OPTIMIZATION FOR A SPECIAL CLASS OF TRAFFIC FLOW MODELS: COMBINATORIAL AND CONTINUOUS APPROACHES

Simone Göttlich and Oliver Kolb

School of Business Informatics and Mathematics
University of Mannheim
D-68131 Mannheim, Germany

Sebastian Kühn

Department of Mathematics
University of Kaiserslautern
D-67663 Kaiserslautern, Germany

(Communicated by Benedetto Piccoli)

Abstract. In this article, we discuss the optimization of a linearized traffic flow network model based on conservation laws. We present two solution approaches. One relies on the classical Lagrangian formalism (or adjoint calculus), whereas another one uses a discrete mixed-integer framework. We show how both approaches are related to each other. Numerical experiments are accompanied to show the quality of solutions.

1. **Introduction.** Modeling, simulation and optimization of traffic flow networks based on partial differential equations have been investigated intensively during the last years, see for instance for an overview [4, 5, 10, 11, 12, 13, 17, 23, 24, 26, 32, 35, 36].

For optimization purposes, different applications such as optimal routing of traffic at intersections [17, 23, 24, 36], traffic light control [21], evacuation planning [20, 36] and air traffic control [1] are of interest. Since in all problems the underlying optimization problem is restricted by partial differential equations, relaxed models with simplified dynamics have been investigated. In this context, two different solution approaches emerge. On the one hand, continuous optimization techniques have been successfully applied to compute optimal solutions. The first order optimality system is derived and solved by a descent type method [30, 39]. On the other hand, suitable discretizations of the dynamics lead to network flow models that have been widely considered in the field of combinatorial optimization [3].

In fact, there exist a few research results comparing both optimization tools, see [8, 14, 15, 17, 42]. One might assume that a direct relation between continuous and discrete optimization techniques exist. This is especially the case when the governing dynamics in the network are (closely) linear. Appropriate numerical discretizations can be chosen, such that the original optimization problem either

---

leads to numerically solving the finite-dimensional optimality system, i.e. the so-called *discretize-then-optimize approach*, or the interpretation as a mixed-integer programming model (MIP). However, it remains the question of detecting local or global optima. We know from the theory of linear programming, in particular the strong duality theorem [37], that under certain circumstances a global optimum can be reached. This is usually not the case for the adjoint calculus. The solution of the optimality system via gradient methods often gets stuck in local optima.

In this work we intend to close the gap between the two solution procedures that first appear to be different. To do so, we start in section 2 with the traffic flow network model where the evolution of traffic density on roads is governed by the linearized Lighthill-Whitham-Richards (LWR) equations, see [9]. For the coupling conditions at the intersections, we stick to the ones presented by Coclite-Garavello-Piccoli [5]. In a next step, we discretize the full network model in space and time and formally derive the adjoint equations and the mixed-integer model (MIP) as well, see sections 3-5. In section 6, we point out the equivalence of both approaches by comparing the dual variables of the MIP and the adjoint variables for a linear network formally and numerically. Furthermore, a more complex network is treated numerically.

2. **Traffic flow network model.** In this section, we briefly review the main modeling issues, see [5, 23, 26].

Definition 2.1 (Traffic Flow Network). A traffic flow network is a finite, connected directed graph denoted by $\mathcal{G} = (V, E)$, where the edges E correspond to roads and vertices V to junctions. Each edge $e \in E$ is associated with an interval $[a_e, b_e]$ and x represents the position on each edge. For a fixed vertex $v \in V$, $\delta^-(v)$ denotes the set of edges leading into vertex v and $\delta^+(v)$ is the set of edges leading out of vertex v.

On each edge $e \in E$ the dynamics of traffic flow is described by the well-known Lighthill-Whitham-Richards equations [35] for the density $\rho_e(x, t)$, $x \in [a_e, b_e]$, $t \geq 0$. Thus the following equations are assumed to hold on each edge e:

$$\partial_t \rho_e(x, t) + \partial_x f_e(\rho_e(x, t)) = 0 \quad \forall e \in E, x \in (a_e, b_e), t \geq 0$$
$$\rho_e(x, 0) = \tilde{\rho}_e \tag{1}$$

where $\tilde{\rho}_e$ are constant functions, i.e. we assume Riemann initial data. As flow function f_e we consider a symmetric hat function

$$f_e(\rho_e) = \begin{cases} v_e \rho_e & 0 \leq \rho_e \leq \sigma_e \\ f_e(\sigma_e) - v_e (\rho_e - \sigma_e) & \sigma_e < \rho_e \leq \rho_{\max,e} \end{cases} \tag{2}$$

borrowed from [7, 9, 38], also sometimes called Newell-Daganzo-Flux.

Remark 1. The hat function (2) is only piecewise differentiable on the intervals $[0, \sigma_e)$ and $(\sigma_e, \rho_{\max,e}]$ with the property $f'_e(\rho_e) = \pm v_e$. Since differentiability is needed for the adjoint calculus performed in section 4, the function has to be smoothed in an appropriate way.

Looking for a network solution, coupling conditions have to be posed for the problem (1) to be well-defined. For an ingoing edge $e \in \delta^-(v)$ the density at a vertex v is denoted by $\bar{\rho}_e(t) = \rho_e(x = b_e, t)$ and an outgoing edge $\tilde{e} \in \delta^+(v)$ is

referred to by $\bar{\rho}_{\tilde{e}}(t) = \rho_{\tilde{e}}(x = a_{\tilde{e}}, t)$. Then, to guarantee flow conservation for all inner vertices $v \in V$ with $|\delta^{\pm}(v)| > 0$, the following condition must be fulfilled:

$$\sum_{e \in \delta^-(v)} f_e(\bar{\rho}_e(t)) = \sum_{\tilde{e} \in \delta^+(v)} f_{\tilde{e}}(\bar{\rho}_{\tilde{e}}(t)) \quad \forall t \geq 0. \tag{3}$$

Due to the choice of constant initial data as in [5, 23, 26], the computation of the coupling condition may be done using the theory about Riemann problems [27]. This means, to get admissible solutions at junctions, we look for waves of negative speed for incoming edges and waves of positive speed for outgoing edges. Thus, we impose the following restrictions to the density values at the vertex

$$
\begin{aligned}
\bar{\rho}_e &\in [\sigma, \rho_{\max,e}] & \rho_e(b_e) &\geq \sigma_e & e &\in \delta^-(v) \\
\bar{\rho}_e &\in \{\rho_e(b_e)\} \cup (\tau(\rho_e(b_e)), \rho_{\max,e}] & \rho_e(b_e) &\leq \sigma_e & e &\in \delta^-(v) \\
\bar{\rho}_{\tilde{e}} &\in [0, \sigma_{\tilde{e}}] & \rho_{\tilde{e}}(a_{\tilde{e}}) &\leq \sigma_{\tilde{e}} & \tilde{e} &\in \delta^+(v) \\
\bar{\rho}_{\tilde{e}} &\in [0, \tau(\rho_{\tilde{e}}(a_{\tilde{e}}))) \cup \{\rho_{\tilde{e}}(a_{\tilde{e}})\} & \rho_{\tilde{e}}(a_{\tilde{e}}) &\geq \sigma_{\tilde{e}} & \tilde{e} &\in \delta^+(v)
\end{aligned} \tag{4}
$$

where for each $\rho \neq \sigma$, $\rho \in [0, \rho_{\max}]$ the value $\tau(\rho)$ is the unique number $\tau(\rho) \neq \rho$, s.t. $f(\tau(\rho)) = f(\rho)$.

Resulting from (4) bounds for the density values at the boundaries of a road are known and the flow going in or out of an edge can be bounded from above by the following functions:

$$\mathrm{de}_e(\rho_e(a_e)) = maxFlowIngoing(\rho_e(a_e)) = f_e(\sigma_e) \cdot \min\left\{1, 2 - \frac{\rho_e(a_e)}{\sigma_e}\right\}, \tag{5a}$$

$$\mathrm{su}_e(\rho_e(b_e)) = maxFlowOutgoing(\rho_e(b_e)) = f_e(\sigma_e) \cdot \min\left\{\frac{\rho_e(b_e)}{\sigma_e}, 1\right\}, \tag{5b}$$

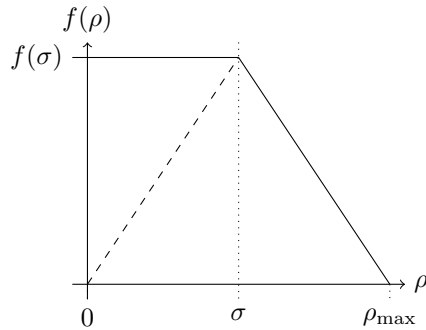being illustrated in figures 1, 2, respectively.

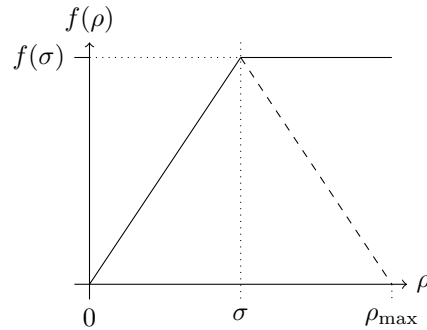FIGURE 1. Maximal flow entering (demand of) edge e, compare (5a).

