JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 126, No. 3, pp. 589–616, September 2005 (© 2005)
DOI: 10.1007/s10957-005-5499-z

Optimal Control for Traffic Flow Networks

M. Gugat,[1] M. Herty,[2] A. Klar,[3] and G. Leugering[4]

Communicated by H. J. Pesch

Abstract. We consider traffic flow models for road networks where the flow is controlled at the nodes of the network. For the analytical and numerical optimization of the control, the knowledge of the gradient of the objective functional is useful. The adjoint calculus introduced below determines the gradient in two ways. We derive the adjoint equations for the continuous traffic flow network model and derive also the adjoint equations for a discretized model. Numerical examples for the solution of problems of optimal control for traffic flow networks are presented.

Key Words. Traffic flows, networks, hyperbolic systems, adjoint systems.

1. Introduction

Modelling and simulation of traffic flow on highways has been investigated intensively during the last years; see for example Refs. 1–6. For the present investigation, we are interested in traffic flow models for road networks using models based on hyperbolic partial differential equations, so-called fluidodynamic models of traffic flow. Two such models based upon a single conservation law were introduced by Holden/Risebro in Ref. 7 and by Coclite/Piccoli in Ref. 8. To model the traffic flow in a network of roads, conditions that govern the flow at the junction are necessary. With the conditions from Ref. 8, a unique solution is determined on the network. For

[1]Lecturer, Institut für Angewandte Mathematik, FAU Erlangen-Nürnberg, Erlangen, Germany.
[2]Research Assistant, Fachbereich Mathematik, TU Darmstadt, Darmstadt, Germany. This author was supported by Deutsche Forschungsgemeinschaft (DFG), Grant KL 1105/5.
[3]Professor, Fachbereich Mathematik, TU Darmstadt, Darmstadt, Germany.
[4]Professor, Insitut für Angewandte Mathematik, FAU Erlangen-Nürnberg, Erlangen, Germany.

different conditions at the junctions, we refer also to the traffic engineering literature; see for example Refs. 9–10.

To support decision making in traffic management, problems of optimal control of the flow in traffic networks are of interest. In this paper, such optimization problems are studied. The main focus is on the evaluation of the gradient of the objective functions that appear in these problems. This is done with an adjoint system using first the continuous pde model and second a discretized version of the pde model. A comparison of both approaches is given. Adjoint systems for conservation laws are derived also in Refs. 11–14. For the application of the adjoint systems to shallow-water equations on networks, we refer to Refs. 15–18.

In this context, controllability problems have been considered in Refs. 18–20. The adjoint systems are used for the solution of optimization problems for the flow in networks of open channels that have a structure similar to that of the problems studied in this paper. In contrast to the models for traffic flow studied here, the shallow water equations are a system of two conservation laws which allows information to travel both upstream and downstream at the same time. Models using systems of two conservation laws have been proposed also for traffic flow; see Ref. 4. Similar problems appear also in the management of networks of gas pipelines.

2. Continuous Model of Traffic Flow and Adjoint Equations

2.1. Continuous Model of Traffic Flow. The Coclite/Piccoli model for traffic flow networks is described in detail in Ref. 8. We give only a brief review here.

Our network has a total of J roads. For $j \in \{1, \dots, J\}$, road j corresponds to an interval $[a_j, b_j]$. Let $\rho_j(x,t)$ denote the density of cars on road j at the point $x \in [a_j, b_j]$ and time $t \in [0, T]$. The roads correspond to the edges of a graph, and the junctions where the roads are connected correspond to the nodes of this graph. For each $j \in \{1, \dots, J\}$, the unidirectional heavy traffic flow on road j is modelled by the following Lighthill-Whitham equations (see Ref. 21):

$$\partial_t \rho_j(x,t) + \partial_x f_j(\rho_j(x,t)) = 0, \quad \forall x \in [a_j, b_j], t \in [0, T], \tag{1a}$$

$$\rho_j(x,0) = \rho_{j,0}(x), \quad \forall x \in [a_j, b_j], \tag{1b}$$

with flux

$$f_j(\rho) := \rho U_j(\rho),$$

where U_j is the velocity. By assumption, the velocity U_j is a continuously differential function of only the density. As in Refs. 7–8, we assume that

f_j satisfies

$$f_j(0) = f_j(\rho_j^{\max}) = 0$$

and that there exists $\sigma_j \in (0, \rho_j^{\max})$ such that

$$f_j'(\sigma_j) = 0 \quad \text{and} \quad (\rho - \sigma_j) f_j'(\rho) < 0, \quad \forall \rho \neq \sigma_j, \tag{2}$$

with f_j concave. This condition is fulfilled by any reasonable model for the fundamental diagram; see for example Ref. 22. it implies that we have fixed the positive direction of flow on each road of the network.

We assume boundary conditions for roads which enter or leave the network given in the sense of Bardos, LeRoux, and Nedelec (Ref. 23). We describe briefly the coupling conditions at the nodes as stated in Coclite/Piccoli (Ref. 8). Similar considerations can be found in Holden/Risebro (Ref. 7).

We consider a single junction with n roads labeled by $j = 1, \ldots, n$, with end b_j at the junction, and m roads labeled by $j = n+1, \ldots, n+m$, with end a_j at the junction. To guarantee the conservation of the numbers of cars, the following conditions is prescribed at the junction:

$$\sum_{j=1}^{n} f_j(\rho_j(b_j, t)) = \sum_{j=n+1}^{n+m} f_j(\rho_j(a_j, t)), \quad \forall t \geq 0. \tag{3}$$

However, this condition does not suffice to determine a unique solution on the network.

Coclite and Piccoli introduce a matrix $A \in R^{m \times n}$, where

$$(A)_{ji} = a_{ji}, \quad j \in \{n+1, \ldots, n+m\}, \quad i \in \{1, \ldots, n\},$$

describes the percentages of drivers who want to drive from road i to road j. The matrix A is assumed to fulfill the following assumptions:

$$a_{ij} \neq a_{ji'}, \forall i \neq i', \quad 0 < a_{ji} < 1, \quad \sum_{j=n+1}^{n+m} a_{ji} = 1, \quad \forall i \in \{1, \ldots, n\}. \tag{4}$$

We state the boundary conditions at a junction for weak solution as in Ref. 8.

A weak solution at a junction is a collection of functions $\rho_j : [0, \infty) \times [a_j, b_j] \to \mathbb{R}$, for $i = 1, \ldots, n+m$, such that

$$\sum_{i=1}^{n+m} \int_0^\infty \int_{a_j}^{b_j} (\rho_i \partial_t \phi_i + f(\phi_i) \partial_x \phi_i) dx\, dt = 0, \tag{5}$$

for each $\phi_i, i=1,\dots,n+m$, smooth and having compact support in $\mathbb{R}\times(0,\infty)$ and smooth across the junction, i.e.,

$$\phi_i(b_i,\cdot)=\phi_j(a_j,\cdot),$$
$$\partial_x\phi_i(b_i,\cdot)=\partial_x\phi_j(a_j,\cdot),\quad i=1,\dots,n \text{ and } j=n+1,\dots,n+m.$$

Note, that (5) implies (3) if the functions ρ_i are sufficiently regular. If $\rho_j(t,\cdot)$ are functions of bounded variation, we assume additionally that the following properties are satisfied:

$$f(\rho_j(a_j+,\cdot))=\sum_{i=1}^{n}\alpha_{ji}f(\rho_i(b_i-,\cdot)),\quad j=n+1,\dots,n+m, \tag{6}$$

$$\sum_{i=1}^{n}f(\rho_i(b_i-,\cdot))+\sum_{i=n+1}^{n+m}f(\rho_i(a_i+,\cdot)) \text{ is maximal w. r. t. (6).} \tag{7}$$

Following Coclite/Piccoli, the Cauchy problem for given initial data $\bar{\rho}_j$ and boundary date ψ_j possesses a solution in the following sense.

Definition 2.1. See Definition 2.2 in Ref. 8. Given $\bar{\rho}_j:[a_j,b_j]\to\mathbb{R}$, $j=1,\dots,J$, and possibly ψ_j functions of L^∞, a collection of functions $\rho=(\rho_1,\dots,\rho_J)$, with $\rho_j:[a_j,b_j]\times[0,\infty)$ in $C([0,\infty);L^1_{loc}([a_j,b_j]))$, is an admissible solution if ρ_j is a weak entropic solution to (1), $\rho_j(\cdot,0)=\bar{\rho}_j, \rho_j(b_j,t)=\psi_j$ in the sense of Ref. 23, and such that, at each junction, ρ is a weak solution in the sense of (5). If $\rho_j(\cdot,t)$ is of bounded variation, we require additionally that ρ satisfies (6) and (7).

The following result concerning existence and uniqueness is known.

Theorem 2.1. See Theorem 8.2 of Ref. 8. Consider a flux function f satisfying (2) and a road network in which all the junctions have at most two ingoing roads and two outgoing roads. Let $\bar{\rho}=(\rho_1,\dots,\rho_J)$ be the initial date in L^1_{loc} and let $T>0$ be fixed. Then, there exists a unique admissible solution $\rho=(\rho_1,\dots,\rho_J), \rho_j:[a_j,b_j]\times[0,T]\to\mathbb{R}$, with $\rho(\cdot,0)=\bar{\rho}$.

