Globally Optimal Solutions to the On-ramp Metering Problem - Part I

Gabriel Gomes and Roberto Horowitz
Department of Mechanical Engineering
University of California at Berkeley
Email: gomes@me.berkeley.edu , horowitz@me.berkeley.edu

***Abstract*— A mathematical programming approach to the freeway on-ramp metering problem is formulated. The objective function is a linear combination of mainline and on-ramp flows, termed the *Generalized Total Travel Time*. The underlying freeway model – the Asymmetric Cell Transmission Model (ACTM) – is similar to the original Cell Transmission Model (CTM), except that the merge law of the CTM has been replaced with additional terms weighted by *influence parameters*. It is shown that an appropriate selection of the model parameters and boundary conditions guarantees a physically reasonable evolution of the ACTM. It is also shown that the resulting nonlinear optimization problem can be solved globally, by solving an equivalent linear program, whenever the cost weights are generated by a proposed numerical algorithm.**

I. Introduction

This paper considers the problem of regulating access flows to freeways as a means of reducing the effects of recurrent congestion. The many possible approaches to this problem can be classified in several ways: traffic-responsive versus open-loop, local versus system-wide, static versus dynamic, network versus freeway models, considering versus ignoring the effects of diversion. The approach presented here falls, in all cases, in the latter category.

One of the first applications of mathematical programming to the problem of on-ramp control was by Wattleworth in 1965 [1]. This early formulation was based on a static model of traffic behavior, whereby the flows at any cross-section in the system could be expressed as the sum of the flows entering the freeway upstream of that location, scaled by a known proportion of vehicles that did not exit at any upstream off-ramp. This density-less model allowed the formulation of a linear program, since it avoided the important non-linearity in freeway traffic behavior - the relationship between flow and density known as the *fundamental diagram*.

Many subsequent contributions have built upon the original formulation by Wattleworth. Yuan and Kreer [2] proposed a quadratic cost to replace Wattleworth's linear maximization of on-ramp flows, in order to achieve a more equitable distribution of the control effort. Chen et al. [3] suggested the use of Total Travel Distance as the objective. Wang and May [4] discussed several other enhancements, and extended the model to consider the effects of voluntary diversion to surface streets. Later authors further extended the model to capture the entire *corridor*, which comprises both the freeway and an alternative parallel route that allows drivers some flexibility in their choice of freeway access points. Payne and Thompson [5] considered "Wardrop's first principle" as dictating the selection of routes by drivers, coupled with an on-ramp control formulation similar to Wattleworth's, and solved it with a suboptimal dynamic programming algorithm. Iida et al. [6] posed a similar problem, and employed a heuristic numerical method consisting of iterated solutions of two linear programs (control and assignment).

Another more recent enhancement has been the consideration of *dynamic models*. Most problem formulations using dynamic models have reverted to the simpler situation, where the effect of on-ramp control on access point selection is not considered. Examples include Lei [7], Kotsialos [8], and Hegyi [9]. In these three cases, the numerical method used to solve the resulting nonlinear optimization problem was gradient-based, and therefore provided only local solutions.

The approach presented here is based on an observation stemming from two facts. The first fact, shown in [10], is that minimizing the Total Travel Time is equivalent to maximizing a weighted sum of flows. The second fact is the specification by the LWR theory [11] of a *concave* fundamental diagram. Relaxing the equality constraint imposed by the fundamental diagram to a "$\leq$" constraint therefore results in a convex problem. Because travel time is favored by higher flows, it is not unreasonable to expect the solution of this convex problem to "naturally" seek the upper boundary (i.e. the fundamental diagram).

This idea of relaxing the flow constraint has been suggested previously by Papageorgiou in [12] and Ziliaskopoulos in [13]. However [12] wrongly asserted that the solution to the relaxed problem would always fall on the upper boundary as long as positive and sufficiently large cost weights were assigned to the mainline and on-ramp flows. To see why this is not true, consider a flow $f_i[k]$ from a section i into a *congested* section $i+1$ during time interval k (see Figure 1). An increase in $f_i[k]$ produces an increase in the density of $i+1$ at time $k+1$, which in turn causes $f_i[k+1]$ to fall, because $i+1$ is congested. Thus, the initial increase to $f_i[k]$ will be favored by the objective (weighted maximization of flows) only if its positive effect outweighs the negative effect of decreasing $f_i[k+1]$. One of the findings of this paper is that, in addition to positivity, the cost weights must also decrease in time.

In this paper we first introduce a model for a freeway with on-ramp control. The model is similar to Daganzo's cell transmission model (CTM) [14], in that intercellular flows are computed as the minimum of what can be sent by the upstream cell and what can be received by the downstream cell. The important distinction is in the treatment of merging flows: to make the model more amenable to mathematical optimization, we have replaced the merge law of the CTM with extra terms in the $\min\{\}$ functions, weighted by *influence parameters* $(\alpha_i, \xi_i, \gamma_i)$. It is shown that an appropriate selection of the parameters and boundary conditions guarantees that the evolution of this model will remain within certain implicit constraints (Eq. (8)). This result is used to ensure that the results are reasonable. Two optimization problems, one nonlinear and one linear, are then formulated. The solutions to these two problems are generally not the same. However, it is found that cost weights can be found that render them equivalent. A numerical method for generating such cost weights is then designed and tested.