FIGURE 2. Maximal flow leaving (supply by) edge e, compare (5b).

However, condition (3) is not sufficient to obtain unique solutions. Thus, analogously to Coclite-Garavello-Piccoli [5] we define a time-dependent distribution matrix $\alpha = (\alpha_{e\tilde{e}}(t))$ for all dispersing vertices $v \in V$ (i.e. $|\delta^+(v)| > 1$), where $0 \leq \alpha_{e\tilde{e}} \leq 1$ is the percentage of flow from edge e going to edge $\tilde{e}$ and $\sum_{\tilde{e} \in \delta^+(v)} \alpha_{e\tilde{e}} = 1$

for all $e \in \delta^-(v)$. We get

$$f_{\tilde{e}}\left(\bar{\rho}_{\tilde{e}}(t)\right) = \sum_{e \in \delta^-(v)} \alpha_{e\tilde{e}} f_e\left(\bar{\rho}_e(t)\right) \ \forall \tilde{e} \in \delta^+(v) . \tag{6}$$

As explained in [5, 26], this still leaves a degree of freedom. Assuming that all drivers like to move forward as fast as possible, the throughput at each vertex v shall be maximized. Therefore we solve the following optimization problem at each vertex

$$\begin{aligned}
\max \ & \sum_{e \in \delta^-(v)} f_e\left(\bar{\rho}_e(t)\right) \\
\text{s.t. } & (3) \text{ and } (6) \\
& 0 \leq f_e\left(\bar{\rho}_e\right) \leq \mathrm{su}_e\left(\bar{\rho}_e\right) \ \forall e \in \delta^-(v) \\
& 0 \leq f_{\tilde{e}}\left(\bar{\rho}_{\tilde{e}}\right) \leq \mathrm{de}_{\tilde{e}}\left(\bar{\rho}_{\tilde{e}}\right) \ \forall \tilde{e} \in \delta^+(v) .
\end{aligned} \tag{7}$$

Remark 2. For simulation purposes the distribution matrix α is predefined and we consider the traffic flow network model consisting of the equations

$$\left. (1), (3), (4), (7) \right\} \tag{8}$$

However, considering the optimal routing of cars through the network, the entries of the matrix α will be the control parameters.

By imposing certain conditions on the matrix α and restricting to networks without merging vertices, Coclite-Garavello-Piccoli [5] proved the existence of a unique solution for the optimization problem (7) (Theorem 3.2 in [5]) and the whole traffic flow network problem (8). For a merging junction with two ingoing and one outgoing edge the unique solvability of the maximization problem (7) is shown, among others, by Klar-Herty [23] by proposing FIFO-conditions for the flow through the vertex or Daganzo [6] and Göttlich-Herty-Ziegler [19] by defining a priority road.

In our study we restrict to three types of junctions presented in figure 3. The solution of the coupling conditions is explicitly given in equations (9)–(11).
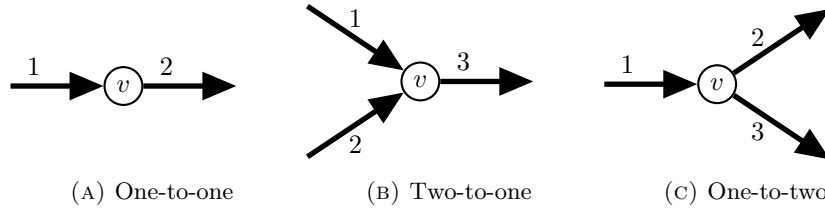

(A) One-to-one   (B) Two-to-one   (C) One-to-two

FIGURE 3. Junction types considered. Types 3(a) and 3(c) are solved using the techniques described by Coclite-Garavelo-Picolli [5], whereas for type 3(b) we use the approach of Klar-Herty [23].

For a simple junction connecting two edges (cf. fig. 3(a)) the unique solution of the coupling is then

$$f_1(\bar{\rho}_1) = f_2(\bar{\rho}_2) = \min\left\{\mathrm{su}_1(\rho_1(b_1)), \mathrm{de}_2(\rho_2(a_2))\right\} . \tag{9}$$

Considering a merging junction (cf. fig. 3(b)), the approach of Klar-Herty [23] yields the optimal solution

$$f_1(\bar{\rho}_1) = \begin{cases} \min \left\{ \begin{array}{l} \mathrm{su}_1(\rho_1(b_1)) \\ \max \left\{ \begin{array}{l} \mathrm{de}_3(\rho_3(a_3)) - \mathrm{su}_2(\rho_2(b_2)) \\ \mathrm{de}_3(\rho_3(a_3))/2 \end{array} \right\} \end{array} \right\} & \begin{array}{l} \mathrm{su}_1(\rho_1(b_1)) + \mathrm{su}_2(\rho_2(b_2)) \\ > \mathrm{de}_3(\rho_3(a_3)) \end{array} \\ \mathrm{su}_1(\rho_1(b_1)) & \mathrm{else} \end{cases}$$

(10a)

$$f_2(\bar{\rho}_2) = \begin{cases} \min \left\{ \begin{array}{l} \mathrm{su}_2(\rho_2(b_2)) \\ \max \left\{ \begin{array}{l} \mathrm{de}_3(\rho_3(a_3)) - \mathrm{su}_1(\rho_1(b_1)) \\ \mathrm{de}_3(\rho_3(a_3))/2 \end{array} \right\} \end{array} \right\} & \begin{array}{l} \mathrm{su}_1(\rho_1(b_1)) + \mathrm{su}_2(\rho_2(b_2)) \\ > \mathrm{de}_3(\rho_3(a_3)) \end{array} \\ \mathrm{su}_2(\rho_e(b_2)) & \mathrm{else} \end{cases}$$

(10b)

$$f_3(\bar{\rho}_3) = \begin{cases} f_1(\bar{\rho}_1) + f_2(\bar{\rho}_2) & \mathrm{su}_1(\rho_1(b_1)) + \mathrm{su}_2(\rho_2(b_2)) > \mathrm{de}_3(\rho_3(a_3)) \\ f_1(\bar{\rho}_1) + f_2(\bar{\rho}_2) & \mathrm{else} \end{cases}$$

(10c)

to (7).

The coupling at the dispersing junction (cf. fig. 3(c)) is solved using the conditions of Coclite-Garavello-Piccoli [5]

$$f_1(\bar{\rho}_1) = f_2(\bar{\rho}_2) + f_3(\bar{\rho}_3) \tag{11a}$$

$$f_2(\bar{\rho}_2) = \min\{\alpha_{1,2}\,\mathrm{su}_1(\rho_1(b_1)), \mathrm{de}_2(\rho_2(a_2))\} \tag{11b}$$

$$f_3(\bar{\rho}_3) = \min\{\alpha_{1,3}\,\mathrm{su}_1(\rho_1(b_1)), \mathrm{de}_3(\rho_3(a_3))\}. \tag{11c}$$

2.1. **Discretization.** For the numerical solution of equation (1), there exist several standard discretization schemes. Since the focus of this work is on optimization and the comparison of two different optimization approaches, we apply the simple staggered Lax-Friedrichs scheme as in [21, 29, 33]. In particular, this scheme does not necessitate the solution of Riemann problems in the interior of the computational domain and can thus be handled (easier) within both considered optimization approaches. We apply a staggered Lax-Friedrichs scheme on the grid $x_{e,j} = a_e + j\Delta x_e = a_e + jL_e/Nx_e$, where $L_e = b_e - a_e$ denotes the length of edge e and Nx_e the number of cells. The time horizon $[0, T]$ is divided into parts of equal length, $t_n = n\Delta t = nT/Nt$, where the time step Δt has to satisfy the CFL-condition

$$\Delta t \leq \min_{e \in E} \left\{ \frac{\Delta x_e}{2v_e} \right\}. \tag{12}$$

Then, for a fixed edge (neglecting the index), with $\lambda = \Delta t/\Delta x$ the evolution of the discretized density reads ($j \in \{2, \ldots, Nx - 1\}$, $n \in \{0, \ldots, Nt - 1\}$)

$$\rho_{0.5}^{n+1} = \tfrac{1}{4}(3\rho_{0.5}^n + \rho_{1.5}^n) - \tfrac{\lambda}{2}\left[f(\rho_{1.5}^n) + f(\rho_{0.5}^n) - 2f(\rho_0^n)\right] \tag{13a}$$

$$\rho_{j-0.5}^{n+1} = \tfrac{1}{4}(\rho_{j-1.5}^n + 2\rho_{j-0.5}^n + \rho_{j+0.5}^n) - \tfrac{\lambda}{2}[f(\rho_{j+0.5}^n) - f(\rho_{j-1.5}^n)] \tag{13b}$$

$$\rho_{Nx-0.5}^{n+1} = \tfrac{1}{4}(\rho_{Nx-1.5}^n + 3\rho_{Nx-0.5}^n) - \tfrac{\lambda}{2}[2f(\rho_{Nx}^n) - f(\rho_{Nx-0.5}^n) - f(\rho_{Nx-1.5}^n)]. \tag{13c}$$

The initial conditions $\rho_{j-0.5}^0$ ($j \in \{1, \ldots, Nx\}$) are calculated using

$$\rho_{j-0.5}^0 = \frac{1}{\Delta x} \int\limits_{x_{j-1}}^{x_j} \tilde{\rho}(x)\,\mathrm{d}x. \tag{14}$$

The values $f(\rho_0^n)$ and $f(\rho_{Nx}^n)$ result from the coupling conditions (4) and are affected by $\rho_{0.5}^n$ and $\rho_{Nx-0.5}^n$, respectively. So for instance, at a vertex with one entering edge e and one leaving edge $\tilde{e}$, we get

$$f(\rho_{e,Nx_e}^n) = f(\rho_{\tilde{e},0}^n) = \min\left\{ \mathrm{su}_e(\rho_{e,Nx_e-0.5}^n), \mathrm{de}_{\tilde{e}}(\rho_{\tilde{e},0.5}^n) \right\} . \tag{15}$$

Remark 3. In equation (15) another advantage of the staggered Lax-Friedrichs scheme can be seen: The flow values at the boundaries needed in (13a), (13c) are directly computed from the coupling conditions. In contrast to other schemes, like the standard Lax-Friedrichs scheme, no computation of densities has to be done. This fact speeds up the computation and saves a lot of technical effort in the optimization part (see sections 4 and 5).