The main step in the proof of existence and uniqueness of admissible weak solutions is the consideration of constant initial data $\rho_{j,0}$ and for one junction only. An admissible solution can be constructed as follows. For each road j, we introduce an intermediate state $\bar{\rho}_j\in\mathbb{R}$, $j=1,\dots,n+m$. The function $\rho_j(x,t)$ solving problem (1) and (3) is given as solution to a Riemann problem on each road j. For incoming roads, the initial conditions for the Riemann problem are

$$\rho_j(x,0)=\begin{cases}\rho_{j,0}, & x\le b_j,\\ \bar\rho_j, & x>b_j.\end{cases} \tag{8}$$

and similarly for the outgoing roads. Hereby, we impose certain restrictions to the values $\bar\rho_j$. We assume $\bar\rho_j$ to be independent of time; that is, all the waves of the Riemann problems have to emerge from the junction. Depending on the value of $\rho_{j,0}$, this implies restrictions to $\bar\rho_j$, specifically,

$$\bar\rho_j\in[\sigma_j,\rho_j^{\max}],\qquad \rho_{j,0}\ge\sigma_j,\quad j=1,\dots,n, \tag{9a}$$

$$\bar\rho_j\in\{\rho_{j,0}\}\cup(\tau_j(\rho_{j,0}),\rho_j^{\max}],\qquad \rho_{j,0}\le\sigma_j,\quad j=1,\dots,n, \tag{9b}$$

$$\bar\rho_j\in[0,\sigma_j],\qquad \rho_{j,0}\le\sigma_j,\quad j=n+1,\dots,n+m, \tag{9c}$$

$$\bar\rho_j\in[0,\tau_j(\rho_{j,0}))\cup\{\rho_{j,0}\},\qquad \rho_{j,0}\ge\sigma_j,\quad j=n+1,\dots,n+m, \tag{9d}$$

where, for each j and $x\ne\sigma_j$, the value $\tau_j(x)$ is the unique number $\tau_j(x)\ne x$ such that $f_j(x)=f_j(\tau_j(x))$.

We impose more constraints to $\bar\rho_j$, which corresponds to the drivers' intentions modeled in the matrix A,

$$f_j(\bar\rho_j)=\sum_{i=1}^{n}a_{ji}f_i(\bar\rho_i),\quad \forall j=n+1,\dots,n+m. \tag{10}$$

The assumptions on A imply that these conditions guarantee also that $\rho_j(x,t)$ fulfills the coupling condition (3).

Let the function E measuring the flux be defined as follows:

$$E(\bar\rho_1,\dots,\bar\rho_{n+m})=\sum_{j=1}^{n+m}f_j(\bar\rho_j). \tag{11}$$

The following problem has a unique solution (see Theorem 3.1 in Ref. 8.):

$$\max E,\quad \text{s.t. (10) and (9).} \tag{12}$$

By (12), the values $\bar\rho_j$, $j=1,\dots,n+m$, are determined uniquely and we obtain a solution $\rho_j(x,t)$ for constant initial data by solving the $n+m$ Riemann problems. The solution to a problem with nonconstant initial data is obtained by wave tracking or front tracking; see Refs. 24–25. This procedure applies also to networks with more than one junction; see Refs. 7–8.

The motivation to model the flow through a junction by (12) is that, while respecting their intentions, the drivers choose their path in such a way that the total flux is maximized.

2.2. Continuous Adjoint Equations. For the subsequent derivation, we assume a fixed positive direction of the flow. In the context of Coclite/Piccoli, this is equivalent to imposing the following restriction on the densities ρ_j:

$$\rho_j(x,t) < \sigma_j, \quad \forall_j \in \{1,\dots,J\}, x \in [a_j, b_j], t \in [0,T]. \tag{13}$$

This condition is satisfied if we consider inflow problems to networks with suitable large maximum flux values on each road. The condition is violated in the case of backward going shock waves. We exclude also network graphs with a loop, as traffic along a loop will eventually end up at the same branching node, which is an unrealistic situation in this context.

For simplicity, we assume that the network is initially empty; that is,

$$\rho_j(x,0) = 0, \quad \forall j \in \{1,\dots,J\},\ x \in [a_j, b_j].$$

The more general case can be treated equally well.

Finally, for the sake of simpler notation, we restrict our attention to graphs with junctions of edge degree three. Again, the more general case can be handled also with little extra effort. There are two types of such junctions. The first junction type has an ingoing street m with the end b_m at the junction and two outgoing streets labeled r, s with ends a_r, a_s at the junction. At such junction, a real-valued control $\alpha_m \in \mathbb{R}$ is applied. This control is the flux distribution factor a_{mr} in the Coclite/Piccoli context. At the second type of junctions, the roads p and q, with ends b_p and b_q at the junction, merge to road r with end a_r at the junction; the traffic flow is not controlled, since it is uniquely determined by the conservation of cars. Furthermore, there are inflow arcs and outflow arcs to the network.

In the case of junctions with a total of three roads, the described coupling conditions of Coclite/Piccoli can be expressed in terms of the boundary conditions at $x = a_j$ for all roads j in the network. To be more specific, we may write the following conditions with functions u_j that depend on the functions f_j and describe the solution of problem (12) as a function of the parameters $\rho_m(b_m, t)$, α_m for junctions of the first type and as a function of $\rho_p(b_p, t)$, $\rho_q(b_q, t)$ at junctions of the second type. See Figure 1 for the notation. In Section 4, we state explicitly the definition of u_j.

(A1) Boundary Values for Junctions of the First Type:

$$f_r(\rho_r(a_r, t)) = \alpha_m f_m(\rho_m(b_m, t)), \tag{14a}$$

$$f_s(\rho_s(a_s, t)) = (1-\alpha_m) f_m(\rho_m(b_m, t)), \tag{14b}$$

$$\rho_r(a_r, t) = u_r(\rho_m(b_m, t), \alpha_m), \tag{14c}$$

$$\rho_s(a_s, t) = u_s(\rho_m(b_m, t), \alpha_m). \tag{14d}$$

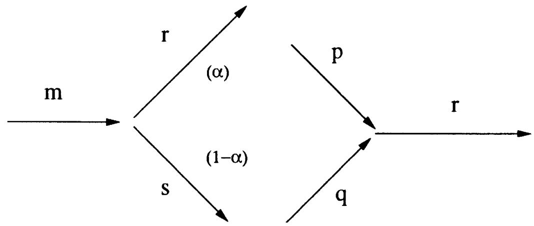

Fig. 1. Labeling of the roads connected to the junction (first type on the left, second type on the right).

(A2) Boundary Values for Junctions of the Second Type:

$$f_r(\rho_r(a_r,t)) = f_p(\rho_p(b_p,t)) + f_q(\rho_q(b_q,t)), \tag{15a}$$

$$\rho_r(a_r,t) = u_r(\rho_p(b_p,t), \rho_q(b_q,t)). \tag{15b}$$

(A3) Boundary Values for Road s with a Free Boundary Node at End a_s:

$$\rho_s(a_s,t) = u_s(t). \tag{16}$$

Remark 2.1. The functions u_r and u_s can be derived from (9), (11) and can be given explicitly, since we assume (13). To be more precise, we have that the initial network is empty. In a controlled junction, for the solution of (12), on account of the restriction $\rho_j < \sigma_j$, we have

$$f_m(\bar{\rho}_m) = f_m(\rho_m),$$
$$f_r(\bar{\rho}_r) = \alpha_m f(\rho_m),$$
$$f_s(\bar{\rho}_s) = (1-\alpha_m) f_m(\rho_m).$$

Since we assumed (13), we can invert $f_{r,s,m}(\cdot)$ and obtain an explicit formulation for the boundary values $\bar{\rho}_m, \bar{\rho}_r, \bar{\rho}_s$. Thus, the node conditions of Coclite/Piccoli (Ref. 8.) can be written as

$$u_r = u(\rho_m, \alpha), \quad u_s = u(\rho_m, 1-\alpha).$$

Also, if the function $\rho_m(b_m,t)$ for $t>0$ is smooth, we obtain a smooth dependence of $u_{r,s}$ on the control α_m. The derivation is similar in the case of a merging junction.

For example, let

$$f_r(\rho) = f_s(\rho) = 4\rho(1-\rho).$$

This yields

$$\bar{\rho}_r = \left(1 - \sqrt{1 - \alpha f(\rho_m)}\right)/2,$$
$$\bar{\rho}_s = \left(1 - \sqrt{1 - (1-\alpha) f(\rho_m)}\right)/2,$$

and

$$u(y, x) = (1 - \sqrt{1 - x f(y)})/2. \tag{17}$$

For the merging junction, we have

$$f(\bar{\rho}_r) = f(\rho_p) + f(\rho_q);$$

since $\bar{\rho}_r \in [0, \sigma_r]$, we obtain for our example

$$\begin{aligned} u_r &= u(\rho_p, \rho_q) \\ &= \left(1 - \sqrt{1 - f(\rho_p) - f(\rho_q)}\right)/2. \end{aligned} \tag{18}$$

Assume that we have K controls, that is, K junctions of the first type. The objective functional is of the type

$$F(\alpha_1, \ldots, \alpha_K) = \sum_{j=1}^{J} \int_0^T \int_{a_j}^{b_j} \mathcal{F}_j(\rho_j(x,t))\,dx\,dt, \tag{19}$$

with given smooth functions $\mathcal{F}_j$: $\mathbb{R} \to \mathbb{R}$. Consider the optimization problem

$$\min_{\alpha_1,\ldots,\alpha_K} \quad F(\alpha_1, \ldots, \alpha_K), \tag{20a}$$
$$\text{s.t.} \quad 0 \le \alpha_i \le 1, \tag{20b}$$

and such that $\rho_j(x,t)$ is a solution of (1), (14), (15), (16).

