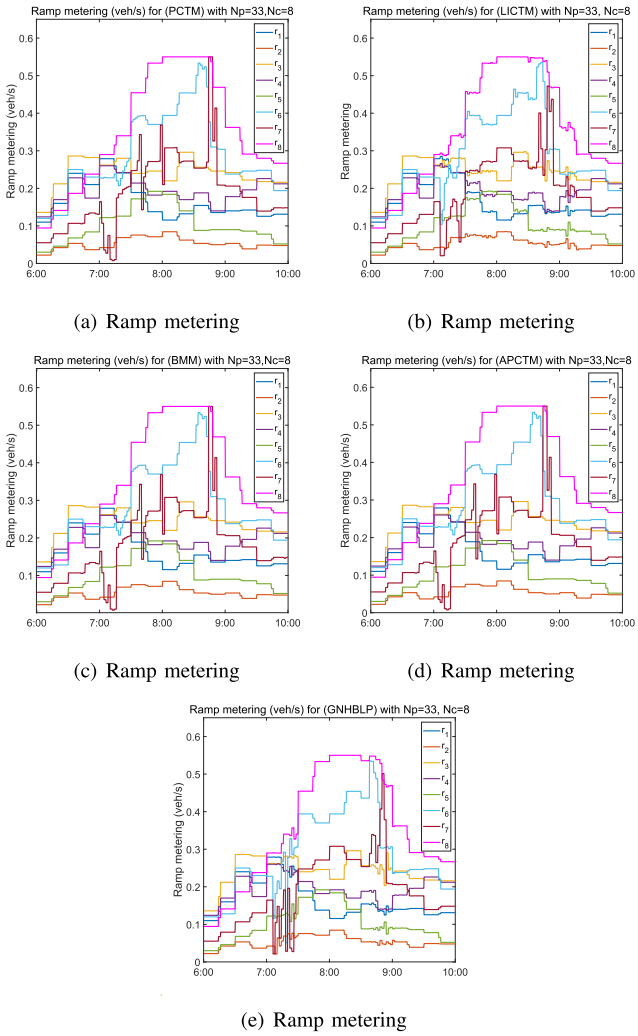


(a) Flow (b) Flow

(c) Flow (d) Flow

(e) Flow (f) Flow

Fig. 6. Flow for actually being measured, PCTM, LICTM, BMM, APCTM and GNHBLP.



(a) Ramp metering (b) Ramp metering

(c) Ramp metering (d) Ramp metering

(e) Ramp metering

Fig. 7. Ramp metering for PCTM, LICTM, BMM, APCTM and GNHBLP.

GNHBLP can solve the flow holding-back problem. However, the performance of PCTM, BMM and APCTM is better than that of GNHBLP. The experimental results agree with the theoretical analysis.

We show the ramp metering and queue length of the five models (PCTM, LICTM, BMM, APCTM and GNHBLP) in Fig. 7 and Fig. 8. From Figs. 7(a), 7(c) and 7(d), we see that the solution of our model APCTM converges to those of PCTM and BMM when the parameter ϵ is appropriately chosen. Due to the flow holding-back problem arising from linear relaxation, the solution of LICTM shows a different trend. GNHBLP only solves the flow holding-back problem to a certain extent.

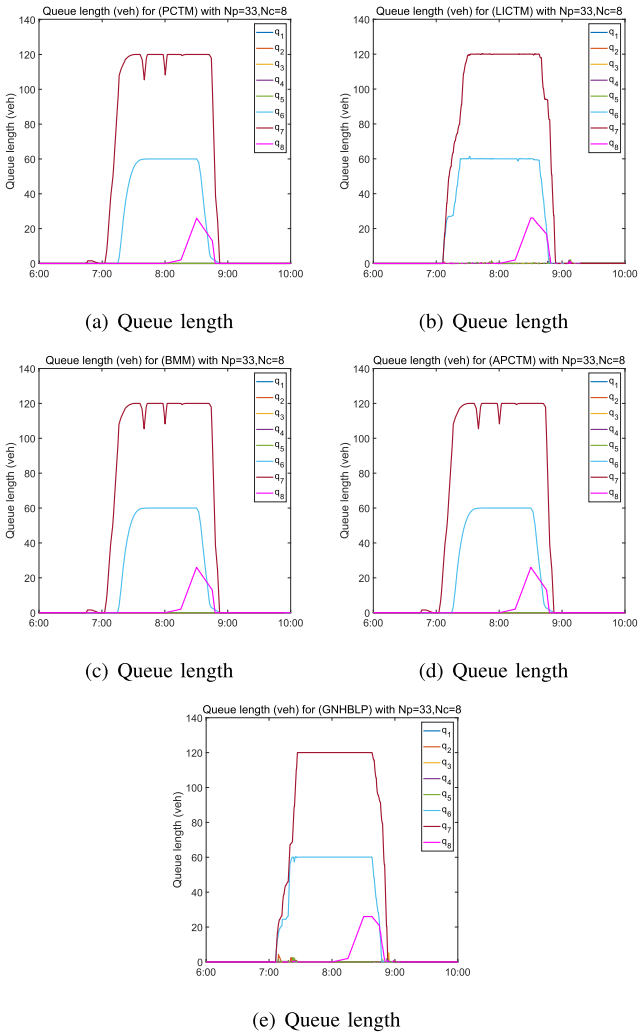
We compare the sizes of the fives models (PCTM, LICTM, BMM, APCTM, GNHBLP) with different prediction horizon N_p in Table III, where N_v and N_c denote, respectively, the numbers of variables and constraints. We find that PCTM and APCTM have considerably fewer numbers of variables and constraints than LICTM, BMM and GNHBLP. Even though PCTM and APCTM have the same numbers of variables and constraints, we find that the time cost of APCTM is far less