Note that the boundaries of the network have to be treated separately. At the inflow boundary with a desired inflow rate $f_e^{\mathrm{in}}(t_n)$ the actual inflow to the edge e is given by

$$f(\rho_{e,0}^n) = \min\left\{ f_e^{\mathrm{in}}(t_n), \mathrm{de}(\rho_{e,0.5}^n) \right\} , \tag{16}$$

i.e. either the desired inflow rate is allowed to enter the edge or the edge is already jammed and thus the inflow rate is reduced to the amount of flow being able to be processed. At the outflow boundary at edge e we have

$$f(\rho_{e,Nx}^n) = f(\rho_{e,Nx-0.5}^n), \tag{17}$$

which ensures that the flow reaching the end of an outflow edge is able to leave the network without being stopped (Neumann boundary condition).

3. **Optimization problem.** For the traffic flow control in a road network, it is relevant to have an upper bound on the maximal flow in a certain time horizon. However, for evacuation purposes [20], a lower bound on the minimal evacuation time is desired. In both cases, these bounds may be calculated by the traffic flow network model (8) where at dispersing junctions the drivers/evacuees are assigned to the connected roads in an optimal way, i.e. control of the distribution matrix α. In this optimization problem the drivers do not choose their routes themselves but are directed to a special route. Therefore they do not act selfish and a system optimum is achieved rather than a Nash equilibrium [2]. In order to achieve such a system optimum, we solve the following optimization problem:

$$\begin{aligned} \max\ &J(\rho) \\ \mathrm{s.t.}\ &(13),\ (14),\ (15),\ (16),\ (17) \end{aligned} \tag{18}$$

where $J(\rho)$ is the objective function depending on the density of drivers.

There are actually two ways to solve the optimization problem (18). On the one hand, we can apply an adjoint approach based on the Lagrangian formalism, which is explained in section 4. This procedure is the natural choice for PDE-restricted optimization problems. On the other hand, we use a mixed integer programming (MIP) approach, see section 5. This means a reformulation of the traffic flow model using suitable discretizations. Although both approaches initially seem to be different, they are closely related to each other.

The goal of the remaining article is to show the connection between the adjoint and the mixed integer approach as illustrated in figure 4.

Starting from the discretized traffic flow network model and the corresponding optimization problem (18), we follow two strategies to compute an optimal solution. In a first step, (**A** in figure 4), the first order optimality system (KKT-system) based
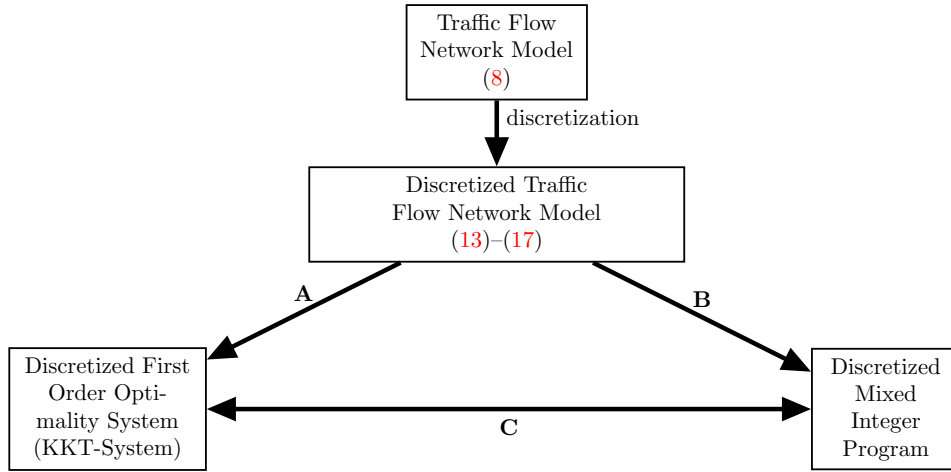
FIGURE 4. Sketch of the interaction between the discrete first order optimality system (section 4) and the mixed integer program (MIP) model (section 5) for the traffic flow network model with hat-function (2).

on the discretization introduced in section 2.1 is formally derived, i.e. the forward equations, the backward equations (adjoints) and the gradient equations. This procedure is called *discretize-then-optimize approach*, see [8, 22]. From a numerical point of view, gradient type methods are used to solve the KKT-system.

Alternatively, (**B** in figure 4), a special discretization of the problem (18) yields a *linear* mixed integer program (MIP), see [8, 15, 16]. In the discretization of (1) with (2), one has to distinguish whether the flux derivative is positive ($f'(\rho_e) > 0$) or negative ($f'(\rho_e) < 0$). Therefore binary variables will enter the problem. Further differences to the adjoint approach appear in the representation of the coupling conditions (3). The entire optimization model resulting in the MIP approach can be solved using Branch-and-Bound-techniques, e.g. CPLEX [28].

The connection between the two approaches (**C** in figure 4) will be illustrated in section 6. The idea is to consider the dual problem of the MIP and to compare the resulting dual variables with the adjoint variables.

To refine our ideas we present the following example for a LP.

Example 1. We consider a linear program of the form

$$\min_{x} J(x) = c^T x$$
$$\text{s.t. } Ax = b \tag{19}$$
$$x \in \mathbb{R}^n$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ and $m, n \in \mathbb{N}$. The dual of (19) is given by

$$\max_{\Phi} b^T \Phi$$
$$\text{s.t. } A^T \Phi = c$$
$$\Phi \in \mathbb{R}^m.$$

Furthermore the Lagrangian formulation of the linear program (19) with Lagrangian multiplier φ is given by

$$\min_x \sup_\varphi J(x) - \varphi^T(Ax - b)$$

$$\text{s.t. } x \in \mathbb{R}^n, \varphi \in \mathbb{R}^m.$$

This leads to the following KKT-system (state and gradient conditions)

$$Ax = b \tag{20a}$$

$$\nabla J(x) - A^T \varphi = 0 \tag{20b}$$

for (19).

Equation (20b) can be interpreted as the adjoint equation of (19). Since the objective function $J(x) = c^T x$ is linear, we get from the adjoint equation (20b)

$$A^T \varphi = c.$$

We directly see that the dual variables Φ and the adjoint variables φ obey the same restrictions.

4. **Adjoint equations.** Adjoint calculus has been used for optimization purposes in a wide variety of applications, among others traffic flow problems [22, 23], supply chains [8], optimal control of gas and water supply networks [14, 25, 33, 34]. In this section we present an approach to solve the optimization problem (18), i.e. finding optimal distribution rates at each vertex subject to the considered objective function, by explicitly solving the so-called discrete adjoint equations, which are part of the discrete first order optimality system.

Within this section we start with a description of the general way to approach the adjoint equations. Then we dedicate a section to the application of the general approach to the different parts of the traffic flow network model, beginning with propagation (eq. (1) and its discretization (13)). We go on with the coupling conditions for a simple junction (eq. (15)) and close with the treatment of boundary conditions (eqs. (16) and (17)).