In order to determine the gradient $\nabla F(\alpha)$, we use the following adjoint system:

$$\partial_t \mu_j(x,t) + f_j'(\rho_j(x,t))\partial_x \mu_j(x,t) = \mathcal{F}_j'(\rho_j(x,t)), \quad \forall x \in [a_j, b_j],\ t \in [0, T], \tag{21a}$$
$$\mu_j(b_j, t) = u_j^*(t), \qquad \forall t \in [0, T], \tag{21b}$$
$$\mu_j(x, T) = 0, \qquad \forall x \in [a_j, b_j]. \tag{21c}$$

The functions $u_j^*(t)$ are the adjoint boundary and junction conditions and are given by the following relations.

(B1) Boundary Values for Junctions of the First Type. For this type of junction,

$$f_m'(\rho_m(b_m,t))u_m^*(t)=\partial_\rho u_r(\rho_m(b_m,t),\alpha_m)f_r'(\rho_r(a_r,t))\mu_r(a_r,t)$$
$$+\partial_\rho u_s(\rho_m(b_m,t),\alpha_m)f_s'(\rho_s(a_s,t))\mu_s(a_s,t), \tag{22}$$

or equivalently

$$u_m^*(t)=\alpha_m\mu_r(a_r,t)+(1-\alpha_m)\mu_s(a_s,t).$$

(B2) Boundary Values for Junctions of the Second Type. For this type of junction,

$$f_p'(\rho_p(b_p,t))u_p^*(t)=\partial_{\rho_p}u_r(\rho_p(b_p,t),\rho_q(b_q,t))f_r'(\rho_r(a_r,t))\mu_r(a_r,t),$$
$$f_q'(\rho_q(b_q,t))u_q^*(t)=\partial_{\rho_q}u_r(\rho_p(b_p,t),\rho_q(b_q,t))f_r'(\rho_r(a_r,t))\mu_r(a_r,t),$$

or equivalently

$$u_p^*(t)=\mu_r(a_r,t),\quad u_q^*(t)=\mu_r(a_r,t).$$

(B3) Boundary Values for the Outflow Arc n_0. That is,

$$u_{n_0}^*(t)=0.$$

Note that $f_j'\neq 0$ by the assumption on the flow direction, hence $u_j^*(t)$ is well-defined. Then, the gradient $\nabla F(\alpha)$ is given by

$$\partial_{\alpha_m}F(\alpha_1\dots,\alpha_K)=\int_0^T -\mu_r(a_r,t)f_r'(u_r(\rho_m(b_m,t),\alpha_m))\partial_\alpha u_r(\rho_m(b_m,t),\alpha_m)$$
$$-\mu_s(a_s,t)f_s'(u_s(\rho_m(b_m,t),\alpha_m))\partial_\alpha u_s(\rho_m(b_m,t),\alpha_m)dt$$
$$=\int_0^T[\mu_s(a_s,t)-\mu_r(a_r,t)]f_m(\rho_m(b_m,t))dt. \tag{23}$$

Remark 2.2. In order to simplify the computations, note that the adjoint equation can be transformed to a forward equation,

$$v_t(x,t)-f_j'(\rho_j(x,T-t))v_x(x,t)=0,\quad \forall x\in[a_j,b_j],\ \ t\in[0,T],$$
$$v(b_j,t)=u_j^*(T-t),\ \ \forall t\in[0,T],$$
$$v(x,0)=0,\ \ \forall x\in[a_j,b_j].$$

We assume classical solutions $\rho_j, \mu_j, j = 1, \ldots, J$, to the forward and adjoint equation. We assume $\rho_{j,0} = 0$ and a circle-free network with one inflow arc. Then, the adjoint system can be derived by the following considerations. We introduce the Lagrange function L depending on $\alpha = (\alpha_k)_k,\ k = 1, \ldots, K,\ \rho = (\rho_j)_j$, and $\mu = (\mu_j)_j, j = 1, \ldots, J$,

$$L(\alpha, \rho, \mu) = \sum_{j=1}^{J} \int_0^T \int_{a_j}^{b_j} \mathcal{F}_j(\rho_j) dx\, dt + \sum_{j=1}^{J} \int_0^T \int_{a_j}^{b_j} \mu_j (\partial_t \rho_j + \partial_x f_j(\rho_j)) dx\, dt.$$

Linearization of $L(\alpha, \rho, \mu)$ and partial integration yields

$$\begin{aligned} \partial_{\alpha_m} L(\alpha, \rho, \mu) = & \sum_{j=1}^{J} \int_0^T \int_{a_j}^{b_j} \left(\mathcal{F}_j'(\rho_j) - \mu_t - f'(\rho_j)\mu_x \right) \partial_{\alpha_m} \rho_j \, dx\, dt \\ & + \sum_{j=1}^{J} \int_{a_j}^{b_j} \mu_j(\cdot, T) \partial_{\alpha_m} \rho_j(\cdot, T) - \mu_j(\cdot, 0) \partial_{\alpha_m} \rho_{j,0} \, dx \\ & + \sum_{j=1}^{J} \int_0^T \mu(b_j, \cdot) f_j'(\rho_j(b_j, \cdot)) \partial_{\alpha_m} \rho_j(b_j, \cdot) \\ & \qquad - \mu(a_j, \cdot) \partial_{\alpha_m} f_j(\tilde{v}_j(\rho, \alpha)) dt. \end{aligned}$$

Herein, $\tilde{v}_j$ denotes the boundary value for road j at a_j. The exact formulation depends on the type of junctions to which road j is connected. For example, let $\tilde{v}_j$ for $j = r$ be given by

$$\tilde{v}_r = u_r(\rho_m(b_m, t), \alpha_m).$$

Then, we obtain for the last term of the above equality

$$\begin{aligned} \int_0^T & \big[-\mu_r(a_r, \cdot) f_r'(u_r(\rho_m(b_m, \cdot), \alpha_m)) \partial_{\rho_r} u_r(\rho_m(b_m, \cdot), \alpha_m) \partial_{\alpha_m} \rho_m(b_m, \cdot) \\ & - \mu_r(a_r, \cdot) f_r'(u_r(\rho_m(b_m, \cdot), \alpha_m)) \partial_\alpha u_r(\rho_m(b_m, \cdot), \alpha_m) \big] dt. \end{aligned} \tag{24}$$

We use

$$f_r(\rho_r(a_r, t)) = \alpha_m f_m(\rho_m(b_m, t))$$

to simplify the second term,

$$f_r'(u_r(\rho_m(b_m, t), \alpha_m)) \partial_\alpha u_r(\rho_m(b_m, t), \alpha_m) = f_m(\rho_m(t, b_m)).$$

The first term of (24) appears in the boundary condition for $\mu_m(t, b_m)$ as seen in equation (22).

We proceed similarly for the other possible boundary terms. From the above equations, we obtain the given adjoint system of equations and the representation of the gradient (23).

2.3. Componentwise Convexity. The following result is important for the analysis of the optimization problem (20), as it gives sufficient conditions for the componentwise convexity of the objective function F. For the result, it is essential that the considered network does not contain directed circles, that is, that a car cannot go through the same road twice. This implies that the control at a junction influences only the traffic flow behind the junction and not the flow that enters the junction.

Theorem 2.2. Assume that the functions $\mathcal{F}_j, j \in \{1, \dots, J\}$, are all convex and that, for a given control vector $(\alpha_1, \dots, \alpha_K) \in [0, 1]^K$, the corresponding adjoint solutions $\mu_j(\cdot, t)$ are nondecreasing for all $t \in (0, T)$. Let $k \in \{1, \dots, K\}$. Then, for all $\beta_k \in [0, 1]$, the following inequality holds:

$$F(\alpha_1, \dots, \beta_k, \dots, \alpha_K) \geq F(\alpha_1, \dots, \alpha_k, \dots, \alpha_K) + (\beta_k - \alpha_k)\partial_{\alpha_k} F(\alpha_1, \dots, \alpha_k, \dots, \alpha_K).$$

Hence, if $\partial_x \mu$ is always positive, the function F is componentwise convex with respect to each component.

Proof. Let $H(\beta_K) = F(\alpha_1, \dots, \beta_k, \dots, \alpha_K)$, let ρ^α denote the system solution with control vector $(\alpha_1, \dots, \alpha_K)$, and let ρ^β denote the system solution with control vector $(\alpha_1, \dots, \beta_k, \dots, \alpha_K)$. Then, we have

$$\begin{aligned} H(\beta_k) - H(\alpha_k) = & \sum_{j=1}^{J} \int_0^T \int_{a_j}^{b_j} \left[\mathcal{F}_j(\rho_j^\beta(x, t)) - \mathcal{F}_j\left(\rho_j^\alpha(x, t)\right) \right] dx dt \\ & + \sum_{j=1}^{J} \int_0^T \int_{a_j}^{b_j} \mu_j(x, t) \Big[\partial_t(\rho_j^\beta(x, t) - \rho_j^\alpha(x, t)) \\ & + \partial_x (f_j(\rho_j^\beta(x, t)) - f_j(\rho_j^\alpha(x, t))) \Big] dx dt \\ \geq & \sum_{j=1}^{J} \int_0^T \int_{a_j}^{b_j} \left[\mathcal{F}'_j(\rho_j^\alpha(x, t))(\rho_j^\beta(x, t) - \rho_j^\alpha(x, t)) \right] dx dt \end{aligned}$$

$$
\begin{aligned}
&-\sum_{j=1}^{J}\int_0^T\int_{a_j}^{b_j}\partial_t\mu_j(x,t)\left[\rho_j^\beta(x,t)-\rho_j^\alpha(x,t)\right]dxdt\\
&-\sum_{j=1}^{J}\int_0^T\int_{a_j}^{b_j}\partial_x\mu_j(x,t)\left[f_j(\rho_j^\beta(x,t))-f_j(\rho_j^\alpha(x,t))\right]dxdt\\
&+\sum_{j=1}^{J}\int_0^T\mu_j(x,t)\left[f_j(\rho_j^\beta(x,t))-f_j(\rho_j^\alpha(x,t))\right]\Big|_{x=aj}^{b_j}dt=:S,
\end{aligned}
$$

where the inequality follows from the convexity of the functions $\mathcal{F}_j$. Then, the concavity of the functions f_j and the monotonicity of $\mu(\cdot,t)$ imply

$$
\begin{aligned}
S\geq&\sum_{j=1}^{J}\int_0^T\int_{a_j}^{b_j}\left[\rho_j^\beta(x,t)-\rho_j^\alpha(x,t)\right]\\
&\times\left[\mathcal{F}_j'(\rho_j^\alpha(x,t))-\partial_t\mu_j(x,t)-\partial_x\mu_j(x,t)f_j'(\rho_j^\alpha(x,t))\right]dxdt\\
&+\sum_{j=1}^{J}\int_0^T\mu_j(x,t)\left[f_j(\rho_j^\beta(x,t))-f_j(\rho_j^\alpha(x,t))\right]\Big|_{x=a_j}^{b_j}dt\\
=&\sum_{j=1}^{J}\int_0^T\mu_j(x,t)\left[f_j(\rho_j^\beta(x,t))-f_j(\rho_j^\alpha(x,t))\right]\Big|_{x=a_j}^{b_j}dt.
\end{aligned}
$$