TABLE III

THE SIZES OF VARIABLES AND CONSTRAINTS OF THE FIVE MODELS WITH DIFFERENT PREDICTIVE HORIZON N_p

Model	$N_p = 33$		$N_p = 121$		$N_p = 481$	
	N_v	N_c	N_v	N_c	N_v	N_c
PCTM	256	2244	960	8160	3840	30784
LICTM	1088	5610	4080	20570	16320	81770
BMM	4352	9834	16320	36058	65280	143338
APCTM	256	2244	960	8160	3840	30784
GNHBLP	1089	5611	4081	20571	16321	81771

than that required by PCTM model (Table IV). In particular, the trend is more evident when the number of the variables increases. BMM equivalently represents the 'min' operators by introducing a large number of auxiliary variables, so the CPU time is more than that required by APCTM. Since LICTM and GNHBLP are linear program problems, the computational time will obviously be less than that of APCTM. However, they are relaxation problems of the original problem, so the total delay and ramp delay obtained tend to be longer.



(a) Queue length (b) Queue length

(c) Queue length (d) Queue length

(e) Queue length

Fig. 8. Queue length for PCTM, LICTM, BMM, APCTM and GNHBLP.

TABLE IV

THE CPU TIME (IN SECONDS) REQUIRED FOR APPLICATION (58) WITH DIFFERENT PREDICTIVE HORIZON N_p

Model	$N_p = 33$	$N_p = 121$	$N_p = 481$
	CPU time	CPU time	CPU time
PCTM	2.0598	35.8489	10938.5695
LICTM	0.2289	2.0629	22.2301
BMM	1.3830	20.7851	2686.0885
APCTM	0.8402	8.3765	767.0149
GNHBLP	0.2521	2.3057	22.4705

For comparison, we summarize the total delay, ramp delay and total CPU time (in seconds) of the above five models in Table V. Clearly, we can see that the total delay of APCTM is better than LICTM and GNHBLP because our model APCTM does not have a flow holding-back problem. In Table V, we can also see that the total delay and ramp delay of APCTM are, respectively, 99.9965% and 99.8162% of those of PCTM. Compared to the case with no control, our model can reduce the total delay by 55.6279%. Compared with the LICTM, our approximate model reduces the total

TABLE V

TOTAL DELAY, RAMP DELAY AND TOTAL CPU TIME (IN SECONDS) REQUIRED FOR APPLICATION (58) WITH NP=33, NC=8

Model	Total delay	Ramp delay	Total CPU time
No control	1074.6552h	0h	0
PCTM	476.8305h	276.6920h	345.6726
LICTM	486.8045h	282.4111h	51.3336
BMM	476.8338h	276.6953h	217.5161
APCTM	476.8471h	276.7086h	119.1033
GNHBLP	480.2179h	279.5597h	53.6784

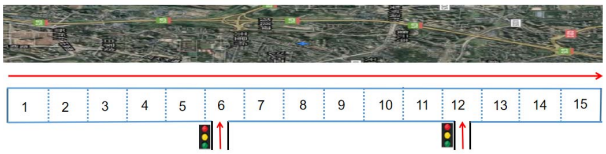


Fig. 9. Road graph.

delay and ramp delay by 2.0455% and 2.0192%, respectively. Furthermore, the total delay and ramp delay of APCTM are also better than those of GNHBLP. The total CPU time taken by APCTM is 119.1033s. Since in the real application, the control sample time is 2 minutes, it means that the CPU time taken by APCTM can meet the requirements of online control applications. By comparison, we can find the total CPU time of PCTM and BMM are, respectively, 2.6856 and 1.8263 times of that of APCTM. Compared with LICTM and GNHBLP, APCTM requires more computational time due to the involvement of nonlinearity. On the other hand, the total delay and ramp delay are shorter when compared with those of LICTM and GNHBLP.