4.1. **First-discretize adjoint approach.** The discretized equations (13), (14), (15), (16) and (17) are of the form $C(\rho, \alpha) = 0$. So, the optimization tasks we consider are of the form

$$\min J(\rho)$$

$$\text{s.t. } C(\rho, \alpha) = 0 \tag{21}$$

$$\alpha_{\min} \leq \alpha \leq \alpha_{\max}$$

with lower and upper bounds $\alpha_{\min}, \alpha_{\max}$ for the distribution rates α. For a given control α, the state equations $C(\rho, \alpha) = 0$ uniquely determine the state variables ρ. Therefore, the state variables can be considered as a function of the control variables, i.e. $\rho = \rho(\alpha)$, which is implicitly given via the state equations, $C(\rho(\alpha), \alpha) = 0$. The resulting so-called *reduced* optimization problem reads

$$\min J(\rho(\alpha))$$

$$\text{s.t. } \alpha_{\min} \leq \alpha \leq \alpha_{\max} \tag{22}$$

To solve (22) with derivative-based optimization techniques, one needs to compute the total derivative of J with respect to α. Formally, this leads to

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} J(\rho(\alpha)) = -\varphi^T \frac{\partial}{\partial \alpha} C(\rho(\alpha), \alpha)$$

with the so-called *adjoint state* φ being the solution of the *adjoint equation*

$$\left(\frac{\partial}{\partial\rho}C(\rho(\alpha),\alpha)\right)^T \varphi = \left(\frac{\partial}{\partial\rho}J(\rho(\alpha))\right)^T. \tag{23}$$

Rewriting (23) in the form

$$\frac{\partial}{\partial\rho}J(\rho(\alpha)) - \varphi^T\frac{\partial}{\partial\rho}C(\rho(\alpha),\alpha) = 0$$

directly shows that (23) is part of the KKT system of (21) with Lagrangian multipliers φ. Note that we need to smoothen the kink in the hat function (2) as well as the minima in the supply/demand functions (5) and in the coupling conditions (9)–(11) to make the state equations C differentiable with respect to ρ.

Since we consider a time-dependent problem, the state equations and therewith the adjoint equation (23) have a very special structure. In general, this can be easily exploited to reduce the computational effort in practice (cf. [33, 34]) and will be used below to explicitly solve the adjoint equation for the comparison with the dual variables of the MIP formulation.

4.2. **Propagation.** In this section, we explicitly give the solution of the adjoint system (23) for the adjoint variables corresponding to the staggered Lax-Friedrichs discretization presented in section 2.1. Neglecting the index for the underlying edge, we consider the adjoint variables $\varphi_{j-0.5}^{n+1}$ for (13a)–(13c) (with $j \in \{1,\ldots,Nx\}$, $n \in \{0,\ldots,Nt-1\}$). Note that we have to subtract the densities on the left-hand sides of (13a)–(13c) to bring these state equations into the right form ($C(\rho(\alpha),\alpha) = 0$).

A short computation yields the adjoint initial conditions

$$\varphi_{j-0.5}^{Nt} = -\frac{\partial}{\partial\rho_{j-0.5}^{Nt}}J(\rho(\alpha)) \tag{24}$$

at the final time (for $j \in \{1,\ldots,Nx\}$). For the other times ($n \in \{0,\ldots,Nt-1\}$), we get for the (spatially) inner points ($j \in \{2,\ldots,Nx-1\}$)

$$\varphi_{j-0.5}^n = \tfrac{1}{4}\left(\varphi_{j-1.5}^{n+1} + 2\varphi_{j-0.5}^{n+1} + \varphi_{j+0.5}^{n+1}\right)$$
$$+ \tfrac{\lambda}{2}f'(\rho_{j-0.5}^n)\left(\varphi_{j+0.5}^{n+1} - \varphi_{j-1.5}^{n+1}\right) - \frac{\partial}{\partial\rho_{j-0.5}^n}J(\rho(\alpha)) \tag{25}$$

corresponding (as expected) to the adjoint PDE

$$\frac{\partial}{\partial t}\varphi + f'(\rho)\frac{\partial}{\partial x}\varphi = \frac{\partial}{\partial\rho}J(\rho(\alpha)). \tag{26}$$

4.3. **Coupling and boundary conditions.** For a simple junction connecting the end of edge e to the beginning of edge $\tilde{e}$ we consider the state variables (negative cell indices corresponding to the last cells of the left edge, i.e. $\rho_{-2.5}^n = \rho_{e,Nx-2.5}^n$, and the positive ones to the first cells of the right edge, i.e. $\rho_{2.5}^n = \rho_{\tilde{e},2.5}^n$, cf. figure 5)

$$\{\ldots,\rho_{-2.5}^n,\rho_{-1.5}^n,\rho_{-0.5}^n,f(\rho_0^n),\rho_{0.5}^n,\rho_{1.5}^n,\rho_{2.5}^n,\ldots\}$$

for $n \in \{0,\ldots,Nt\}$ and the adjoint variables

$$\{\ldots,\varphi_{-2.5}^n,\varphi_{-1.5}^n,\varphi_{-0.5}^n,\tilde{\varphi}_0^n,\varphi_{0.5}^n,\varphi_{1.5}^n,\varphi_{2.5}^n,\ldots\}$$

for the initial conditions (14), the scheme (13a)–(13c) and the coupling condition (cf. (15))

$$\min\left\{\mathrm{su}_e(\rho_{-0.5}^n),\mathrm{de}_{\tilde{e}}(\rho_{0.5}^n)\right\} - f(\rho_0^n) = 0.$$
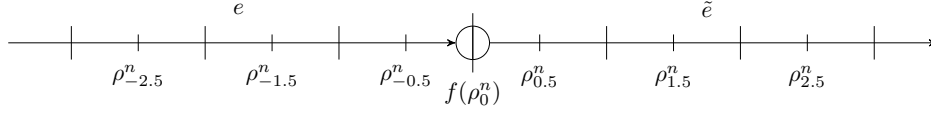
FIGURE 5. Discretization points and variables at a simple junction connecting edges e and $\tilde{e}$.

A short computation yields for the adjoint variables at the coupling node

$$\varphi^n_{-0.5} = \tfrac{1}{4}\left(\varphi^{n+1}_{-1.5} + 3\,\varphi^{n+1}_{-0.5}\right) + \tfrac{\lambda}{2} f'_e(\rho^n_{-0.5})\left(\varphi^{n+1}_{-0.5} - \varphi^{n+1}_{-1.5}\right) - \frac{\partial}{\partial \rho^n_{-0.5}} J(\rho(\alpha))$$

$$+\partial_1 \min\Big\{\mathrm{su}_e(\rho^n_{-0.5}), \mathrm{de}_{\tilde{e}}(\rho^n_{0.5})\Big\}\cdot \mathrm{su}'_e(\rho^n_{-0.5})\left(\lambda\left(\varphi^{n+1}_{0.5} - \varphi^{n+1}_{-0.5}\right) - \frac{\partial}{\partial f(\rho^n_0)} J(\rho(\alpha))\right),$$

$$\tilde{\varphi}^{n+1}_0 = \lambda\left(\varphi^{n+1}_{0.5} - \varphi^{n+1}_{-0.5}\right) - \frac{\partial}{\partial f(\rho^n_0)} J(\rho(\alpha), \alpha) \tag{27}$$

and

$$\varphi^n_{0.5} = \tfrac{1}{4}\left(3\,\varphi^{n+1}_{0.5} + \varphi^{n+1}_{1.5}\right) + \tfrac{\lambda}{2} f'_{\tilde{e}}(\rho^n_{0.5})\left(\varphi^{n+1}_{1.5} - \varphi^{n+1}_{0.5}\right) - \frac{\partial}{\partial \rho^n_{0.5}} J(\rho(\alpha))$$

$$+\partial_2 \min\Big\{\mathrm{su}_e(\rho^n_{-0.5}), \mathrm{de}_{\tilde{e}}(\rho^n_{0.5})\Big\}\cdot \mathrm{de}'_{\tilde{e}}(\rho^n_{0.5})\left(\lambda\left(\varphi^{n+1}_{0.5} - \varphi^{n+1}_{-0.5}\right) - \frac{\partial}{\partial f(\rho^n_0)} J(\rho(\alpha))\right),$$

where ∂_1 and ∂_2 are the partial derivatives with respect to the first and second argument of the min-function, respectively. The derivatives of the (non-smooth) functions su_e and $\mathrm{de}_{\tilde{e}}$ defined in section 2, (5) are given by

$$\mathrm{su}'_e(\rho) = \begin{cases} f'_e(\rho) = v_e & \text{for } \rho < \sigma_e \\ 0 & \text{for } \rho > \sigma_e \end{cases}$$

and

$$\mathrm{de}'_{\tilde{e}}(\rho) = \begin{cases} 0 & \text{for } \rho < \sigma_{\tilde{e}} \\ f'_{\tilde{e}}(\rho) = -v_{\tilde{e}} & \text{for } \rho > \sigma_{\tilde{e}} \end{cases}.$$

Obviously,

$$\partial_1 \min\Big\{\mathrm{su}_e(\rho^n_{-0.5}), \mathrm{de}_{\tilde{e}}(\rho^n_{0.5})\Big\} \cdot \mathrm{su}'_e(\rho^n_{-0.5}) = X_1 \cdot f'_e(\rho^n_{-0.5})$$

with

$$X_1 = \begin{cases} 1 & \text{if } \rho^n_{-0.5} < \sigma_e \quad \text{and} \quad \mathrm{su}_e(\rho^n_{-0.5}) < \mathrm{de}_{\tilde{e}}(\rho^n_{0.5}) \\ 0 & \text{if } \rho^n_{-0.5} > \sigma_e \quad \text{or} \quad \mathrm{su}_e(\rho^n_{-0.5}) > \mathrm{de}_{\tilde{e}}(\rho^n_{0.5}) \end{cases}, \tag{28}$$

and thus

$$\varphi^n_{-0.5} = \tfrac{1}{4}\left(\varphi^{n+1}_{-1.5} + 3\,\varphi^{n+1}_{-0.5}\right) + \tfrac{\lambda}{2} f'_e(\rho^n_{-0.5})\left(2X_1\varphi^{n+1}_{0.5} + (1 - 2X_1)\,\varphi^{n+1}_{-0.5} - \varphi^{n+1}_{-1.5}\right)$$

$$- X_1 \cdot f'_e(\rho^n_{-0.5})\frac{\partial}{\partial f(\rho^n_0)} J(\rho(\alpha)) - \frac{\partial}{\partial \rho^n_{-0.5}} J(\rho(\alpha)). \tag{29}$$