Let $\kappa\in\{1,\ldots,K\}$. For the κth junction of the first type, we use the notation $m(\kappa), r(\kappa), s(\kappa)$ as numbers for the adjacent edges and

$$
\begin{aligned}
&f_{m(\kappa)}^\beta(t)=f_{m(\kappa)}^\beta\left(\rho_{m(\kappa)}^\beta(b_{m(\kappa)},t)\right),\\
&f_{r(\kappa)}^\beta(t)=f_{r(\kappa)}^\beta\left(\rho_{r(\kappa)}^\beta(a_{r(\kappa)},t)\right),\\
&f_{s(\kappa)}^\beta(t)=f_{s(\kappa)}^\beta\left(\rho_{s(\kappa)}^\beta(a_{s(\kappa)},t)\right).
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
H(\beta_k)-H(\alpha_k) \geq & \sum_{\kappa=1}^{K}\int_0^T \mu_{m(\kappa)}(b_{m(\kappa)},t)(f_{m(\kappa)}^{\beta}(t)-f_{m(\kappa)}^{\alpha}(t))dt \\
& +\sum_{\kappa=1}^{K}\int_0^T \Big[-\mu_{r(\kappa)}(a_{r(\kappa)},t)\left(f_{r(\kappa)}^{\beta}(t)-f_{r(\kappa)}^{\alpha}(t)\right) \\
& \qquad -\mu_{s(\kappa)}(a_{s(\kappa)},t)\left(f_{s(\kappa)}^{\beta}(t)-f_{s(\kappa)}^{\alpha}(t)\right)\Big]dt.
\end{aligned}
$$

The last equation follows from the junction conditions (14) and the corresponding adjoint junction conditions, which make the terms corresponding to the boundary nodes and junctions of the second type vanish.

Using the conditions for junctions of the first type (14), (22), we obtain

$$
\begin{aligned}
H(\beta_k)-H(\alpha_k) \geq & \sum_{\kappa=1}^{K}\int_0^T \big[\alpha_{m(\kappa)}\mu_{r(\kappa)}(a_{r(\kappa)},t) \\
& \qquad +(1-\alpha_{m(\kappa)})\mu_{s(\kappa)}(a_{s(\kappa)},t)\big](f_{m(\kappa)}^{\beta}(t)-f_{m(\kappa)}^{\alpha}(t))dt \\
& +\sum_{\kappa=1}^{K}\int_0^T -\mu_{r(\kappa)}(a_{r(\kappa)},t)\Big[\beta_{m(\kappa)}f_{m(\kappa)}^{\beta}(t)-\alpha_{m(\kappa)}f_{m(\kappa)}^{\alpha}(t)\Big]dt \\
& -\int_0^T \mu_{s(\kappa)}(a_{s(\kappa)},t)\Big[(1-\beta_{m(\kappa)})f_{m(\kappa)}^{\beta}(t)-(1-\alpha_{m(\kappa)})f_{m(\kappa)}^{\alpha}(t)\Big]dt \\
= & \sum_{\kappa=1}^{K}\int_0^T (\alpha_{m(\kappa)}-\beta_{m(\kappa)})f_{m(\kappa)}^{\beta}(t)\big[\mu_{r(\kappa)}(a_{r(\kappa)},t)-\mu_{s(\kappa)}(a_{s(\kappa)},t)\big]dt \\
= & (\beta_k-\alpha_k)\int_0^T \big[\mu_{s(k)}(a_{s(k)},t)-\mu_{r(k)}(a_{r(k)},t)\big]f_{m(k)}^{\beta}(t)dt \\
= & (\beta_k-\alpha_k)\int_0^T \big[\mu_{s(k)}(a_{s(k)},t)-\mu_{r(k)}(a_{r(k)},t)\big]f_{m(k)}^{\alpha}(t)dt \\
= & (\beta_k-\alpha_k)H'(\alpha_k),
\end{aligned}
$$

where the last equality follows from (23). This yields the assertion. □

Lemma 2.1. Assume that, for all $j\in\{1,\ldots,J\}$, the function $\mathcal{F}_j$ is convex and that, for all $t\in[0,T]$, the function $\rho_j(\cdot,t)$ is decreasing on the interval $[a_j,b_j]$. Then, for all $t\in[0,T]$, function $\mu_j(\cdot,t)$ is increasing on the interval $[a_j,b_j]$.

Hence, Theorem 2.2 implies that the function F is componentwise convex.

Proof. Let $t\in[0,T]$, $j\in\{1,\ldots,J\}$, and $a_j<x_1<x_2<b_j$. We consider the characteristic curve

$$C_1=\{(\xi_1(s),s):s\geq t\}$$

that starts from the point (x_1,t) and the characteristic curve

$$C_2=\{(\xi_2(s),s):s\geq t\}$$

that starts from the point (x_2,t). Both curves end either at a point in $[a_j,b_j]\times\{T\}$, where the value zero is prescribed for μ_j, or at a point in $\{b_j\}\times[0,T]$, where the value of μ_j is prescribed by the boundary conditions.

Since $\mathcal{F}_j(x)$ is convex, the derivative $\mathcal{F}_j'$ is increasing.

If ρ_j is decreasing, this implies that, at any time $s\in[0,T]$, we have the inequality

$$\mathcal{F}_j'(\rho_j(\xi_1(s),s)\geq\mathcal{F}_j'(\rho_j(\xi_2(s),s). \tag{25}$$

So, on the way along the curve C_1 to a point where the value of μ_j is prescribed, μ_j grows faster than along the corresponding way on the curve C_2. Moreover, the time that elapses from t until the curve C_1 meets such a point is at least as long as the corresponding time for the curve C_2.

These ways can also be continued through the junctions respecting the corresponding junction conditions until a point is reached, where the value zero is prescribed for μ_j. This continuation through the nodes makes the difference between a single road and a network. For this step, it is essential that the node conditions preserve order. This is obvious for junctions of the second type. At a junction of the first type, a kind of bifurcation of the characteristic curves occurs in the sense that, on each of the two outgoing edges, a characteristic curve has to be taken into account. On each of edges, our above argument based upon (25) holds. Since the adjoint node conditions (22) preserve order, in the sense that $\mu_r^1\leq\mu_r^2$ and $\mu_s^1\leq\mu_s^2$ imply

$$\mu_m^1=\alpha_m\mu_r^1+(1-\alpha_m)\mu_s^1\leq\alpha_m\mu_r^2+(1-\alpha_m)\mu_s^2=\mu_m^2,$$

the continuation through the nodes is possible. Thus, we have

$$\mu_j(x_1,t)\leq\mu_j(x_2,t).$$

□

The following lemma gives sufficient conditions for the strict componentwise convexity of the objective function F of the optimization problem (20).

Lemma 2.2. Let the assumptions of Theorem 2.2 hold. In addition, assume that, for the node k where the flow is controlled by the value of α_k, we have $f^{\alpha}_{m(k)}(t) > 0$ for a time interval $[T_1, T_2] \subset [0, T]$, with $T_1 < T_2$, and that the functions $f_{r(k)}$ and $f_{s(k)}$ are strictly concave. Assume that, $\forall_{\alpha_k} \in [0, 1]$, there is $0 < \delta_1 < \delta_2$ such that, for the corresponding adjoint solution μ, we have

$$\partial_x \mu_{r(k)}(x, t) > 0, \quad \text{for } (x, t) \in a_{r(k)} + \delta_1, (a_{r(k)} + \delta_2) \times (T_1, T_2), \tag{26a}$$

or

$$\partial_x \mu_{s(k)}(x, t) > 0, \quad \text{for } (x, t) \in a_{s(k)} + \delta_1, (a_{s(k)} + \delta_2) \times (T_1, T_2). \tag{26b}$$

Then, the function F is strictly convex with respect to the kth component.

The assumption on $\partial_x \mu_{r(k)}$ is in fact an assumption on $\mathcal{F}_{r(k)}$, since they are coupled by the equation. In most application, $\mathcal{F}_j$ is the identity, hence not strictly convex. But due to the strict concavity f_j, the function F is componentwise strictly convex.