C. Example 2

1) Scenario: The example is built based on the actual situation of the Freeway in the vicinity of Shapingba, Chongqing in China. The section of freeway under consideration is divided into $I = 15$ cells with $J = 2$ on-ramps. Cells $i = 1, 2, \ldots, 9$ have 4 lanes and cells $i = 10, 11, \ldots, 15$ have 3 lanes. Fig. 9 shows the information of road map. For this example, we use SUMO to simulate the traffic flow dynamic from 6:00 am to 10:00 am to obtain the required data. Mainline demand and on-ramp demands are presented in Fig. 10. The values of parameters are listed in Table VI. In this example, we let $Np = 17$, i.e., 2 minutes and 40 seconds as predictive horizon, and $Nc = 6$, i.e., 1 minute for control sample time, which can meet the above requirements.

2) Experiment Results: We compare the sizes and CPU time (in seconds) of the five models (PCTM, LICTM, BMM, APCTM, GNHBLP) for each iteration with $Np = 17$, $Nc = 6$ in Table VII. Similar to Example 1, PCTM and APCTM have fewer numbers of variables and constraints than those of LICTM, BMM and GNHBLP. APCTM can be solved more efficiently than PCTM and BMM, but not better than LICTM and GNHBLP.

TABLE VI
MODEL PARAMETERS

Name of parameter	Value	Name of parameter	Value
$v_i, i = 1, \ldots, 15$	22.2222m/s	$q_{max,j}, j = 1$	60veh
$w_i, i = 1, 2, \ldots, 9$	6.0206m/s	$q_{max,j}, j = 2$	70veh
$w_i, i = 10, 11, \ldots, 15$	11.6436m/s	$r_{max,j}, j = 1$	1veh/s
$C_i, i = 1, 2, \ldots, 9$	2.3685veh/s	$r_{max,j}, j = 2$	0.55veh/s
$C_i, i = 10, 11, \ldots, 15$	2.1916veh/s	Δt	10s
$\rho_{max,i}, i = 1, 2, \ldots, 9$	0.5veh/m	ϵ	0.00000001
$\rho_{max,i}, i = 10, 11, \ldots, 15$	0.2868veh/m		

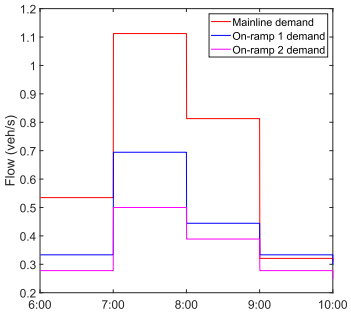


Fig. 10. Road graph.

TABLE VII
THE SIZES AND CPU TIME (IN SECONDS) REQUIRED FOR APPLICATION (58) WITH NP=17, NC=6

Model	N_v	N_c	CPU time
PCTM	32	578	0.3502
LICTM	272	1564	0.0609
BMM	1200	2550	0.2686
APCTM	32	578	0.1736
GNHBLP	273	1565	0.0641

The total delay, ramp delay and total CPU time (in seconds) of the five models are summarized in Table VIII. From Table VIII, we clearly observe that the total delay of APCTM is better than those of LICTM and GNHBLP, because APCTM can solve the flow holding-back problem if the parameter ϵ is appropriately chosen. In addition, we can also observe that the total delay and ramp delay of our APCTM are, respectively, 99.9775% and 99.678% of those obtained using PCTM. Compared with the case of no control, APCTM can reduce the total delay by 52.4996%, but the ramp delay increases by 18.6882%. Compared with LICTM and GNHBLP, APCTM reduces the total delay by 1.7654% and 0.872%, respectively. As for the total CPU time, PCTM and BMM are, respectively, 2.0173 and 1.6762 times of that of APCTM, but LICTM and GNHBLP take shorter times than that of APCTM. In this example, the total CPU time taken by

TABLE VIII
TOTAL DELAY, RAMP DELAY AND TOTAL CPU TIME (IN SECONDS) REQUIRED FOR APPLICATION (58) WITH NP=17, NC=6

Name of model	Total delay	Ramp delay	Total CPU time
No control	520.7225h	64.0275h	0
PCTM	247.2868h	75.9776h	78.3409
LICTM	251.7902h	81.0918h	12.8565
BMM	247.2896h	75.9931h	65.0922
APCTM	247.3452h	76.2231h	38.8343
GNHBLP	249.5209h	77.4397h	13.8184

APCTM is 38.8343s. Since in the real application, the control sample time is 1 minute, it means that the CPU time taken by APCTM can also meet the requirements of online control applications.

Based on the results obtained for Example 1 and Example 2, we can conclude that the total delay and ramp delay of APCTM is better than those of LICTM and GNHBLP. Furthermore, the CPU time of APCTM is at least 2 and 1.5 times of those of PCTM and BMM. Therefore, APCTM is efficient in terms of computational time and the reduction of total delay and ramp delay.