II. NOTATION

The freeway is partitioned into I sections, each containing at most one on-ramp and/or one off-ramp. The sections are identified with indices beginning with zero at the upstream end, and increasing sequentially downstream. Time is divided into K discrete intervals of duration Δt. Figure 1 illustrates the model variables.

<u>Sets</u>

$\mathcal{I}$: set of all freeway sections. $\mathcal{I} = \{0 \ldots \mathrm{I}-1\}$
$\mathcal{K}$: set of time intervals. $\mathcal{K} = \{0 \ldots \mathrm{K}-1\}$
$\mathcal{E}n$: set of sections with on-ramps. $\mathcal{E}n \subseteq \mathcal{I}$
$\mathcal{E}n^+$: set of sections with metered on-ramps. $\mathcal{E}n^+ \subseteq \mathcal{E}n$

<u>Variables</u> (All normalized to vehicle units)

$\rho_i[k]$: vehicles in section i at time $k\Delta t$.
$l_i[k]$: vehicles queueing in on-ramp i at time $k\Delta t$.
$f_i[k]$: vehicles going from i to $i+1$ during interval k.
$r_i[k]$: vehicles entering i from an on-ramp during k.
$r_i^c[k]$: metering rate for on-ramp i.
$d_i[k]$: demand for on-ramp i.
$s_i[k]$: vehicles using off-ramp i during interval k.
$\beta_i[k]$: dimensionless split ratio for off-ramp i.

$\beta_i[k]$ is defined $\forall\ i \in \mathcal{I}$, and is set to 0 whenever i does not contain an off-ramp. We also define an on-ramp indicator δ_i:

$$\delta_i \triangleq \begin{cases} 1 & \text{if} \quad i \in \mathcal{E}n \\ 0 & \text{else} \end{cases}$$

<u>Model parameters</u>

v_i	: normalized freeflow speed	$\in [0,1]$
w_i	: congestion wave speed	$\in [0,1]$
$\bar{\rho}_i$	: jam density	[veh]
$\bar{f}_i$	: mainline capacity	[veh]
$\bar{s}_i$	: off-ramp capacity	[veh]
α_i, γ_i , ξ_i	: influence parameters	$\in [0,1]$

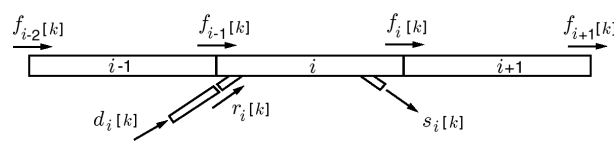

Fig. 1. Interpretation of model variables

III. THE ACTM FREEWAY MODEL

The five components of the ACTM are given by Eqs. (1) through (5). Similarly to [12] and in contrast to [13], the model assumes that the off-ramp split ratios are known and control independent.

<u>Off-ramp flows</u> $\quad \forall\ i \in \mathcal{I}\ ,\ k \in \mathcal{K}$:

$$s_i[k] \ = \ \beta_i[k] (s_i[k] + f_i[k])$$
$$\therefore \quad s_i[k] = \frac{\beta_i[k]}{1-\beta_i[k]} f_i[k] = \frac{\beta_i[k]}{\bar{\beta}_i[k]} f_i[k] \tag{1}$$

where $\beta_i[k] \in [0,1]$, and $\bar{\beta}_i[k] \triangleq 1 - \beta_i[k]$ has been defined to simplify the equations.

<u>Mainline conservation</u> $\quad \forall\ i \in \mathcal{I}\ ,\ k \in \mathcal{K}$:

$$\rho_i[k+1] = \rho_i[k] + f_{i-1}[k] + \delta_i r_i[k] - f_i[k] - s_i[k] \tag{2}$$

with initial and boundary conditions $\rho_i[0]$ and $f_{-1}[k]$.

<u>On-ramp conservation</u> $\quad \forall\ i \in \mathcal{E}n\ ,\ k \in \mathcal{K}$:

$$l_i[k+1] \ = \ l_i[k] \ + \ d_i[k] \ - \ r_i[k] \tag{3}$$

with initial and boundary conditions $l_i[0]$ and $d_i[k]$.

Mainline flows $\quad \forall\, i \in \mathcal{I}\ ,\ k \in \mathcal{K}$:

$$f_i[k] = \min\Big\{\ v_i(\rho_i[k] + \delta_i\gamma_i r_i[k]) - s_i[k]\ ;\ \bar{f}_i\ ; \tag{4}$$
$$w_{i+1}(\bar{\rho}_{i+1} - \rho_{i+1}[k]) - \delta_{i+1}\alpha_{i+1} r_{i+1}[k]\ ;\ \frac{\bar{\beta}_i[k]}{\beta_i[k]}\bar{s}_i\ \Big\}$$

Mainline flow $f_i[k]$ is determined in Eq. (4) by one of four terms: freeflow, mainline capacity, congestion, or off-ramp capacity.