Analogously, with

$$X_2 = \begin{cases} 1 & \text{if } \rho^n_{0.5} > \sigma_{\tilde{e}} \quad \text{and} \quad \mathrm{su}_e(\rho^n_{-0.5}) > \mathrm{de}_{\tilde{e}}(\rho^n_{0.5}) \\ 0 & \text{if } \rho^n_{0.5} < \sigma_{\tilde{e}} \quad \text{or} \quad \mathrm{su}_e(\rho^n_{-0.5}) < \mathrm{de}_{\tilde{e}}(\rho^n_{0.5}) \end{cases}, \tag{30}$$

we get

$$\varphi_{0.5}^n = \tfrac{1}{4}\left(3\,\varphi_{0.5}^{n+1} + \varphi_{1.5}^{n+1}\right) + \tfrac{\lambda}{2} f_{\tilde{e}}'(\rho_{0.5}^n)\left(\varphi_{1.5}^{n+1} + (2X_2 - 1)\varphi_{0.5}^{n+1} - 2X_2\varphi_{-0.5}^{n+1}\right)$$
$$- X_2 \cdot f_{\tilde{e}}'(\rho_{0.5}^n)\frac{\partial}{\partial f(\rho_0^n)}J(\rho(\alpha)) - \frac{\partial}{\partial \rho_{0.5}^n}J(\rho(\alpha)). \quad (31)$$

Similar to the coupling conditions, adjoint boundary conditions can be derived. Let us begin with an inflow boundary: For a desired ingoing flow rate f_{in}^n, the resulting inflow is given by (compare (16))

$$f(\rho_0^n) = \min\left\{f_{\text{in}}^n, \text{de}_e(\rho_{0.5}^n)\right\}.$$

With

$$X_{\text{in}} = \begin{cases} 1 & \text{if} \quad \text{de}_e(\rho_{0.5}^n) < f_{\text{in}}^n \quad \text{and} \quad \rho_{0.5}^n > \sigma_e \\ 0 & \text{if} \quad \text{de}_e(\rho_{0.5}^n) > f_{\text{in}}^n \quad \text{or} \quad \rho_{0.5}^n < \sigma_e \end{cases}$$

one gets

$$\varphi_{0.5}^n = \tfrac{1}{4}\left(3\,\varphi_{0.5}^{n+1} + \varphi_{1.5}^{n+1}\right) + \tfrac{\lambda}{2} f_e'(\rho_{0.5}^n)\left(\varphi_{1.5}^{n+1} + 2(X_{\text{in}} - 1)\varphi_{0.5}^{n+1}\right)$$
$$- X_{\text{in}} \cdot f_e'(\rho_{0.5}^n)\frac{\partial}{\partial f(\rho_0^n)}J(\rho(\alpha)) - \frac{\partial}{\partial \rho_{0.5}^n}J(\rho(\alpha)). \quad (32)$$

At outflow boundaries, we have (compare (17))

$$f(\rho_{Nx}^n) = f(\rho_{Nx-0.5}^n).$$

Here, we get from the adjoint system

$$\varphi_{Nx-0.5}^n = \tfrac{1}{4}\left(3\,\varphi_{Nx-0.5}^{n+1} + \varphi_{Nx-1.5}^{n+1}\right) - \tfrac{\lambda}{2} f_{\tilde{e}}'(\rho_{Nx-0.5}^n)\left(\varphi_{Nx-0.5}^{n+1} + \varphi_{Nx-1.5}^{n+1}\right)$$
$$- f_{\tilde{e}}'(\rho_{Nx-0.5}^n)\frac{\partial}{\partial f(\rho_{Nx}^n)}J(\rho(\alpha)) - \frac{\partial}{\partial \rho_{Nx-0.5}^n}J(\rho(\alpha)). \quad (33)$$

In the following, we summarize the discrete adjoint equations given above for a better referencing

$$\left.\begin{array}{r} \text{initialisation: (24)} \\ \text{propagation: (25)} \\ \text{coupling at a node: (29), (31)} \\ \text{in- and outflow: (32), (33)} \end{array}\right\} \quad (34)$$

5. **Mixed Integer Program (MIP).** As already announced, an alternative way to solve the optimal control problem (18) might be the derivation of a mixed integer program (MIP). To do so, we proceed similarly as in section 4. We discretize the constraints of (18) in a straightforward way following the ideas in [8, 16, 17, 18, 21]. We start with the reformulation of the hat function (2) introducing binary variables. Then we go on with the discretization of (1) using the staggered Lax-Friedrichs scheme, cf. section 2.1. The coupling conditions (3) and (7) are rewritten accordingly. We close the section with a summary of the complete mixed integer program. The latter can be solved using a commercial solver as for instance CPLEX [28].

5.1. **Reformulation of the hat function.** The hat function (2) is a piecewise linear, symmetric function. This can be understood as a composition of two linear parts where a binary variable $\kappa \in \mathbb{B}$ indicates whether the density ρ is less ($\kappa = 1$) or greater ($\kappa = 0$) than the breakpoint σ. This is guaranteed by

$$0 \leq \left(\kappa - \frac{1}{2} \right) (\sigma - \rho). \tag{35}$$

Thus the binary variable κ enables us to write the hat function (2) in one line only

$$f(\rho) = \left\{ \begin{array}{ll} v\rho; & 0 \leq \rho \leq \sigma; \quad \kappa = 1 \\ f(\sigma) - v(\rho - \sigma); & \sigma \leq \rho \leq \rho_{\max}; \quad \kappa = 0 \end{array} \right.$$
$$= \kappa v\rho + (1 - \kappa)(f(\sigma) - v(\rho - \sigma)).$$

However, for a standard MIP, we need all constraints to be linear. We introduce an auxiliary variable $\tilde{\kappa}$ representing the product $\kappa\rho$. Following [18] we get some additional technical constraints and may rewrite the hat function as a linear function in the variables ρ, κ and $\tilde{\kappa}$

$$f(\rho) = 2v\tilde{\kappa} - v\rho + 2f(\sigma) - 2f(\sigma)\kappa. \tag{36}$$

5.2. **Propagation.** As described in section 2.1, the discretization of the transport equation (1) is done applying a staggered Lax-Friedrichs scheme (13). In order to deduce a mixed integer formulation, we use this discretization with $\lambda = \min\{1/2v_e\}$ being the largest possible time step. This yields (omitting the index e for the edge)

$$\rho_{0.5}^{n+1} = \frac{1}{4}(3\rho_{0.5}^n + \rho_{1.5}^n) - \frac{\lambda}{2}[f(\rho_{1.5}^n) + f(\rho_{0.5}^n) - 2f_1^n] \tag{37a}$$

$$\rho_{j-0.5}^{n+1} = \frac{1}{4}(\rho_{j-1.5}^n + 2\rho_{j-0.5}^n + \rho_{j+0.5}^n) - \frac{\lambda}{2}[f(\rho_{j+0.5}^n) - f(\rho_{j-1.5}^n)] \tag{37b}$$

$$\rho_{Nx-0.5}^{n+1} = \frac{1}{4}(\rho_{Nx-1.5}^n + 3\rho_{Nx-0.5}^n) - \frac{\lambda}{2}[2f_2^n - f(\rho_{Nx-0.5}^n) - f(\rho_{Nx-1.5}^n)], \tag{37c}$$

where the index j describes the spatial discretization point and n the time step considered. The variable f_1^n denotes the inflow into the edge, i.e. $f(\rho_0^n) = f(\rho(a, t_n))$ and f_2^n describes the outflow of the edge, i.e. $f(\rho_{Nx}^n) = f(\rho(b, t_n))$. The expression $f(\rho_j^n)$ is an abbreviation for the expression of (36) for each j and n.

Thus, we get the set of equations (37) for each time step n, each discretization point j and each edge e, making a total of $(Nt - 1) \times Nx \times |E|$ constraints.

Note that the initial conditions of (1) have to be discretized as well:

$$\rho_{e,j}^0 = \frac{1}{\Delta x_e} \int_{x_{j-1}}^{x_j} \tilde{\rho}_e(x) \, dx = \tilde{\rho}_e \tag{38}$$

for all discretization points j and edges $e \in E$. Accordingly the variables describing the in- and outflow to the edges are initialized by

$$f_{e,i}^0 = f(\tilde{\rho}_e). \tag{39}$$

5.3. **Coupling and boundary conditions.** For the coupling at a simple junction connecting the end of edge e to the beginning of edge $\tilde{e}$ we need to consider the flow conservation (3) yielding

$$f_{e,2}^n = f_{\tilde{e},1}^n \qquad \forall 0 \le n \le Nt. \tag{40}$$

Furthermore we need to implement the constraint for maximal throughput (9) in order to guarantee a unique solution

$$f_{e,2}^n = \min \left\{ \mathrm{su}_{e,2}^n, \mathrm{de}_{\tilde{e},1}^n \right\}. \tag{41}$$

This minimum is included in the MIP by introducing a binary variable deciding on which term is smaller. Hereby we follow the approach in Göttlich et al. [18, Proposition 2.2]. The equations (5) are rewritten in the same way by making use of the binary variable κ, which has already been defined in (35).