VI. CONCLUSION

This paper proposed a novel approximate optimization model based on the CTM to overcome the flow holding-back problem caused by unrealistic fundamental diagrams if inequality convex relaxation is used directly. Theoretical analysis showed that the solution obtained from our model converges to the solution of the original CTM as ϵ approaches to zero. To obtain a feedback control law, a customized MPC approach is designed under the framework of our proposed optimization problem. Experimental results showed that our approach is better than the existing methods. Note that the scale of the numerical experiments in this paper is rather limited due to the availability of data. Conducting large-scale numerical applications in future is necessary. For example, we can take the entire Kwinana Freeway as a test bed to verify the effectiveness of our proposed model. In addition, capacity drop is not considered when on-ramp is merging to the mainline in this paper.

APPENDIX

In this appendix, we will give some auxiliary lemmas, which are used in the convergence analysis. Lemma 1 is to estimate upper and lower bounds of $f_{i,t} - f^{\epsilon}_{i,t}$, which play an essential role in the proofs of Lemma 2 and Theorem 1.

Lemma 1: Consider (23), (24) and (29). Then, for all $\epsilon > 0$, it holds that $-\frac{\epsilon}{4} \le f_{i,t} - f^{\epsilon}_{i,t} \le \frac{3\epsilon}{4}$, and $\lim_{\epsilon \to 0} f^{\epsilon}_{i,t} = f_{i,t}$.

Proof: First, we estimate an upper bound of $f_{i,t} - f^{\epsilon}_{i,t}$ for any i and t. By (27) and (29), it gives

$$\begin{aligned} f_{i,t} - f^{\epsilon}_{i,t} &= f_{i,t} - f^{\epsilon\epsilon}_{i,t} + f^{\epsilon\epsilon}_{i,t} - f^{\epsilon}_{i,t} \\ &= \frac{1}{2}[f^{D}_{i,t} + f^{S}_{i+1,t} - |f^{D}_{i,t} - f^{S}_{i+1,t}| - (f^{D\epsilon}_{i,t} \end{aligned}$$

$$+ f^{S\epsilon}_{i+1,t} - |f^{D\epsilon}_{i,t} - f^{S\epsilon}_{i+1,t}|)] + f^{\epsilon\epsilon}_{i,t} - f^{\epsilon}_{i,t}$$
$$\leq \frac{\epsilon}{2} + \frac{1}{2}(|f^{D\epsilon}_{i,t} - f^{S\epsilon}_{i+1,t}| - |f^{D}_{i,t} - f^{S}_{i+1,t}|), \quad (59)$$

where the last inequality is due to (25), (26) and (30). We next discuss the positivity and negativity of $f^{D\epsilon}_{i,t} - f^{S\epsilon}_{i+1,t}$ and $f^{D}_{i,t} - f^{S}_{i+1,t}$. Let

$$a = f^{D\epsilon}_{i,t} - f^{S\epsilon}_{i+1,t},\ b = f^{D}_{i,t} - f^{S}_{i+1,t}. \quad (60)$$

1): When $a \geq 0,\ b \geq 0$, we obtain

$$|a| - |b| = f^{D\epsilon}_{i,t} - f^{S\epsilon}_{i+1,t} - f^{D}_{i,t} + f^{S}_{i+1,t} \leq \frac{\epsilon}{4},$$

and thus

$$f_{i,t} - f^{\epsilon}_{i,t} \leq \frac{5\epsilon}{8}. \quad (61)$$

2): When $a < 0,\ b < 0$, we have

$$|a| - |b| = f^{S\epsilon}_{i+1,t} - f^{D\epsilon}_{i,t} + f^{D}_{i,t} - f^{S}_{i+1,t} \leq \frac{\epsilon}{4},$$

and thus

$$f_{i,t} - f^{\epsilon}_{i,t} \leq \frac{5\epsilon}{8}. \quad (62)$$

3): When $a \geq 0,\ b < 0$, we have

$$|a| - |b| = f^{D\epsilon}_{i,t} - f^{S\epsilon}_{i+1,t} + f^{D}_{i,t} - f^{S}_{i+1,t} \leq \frac{\epsilon}{2},$$

and thus

$$f_{i,t} - f^{\epsilon}_{i,t} \leq \frac{3\epsilon}{4}. \quad (63)$$

4): When $a < 0,\ b \geq 0$, we obtain

$$|a| - |b| = f^{S\epsilon}_{i+1,t} - f^{D\epsilon}_{i,t} - f^{D}_{i,t} + f^{S}_{i+1,t} \leq \frac{\epsilon}{2},$$

and thus

$$f_{i,t} - f^{\epsilon}_{i,t} \leq \frac{3\epsilon}{4}. \quad (64)$$

By (61), (62), (63) and (64), it follows that

$$f_{i,t} - f^{\epsilon}_{i,t} \leq \frac{3\epsilon}{4}. \quad (65)$$

Next, we estimate a lower bound of $f_{i,t} - f^{\epsilon}_{i,t}$ for any i and t. By (27) and (29), it gives

$$f_{i,t} - f^{\epsilon}_{i,t} \geq \frac{1}{2}(|f^{D\epsilon}_{i,t} - f^{S\epsilon}_{i+1,t}| - |f^{D}_{i,t} - f^{S}_{i+1,t}|). \quad (66)$$

Similarly, we study the positivity and negativity of $f^{D\epsilon}_{i,t} - f^{S\epsilon}_{i+1,t}$ and $f^{D}_{i,t} - f^{S}_{i+1,t}$.