Proof. Let $\alpha_k \neq \beta_k$ be given. We consider the case that (26) holds. At the node k, we have

$$f^{\alpha}_{r(k)}(t) = \alpha_k f^{\alpha}_{m(k)}(t) \neq \beta_k f^{\beta}_{m(k)}(t) = f^{\beta}_{r(k)}(t), \quad \text{for all } t \in [T_1, T_2].$$

Hence, we have

$$\rho^{\alpha}_{r(k)}(x, t) \neq \rho^{\beta}_{r(k)}(x, t),$$

for (x, t) in a subset of positive measure of $[a_{r(k)} + \delta_1, a_{r(k)} + \delta_2] \times [T_1, T_2]$. Hence, the strict concavity of $f_{r(k)}$ implies that, for $j = r(k)$, we have the strict inequality

$$\begin{aligned} &-\int_0^T \int_{a_j}^{b_j} \partial_x \mu_j(x, t) \left[f_j \left(\rho^{\beta}_j(x, t) \right) - f_j \left(\rho^{\alpha}_j(x, t) \right) \right] dx\, dt \\ &> -\int_0^T \int_{a_j}^{b_j} \partial_x \mu_j(x, t) f'_j \left(\rho^{\alpha}_j(x, t) \right) \left[\rho^{\beta}_j(x, t) - \rho^{\alpha}_j(x, t) \right] dx\, dt, \end{aligned}$$

which yields the strict convexity of H as in the proof of Theorem 2.2. □

Remark 2.3. The result about componentwise convexity is of interest, since in our road network each component of the decision variable has meaning as the control variable at the corresponding junction. Consider the problem of optimizing the controlled flow through a single junction, while leaving the other controls fixed. If the assumptions for strict componentwise convexity hold, this problem has a unique solution α_{opt} and the objective function is strictly increasing as a function of this single control variable on the interval $[\alpha_{\text{opt}}, 1]$ and is strictly decreasing on the interval $[0, \alpha_{\text{opt}}]$. In other words, if the result about componentwise convexity is applicable, the objective function is very well-behaved as a function of each single control variable.

3. Discretized Model and Discrete Adjoint Equations

We use upwind discretization of the above equations and derive a discrete adjoint scheme. We use the same assumptions as above and the same notation for junctions and roads. Let $1 \le i \le I$ and $1 \le n \le N$ denote the discretization points of the conservation law, where

$$h = L/I, \quad \tau = T/N, \quad \lambda = \tau/h$$

satisfy the CFL condition. Then, the discretized equations are

$$\rho_{j,i}^{n+1} = \rho_{j,i}^{n} - \lambda \left(f_j(\rho_{j,i}^{n}) - f_j(\rho_{j,i-1}^{n}) \right), \quad i = 2, \ldots, I \text{ and } n = 1, \ldots, N-1, \tag{27a}$$

$$\rho_{j,i}^{1} = 0, \qquad i = 1, \ldots, I, \tag{27b}$$

with the boundary conditions for $\rho_{j,i}^{n+1}, n = 1, \ldots, N-1$.

(C1) Boundary Conditions for Junctions of the First Type. For this type of junction,

$$f_r\left(\rho_{r,1}^{n+1}\right) = \alpha_m f_m\left(\rho_{m,I}^{n}\right), \quad f_s\left(\rho_{s,1}^{n+1}\right) = (1 - \alpha_m) f_m\left(\rho_{m,I}^{n}\right), \tag{28a}$$

$$\rho_{r,1}^{n+1} = u_r\left(\rho_{m,I}^{n}, \alpha_m\right), \quad \rho_{s,1}^{n+1} = u_s\left(\rho_{m,I}^{n}, \alpha_m\right). \tag{28b}$$

(C2) Boundary Conditions for Junctions of the Second Type. For this type of junction,

$$f_r\left(\rho_{r,I}^{n+1}\right) = f_p\left(\rho_{p,I}^{n}\right) + f_q\left(\rho_{q,I}^{n}\right), \tag{29a}$$

$$\rho_{r,1}^{n+1} = u_r(\rho_{p,I}^{n}, \rho_{q,I}^{n}). \tag{29b}$$

(C3) Boundary Conditions for Roads with a Free Boundary at a_s,

$$\rho_{s,1}^{n+1} = u_s^{n}. \tag{30}$$

The discretized functional with the controls $\alpha_k, k=1,\dots,K$, is given by

$$F(\alpha_1,\dots,\alpha_k)=\tau h\sum_{j=1}^{J}\sum_{n=2}^{N}\sum_{i=1}^{I-1}\mathcal{F}_j(\rho_{j,i}^n). \tag{31}$$

The gradient ∇F is derived via the discrete adjoint system for $\mu_{j,i}^n$,

$$\mu_{j,i}^{n-1}=\mu_{j,i}^n-\tau\mathcal{F}_j'(\rho_{j,i}^n)-\lambda f_j'(\rho_{j,i}^n)(\mu_{j,i}^n-\mu_{j,i+1}^n),$$
$$i=1,\dots,I-1 \text{ and } n=2,\dots,N, \tag{32a}$$
$$\mu_{j,i}^N=0, \qquad i=1,\dots,I. \tag{32b}$$

(D1) Boundary for Junctions of the First Type, $n=2,\dots,N-1$. For this type of junction,

$$\begin{aligned}\lambda f_m'(\rho_{m,I}^n)\mu_{m,I}^n=\mu_{m,I}^n-\mu_{m,I}^{n-1}&\\ +\partial_p u_s(\alpha_m,\rho_{m,I}^n)&\left(\mu_{s,1}^n-\mu_{s,1}^{n+1}+\lambda f_s'(\rho_{s,1}^{n+1})\mu_{s,1}^{n+1}\right)\\ +\partial_p u_r(\alpha_m,\rho_{m,I}^n)&\left(\mu_{r,1}^n-\mu_{r,1}^{n+1}+\lambda f_r'(\rho_{r,1}^{n+1})\mu_{r,1}^{n+1}\right),\end{aligned}$$

or equivalently,

$$\begin{aligned}&f_m'(\rho_{m,I}^n)\left(\mu_{m,I}^n-(1-\alpha_m)\mu_{s,1}^{n+1}-\alpha_m\mu_{r,1}^{n+1}\right)\\ &=h\Big[\left(\mu_{m,I}^n-\mu_{m,I}^{n-1}\right)/\tau+\partial_p u_s(\alpha_m,\rho_{m,I}^n)(\mu_{s,1}^n-\mu_{s,1}^{n+1})/\tau\\ &\qquad+\partial_p u_r(\alpha_m,\rho_{m,I}^n)\left(\mu_{r,1}^n-\mu_{r,1}^{n+1}\right)/\tau\Big].\end{aligned}$$

(D2) Boundary Values for Junctions of the Second Type, $n=2,\dots,N-1$. For this type of junction,

$$\begin{aligned}\lambda f_p'(\rho_{p,I}^n)\mu_{p,I}^n&=\mu_{p,I}^n-\mu_{p,I}^{n-1}+\partial_{\rho_p}u_r\left(\rho_{p,I}^n,\rho_{q,I}^n\right)\\ &\quad\times\left[\mu_{r,1}^n-\mu_{r,1}^{n+1}+\lambda f_r'(\rho_{r,1}^{n+1})\mu_{r,1}^{n+1}\right],\\ \lambda f_q'(\rho_{q,I}^n)\mu_{q,I}^n&=\mu_{q,I}^n-\mu_{q,I}^{n-1}+\partial_{\rho_q}u_r\left(\rho_{p,I}^n,\rho_{q,I}^n\right)\\ &\quad\times\left[\mu_{r,1}^n-\mu_{r,1}^{n+1}+\lambda f_r'(\rho_{r,1}^{n+1})\mu_{r,1}^{n+1}\right],\end{aligned}$$

or equivalently,

$$
\begin{aligned}
f_p'(\rho_{p,I}^n)\left(\mu_{p,I}^n-\mu_{r,1}^{n+1}\right) &= h\Big[\left(\mu_{p,I}^n-\mu_{p,I}^{n-1}\right)/\tau \\
&\quad +\partial_{\rho_p}u_r\left(\rho_{p,I}^n,\rho_{q,I}^n\right)\left(\mu_{r,1}^n-\mu_{r,1}^{n+1}\right)/\tau\Big],\\
f_q'(\rho_{q,I}^n)\left(\mu_{q,I}^n-\mu_{r,1}^{n+1}\right) &= h\Big[\left(\mu_{q,I}^n-\mu_{q,I}^{n-1}\right)/\tau \\
&\quad +\partial_{\rho_q}u_r\left(\rho_{p,I}^n,\rho_{q,I}^n\right)\left(\mu_{r,1}^n-\mu_{r,1}^{n+1}\right)/\tau\Big].
\end{aligned}
$$

(D3) Boundary Conditions for Roads with a Free Boundary at b_s.

$$\mu_{s,I}^n=0,\quad n=2,\ldots,N. \tag{33}$$

The discretized gradient is given by

$$
\begin{aligned}
&\partial\alpha_m F(\alpha_1,\ldots,\alpha_K)\\
&= h\sum_{n=1}^{N-1} -\partial_\alpha u_r(\rho_{m,I}^n,\alpha_m)\left[\mu_{r,1}^n-\mu_{r,1}^{n+1}+\lambda f_r'(\rho_{r,1}^{n+1})\mu_{r,1}^{n+1}\right]\\
&\quad -\partial_\alpha u_s(\rho_{m,I}^n,\alpha_m)\left[\mu_{s,1}^n-\mu_{s,1}^{n+1}+\lambda f_s'(\rho_{s,1}^{n+1})\mu_{s,1}^{n+1}\right]\\
&= \tau\sum_{n=1}^{N-1} f_m(\rho_{m,I}^n)\left(\mu_{s,1}^{n+1}-\mu_{r,1}^{n+1}\right)\\
&\quad -h\Big[\partial_\alpha u_r(\rho_{m,I}^n,\alpha_m)(\mu_{r,1}^n\mu_{r,1}^{n+1})/\tau\\
&\qquad +\partial_\alpha u_s(\rho_{m,I}^n,\alpha_m)(\mu_{s,1}^n-\mu_{s,1}^{n+1})+/\tau\Big].
\end{aligned} \tag{34}
$$

Remark 3.1. The coupling conditions at the junctions coincide with the conditions derived in Section 2, except for a term of order $O(h)$. Also, the discrete adjoint gradient (34) coincides with the continuous adjoint gradient (23), except for a term or order $O(h)$. Equation (32) is an upwind discretization of (21). Thus, taking formally the limit h, $\tau\to 0$, we obtain the same equations as in Section 2.