Remark 4. The coupling of more than two edges in a junction is a straightforward extension of equation (41) adding technical constraints and four more variables per junction.

For the boundaries of the network, boundary conditions have to be specified. First we consider the inflow boundaries, i.e. all edges going into the network, $e \in E^{\mathrm{in}}$. For these edges the inflow variable $f_{e,1}$ has to be calculated by

$$f_{e,1}^n = f_e^{\mathrm{in}}(t_n) \tag{42}$$

for all $0 \le n \le Nt$, where $f_e^{\mathrm{in}}(t)$ is a function describing the inflow to this edge. For the edges leading out of the network, we assume that all flow is leaving the network. Thus we set

$$f_{e,2}^n = f(\rho_{e,Nx-0.5}^n). \tag{43}$$

Summarizing, the constraints of the mixed integer formulation of (18) are

$$\left.\begin{array}{r} \text{initialisation: } (38),(39) \\ \text{propagation: } (37) \\ \text{coupling at a node: } (40),\ (41) \\ \text{in- and outflow: } (42),\ (43) \end{array}\right\}. \tag{44}$$

Remark 5. Note that the controls α are only implicitly given in the MIP (44). They will be computed from the optimized flow values $f_{e,i}^n$ in a post processing step.

6. **Comparison and numerical experiments.**

6.1. **Dual problem and connection to discrete adjoints.** We start with a formal comparison of the two optimization approaches. For simplicity, we stick to the situation depicted in figure 3(a). The computations are numerically validated in subsection 6.2 and extended in subsection 6.3.

For our investigations, we stick to linear objective functions J, e.g. the maximization of outflow. In order to dualize the MIP we introduce a dual variable Φ for each constraint of (44). Applying complementary slackness conditions, the number of dual variables reduces and simplifies the problem. We mainly distinguish between

the free flow ($f'(\rho) > 0$) or the heavy traffic regime ($f'(\rho) < 0$). This leads to the dual problem of the MIP (44)

Propagation:

$$\Phi_{e,0.5}^{n} = \frac{1}{4}\left(3\Phi_{e,0.5}^{n+1} + \Phi_{e,1.5}^{n+1}\right) + \frac{\lambda v}{2}\left(2\tilde{X}_{e,2}^{n+1} - 1\right)\left(\Phi_{e,1.5}^{n+1} - \Phi_{e,0.5}^{n+1}\right)$$
$$+ v\tilde{X}_{e,2}^{n+1}\Phi_{\text{coup}}^{n+1} - \frac{\partial J(\rho)}{\partial \rho_{e,0.5}^{n+1}}, \tag{45a}$$

$$\Phi_{e,j-0.5}^{n} = \frac{1}{4}\left(\Phi_{e,j-1.5}^{n+1} + 2\Phi_{e,j-0.5}^{n+1} + \Phi_{e,j+0.5}^{n+1}\right)$$
$$+ \frac{\lambda v}{2}\left(2\kappa_{e,j-0.5}^{n+1} - 1\right)\left(\Phi_{e,j+0.5}^{n+1} - \Phi_{e,j-1.5}^{n+1}\right) - \frac{\partial J(\rho)}{\partial \rho_{e,j-0.5}^{n+1}}, \tag{45b}$$

$$\Phi_{e,Nx-0.5}^{n} = \frac{1}{4}\left(\Phi_{e,Nx-1.5}^{n+1} + 3\Phi_{e,Nx-0.5}^{n+1}\right)$$
$$+ \frac{\lambda v}{2}\left(2\tilde{X}_{e,1}^{n+1} - 1\right)\left(\Phi_{e,Nx-0.5}^{n+1} - \Phi_{e,Nx-1.5}^{n+1}\right) \tag{45c}$$
$$+ v\tilde{X}_{e,1}^{n+1}\Phi_{\text{coup}}^{n+1} - \frac{\partial J(\rho)}{\partial \rho_{e,Nx-0.5}^{n+1}}.$$

Coupling at a node:

$$\Phi_{\text{coup}}^{n+1} = \lambda\left(\Phi_{\tilde{e},0.5}^{n+1} - \Phi_{e,Nx-0.5}^{n+1}\right) - \frac{\partial J(\rho)}{\partial f_{e,2}^{n+1}}. \tag{45d}$$

Initialisation:

$$\Phi_{e,j-0.5}^{Nt} = -\frac{\partial J(\rho)}{\partial \rho_{e,j-0.5}^{Nt}}. \tag{45e}$$

In- and outflow:

$$\Phi_{e,0.5}^{n} = \frac{1}{4}\left(3\Phi_{e,0.5}^{n+1} + \Phi_{e,1.5}^{n+1}\right) + \frac{\lambda v}{2}\left(\Phi_{e,1.5}^{n+1} - \Phi_{e,0.5}^{n+1}\right) - \frac{\partial J(\rho)}{\partial \rho_{e,0.5}^{n+1}}, \tag{45f}$$

$$\Phi_{e,-0.5}^{n} = \frac{1}{4}\left(\Phi_{e,Nx-1.5}^{n+1} + 3\Phi_{e,Nx-0.5}^{n+1}\right)$$
$$- \frac{\lambda v}{2}\left(\Phi_{e,Nx-0.5}^{n+1} + \Phi_{e,Nx-1.5}^{n+1}\right) - v\frac{\partial J(\rho)}{\partial f_{e,2}^{n+1}} - \frac{\partial J(\rho)}{\partial \rho_{e,Nx-0.5}^{n+1}}. \tag{45g}$$

with $\tilde{X}_{e,1}^{n} \in \mathbb{B}$ and $\tilde{X}_{e,2}^{n} \in \mathbb{B}$ being the counter parts of the binary variables defined in equations (28) and (30), respectively. They take their values according to the optimal primal solution being in the according regimes. Obviously, the equations of (45) correspond to the adjoint problem in the following way:

	Dual MIP	Adjoints
Propagation	(45a), (45b), (45c)	(31), (25), (29)
Coupling	(45d)	(27)
Initialization	(45e)	(24)
In- and Outflow	(45f), (45g)	(32), (33)

For the initialization and the coupling the restrictions directly correspond. For all other constraints to be compared we note that $f'(\rho) = \pm v$ due to the use of the hat function. Inserting $\tilde{X}_{e,i}^{n}$ also decides on the sign of the velocity v and thus the comparison is straightforward again.

Note that for the numerical solution of (45), we first solve the primal MIP (44) and then insert the values for $\tilde{X}^n_{e,i}$ of the optimal solution. The numerical results can be found in subsection 6.2. Our investigations can be extended to general network topologies including junctions of merging and dispersing type. However, we omit the details of this representation in this work and solely refer to a numerical study in section 6.3.

Next, we solve the optimization problem (18) numerically with the adjoint approach presented in section 4 (to compute derivatives for the reduced problem), which is done within the software package ANACONDA [33], and applying DONLP2 [40, 41] as well as the mixed integer program (MIP) presented in section 5 applying CPLEX [28]. For the latter we employ a barrier procedure as starting algorithm (as it performs better than the default setting of CPLEX) and a tolerance gap of 0.

6.2. **Chain of roads.** We present numerical examples for two different network scenarios. For all settings we give the properties of edges (maximal density $\rho_{\max}$, maximal flow $f(\sigma)$, velocity v and length $L := b - a$) in tables 7, 10 and display the corresponding graphs in figures 6, 9. Furthermore, we choose a spatial discretization of $Nx = 4$, i.e. $\Delta x = L/Nx$, and satisfy the corresponding CFL-condition with equality yielding $\Delta t = L/8v$. The initial values as well as the inflows are given for each example separately.

In the first example, the chain of roads, the focus is on the numerical equivalence of the adjoint variables and the dual variables of the mixed integer program (MIP). Note that there are no control variables in this example, i.e. no degrees of freedom for optimization. Nevertheless, the objective function influences the adjoint and dual variables.

The second example deals with a more complex setting, which can be considered as a straightforward extension of our previous results. In contrast to the work of Fügenschuh et al. [17], backtravelling waves enter into the adjoint and MIP problem. To the best of our knowledge, this effect has not been treated in the network context so far.

We come to the first example: a chain of roads, i.e. a vertex with one ingoing and one outgoing edge (see fig. 6). The properties of the edges are given in table 7.
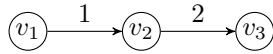

edge nr.	$\rho_{\max}$	$f(\sigma)$	v	L
1	20	10	1	1
2	12	6	1	1

FIGURE 6. A chain of roads.

FIGURE 7. Properties of edges.