1): When $a \geq 0,\ b \geq 0$,

$$|a| - |b| = f^{D\epsilon}_{i,t} - f^{S\epsilon}_{i+1,t} - f^{D}_{i,t} + f^{S}_{i+1,t} \geq -\frac{\epsilon}{4},$$

and thus

$$f_{i,t} - f^{\epsilon}_{i,t} \geq -\frac{\epsilon}{8}. \quad (67)$$

2): When $a < 0,\ b < 0$,

$$|a| - |b| = f^{S\epsilon}_{i+1,t} - f^{D\epsilon}_{i,t} + f^{D}_{i,t} - f^{S}_{i+1,t} \geq -\frac{\epsilon}{4},$$

and thus

$$f_{i,t} - f^{\epsilon}_{i,t} \geq -\frac{\epsilon}{8}, \quad (68)$$

3): When $a \geq 0,\ b < 0$,

$$|a| - |b| = f^{D\epsilon}_{i,t} - f^{S\epsilon}_{i+1,t} + f^{D}_{i,t} - f^{S}_{i+1,t} \geq -\frac{\epsilon}{2},$$

and thus

$$f_{i,t} - f^{\epsilon}_{i,t} \geq -\frac{\epsilon}{4}. \quad (69)$$

4): When $a < 0,\ b \geq 0$,

$$|a| - |b| = f^{S\epsilon}_{i+1,t} - f^{D\epsilon}_{i,t} - f^{D}_{i,t} + f^{S}_{i+1,t} \geq -\frac{\epsilon}{2},$$

and thus

$$f_{i,t} - f^{\epsilon}_{i,t} \geq -\frac{\epsilon}{4}. \quad (70)$$

By (67), (68), (69) and (70), we have

$$f_{i,t} - f^{\epsilon}_{i,t} \geq -\frac{\epsilon}{4}. \quad (71)$$

Due to (65) and (71), we obtain the first part of the result. The second part of the result is obvious when $\epsilon \to 0$, and thus the proof is completed. ■

Lemma 2: Let $r^{*\epsilon}$ be an optimal solution to Problem APCTM. *Then, there exists a subsequence of $\{r^{*\epsilon}\}$, which is again denoted by the original sequence, and a control vector $\overline{r} \in F$ such that*

$$\lim_{\epsilon \to 0} ||r^{*\epsilon} - \overline{r}|| = 0, \quad (72)$$
$$\lim_{\epsilon \to 0} \rho^{*\epsilon}_{i,t} = \overline{\rho}_{i,t},\ \forall i, t. \quad (73)$$

Similarly, let r^* be an optimal solution to Problem PCTM. *Then, there exists a sequence $\{\overline{r}^{\epsilon}\}$, which is again denoted by the original sequence, such that*

$$\lim_{\epsilon \to 0} ||\overline{r}^{\epsilon} - r^*|| = 0, \quad (74)$$
$$\lim_{\epsilon \to 0} \overline{\rho}^{\epsilon}_{i,t} = \rho^{*}_{i,t},\ \forall i, t. \quad (75)$$

Proof: Note that V is a compact subset of $\mathbb{R}^J$. Since $\{r^{*\epsilon}\}$ as a sequence in ϵ is in V, it is clear that there exists a subsequence, which is again denoted by the original sequence, and a control vector $\overline{r} \in F$ such that

$$\lim_{\epsilon \to 0} ||r^{*\epsilon} - \overline{r}|| = 0. \quad (76)$$

We prove $\lim_{\epsilon \to 0} \rho^{*\epsilon}_{i,t} = \overline{\rho}_{i,t}$ by exploiting the mathematical induction. The result is true when $t = 1$ for each i, because $\lim_{\epsilon \to 0} \rho^{*\epsilon}_{i,1} = \rho_{i,1}$. Next, we prove that the result is true when $t = 2$ for any $i = 1, \ldots, I$. Based on Equations (2) and (32), we have, for each i,

$$\overline{\rho}_{i,2} = \overline{\rho}_{i,1} + \frac{\Delta t}{\Delta x_i} \times (\overline{f}_{i-1,1} - \overline{f}_{i,1} + \overline{r}_{i,1} - s_{i,1}),$$