On-ramp flows $\quad \forall\, i \in \mathcal{E}n\ ,\ k \in \mathcal{K}$:

$$r_i[k] = \tag{5}$$
$$\begin{cases} \min\{\ l_i[k] + d_i[k]\ ;\ \xi_i(\bar{\rho}_i - \rho_i[k])\ ;\ r_i^c[k]\ \} & i \in \mathcal{E}n^+ \\ \min\{\ l_i[k] + d_i[k]\ ;\ \xi_i(\bar{\rho}_i - \rho_i[k])\ \} & i \in \mathcal{E}n \backslash \mathcal{E}n^+ \end{cases}$$

On-ramp flow $r_i[k]$ is determined in Eq. (5) by one of three terms: demand, mainline congestion, or ramp control. Using (1) to eliminate $s_i[k]$ in (2) and (4):

$$\rho_i[k+1] = \rho_i[k] + f_{i-1}[k] + \delta_i r_i[k] - f_i[k]/\bar{\beta}_i[k] \tag{6}$$

$$f_i[k] = \min\Big\{\ \bar{\beta}_i[k] v_i(\rho_i[k] + \delta_i\gamma_i r_i[k])\ ;\ \bar{f}_i\ ; \tag{7}$$
$$w_{i+1}(\bar{\rho}_{i+1} - \rho_{i+1}[k]) - \delta_{i+1}\alpha_{i+1} r_{i+1}[k]\ ;\ \frac{\bar{\beta}_i[k]}{\beta_i[k]}\bar{s}_i\ \Big\}$$

The complete traffic model consists of equations (3), (5), (6), and (7). α_i, γ_i, and ξ_i respectively dictate the influence of downstream and upstream on-ramp flows on mainline flow, and the influence of mainline density on on-ramp flow.

A. *Theorem*

The following theorem provides conditions on the model parameters and boundary conditions that ensure a physically reasonable evolution of the model.

$$\left\{ \begin{array}{c} \rho_i[0] \in [0, \bar{\rho}_i]\ ;\ v_i, w_i, \alpha_i, \gamma_i \in [0,1] \\ l_i[0], \bar{f}_i, \bar{s}_i, \beta_i[k], d_i[k], f_{-1}[k], r_i^c[k] \geq 0 \\ \xi_i \in [0, \min(\frac{w_i}{\alpha_i}, \frac{1-w_i}{1-\alpha_i})] \end{array} \right\} \begin{array}{l} \forall\, k \in \mathcal{K} \\ \forall\, i \in \mathcal{I} \end{array}$$
$$\Downarrow$$
$$\left\{ \begin{array}{c} \rho_i[k] \in [0, \bar{\rho}_i] \\ l_i[k]\ ,\ f_i[k]\ ,\ r_i[k] \geq 0 \end{array} \right\} \begin{array}{l} \forall\, k \in \mathcal{K} \\ \forall\, i \in \mathcal{I} \end{array} \tag{8}$$

Proof

The proof is by induction. Assuming that $\rho_i[k] \in [0, \bar{\rho}_i]$ and $l_i[k] \geq 0$ for some k, we show that it holds for $k+1$. First, from (5), with $l_i[k] \geq 0$, $d_i[k] \geq 0$, $\xi_i \geq 0$, $\rho_i[k] \leq \bar{\rho}_i$, $r_i^c[k] \geq 0$, it follows that $r_i[k] \geq 0$. To show $f_i[k] \geq 0$, we need to check that each of the four terms in (7) is positive. The only non-obvious one is the second. However, $\xi_{i+1} \leq \frac{w_{i+1}}{\alpha_{i+1}}$ implies:

$$\xi_{i+1}(\bar{\rho}_{i+1} - \rho_{i+1}[k]) \leq \frac{w_{i+1}}{\alpha_{i+1}}(\bar{\rho}_{i+1} - \rho_{i+1}[k])$$
$$\Rightarrow\quad \delta_{i+1} r_{i+1}[k] \leq \frac{w_{i+1}}{\alpha_{i+1}}(\bar{\rho}_{i+1} - \rho_{i+1}[k])$$
$$\Rightarrow\quad w_{i+1}(\bar{\rho}_{i+1} - \rho_{i+1}[k]) - \delta_{i+1}\alpha_{i+1} r_{i+1}[k] \geq 0$$

Therefore, $f_i[k] \geq 0$. Using the above, we can deduce $l_i[k+1] \geq 0$ and $\rho_i[k+1] \in [0, \bar{\rho}_i]$:

$$\begin{aligned} l_i[k+1] &= l_i[k] + d_i[k] - r_i[k] \\ &\geq l_i[k] + d_i[k] - (l_i[k] + d_i[k]) \\ &= 0 \end{aligned}$$

$$\begin{aligned} \rho_i[k+1] &= \rho_i[k] + f_{i-1}[k] - f_i[k]/\bar{\beta}_i[k] + \delta_i r_i[k] \\ &\geq \rho_i[k] - f_i[k]/\bar{\beta}_i[k] + \delta_i r_i[k] \\ &\geq \rho_i[k] - \bar{\beta}_i[k] v_i(\rho_i[k] + \delta_i\gamma_i r_i[k])/\bar{\beta}_i[k] + \delta_i r_i[k] \\ &= (1 - v_i)\rho_i[k] + \delta_i(1 - v_i\gamma_i) r_i[k] \\ &\geq 0 \end{aligned}$$