The discrete adjoint system can be derived by considerations similar as in the continuous case. We introduce a Lagrange function L depending on $\alpha=(\alpha_k)_k, k=1,\ldots,K, \rho=(\rho_{j,i}^n)$, and $\mu=(\mu_{j,i}^n)$ with $i=1,\ldots,I$,

$j=1,\dots,J, n=1,\dots,N,$

$$L(\alpha,\rho,\mu)=\tau h\sum_{j=1}^{J}\sum_{n=2}^{N}\sum_{i=1}^{I-1}\mathcal{F}_j(\rho_{j,i}^n)$$
$$+h\sum_{j=1}^{J}\sum_{n=1}^{N-1}\sum_{i=2}^{I}\mu_{j,i}^n\left[\rho_{j,i}^{n+1}-\rho_{j,i}^n+\lambda(f_j(\rho_{j,i}^n)-f_j(\rho_{j,i-1}^n))\right].$$

Linearization and reformulation yields

$$\begin{aligned}
&\partial_{\alpha_m}L(\alpha,\rho,\mu)\\
&=h\sum_{j=1}^{J}\sum_{n=2}^{N}\sum_{i=1}^{I-1}\partial_{\alpha_m}(\rho_{j,i}^n)\left[\tau\mathcal{F}_j'\left(\rho_{j,i}^n\right)+\mu_{j,i}^{n-1}-\mu_{j,i}^n+\lambda f_j'\left(\rho_{j,i}^n\right)\left(\mu_{j,i}^n-\mu_{j,i+1}^n\right)\right]\\
&\quad+h\sum_{j=1}^{J}\sum_{n=2}^{N}\partial_{\alpha_m}\rho_{j,I}^n\left(\mu_{j,I}^{n-1}-\mu_{j,I}^n\right)-\partial_{\alpha_m}\tilde{v}_j^n\left(\mu_{j,1}^{n-1}-\mu_{j,1}^n\right)\\
&\quad+h\lambda\sum_{j=1}^{J}\sum_{i=1}^{I-1}f_j'(\rho_{j,i}^1)\partial_{\alpha_m}\rho_{j,i}^1\left(\mu_{j,i}^1-\mu_{j,i+1}^1\right)\\
&\quad-f_j'(\rho_{j,i}^N)\partial_{\alpha_m}(\rho_{j,i}^N)\left(\mu_{j,i}^N-\mu_{j,i+1}^N\right)\\
&\quad+h\lambda\sum_{j=1}^{J}\sum_{n=2}^{N}\mu_{j,I}^n f_j'(\rho_{j,I}^n)\partial\alpha_m\rho_{j,I}^n-\mu_{j,1}^n f_j'(\tilde{v}_j^n)\partial_{\alpha_m}\tilde{v}_j^n.
\end{aligned}$$

Herein $\tilde{v}_j^n$ denotes the boundary values for ρ_j^n at $i=1$ one road j. The detailed form depends on the junction type. Using the adjoint equation for $\mu_{j,i}^n$ and the initial conditions $\rho_{j,i}^1=\mu_{j,i}^N=0$, the above representation simplifies and we have

$$\partial_{\alpha_m}L(\alpha,\rho,\mu)=h\sum_{j=1}^{J}\sum_{n=2}^{N}\partial_{\alpha_m}\rho_{j,I}^n\left[\mu_{j,I}^{n-1}-\mu_{j,I}^n+\lambda\mu_{j,I}^n f_j'\left(\rho_{j,I}^n\right)\right]$$
$$-\partial_{\alpha_m}\tilde{v}_j^n\left[\mu_{j,1}^{n-1}-\mu_{j,1}^n+\lambda\mu_{j,1}^n f_j'\left(\tilde{v}_j^n\right)\right].$$

As in the continuous case, we state only the boundary terms for one type of junction. The other cases are treated similarly. Assume that $\tilde{v}_j^n$ is given by (28) for $j=r$, i.e.,

$$\tilde{v}_r^n=u_r\left(\rho_{m,I}^{n-1},\alpha_m\right).$$

From the above, we obtain

$$h\sum_{n=2}^{N}-\partial\alpha_m\tilde{v}_r^n\left[\mu_{r,1}^{n-1}-\mu_{r,1}^n+\lambda\mu_{r,1}^n f_j'\left(\tilde{v}_r^n\right)\right]$$
$$=\sum_{n=2}^{N}-\tau f_m\left(\rho_{m,I}^{n-1}\right)\mu_{r,1}^n-h\left[\partial_\alpha u_r\left(\rho_{m,I}^{n-1}\right)\left(\mu_{r,1}^{n-1}-\mu_{r,1}^n\right)\right]$$
$$-h\sum_{n=2}^{N}\partial_{\alpha_m}\rho_{m,I}^{n-1}\partial_\rho u_r\left(\rho_{m,I}^{n-1}\right)\left[\mu_{r,1}^{n-1}-\mu_{r,1}^n+\lambda\mu_{r,1}^n f_j'\left(\tilde{v}_r^n\right)\right].$$

We used

$$\lambda=\tau/h \quad \text{and} \quad f_r(\rho_{r,1}^n)=\alpha_m f_m\left(\rho_{m,I}^n\right)$$

in the derivation. We find the last term in the boundary condition for $\mu_{m,I}^n$ and the first two terms in the representation of the gradient (34).

4. Comparison of Different Approaches

The prototype network consists of seven roads with four junctions. It is drawn in Figure 2. Each road is assumed to have length

$$b_j-a_j=1, \quad j\in\{1,\dots,7\}.$$

The flow is controlled at two points with two independent controls, namely,

$$\alpha_1=\alpha, \quad \alpha_2=\beta.$$

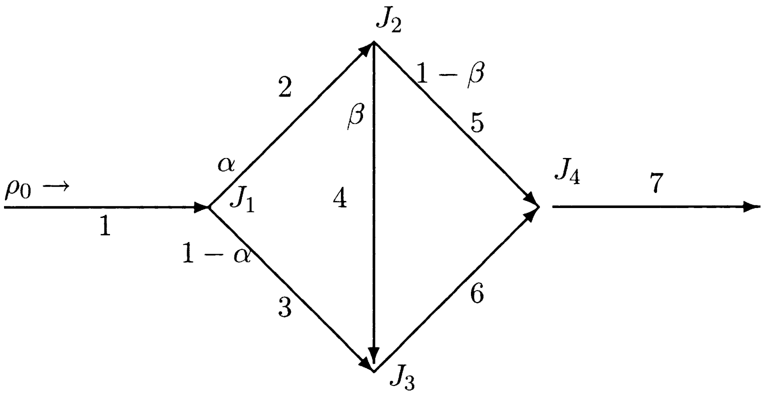

Fig. 2. Prototype of a network.

We used the same flux function for all roads, namely,

$$f_j = f = 4\rho(1-\rho); \tag{35}$$

at a_1, we used the constant density $\rho_o = 0.4$ as boundary value. The node conditions u_r and u_s are as introduced in the previous section. The functional F in (19), given by

$$F(\alpha, \beta) = \sum_{j=1}^{7} \int_0^T \int_0^1 \rho_j(x,t)dx\,dt, \tag{36}$$

measures the total densities in the network. In Ref. 26, it is shown that minimizing this functional is equivalent to steering the network such that the maximal possible outflow is achieved. The time horizon is set to $T = 5.0$. A plot of the objective function is given in Figure 3.

The objective function is not convex. We plot a cut $\mathcal{G}$ of the objective over a line in Figure 4 and its derivative $d_x\mathcal{G}$ in Figure 5. The cut $\mathcal{G}$ is defined by

$$\mathcal{G}(x) := F(\alpha(x), \beta(x)) = F\left(\begin{bmatrix}\alpha_0\\ \beta_0\end{bmatrix} + x\begin{bmatrix}\alpha_1\\ \beta_1\end{bmatrix}\right), \tag{37}$$

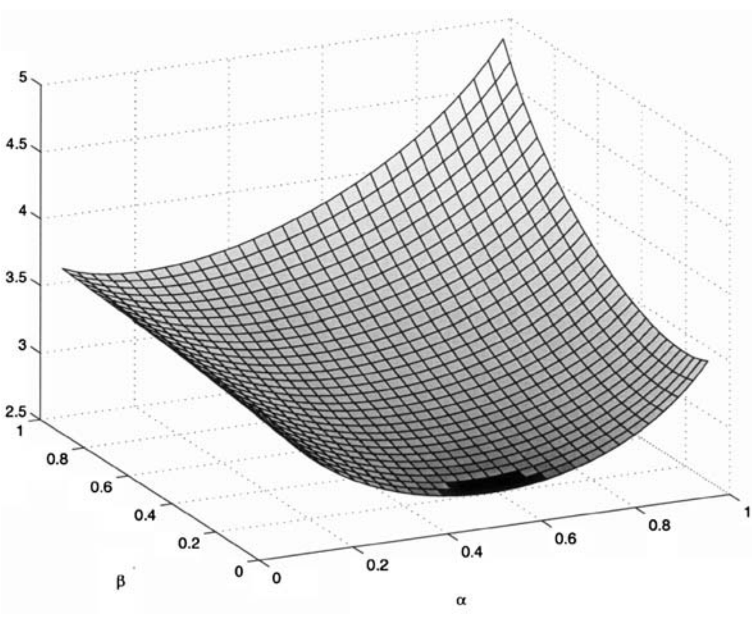

Fig. 3. Functional values for the prototype network.