and

$$\rho^{*\epsilon}_{i,2} = \rho^{*\epsilon}_{i,1} + \frac{\Delta t}{\Delta x_i} \times (f^{\epsilon}_{i-1,1}(\rho^{*\epsilon}_{i-1,1}) - f^{\epsilon}_{i,1}(\rho^{*\epsilon}_{i,1}) + r^{*\epsilon}_{i,1} - s_{i,1}),$$

where $\overline{\rho}_{i,1} = \rho^{*\epsilon}_{i,1} = \rho_{i,1}$ for each i. Using Lemma 1, we obtain for any i

$$\lim_{\epsilon \to 0} f^{\epsilon}_{i,1}(\rho^{*\epsilon}_{i,1}) = \overline{f}_{i,1}. \quad (77)$$

Since $0 < \frac{\Delta t}{\Delta x_i} < 1$, by (76) and (77), we have

$$\lim_{\epsilon \to 0} \rho^{*\epsilon}_{i,2} = \overline{\rho}_{i,2}, \quad \forall i. \tag{78}$$

Now, we assume that the result is true when $t = T - 1$ for any $i = 1, \ldots, I$, i.e.,

$$\lim_{\epsilon \to 0} \rho^{*\epsilon}_{i,T-1} = \overline{\rho}_{i,T-1}. \tag{79}$$

Next, we shall prove that it is also true when $t = T$. Based on equations (2) and (32), we have, for any $i = 1, \ldots, I$,

$$\overline{\rho}_{i,T} = \overline{\rho}_{i,T-1} + \frac{\Delta t}{\Delta x_i} \times (\overline{f}_{i-1,T-1} - \overline{f}_{i,T-1} + \overline{r}_{i,T-1} - s_{i,T-1}), \tag{80}$$

and

$$\rho^{*\epsilon}_{i,T} = \rho^{*\epsilon}_{i,T-1} + \frac{\Delta t}{\Delta x_i} \times (f^{\epsilon}_{i-1,T-1}(\rho^{*\epsilon}_{i-1,T-1}) - f^{\epsilon}_{i,T-1}(\rho^{*\epsilon}_{i,T-1}) + r^{*\epsilon}_{i,T-1} - s_{i,T-1}). \tag{81}$$

Using Lemma 1, we obtain, for any $i = 1, \ldots, I$,

$$\lim_{\epsilon \to 0} f^{\epsilon}_{i,T-1}(\overline{\rho}_{i,T-1}) = \overline{f}_{i,T-1}. \tag{82}$$

Furthermore, $f^{\epsilon}_{i,t}$ is continuously differentiable with respect to each of the components of ρ and r. Thus, it follows from (76) and (79) that

$$\lim_{\epsilon \to 0} f^{\epsilon}_{i,T-1}(\rho^{*\epsilon}_{i,T-1}) = \lim_{\epsilon \to 0} f^{\epsilon}_{i,T-1}(\overline{\rho}_{i,T-1}). \tag{83}$$

The following result follows readily from (82) and (83)

$$\lim_{\epsilon \to 0} f^{\epsilon}_{i,T-1}(\rho^{*\epsilon}_{i,T-1}) = \overline{f}_{i,T-1}. \tag{84}$$

Then, $\lim_{\epsilon \to 0} \rho^{*\epsilon}_{i,T} = \overline{\rho}_{i,T}$ is obvious by virtue of (79), (80), (81) and (84). The second part of the result follows readily by using a similar approach, and thus the details are omitted. The proof is completed. ■

REFERENCES

[1] R. Mohebifard and A. Hajbabaie, "Dynamic traffic metering in urban street networks: Formulation and solution algorithm," *Transp. Res. C, Emerg. Technol.*, vol. 93, pp. 161–178, Aug. 2018.

[2] J. Long and W. Y. Szeto, "Link-based system optimum dynamic traffic assignment problems in general networks," *Oper. Res.*, vol. 67, no. 1, pp. 167–182, Jan. 2019.

[3] A. Hegyi, B. De Schutter, and J. Hellendoorn, "Model predictive control for optimal coordination of ramp metering and variable speed limits," *Transp. Res. C, Emerg. Technol.*, vol. 13, no. 3, pp. 185–209, Jun. 2005.

[4] G. Gomes, R. Horowitz, A. A. Kurzhanskiy, P. Varaiya, and J. Kwon, "Behavior of the cell transmission model and effectiveness of ramp metering," *Transp. Res. C, Emerg. Technol.*, vol. 16, no. 4, pp. 485–513, 2008.

[5] M. Papageorgiou, H. Hadj-Salem, and J.-M. Blosseville, "ALINEA: A local feedback control law for on-ramp metering," *Transp. Res. Rec.*, vol. 1320, pp. 58–64, Jan. 1991.