$$\begin{aligned} \rho_i[k+1] &= \rho_i[k] + f_{i-1}[k] - f_i[k]/\bar{\beta}_i[k] + \delta_i r_i[k] \\ &\leq \rho_i[k] + f_{i-1}[k] + \delta_i r_i[k] \\ &\leq \rho_i[k] + w_i(\bar{\rho}_i - \rho_i[k]) - \delta_i\alpha_i r_i[k] + \delta_i r_i[k] \\ &= (1 - w_i)\rho_i[k] + \delta_i r_i[k](1 - \alpha_i) + w_i\bar{\rho}_i \\ &\leq (1 - w_i)\rho_i[k] + \delta_i\xi_i(\bar{\rho}_i - \rho_i[k])(1 - \alpha_i) + w_i\bar{\rho}_i \\ &= \begin{cases} (1 - w_i)\rho_i[k] + w_i\bar{\rho}_i & \text{if } i \notin \mathcal{E}n \\ (1 - \bar{w}_i)\rho_i[k] + \bar{w}_i\bar{\rho}_i & \text{if } i \in \mathcal{E}n \end{cases} \\ &\leq \bar{\rho}_i \end{aligned}$$

where $\bar{w}_i \triangleq w_i + \xi_i(1 - \alpha_i)$. The last line holds since by assumption both w_i and $\bar{w}_i \in [0,1]$. ■

IV. Problem Formulation

Papageorgiou shows in [10] that *Total Travel Time* (TTT) can be expressed as a weighted sum of off-ramp flows and other control independent terms, with weights decreasing linearly in time so as to favor earlier exiting flows. Similarly, the *Total Travel Distance* (TTD) is also a linear combination of mainline and on-ramp flows. These two popular objective functions are captured by the *Generalized Total Travel Time* (gTTT):

$$\text{gTTT} \triangleq -\sum_{k \in \mathcal{K}} \left[\sum_{i \in \mathcal{I}} a_i[k] f_i[k] + \sum_{i \in \mathcal{E}n} b_i[k] r_i[k] \right] \tag{9}$$

$a_i[k]$ and $b_i[k]$ in (9) are positive *cost weights* that can be selected to obtain a desired objective function. For example, it is shown in [10] that TTT is minimized by setting:

$$a_i[k] = \begin{cases} (\mathrm{K}-k)\ \beta_i[k]/\bar{\beta}_i[k] & i < I-1 \\ (\mathrm{K}-k)\ \big(\beta_i[k]/\bar{\beta}_i[k] + 1\big) & i = I-1 \end{cases} \tag{10}$$
$$b_i[k] = 0$$

That is, $\mathtt{TTT}$ is minimized by maximizing a weighted sum of *off-ramp flows*: $(\mathrm{K}-k)s_{i[k]}$ (the "+1" in the term for section $I-1$ weighs flows leaving through the downstream mainline boundary). $\mathtt{TTD}$ is obtained by setting $a_{i[k]}$ and $b_{i[k]}$ equal to their respective mainline and on-ramp section lengths.

Bounds on the control and state are often required. Explicit bounds on flows and densities are not necessary here due to Eq. (8). Eq. (11) places upper and lower limits on the ramp metering rates.

$$\underline{r}_i^c \;\leq\; r_{i[k]}^c \;\leq\; \bar{r}_i^c \tag{11}$$

An upper limit on the length of the on-ramp queue cannot be enforced for reasons explained in Section V-A. The nonlinear optimization problem $\mathcal{P}_A$ is now stated as follows:

Problem $\mathcal{P}_A$: Given initial and boundary conditions satisfying (8), find

$$\psi_A^* = \arg\min_{\psi\in\Omega_A} \mathtt{gTTT}(\psi) \tag{12}$$

$$\begin{aligned}\Omega_A = \Big\{ \; & \psi = \{\rho_{i[k]}, l_{i[k]}, f_{i[k]}, r_{i[k]}, r^c_{i[k]}\} \;: \\ & \text{Dynamic equations}: (3),(6), \\ & \text{Concave fundamental diagram}: (5),(7) \\ & \text{Control Bounds}: (11) \;\Big\}\end{aligned}$$

Problem $\mathcal{P}_B$ is a linear program formed by relaxing equality constraints (5) and (7).