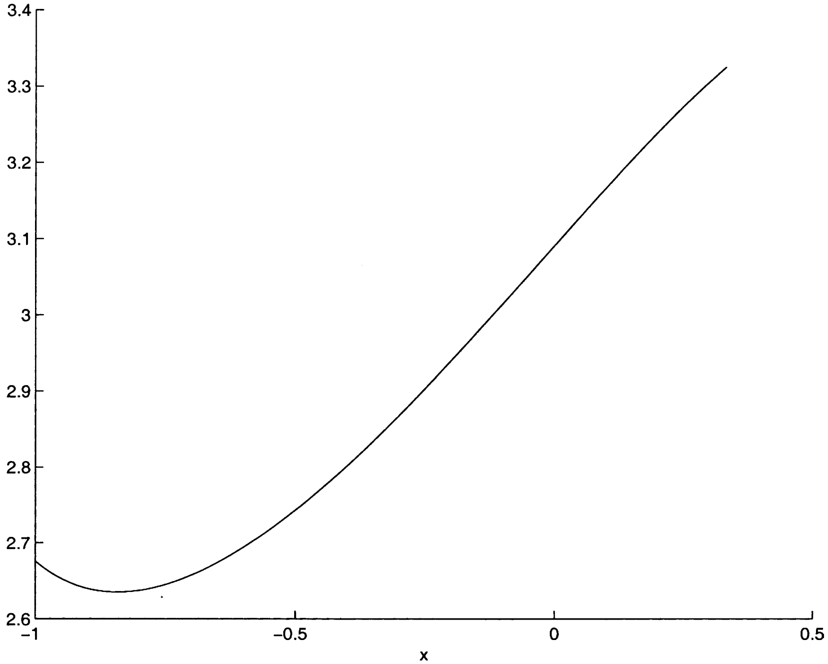

Fig. 4. Functional cut $\mathcal{G}(x)$.

where

$$\alpha_0 = 1/4, \quad \beta_0 = 3/4, \quad \alpha_1 = -1/2, \quad \beta_1 = 3/4, \quad x \in [-1, 1/3].$$

A plot of $\mathcal{G}$ and the objective F is given in Figure 6.

4.1. Solving the Minimization Problem. To solve the minimization problem (20), we used the L-BFGS-B optimization routine of Byrd, Lu, Nocedal, and Zhu (Refs. 27–29). This is an implementation of a quasi–Newton method that can handle box constraints on the control; such methods are described in detail in Ref. 30. In our computations, we imposed the constraint $0 \leq \alpha, \beta \leq 1$. The optimal control is $\alpha_{opt} = 0.5$ and $\beta_{opt} = 0$. This choice implies that road number 4 is empty. We compared three different approaches to calculate the gradient ∇F in Figure 7. We used an upwind discretization for the model with the partial differential equations and for the adjoint system. The objective functional is calculated as in (31). The discretized adjoint approach corresponds to the discretized model and its adjoint. Lastly we plot an approach, where the

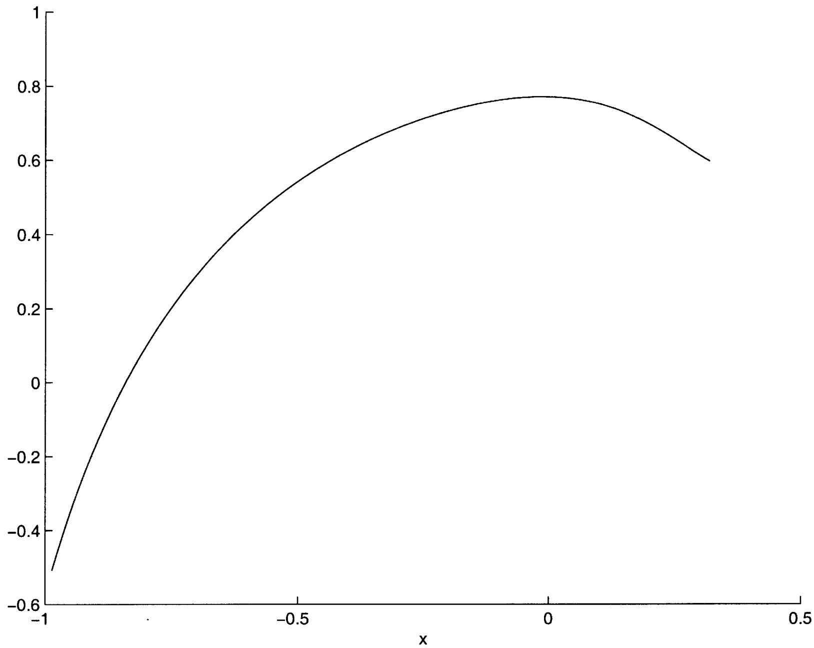

Fig. 5. Derivative of the functional cut $\mathcal{G}(x)$.

partial differential equation is discretized by an upwind scheme and the gradient is calculated by the Richardson extrapolation with six steps. The Richardson extrapolation uses successive calculations of the central differences and guarantees high-order approximations of the derivative. Initially, we set $h = 10^{-1}$. Assuming that the function is smooth, the extrapolation used is of order $O(h^{12})$.

$N = 50$ gripoints are used to discretize each road. The time discretization is such that the CFL condition is satisfied.

Additionally, we compared the gradients obtained by the discretization of the adjoint equation (G_1) and the gradients of the discretized adjoint equations (G_2) in Figure 8. We plot

$$|(\partial_\alpha G_1 - \partial_\alpha G_2)/\max \partial_\alpha G_2| + |(\partial_\beta G_1 - \partial_\beta G_2)/\max \partial_\beta G_2|.$$

There is a major benefit in using the adjoint approach for optimization compared with approximations of the gradient obtained by finite differences. Even using central differences, which is the cheapest reasonable approximation, one needs more computations than with the adjoint

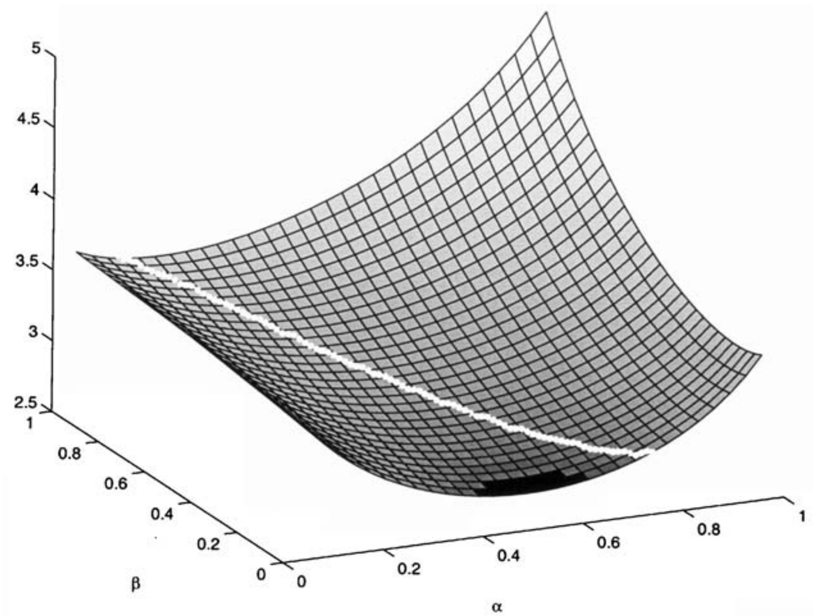

Fig. 6. Functional F and $\mathcal{G}(x)$.

formulas. Let K controls be given. As before, we have $N \times I$ gridpoints for each road $j = 1, \ldots, J$. Assume that, for each gridpoint (n, i), with $n = 1, \ldots, N$, $i = 1, \ldots, I$, on each road j, calculating (27) and (32) [resp. (31) and (34)] is similarly expensive. Let c_1 denote the costs for (27) and let c_2 denote the costs for (31), where we can assume that

$$c_2 \ll c_1. \tag{38}$$

Then, the total costs for evaluating ∇F by central differences are

$$(\text{cost})_{\text{CD}} = 2K(NIJ)c_1 + 2I(KNJ)c_2, \tag{39}$$

while for the adjoint approach it is

$$(\text{cost})_{\text{AA}} = 2(NIJ)c_1 + (KNJ)c_2. \tag{40}$$

The advantage of the adjoint calculus is even higher, since the adjoint equation is linear and for more precise gradients one has to use discrete differences other than central differences.