[6] I. Papamichail and M. Papageorgiou, "Traffic-responsive linked ramp-metering control," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 1, pp. 111–121, Mar. 2008.

[7] D. Zhao, X. Bai, F. Wang, and J. Xu, "DHP method for ramp metering of freeway traffic," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 990–999, Dec. 2011.

[8] N. Geroliminis, A. Srivastava, and P. Michalopoulos, "A dynamic-zone-based coordinated ramp-metering algorithm with queue constraints for Minnesota's freeways," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1576–1586, Dec. 2011.

[9] C. F. Daganzo, "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transp. Res. B, Methodol.*, vol. 28, no. 4, pp. 269–287, Aug. 1994.

[10] C. F. Daganzo, "The cell transmission model, Part II: Network traffic," *Transp. Res. B, Methodol.*, vol. 29, no. 2, pp. 79–93, Apr. 1995.

[11] A. Messmer and M. Papageorgiou, "METANET: A macroscopic simulation program for motorway networks," *Traffic Eng. Control*, vol. 31, no. 8, pp. 466–470, 1990.

[12] T. Bellemans, B. De Schutter, and B. De Moor, "Anticipative model predictive control for ramp metering in freeway networks," in *Proc. ACC*, vol. 5, Jun. 2003, pp. 4077–4082.

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

[14] J. A. Wattleworth, "Peak-period control of a freeway system: Some theoretical considerations," Ph.D. dissertation, Northwestern Univ., Evanston, IL, USA, 1963.

[15] Y. Iida, T. Hasegawa, Y. Asakura, and C. F. Shao, "A formulation of on-ramp traffic control system with route guidance for urban expressway," *IFAC Proc. Volumes*, vol. 23, no. 2, pp. 229–236, Sep. 1990.

[16] G. Gomes and R. Horowitz, "Optimal freeway ramp metering using the asymmetric cell transmission model," *Transp. Res. C, Emerg. Technol.*, vol. 14, no. 4, pp. 244–262, 2006.

[17] A. Muralidharan and R. Horowitz, "Optimal control of freeway networks based on the link node cell transmission model," in *Proc. Amer. Control Conf. (ACC)*, Jun. 2012, pp. 5769–5774.

[18] G. Como, E. Lovisari, and K. Savla, "Convexity and robustness of dynamic network traffic assignment for control of freeway networks," *Transp. Res. B, Methodol.*, vol. 91, pp. 446–465, Sep. 2016.

[19] C. Roncoli, M. Papageorgiou, and I. Papamichail, "Traffic flow optimisation in presence of vehicle automation and communication systems—Part I: A first-order multi-lane model for motorway traffic," *Transp. Res. C, Emerg. Technol.*, vol. 57, pp. 241–259, Aug. 2015.

[20] C. Roncoli, M. Papageorgiou, and I. Papamichail, "Traffic flow optimisation in presence of vehicle automation and communication systems—Part II: Optimal control for multi-lane motorways," *Transp. Res. C, Emerg. Technol.*, vol. 57, pp. 260–275, Aug. 2015.

[21] C. Rosdahl, G. Nilsson, and G. Como, "On distributed optimal control of traffic flows in transportation networks," in *Proc. IEEE Conf. Control Technol. Appl. (CCTA)*, Aug. 2018, pp. 903–908.

[22] D. Bianchi, A. Borri, M. D. Di Benedetto, and A. Ferrara, "Decentralized model predictive control of freeway traffic systems over lossy communication networks," in *Proc. IEEE Int. Conf. Syst., Man Cybern. (SMC)*, Oct. 2019, pp. 1074–1079.

[23] R. X. Zhong, F. F. Yuan, T. L. Pan, A. H. F. Chow, C. J. Chen, and Z. Yang, "Linear complementarity system approach to macroscopic freeway traffic modelling: Uniqueness and convexity," *Transportmetrica A, Transp. Sci.*, vol. 12, no. 2, pp. 142–174, Feb. 2016.

[24] M. Schmitt and J. Lygeros, "An exact convex relaxation of the freeway network control problem with controlled merging junctions," *Transp. Res. B, Methodol.*, vol. 114, pp. 1–25, Aug. 2018.

[25] M. J. Lighthill and G. B. Whitham, "On kinematic waves II. A theory of traffic flow on long crowded roads," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 229, no. 1178, pp. 317–345, 1955.

[26] S. K. Godunov, "A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics," *Matematicheskii Sbornik*, vol. 47, no. 89, pp. 271–306, 1959.

[27] J. P. Lebacque, "The Godunov scheme and what it means for first order traffic flow models," in *Proc. 13th Int. Symp. Transp. Traffic Theory*, Jul. 1996, pp. 24–26.

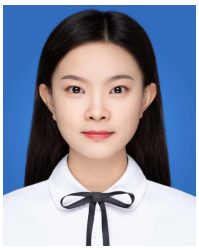
[28] A. H. F. Chow and Y. Li, "Robust optimization of dynamic motorway traffic via ramp metering," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 3, pp. 1374–1380, Jun. 2014.