Problem $\mathcal{P}_B$: Given initial and boundary conditions satisfying (8), find

$$\psi_B^* = \arg\min_{\psi\in\Omega_B} \mathtt{gTTT}(\psi) \tag{13}$$

$$\begin{aligned}\Omega_B = \Big\{ \; & \psi = \{\rho_{i[k]}, l_{i[k]}, f_{i[k]}, r_{i[k]}, r^c_{i[k]}\} \;: \\ & \text{Dynamic equations}: (3),(6), \\ & \text{Linear inequality constraints}: (14)-(19) \\ & \text{Control Bounds}: (11) \;\Big\}\end{aligned}$$

$\forall\, k\in\mathcal{K}, i\in\mathcal{I}$:

$$f_{i[k]} \leq \bar{\beta}_{i[k]} v_i(\rho_{i[k]} + \delta_i\gamma_i r_{i[k]}) \tag{14}$$

$$f_{i[k]} \leq w_{i+1}(\bar{\rho}_{i+1} - \rho_{i+1[k]}) \;-\; \delta_{i+1}\alpha_{i+1} r_{i+1[k]} \tag{15}$$

$$f_{i[k]} \leq \min\left\{ \bar{f}_i \;;\; \frac{\bar{\beta}_{i[k]}}{\beta_{i[k]}} \bar{s}_i \right\} \tag{16}$$

$$\forall\, k\in\mathcal{K}\;,\; i\in\mathcal{E}n: \quad r_{i[k]} \leq l_{i[k]} \;+\; d_{i[k]} \tag{17}$$

$$r_{i[k]} \leq \xi_i(\bar{\rho}_i - \rho_{i[k]}) \tag{18}$$

$$\forall\, k\in\mathcal{K}\;,\; i\in\mathcal{E}n^+: \quad r_{i[k]} \leq r^c_{i[k]} \tag{19}$$

The goal is to find a globally optimal solution to $\mathcal{P}_A$. This is difficult to do because of the non-linearities involved in Eqs. (5) and (7). Problem $\mathcal{P}_B$, on the other hand, is linear, and can therefore be solved globally and efficiently. In the following we will derive conditions on the cost weights $a_{i[k]}$ and $b_{i[k]}$ that make the two problems *equivalent*. Thus, under these conditions, the global and efficient solution to $\mathcal{P}_B$ will also solve $\mathcal{P}_A$.

V. The Cost Weights Synthesis Problem

The goal of the cost weights synthesis (CWS) problem is to find weights $a_{i[k]}$ and $b_{i[k]}$ that render problems $\mathcal{P}_A$ and $\mathcal{P}_B$ *equivalent* ($\mathcal{P}_A \equiv \mathcal{P}_B$) , in the sense that their solution sets are identical:

$$\{\psi^* \text{ solves } \mathcal{P}_A\} \;\Leftrightarrow\; \{\psi^* \text{ solves } \mathcal{P}_B\} \tag{20}$$

For $\mathcal{P}_A \equiv \mathcal{P}_B$, all of the solutions to both $\mathcal{P}_A$ and $\mathcal{P}_B$ must lie in $\Omega_A \cap \Omega_B$, which in this case equals Ω_A. Furthermore, the solutions to $\mathcal{P}_B$ must also solve $\mathcal{P}_A$. Given the first requirement, the second is trivially satisfied since $\mathcal{P}_B$ is a relaxation of $\mathcal{P}_A$. The CWS problem can therefore be stated as: Find weights $a_{i[k]}$ and $b_{i[k]}$ such that all solutions to $\mathcal{P}_B$ are in Ω_A. For this to happen, the weights must be such that any feasible solution *not* in Ω_A ($\psi \in \Omega_B \setminus \Omega_A$) is not a minimizer of $\mathtt{gTTT}$. This requirement can be expressed in terms of feasible perturbations about ψ: $\mathcal{P}_A \equiv \mathcal{P}_B$ if for every $\psi \in \Omega_B \setminus \Omega_A$ there exists a *feasible* perturbation $\boldsymbol{\Delta}$ that improves the cost. Due to the linearity of the objective, improving perturbations are characterized by $\mathtt{gTTT}(\boldsymbol{\Delta})<0$. The formulation of the CWS problem now becomes: Find weights $a_{i[k]}$ and $b_{i[k]}$ such that for all $\psi\in\Omega_B\setminus\Omega_A$ there exists a feasible perturbation $\boldsymbol{\Delta}$ with $\mathtt{gTTT}(\boldsymbol{\Delta}) < 0$.

A. The MWCC perturbation

Points $\psi \in \Omega_B \setminus \Omega_A$ can be categorized according to which nonlinear equality constraint they violate. We define member sets $\mathrm{I}_{\iota\kappa}$ and $\mathrm{II}_{\iota\kappa}$ as:

$$\mathrm{I}_{\iota\kappa} = \{\; \psi\in\Omega_B\setminus\Omega_A \;: \text{Eq. (7) not satisfied with } i=\iota\;,\; k=\kappa \;\}$$

$$\mathrm{II}_{\iota\kappa} = \{\; \psi\in\Omega_B\setminus\Omega_A \;: \text{Eq. (5) not satisfied with } i=\iota\;,\; k=\kappa \;\}$$

There are $\mathrm{I}\times\mathrm{K}$ member sets of the $\mathrm{I}_{\iota\kappa}$ type and $|\mathcal{E}n|\times\mathrm{K}$ member sets of the $\mathrm{II}_{\iota\kappa}$ type. A member of $\mathrm{I}_{\iota\kappa}$ is a point $\psi\in\Omega_B\setminus\Omega_A$ for which Eq. (7) is violated with $i=\iota$ and $k=\kappa$, or equivalently Eqs. (14), (15), and (16) with $i=\iota$ and $k=\kappa$ apply as strict inequalities.