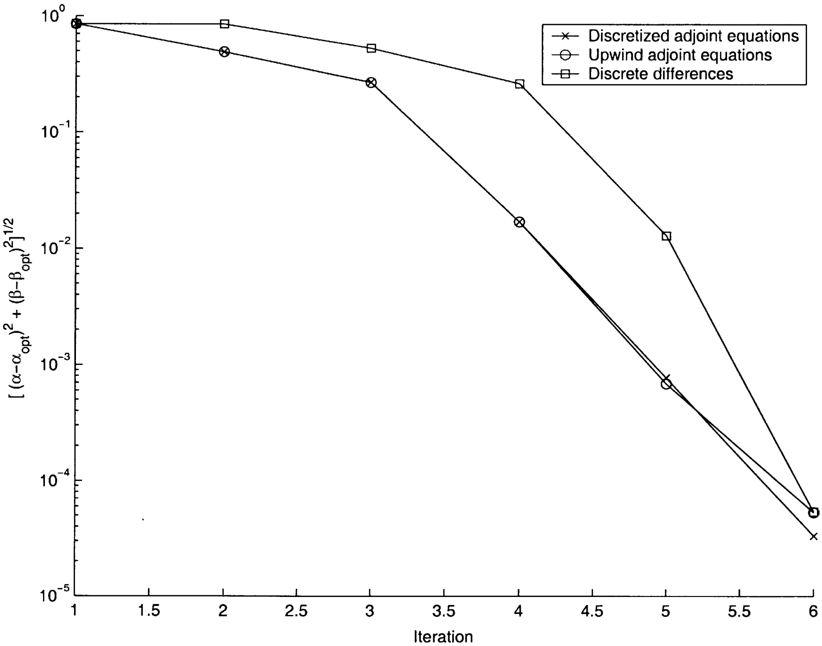

Fig. 7. Convergence results for different approaches.

5. Conclusions

We present an adjoint calculus for traffic flow networks based upon the work of Coclite/Piccoli (Ref. 8) and Holden/Risebro (Ref. 7). We compare the gradients of the discretized adjoint and the discreate adjoint equations on a small network. we show that both approaches yield descent directions which can be used for optimal control.

We give sufficient conditions for the componentwise convexity and for the strict componentwise convexity of the objective function. Our numerical results indicate that the objective is not convex as a function of all variables.

Currently under investigation is the consideration of nonconstant control at the junctions. This is important for the application, since it is reasonable to control a transient traffic by nonconstant controls.

An extension of the scalar models is the modeling of traffic flow on networks with a hyperbolic system of two equations; see Aw/Rascle (Ref. 4). After defining suitable coupling conditions, one can apply a

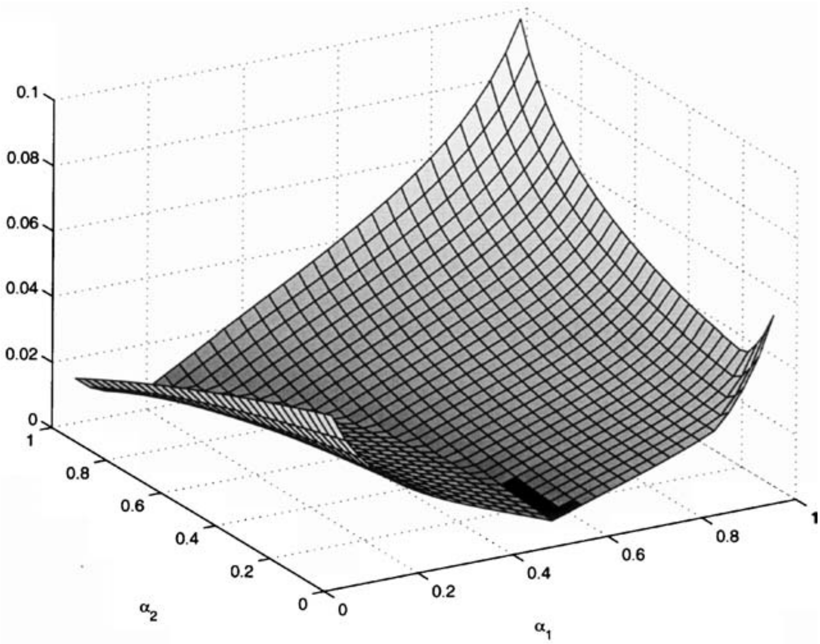

Fig. 8. Difference in the two gradient approximations.

procedure similar to the above to obtain the adjoint equations. This is left for future work.

References

1. WHITHAM, G. B., *Linear and Nonlinear Waves*, Wiley, New York, NY, 1974.
2. PAYNE, H. J., *FREFLO: A Macroscopic Simulation Model of Freeway Traffic*, Transportation Research Record, Vol. 722, pp. 68–75, 1979.
3. KÜHNE, R. D., *Macroscopic Freeway Model for Dense Traffic*, 9th International Symposium on Transportation and Traffic Theory, Edited by N. Vollmuller, VNU Science Press, Utrecht, Netherlands, pp. 21–42, 1984.
4. AW, A., and RASCLE, M., *Resurrection of Second-Order Models of Traffic Flow?*, SIAM Journal on Applied Mathematics, Vol. 60, pp. 916–938, 2000.
5. GÜNTHER, M., KLAR, A., MATERNE, T., and WEGENER, R., *Multivalued Fundamental Diagrams and Stop-and-Go Waves for Continuum Traffic Flow Equations*, Preprint, TU Darmstadt, Dermstodt, Germany.
6. HERTY, M., and KLAR, A., *Modelling and Optimization of Traffic Flow Networks*, SIAM Journal on Statistical and Scientific Computing, 2003 (to appear).

7. Holden, H., and Risebro, N. H., *A Mathematical Model of Traffic Flow on a Network of Unidirectional Roads*, SIAM Journal on Mathematical Analysis, Vol. 4, pp. 999–1017, 1995.
8. Coclite, G. M., and Piccoli, B., *Traffic Flow on a Road Network*, SISSA, Preprint, 2002.
9. Daganzo, C. F., *A Continuum Theory of Traffic Dynamics for Freeways with Special Lanes*, Transportation Research, Vol. 31B, pp. 83–102, 1997.
10. Lebacque, J. P., and Khoshyaran, M. M., *First-Order Macroscopic Traffic Flow Models for Networks in the Context of Dynamic Assignment*, Transportation Planning: State of the Art, Edited by M. Patriksson and M. Labbe, Kluwer Academic Publishers, Dordrecht, Netherlands, 2002.
11. Bardos, C., and Pironneau, O., A *Formalism for the Differentation of Conservation Laws*, Comptes Rendus de l'Academie des Sciences, Paris, Serie I, Vol. 335, pp. 839–845, 2002.
12. Ulbrich, S., *A Sensitivity and Adjoint Calculus for Discontinuous Solutions of Hyperbolic Conservation Laws with Source Terms*, SIAM Journal on Control and Optimization, Vol. 41, pp. 740–797, 2002.
13. Ulbrich, S., *On the Existence and Approximation of Solutions for the Optimal Control of Nonlinear Hyperbolic Conservation Laws*, Optimal Control of Partial Differential Equations, Edited by K. H. Hoffmann et al., Birkhäuser, Basel, Switzerland, pp. 287–299, 1999.
14. Coron, J. M., *Local Controllability of a 1-D Tank Containing a Fluid Modeled by the Shallow Water Equations*, ESAIM: Control Optimization and Calculus of Variations, Vol. 8, pp. 513–554, 2002.
15. Gugat, M., *Problems of Optimal Boundary Control in Flood Management*, Computational Optimization and Applications, 2004 (submitted).
16. Gugat, M., *Nodal Control of Conservation Laws on Networks*, Control and Boundary Analysis, Edited by J. Cagnol and J. P. Zolesio, Marcel Dekker, New York, NY, 2004.
17. Gugat, M., Leugering, G., Schittkowski, K., and Schmidt, E. J. P. G., *Modelling, Stabilization, and Control of Flow in Networks of Open Channels*, Online Optimization of Large-Scale Systems, Edited by M. Groetschel, et al., Springer, Berlin, Germany, pp. 251–270, 2001.
18. Leugering, G., and Schmidt, E. J. P. G., *On the Modeling and Stabilization of Flows in Networks of Open Canals*, SIAM Journal on Control and Optimization, Vol. 41, pp. 164–180, 2002.
19. Gugat, M., and Leugering, G., *Global Boundary Controllability of the de St. Venant Equations between Steady States*, Annales de l'Institut Henri Poincaré: Analysise Non Lineaire Vol. 20, pp. 1–11, 2003.
20. Gugat, M., Leugering, G., and Schmidt, E. J. P. G., *Global Controllability between Steady Supercritical Flows in Channel Networks*, Mathematical Methods in the Applied Sciences, Vol. 27, pp. 781–802, 2004.
21. Lighthill, M. J., and Whitham, J. B., *On Kinematic Waves*, Proceedings of the Royal Society of Edinburgh, Vol. 229A, pp. 281–345, 1983.

22. Klar, A., Kuehne, R. D., and Wegener, R., *Mathematical Models for Vehicular Traffic*, Surveys on Mathematics for Industry, Vol. 6, 1996.
23. Bardos, C., LeRoux, A. Y., and Nedelec, J. C., *First-Order Quasilinear Equations with Boundary Conditions*, Communications in Partial Differential Equations, Vol. 4, pp. 1017–1034, 1979.
24. Bressan, A., *Hyperbolic Systems of Conservation Laws*, Oxford University Press, Oxford, UK, 2000.
25. Holden, H., and Risebro, N. H., *Front Tracking for Hyperbolic Conservation Laws*, Springer, New York, NY, 2002.
26. Herty, M., and Klar A., *Simplified Dynamics and Optimization of Large-Scale Traffic Networks*, M3AS, Vol. 14, pp. 579–601, 2004.
27. Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C., *A Limited-Memory Algorithm for Bound-Constrained Optimization*, SIAM Journal on Statistical and Scientific Computing, Vol. 16, pp. 1190–1208, 1995.
28. Zhu, C., Byrd, R. H., Lu, J., and Nocedal, J., *L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization*, Technical Report NAM-11, EECS Department, Northwestern University, 1994.
29. Byrd, R., Nocedal, J., and Schinabel, R., *Representations of Quasi-Newton Matrices and Their Use in Limited-Memory Methods*, Mathematical Programming, Vol. 63, pp. 129–156, 1994.
30. Kelley, C. T., *Iterative Methods for Optimization*, SIAM, Philadelphia, Pennsylvania, 1999.