[29] R. B. Wilson, "A simplicial algorithm for concave programming," Ph.D. dissertation, Dept. Bus. Admin., Harvard Univ., Cambridge, MA, USA, 1963.

[30] Y. Han, M. Ramezani, A. Hegyi, Y. Yuan, and S. Hoogendoorn, "Hierarchical ramp metering in freeways: An aggregated modeling and control approach," *Transp. Res. C, Emerg. Technol.*, vol. 110, pp. 1–19, Jan. 2020.

[31] Y. Zhang, I. I. Sirmatel, F. Alasiri, P. A. Ioannou, and N. Geroliminis, "Comparison of feedback linearization and model predictive techniques for variable speed limit control," in *Proc. 21st Int. Conf. Intell. Transp. Syst.*, Nov. 2018, pp. 3000–3005.

[32] Y. Han, A. Hegyi, Y. Yuan, S. Hoogendoorn, M. Papageorgiou, and C. Roncoli, "Resolving freeway jam waves by discrete first-order model-based predictive control of variable speed limits," *Transp. Res. C, Emerg. Technol.*, vol. 77, pp. 405–420, Apr. 2017.
[33] E. F. Camacho and C. B. Bordons, *Model Predictive Control* (Advanced Textbooks in Control and Signal Processing). London, U.K.: Springer, 2004.
[34] F. Zhu and S. V. Ukkusuri, "A cell based dynamic system optimum model with non-holding back flows," *Transp. Res. C, Emerg. Technol.*, vol. 36, pp. 367–380, Nov. 2013.
[35] Gurobi Optimization. (2012). *Gurobi Optimizer Reference Manual*. [Online]. Available: http://www.gurobi.com



Chuanye Gu received the B.E. and M.E. degrees from the School of Mathematics, Chongqing Normal University, China, in 2015 and 2018, respectively. She is currently pursuing the Ph.D. degree with the Department of Mathematics and Statistics, Curtin University, Australia. Her current research interests include freeway optimization and its application in the transportation area.

Changzhi Wu received the Ph.D. degree in optimization and control from Sun Yat-sen University, China, in 2006. He is currently a Professor with the School of Management, Guangzhou University. He has published more than 100 papers. His current research interests include optimization, intelligent transportation systems, and data analytics.

Kok Lay Teo (Life Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Ottawa, Canada. He was with the Department of Applied Mathematics, University of New South Wales, Australia; the Department of Industrial and Systems Engineering, National University of Singapore, Singapore; and the Department of Mathematics, The University of Western Australia, Australia. In 1996, he joined the Department of Mathematics and Statistics, Curtin University of Technology, Australia, as a Professor. He then took up the position of the Chair Professor of applied mathematics and the Head of the Department of Applied Mathematics, The Hong Kong Polytechnic University, China, from 1999 to 2004. He returned to Curtin University as a Professor and the Head of the Department of Mathematics and Statistics from 2010 to 2015. He was a John Curtin Distinguished Professor at Curtin University from 2011 until his retirement in November 2019. He is a John Curtin Distinguished Emeritus Professor at Curtin University. Currently, he is a Professor and the Associate Dean (Research and Postgraduate Studies) with the School of Mathematical Science, Sunway University, Malaysia. He is also associated with the Tianjin University of Finance and Economics as the Visiting Distinguished Chair Professor with the Coordinated Innovation Center for Computable Modeling in Management Science.

Yonghong Wu received the Ph.D. degree from the University of Wollongong in 1990. He then worked there as a Post-Doctoral Research Fellow. In 1993, he joined Curtin University, WA, Australia. Since then, he has been working with the Department of Mathematics and Statistics, Curtin University, where he is currently a John Curtin Distinguished Professor in mathematics and statistics. He has published more than 250 research articles in prestigious international journals. His research works are well cited internationally, and he is listed by Clarivate Analytics Web of Science among the World's Highly Cited Researchers. His research interests include computational mathematics and applied mathematics.

Song Wang received the B.Sc. degree from the Wuhan University of Hydraulic and Electric Engineering, Wuhan (now Wuhan University), in 1982, and the Ph.D. degree in numerical analysis from Trinity College Dublin in 1989. He had been with a Dublin-Based Hi-Tech Company-Tritech Ltd., The University of New South Wales, Curtin University, and The University of Western Australia, before he moved to Curtin University again as the Head of the Department of Mathematics and Statistics in 2014. Currently, he is a Full Professor at Curtin University. His research interests include scientific computation, numerical optimization and optimal control, optimum design, and computational finance. He is a leading Researcher in these areas and has published many original research papers in well-established international journals. Currently, he is the editor-in-chief of two journals and on the editorial boards of several other international journals.