To each of the $(\mathrm{I}+|\mathcal{E}n|)\times\mathrm{K}$ member sets corresponds a Maximal Worst-Case Causal (MWCC) perturbation; $\bar{\boldsymbol{\Delta}}^{\mathrm{I}}_{\iota\kappa}$ for $\mathrm{I}_{\iota\kappa}$ and $\bar{\boldsymbol{\Delta}}^{\mathrm{II}}_{\iota\kappa}$ for $\mathrm{II}_{\iota\kappa}$. The definition of the MWCC perturbations is given below. The MWCC is a feasible perturbation for all points in its corresponding member set because it is feasible for the *worst-case* point, where all inequality constraints, aside from the one that defines it, are active. It is *maximal* because it selects the largest (least negative) feasible value for each of the Δf's and Δr's. This is done in order to maximize its beneficial effect on the cost.

$\bar{\mathbf{\Delta}}^{\mathrm{I}}_{\iota\kappa}$ is defined for $\kappa\in\mathcal{K},\ \iota\in\mathcal{I}$ as: (21)

$$\Delta\rho_i[k+1] = \Delta\rho_i[k] + \Delta f_{i-1}[k] - \Delta f_i[k]/\bar{\beta}_i[k] + \delta_i \Delta r_i[k]$$
$$\Delta l_i[k+1] = \Delta l_i[k] - \Delta r_i[k]$$
$$\Delta f_i[k] = \begin{cases} 1 & i=\iota\ ,\ k=\kappa \\ \min\Big\{ \bar{\beta}_i[k] v_i \left(\Delta\rho_i[k] + \delta_i\gamma_i\Delta r_i[k]\right)\ ; & k>\kappa \\ \quad -w_{i+1}\Delta\rho_{i+1}[k] - \delta_{i+1}\alpha_{i+1}\Delta r_{i+1}[k]\ ;\ 0 \Big\} & \\ 0 & k<\kappa \end{cases}$$
$$\Delta r_i[k] = \min\big\{\Delta l_i[k]\ ;\ -\xi_i\Delta\rho_i[k]\ ;\ 0\big\}$$
$$\Delta r^c_i[k] = 0$$

$\bar{\mathbf{\Delta}}^{\mathrm{II}}_{\iota\kappa}$ is defined for $\kappa\in\mathcal{K},\ \iota\in\mathcal{E}n$ as: (22)

$$\Delta\rho_i[k+1] = \Delta\rho_i[k] + \Delta f_{i-1}[k] - \Delta f_i[k]/\bar{\beta}_i[k] + \delta_i \Delta r_i[k]$$
$$\Delta l_i[k+1] = \Delta l_i[k] - \Delta r_i[k]$$
$$\Delta f_i[k] = \min\Big\{ \bar{\beta}_i[k] v_i \left(\Delta\rho_i[k] + \delta_i\gamma_i\Delta r_i[k]\right)\ ;\ - w_{i+1}\Delta\rho_{i+1}[k] - \delta_{i+1}\alpha_{i+1}\Delta r_{i+1}[k]\ ;\ 0 \Big\}$$
$$\Delta r_i[k] = \begin{cases} 1 & i=\iota\ ,\ k=\kappa \\ \min\big\{\Delta l_i[k]\ ;\ -\xi_i\Delta\rho_i[k]\ ;\ 0\big\} & k>\kappa \\ 0 & k<\kappa \end{cases}$$
$$\Delta r^c_i[k] = 0$$

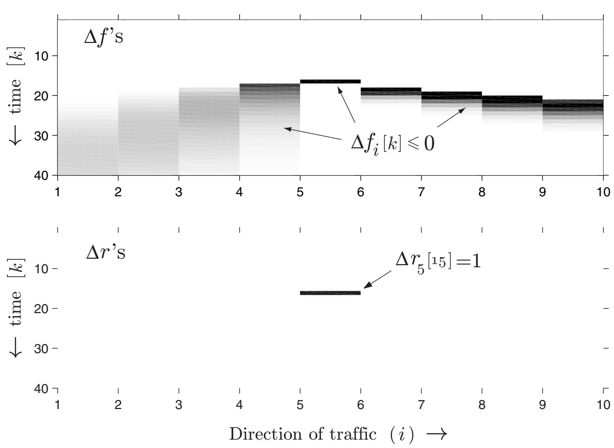

Fig. 2. $\bar{\mathbf{\Delta}}^{\mathrm{II}}_{5,15}$ with parameters of Section VI

A fact not shown here is that a general perturbation such as the MWCC perturbation cannot be defined if any of the following constraints are included in $\mathcal{P}_A$ and $\mathcal{P}_B$: $f_i[k]\geq 0$, $r_i[k]\geq 0$, $0\leq l_i[k]\leq \bar{l}_i$, $0\leq\rho_i[k]\leq\bar{\rho}_i$. This observation underscores the importance of the theorem, which guarantees all except $l_i[k]\leq\bar{l}_i$. It is also the reason why queue length constraints have been omitted.