Since the setting consists of a single junction we use the coupling conditions given in (15) and (40), (41) for the adjoint approach and the mixed integer program (MIP), respectively. We consider the following linear objective function for the MIP,

$$
\max J_1 = -3 \sum_{j=1}^{Nx} \Delta x \sum_{n=1}^{Nt} \Delta t \rho^n_{1,j-0.5} - 4 \sum_{j=1}^{Nx} \Delta x \sum_{n=1}^{Nt} \Delta t \rho^n_{2,j-0.5}
$$
$$
+ \frac{1}{2} \sum_{n=0}^{Nt-1} \Delta t f^n_{1,1} + \frac{3}{2} \sum_{n=0}^{Nt-1} \Delta t f^n_{1,2} + \frac{5}{2} \sum_{n=0}^{Nt-1} \Delta t f^n_{2,2} \,,
$$

(46)

and the corresponding objective function for the adjoint approach. The given objective function can be interpreted in the following way: The density on each edge

shall be minimized, where the second edge has a larger influence than the first edge, and the throughput in each vertex shall be maximized.

For this objective function we compare the solutions obtained with the two approaches, evaluated at the first and the last cell of both edges (figure 8). We observe that the match is perfect, i.e. the adjoint variables (solid line) and the dual variables of the MIP (dots) coincide for all time steps. The maximal error is $5 \cdot 10^{-7}$.
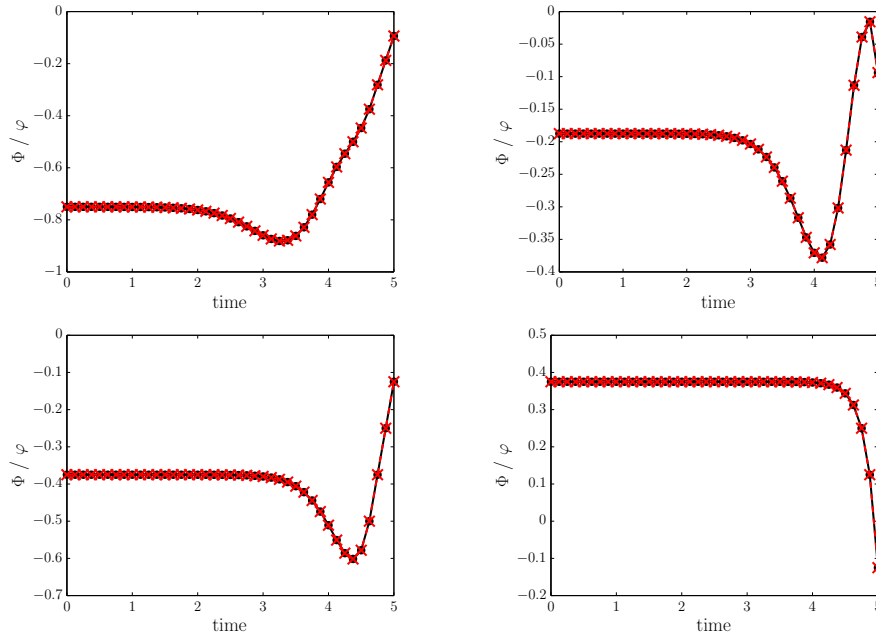

FIGURE 8. Adjoint variables (dashed line, x) compared to the dual variables of the MIP (solid line, dots). Above: edge 1. Below: edge 2. Left: cell 1. Right: cell 4.

6.3. **Diamond network.** Next we study a more complicated network (figure 9). There are both two dispersing (v_2 and v_3) and two merging vertices (v_4 and v_5). The properties of the edges are given in table 10. We stick to this kind of network since we want to show, at least experimentally, that the computations we did in sections 4, 5 still hold true for interlinked networks.

This example is motivated by evacuation dynamics, where the objective is to maximize the total number of rescued evacuees while ensuring a safest possible evacuation. Thus we consider the following objective function, which on the one hand maximizes the throughput and on the other hand penalizes high densities on certain edges

$$\max J_2 = 5 \sum_{n=0}^{Nt-1} \Delta t f_{7,2}^n - \sum_{j=1}^{Nx} \Delta x \sum_{n=1}^{Nt} \Delta t \left(2\rho_{3,j-0.5}^n + \frac{1}{4}\rho_{4,j-0.5}^n + \frac{1}{4}\rho_{6,j-0.5}^n \right). \quad (47)$$

The controls are the distribution rates α_2 and α_3 at the vertices v_2 and v_3, respectively, and similar degrees of freedom in v_4 and v_5 (for backtravelling waves). The inflow is constant, i.e. $f^{in}(t) = 1.5$, and the considered time horizon is $T = 5$. Now,
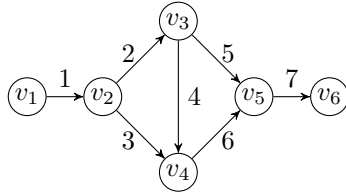
FIGURE 9. The diamond network.

edge nr.	$\rho_{\max}$	$f(\sigma)$	v	L
1	6	3	1	1
2	3	1.5	1	1
3	2	1	1	1
4	1	0.5	1	1
5	1	0.5	1	1
6	2	1	1	1
7	3	1.5	1	1

FIGURE 10. Properties of edges.

we compare both approaches by looking at the optimal solutions (see table 11, figures 12, 13, 14, 15). The mixed integer program (MIP) is solved using CPLEX [28] and the adjoint approach is solved by the software packages ANACONDA [33] and DONLP2 [40, 41] (the latter for the reduced problem).

Comparing the optimal solutions we see that the mixed integer program (MIP) yields (as expected) the best objective value, since the approach guarantees a global optimal solution, but at significantly longer running times. Due to the internal use of the Branch and Bound algorithm, the CPLEX solver tends to prefer bang-bang-solutions, cf. figure 12. This phenomenon has been observed before in [31]. Solving the KKT-system and thus the adjoint equation is much faster and leads to a smoother solution, and in this case not reaching the global optimum, at least not with the same accuracy. In general, this cannot be expected since the optimization problem is not convex and hence the method may get stuck in a local optimum. Using the optimal MIP solution as a starting solution for the adjoint approach improves the results (see table 11). DONLP2 recognizes the optimality of the starting solution and stops after just three iterations.

Alternatively, using the solution of the adjoint approach as a starting solution improves the running time of the mixed integer program (MIP) enormously (about 40 times faster, see table 11), yielding the same solution as before.

	MIP	DONLP2
objective value	31.0628	31.0606
running time [sec]	40463.65	4
objective value with MIP as start	—	31.0628
running time with DONLP2 as start [sec]	1114.80	—

FIGURE 11. Comparison of optimal solutions.

In figure 13 we plot the density in the last cell of edge 7, i.e. the outflow of the network for both solution methods and a uniform distribution at the vertices, i.e. $\alpha_2 = \alpha_3 = 0.5$. Since we present the original solutions without using any starting solutions, the distribution rates and also the outflow slightly differ. This difference immediately vanishes when using the solution of the mixed integer program as a starting solution to the adjoint approach. Finally, we observe that both solution approaches yield an increased output compared to the non-optimal uniform distribution.
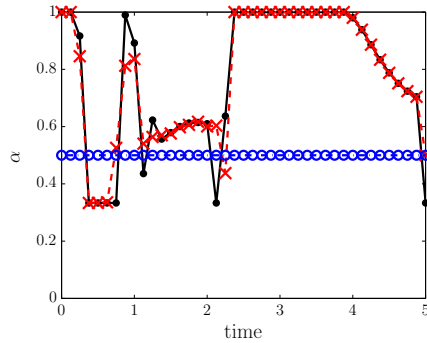
FIGURE 12. Distribution rates at vertex v_2 (solid, dots: MIP; dashed, x: adjoints, dash-dotted, circle: uniform distribution).
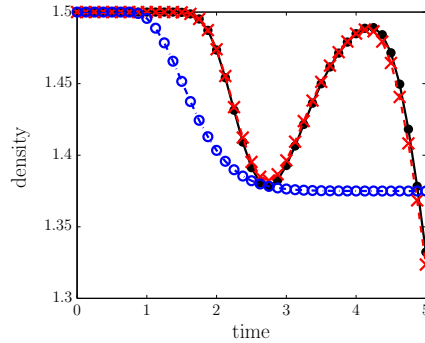

FIGURE 13. Density in the last cell of edge 7 (outflow). Solid, dots: MIP. Dashed, x: adjoints. Dash-dotted, circle: uniform distribution.
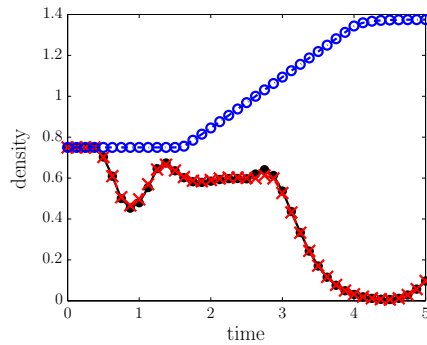

FIGURE 14. Density in the last cell of edge 3. Solid, dots: MIP. Dashed, x: adjoints. Dash-dotted, circle: uniform distribution.
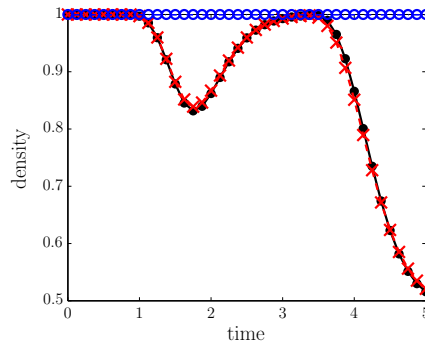

FIGURE 15. Density in the last cell of edge 6. Solid, dots: MIP. Dashed, x: adjoints. Dash-dotted, circle: uniform distribution.