Each of the $(\mathrm{I}+|\mathcal{E}n|)\times\mathrm{K}$ MWCC perturbations can be computed offline given the layout of the freeway, the model parameters, and the off-ramp split ratios. Figure 2 shows a sample MWCC perturbation ($\bar{\mathbf{\Delta}}^{\mathrm{II}}_{5,15}$) with the parameters of Section VI. Here, a unit increase in $r_5[15]$ produces negative waves that propagate forward in time, and upstream and downstream along the freeway.

With all of the MWCC perturbations computed, the CWS problem can be restated as: Find $a_i[k]$ and $b_i[k]$ such that:

$$\mathtt{gTTT}(\bar{\mathbf{\Delta}}^{\mathrm{I}}_{\iota\kappa}) < 0 \qquad \forall\ \kappa\in\mathcal{K}\ ,\ \iota\in\mathcal{I} \tag{23}$$
$$\mathtt{gTTT}(\bar{\mathbf{\Delta}}^{\mathrm{II}}_{\iota\kappa}) < 0 \qquad \forall\ \kappa\in\mathcal{K}\ ,\ \iota\in\mathcal{E}n \tag{24}$$

Because each MWCC perturbation is feasible for every point in its member set, conditions (23) and (24) are sufficient to guarantee non-optimality for all $\psi\in\Omega_B\backslash\Omega_A$, and thus $\mathcal{P}_A\equiv\mathcal{P}_B$.

B. Backstepping Numerical Method

Using Eq. (9), Eqs. (23) and (24) are expressed as $(\mathrm{I}+|\mathcal{E}n|)\times\mathrm{K}$ linear equations of the form:

$$-\sum_{k\in\mathcal{K}}\left[\sum_{i\in\mathcal{I}} a_i[k]\Delta f_i[k] + \sum_{i\in\mathcal{E}n} b_i[k]\Delta r_i[k]\right] = -\epsilon \tag{25}$$

where Δf's and Δr's are components of $\bar{\mathbf{\Delta}}^{\mathrm{I}}_{\iota\kappa}$ or $\bar{\mathbf{\Delta}}^{\mathrm{II}}_{\iota\kappa}$, and ϵ is a positive number. Because the MWCC perturbation is *causal*, the summations in Eq. (25) only contain non-zero Δf's and Δr's for $k\geq\kappa$. They therefore have the following recursive formula:

$$a_\iota[\kappa] = -\sum_{k=\kappa+1}^{\mathrm{K}-1}\left(\sum_{i\in\mathcal{I}} a_i[k]\Delta f_i[k] + \sum_{i\in\mathcal{E}n} b_i[k]\Delta r_i[k]\right) + \epsilon \tag{26}$$

$$b_\iota[\kappa] = -\sum_{k=\kappa+1}^{\mathrm{K}-1}\left(\sum_{i\in\mathcal{I}} a_i[k]\Delta f_i[k] + \sum_{i\in\mathcal{E}n} b_i[k]\Delta r_i[k]\right) - a_\iota[\kappa]\Delta f_\iota[\kappa] + \epsilon \tag{27}$$

Δf's and Δr's in (26) are components of $\bar{\mathbf{\Delta}}^{\mathrm{I}}_{\iota\kappa}$, and Δf's and Δr's in (27) are components of $\bar{\mathbf{\Delta}}^{\mathrm{II}}_{\iota\kappa}$. These equations can be easily solved by setting $a_i[\mathrm{K}-1]$ and $b_i[\mathrm{K}-1]$ to some positive value, and computing the rest sequentially backwards.

VI. EXAMPLE

The CWS problem was solved for a simple test freeway consisting of 10 sections and 40 time intervals. A single on-ramp was placed at $i = 5$. Off-ramps were placed at $i = 4$, 5, and 9 with $\beta_{i[k]} = 0.1$ for all three. $v_i = 0.7$ and $w_i = 0.2$ was used throughout.

Figure 3 shows cost weights resulting from $\alpha_5 = \gamma_5 = 0.2$, and $\xi_5 = 0.06$. A *time decay index* (D) was computed for each sequence $a_i[\cdot]$ and $b_i[\cdot]$ as the number of entries in the sequence that exceeded 10% of the first value, divided by the length of the sequence; for example, $D(a_4[\cdot]) = \text{size}\{a_4[\cdot] \geq 0.1 \times a_4[0]\}/40$. The decay index for a constant sequence is 1.0, and the decay index for a linearly decreasing sequence, such as the TTT weights of Eq. (10), is 0.9. We inspect the decay indices because they give some measure of the similarity of gTTT with TTT. A minimum of 0.9 or higher among the decay indices for all $a_i[\cdot]$'s would suggest that the resulting objective function is "close" to total travel time. The decay indices for the sequences of Figure 3 are $D(a_4[\cdot]) = 0.51$ and $D(b_5[\cdot]) = 0.54$.