In figures 14, 15 we present the density on the last cell of the penalized edges 3 and 6, respectively. Consequently, the MIP and the adjoint method as well result in lower densities compared to the uniform distribution. This is due to the penalization of density on these edges in the objective function (47). The difference between the optimized values and those from the uniform distribution are larger on edge 3 in comparison with edge 6. This is on the one hand caused by the higher penalization of density on edge 3. On the other hand, edge 3 is congested, if the flow is uniformly distributed. This drawback is avoided by both optimization approaches.

REFERENCES

[1] A. M. Bayen, R. L. Raffard and C. Tomlin, Adjoint-based control of a new Eulerian network model of air traffic flow, *IEEE Transactions on Control Systems Technology,* **14** (2006), 804–818.

[2] A. Bressan and K. Han, Optima and equilibria for a model of traffic flow, *SIAM Journal on Mathematical Analysis,* **43** (2011), 2384–2417.

[3] M. Carey and E. Subrahmanian, An approach to modelling time-varying flows on congested networks, *Transportation Research Part B: Methodological,* **34** (2000), 157–183.

[4] A. Cascone, B. Piccoli and L. Rarita, Circulation of car traffic in congested urban areas, *Communications in Mathematical Sciences,* **6** (2008), 765–784.

[5] G. M. Coclite, M. Garavello and B. Piccoli, Traffic flow on a road network, *SIAM Journal on Mathematical Analysis,* **36** (2005), 1862–1886.

[6] C. F. Daganzo, The cell transmission model, part II: Network traffic, *Transportation Research Part B: Methodological,* **29** (1995), 79–93.

[7] C. F. Daganzo, *Fundamentals of Transportation and Traffic Operations,* Pergamon-Elsevier, Oxford, U.K., 1997.

[8] C. D'Apice, S. Göttlich, M. Herty and B. Piccoli, *Modeling, Simulation and Optimization of Supply Chains: A Continuous Approach*, SIAM Book Series on Mathematical Modeling and Computation, 2010.

[9] C. D'Apice, R. Manzo and B. Piccoli, Packet flow on telecommunication networks, *SIAM Journal on Mathematical Analysis,* **38** (2006), 717–740.

[10] C. D'Apice, R. Manzo and B. Piccoli, A fluid dynamic model for telecommunication networks with sources and destinations, *SIAM Journal on Applied Mathematics,* **68** (2008), 981–1003.

[11] C. D'Apice, R. Manzo and B. Piccoli, Modelling supply networks with partial differential equations, *Quarterly of Applied Mathematics,* **67** (2009), 419–440.

[12] C. D'Apice, R. Manzo and B. Piccoli, Existence of solutions to Cauchy problems for a mixed continuum-discrete model for supply chains and networks, *Journal of Mathematical Analysis and Applications,* **362** (2010), 374–386.

[13] C. D'Apice, R. Manzo and B. Piccoli, Optimal input flows for a PDE-ODE model of supply chains, *Communications in Mathematical Sciences,* **10** (2012), 1225–1240.

[14] P. Domschke, B. Geißler, O. Kolb, J. Lang, A. Martin and A. Morsi, Combination of nonlinear and linear optimization of transient gas networks, *INFORMS Journal on Computing*, **23** (2011), 605–617.

[15] A. Fügenschuh, B. Geißler, A. Martin and A. Morsi, The Transport PDE and Mixed-Integer Linear Programming, in *Models and Algorithms for Optimization in Logistics* (eds. C. Barnhart, U. Clausen, U. Lauther and R. H. Möhring), Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, 2009.

[16] A. Fügenschuh, S. Göttlich, M. Herty, A. Klar and A. Martin, A discrete optimization approach to large scale supply networks based on partial differential equations, *SIAM Journal on Scientific Computing*, **30** (2008), 1490–1507.

[17] A. Fügenschuh, M. Herty, A. Klar and A. Martin, Combinatorial and continuous models for the optimization of traffic flows on networks, *SIAM Journal on Optimization*, **16** (2006), 1155–1176.

[18] S. Göttlich, M. Herty, C. Ringhofer and U. Ziegler, Production systems with limited repair capacity, *Optimization,* **61** (2012), 915–948.

[19] S. Göttlich, M. Herty and U. Ziegler, Numerical discretization of Hamilton - Jacobi equations on networks, *Networks and Heterogeneous Networks,* **8** (2013), 685–705.

[20] S. Göttlich, S. Kühn, J.P. Ohst, S. Ruzika and M. Thiemann, Evacuation dynamics influenced by spreading hazardous material, *Networks and Heterogeneous Media,* **6** (2011), 443–464.

[21] S. Göttlich and U. Ziegler, Traffic light control: A case study, *Discrete and Continuous Dynamical Systems Series S,* **7** (2014), 483–501.

[22] M. Gugat, M. Herty, A. Klar and G. Leugering, Optimal control for traffic flow networks, *J. Optimization Theory Appl.,* **126** (2005), 589–616.

[23] M. Herty and A. Klar, Modeling, simulation, and optimization of traffic flow networks, *SIAM Journal on Scientific Computing,* **25** (2003), 1066–1087.

[24] M. Herty and A. Klar, Simplified dynamics and optimization of large scale traffic networks, *Mathematical Models and Methods in Applied Sciences (M3AS)*, **14** (2004), 579–601.

[25] M. Herty and V. Sachers, Adjoint calculus for optimization of gas networks, *Networks and Heterogeneous Media,* **2** (2007), 733–750.

[26] H. Holden and N. H. Risebro, A mathematical model of traffic flow on a network of unidirectional roads, *SIAM Journal on Mathematical Analysis,* **26** (1995), 999–1017.

[27] H. Holden and N. H. Risebro, *Front Tracking for Hyperbolic Conservation Laws,* 2nd edition, Springer, New York, Berlin, Heidelberg, 2002.

[28] IBM ILOG CPLEX Optimization Studio, Cplex version 12, 2010.

[29] G. S. Jiang, D. Levy, C. T. Lin, S. Osher and E. Tadmor, High-Resolution nonoscillatory central schemes with nonstaggered grids for hyperbolic conservation laws, *SIAM Journal on Numerical Analysis,* **35** (1998), 2147–2168.

[30] C. T. Kelley, *Iterative Methods for Optimization,* Society for Industrial and Applied Mathematics, Philadelphia, 1999.

[31] C. Kirchner, M. Herty, S. Göttlich and A. Klar, Optimal control for continuous supply network models, *Networks Heterogenous Media,* **1** (2006), 675–688.

[32] A. Klar, R. D. Kühne and R. Wegener, Mathematical models for vehicular traffic, *Surveys on Mathematics for Industry,* **6** (1996), 215–239.

[33] O. Kolb, *Simulation and Optimization of Gas and Water Supply Networks*, Ph.D thesis, Technische Universität Darmstadt, 2011.

[34] O. Kolb and J. Lang, Simulation and continuous optimization, in *Mathematical Optimization of Water Networks* (eds. A. Martin, K. Klamroth, J. Lang, G. Leugering, A. Morsi, M. Oberlack, M. Ostrowski and R. Rosen), Internat. Ser. Numer. Math., 162, Birkhäuser/Springer Basel AG, Basel, 2012, 17–33.

[35] M. J. Lighthill and G. B. Whitham, On kinematic waves. II. A theory of traffic flow on long crowded roads, *Royal Society of London Proceedings Series A,* **229** (1955), 317–345.

[36] R. Manzo, B. Piccoli and L. Rarita, Optimal distribution of traffic flows at junctions in emergency cases, *European Journal of Applied Mathematics,* **23** (2012), 515–535.

[37] G. L. Nemhauser and L. A. Wolsey, *Integer and Combinatorial Optimization*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, 1988.

[38] G. F. Newell, *Traffic Flow on Transportation Networks*, MIT Press Series in Transportation Studies, MIT Press, Cambridge, MA, USA, 1980.

[39] P. Spellucci, *Numerische Verfahren der Nichtlinearen Optimierung*, Birkhäuser-Verlag, Basel, 1993.

[40] P. Spellucci, A new technique for inconsistent QP problems in the SQP method, *Mathematical Methods of Operations Research,* **47** (1998), 355–400.

[41] P. Spellucci, An SQP method for general nonlinear programs using only equality constrained subproblems, *Mathematical programming,* **82** (1998), 413–448.

[42] D. Sun, I. S. Strub and A. M. Bayen, Comparison of the performance of four Eulerian network flow models for strategic air traffic management, *Networks and Heterogeneous Media,* **2** (2007), 569–595.

E-mail address: goettlich@uni-mannheim.de
E-mail address: kolb@uni-mannheim.de
E-mail address: skuehn@mathematik.uni-kl.de