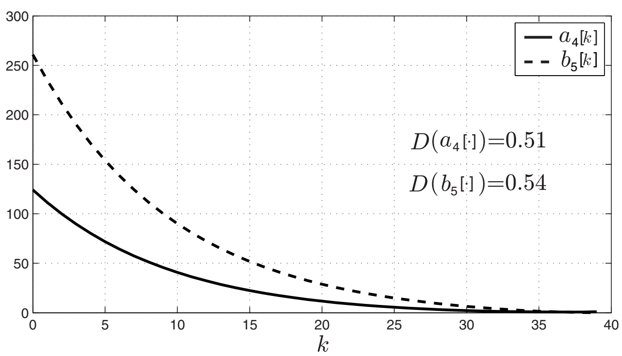

Fig. 3. Cost weights with $\alpha_5 = \gamma_5 = 0.2$ and $\xi_5 = 0.06$

Figure 4 shows $\min_i\{D(a_i[\cdot])\}$ (i.e. the decay index of the fastest decaying mainline weight) in the top window, and $\min_i\{D(b_i[\cdot])\}$ in the bottom window, as functions of ξ_5, and for several values of α_5. The decay index was found to be insensitive to γ_5. The figure shows that the cost weights degrade less quickly, and are therefore more similar to TTT, for smaller values of α_i and ξ_i.

VII. CONCLUSIONS

In this paper we have investigated the possibility of finding a global solution to the on-ramp metering problem by solving a linear program. A variant of the CTM was formulated as the freeway model, with beneficial properties given by Eq. (8). The question of whether the target and relaxed problems are equivalent was transformed into a condition on the weights of the gTTT objective. A numerical algorithm was devised for finding weights that meet the specification. A simple example showed that the cost weights found with the algorithm differ from TTT (Eq.(10)) in two ways: $D(a_i[\cdot]) < 0.9$ (fast decay) and $b_{i[k]} > 0$ (positive on-ramp weights). Part II deals specifically with these problems and applies the technique to a realistic freeway configuration.

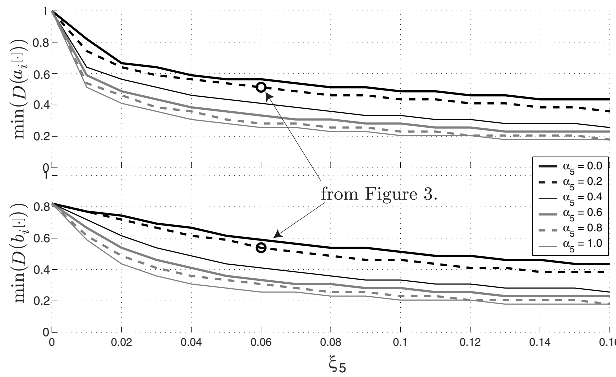

Fig. 4. Fastest decay index vs. ξ_5 and α_5. ($\gamma_5 = 0.2$)

REFERENCES

[1] J. Wattleworth and D. Berry. Peak-period control of a freeway system - Some theoretical investigations. *Highway Research Record*, (89), 1965.

[2] L. Yuan and J. Kreer. Adjustment of freeway ramp metering rates to balance entrance ramp queues. *Trans. Research*, 5, 1971.

[3] C. Chen, J. Cruz, and J. Paquet. Entrance ramp control for travel-rate maximization in expressways. *Trans. Research*, 8, 1974.

[4] J. Wang and A. D. May. Computer model for optimal freeway on-ramp control. *Highway Research Record*, 469, 1973.

[5] H. Payne and W. Thompson. Allocation of freeway ramp metering volumes to optimize corridor performance. In *IEEE Transactions on Automatic Control*, volume 19, 1974.

[6] Y. Iida, T. Hasegawa, Y. Asakura, and C. Shao. A formulation of on-ramp traffic control system with route guidance for urban expressway. *Communications in Transportation*, 1989.

[7] Y. Lei. Model and technology of the on-ramp control over freeway traffic flow. *Transportation systems : Theory and application of advanced technology*, 2, 1995.

[8] A. Kotsialos, M. Papageorgiou, M. Mangeas, and H. Hadj-Salem. Coordinated and integrated control of motorway networks via nonlinear optimal control. *Transportation Research. Part C, Emerging technologies*, 10(1), 2002.

[9] A. Hegyi, B. De Schutter, H. Hellendoorn, and T. Van Den Boom. Optimal coordination of ramp metering and variable speed control - An MPC approach. In *Proceedings of the American Control Conference*, 2002.

[10] M. Papageorgiou and A. Kotsialos. Freeway ramp metering: An overview. In *IEEE Intelligent Transportation Systems*, 2000.

[11] M. Lighthill and G. Whitham. On kinematic waves II. A theory of traffic flow on long crowded roads. *Proceedings Royal Society of London, Part A*, 229(1178), 1955.

[12] M. Papageorgiou. An integrated control approach for traffic corridors. *Transportation Research. Part C*, 3(1), 1995.

[13] A. Ziliaskopoulos. A linear programming model for the single destination system optimum dynamic traffic assignment problem. *Transportation Science*, 34(1), 2000.

[14] C. Daganzo. The cell transmission model: A dynamics representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research. Part B.*, 28(4), 1994.