NETWORKS AND HETEROGENEOUS MEDIA
©American Institute of Mathematical Sciences
Volume **1**, Number **4**, December **2006**

Website: http://aimSciences.org

pp. **675–688**

OPTIMAL CONTROL FOR CONTINUOUS SUPPLY NETWORK MODELS

Claus Kirchner
Technische Universität Kaiserslautern
Fachbereich Mathematik, Postfach 3049
D-67653 Kaiserslautern, Germany

Michael Herty
Technische Universität Kaiserslautern
Fachbereich Mathematik, Postfach 3049
D-67653 Kaiserslautern, Germany

Simone Göttlich
Technische Universität Kaiserslautern
Fachbereich Mathematik, Postfach 3049
D-67653 Kaiserslautern, Germany

Axel Klar
Fraunhofer Institute ITWM, Fraunhofer Platz 1
D-67663 Kaiserslautern, Germany

Abstract. We consider a supply network where the flow of parts can be controlled at the vertices of the network. Based on a coarse grid discretization provided in [6] we derive discrete adjoint equations which are subsequently validated by the continuous adjoint calculus. Moreover, we present numerical results concerning the quality of approximations and computing times of the presented approaches.

1. **Introduction.** Different approaches for the simulation of continuous supply chain models using partial differential equations have been introduced during the last years; see for example [2, 3, 4, 5]. For the purpose of this paper, we are interested in supply chain models for networks which are mainly derived in [2] and extended in [7, 8]. The latter include the formulation of coupling conditions at intersections by introducing time-dependent queues governed by the mass-flux in the network.

An important aspect in supply chain decision making are optimization problems, for example maximizing output of a production process or minimizing used buffers. These optimization problems can be formulated on a continuous level with constraints consisting of partial or ordinary differential equations. Then, usually an adjoint calculus is used for efficient computation of the optimal control. This

2000 *Mathematics Subject Classification.* Primary: 90B10; Secondary: 65Mxx.

Key words and phrases. Supply chains, conservation laws, networks, adjoint calculus.

This work was supported by the University of Kaiserslautern Excellence Cluster 'Dependable Adaptive Systems and Mathematical Modeling' and the "Graduiertenkolleg Mathematik und Praxis" of the University of Kaiserslautern.

C. Kirchner was partly funded by a fellowship from the DAAD ("DAAD Doktorandenstipendium D/05/43739").

approach has been successfully applied in different areas. Among the variety of literature we only mention some examples, like optimal control of fluid flows [10], optimal semiconductor design [11] or general initial value control of hyperbolic equations [16, 17]. For optimal control in the context of networks we refer to [9] for an adjoint calculus in the context of traffic flow networks. To compute the optimal control, the continuous optimality system is discretized and usually solved by a descent type method, see [15, 13, 14]. This approach can therefore be seen as *optimize–then–discretize*. Alternatively, we can proceed by first discretizing the constraints and cost functional and then optimize the finite–dimensional optimization problem; this strategy is known as *discretize–then–optimize*. For our supply network model the discretization can be chosen, such that the optimization problem is in fact a mixed–integer programming problem, see [6] and section 4, below. This is mainly due to the fact that the governing dynamics in the supply network are linear in the state (but not in the control) variables. In [6] further extensions to the mixed–integer problem have been investigated, e.g., finite size buffers, inflow profile optimization or processor shut–down due to maintenance. In this work, we derive the continuous optimality system and show that the mixed–integer formulation is also a valid discretization of the discretized continuous optimality system, i.e., both approaches *discretize–then–optimize* and *optimize–then–discretize* lead to the same continuous optimal control if the discretization width tends to zero. Furthermore, we investigate the numerical properties of the two approaches by comparing computing times for the solution to the mixed–integer model with a steepest descent method based on the adjoint equations.

The paper is organized as follows. In section 2 we review the supply chain model and introduce the continuous optimal control problem that is to be investigated. In section 3 we recall the discretization leading to the mixed–integer model, see [6]. Then we derive both the optimality system for the discrete and continuous model. The numerical results are presented in section 4 and contain a validation of our adjoint calculus by comparison with finite–difference approximations and optimization results of the adjoint–based method and the mixed–integer problem.

2. Modeling supply networks.

In the following we consider a directed graph $(\mathcal{V}, \mathcal{A})$ consisting of a set of arcs $\mathcal{A}$ and a set of vertices $\mathcal{V}$. Each arc correponds to one processor (or supplier). The length of the processor corresponding to arc $e \in \mathcal{A}$ is given by the interval $L^e = [a^e, b^e]$. The maximal processing capacity μ^e and the processing velocity v^e of each processor are constant parameters on each arc. According to the assumption that each processor possesses a queue, we locate a queue at the vertex v in front of the processor. For a fixed vertex v, the set of ingoing arcs is denoted by δ_v^- and the set of outgoing arcs by δ_v^+. In the case of more than one outgoing arc, we introduce distribution rates $A^{v,e}(t)$, $v \in \mathcal{V}_d$ where $\mathcal{V}_d \subset \mathcal{V}$ denotes the set of dispersing junctions. Those rates describe the distribution of incoming parts among the outgoing processors and are later subject to optimization. The functions $A^{v,e}$ are required to satisfy $0 \leq A^{v,e}(t) \leq 1$ and $\sum_{e \in \delta_v^+} A^{v,e}(t) = 1$ for all times $t > 0$.

As a next step, we briefly recall the continuous supply network model and its corresponding optimal control problem. For more details we refer to [7, 8]. The continuous supply network model consists of a coupled system of partial and ordinary differential equations. Here, the transport inside each processor e is governed

by a simple advection equation:

$$\partial_t \rho^e + \partial_x f^e(\rho^e) = 0, \tag{1}$$

where

$$f^e(\rho^e) = v^e \rho^e \tag{2}$$

for some given velocity v^e and where the density of the parts is given by ρ^e on each arc e. Whenever a processor is connected to another processor of possibly different maximal capacity μ^e, we introduce a buffering zone for the incoming but not yet processed parts. To describe the buffering we introduce the time–dependent function $q^e(t)$ describing the load of the buffer or queue. The dynamics of the buffering is governed by the difference of all incoming and outgoing parts at the connection point: If the queue is empty, the outgoing flux is either a percentage of the sum of all incoming fluxes given by $A^{v,e}(t)$ or the maximal processing capacity. In the first case the queue remains empty, in the second case the queue increases. Last, if the queue is full, the queue is always reduced with a capacity determined by the distribution rates $A^{v,e}$ and the capacities of the connected arcs, see below for the mathematical statement. Finally, we introduce a measure for the performance of the supply chain network. A general cost functional is given for example by (3a). This particular choice of the cost functional aims at the minimization of the size of queues and the number of parts in the network. However, other choices are possible; in subsection 4.2 we present an example in which we just want to maximize the output of a particular supply network.

Summarizing, (3) constitutes a constrained optimal control problem where the constraints are given by linear transport and ordinary differential equations. The controls are the distribution rates $A^{v,e}$ and the dependent states are the vectors $\vec{\rho}^e := (\rho^e)_{e\in\mathcal{A}}$ and $\vec{A}^v := (A^{v,e})_{e\in\,\delta_v^+}$.

$$\min_{A^{v,e}(t),v\in\mathcal{V}_d} \sum_{e\in\mathcal{A}} \int_0^T \int_{a^e}^{b^e} f^e(\rho^e(x,t))\,dx\,dt + \int_0^T q^e(t)\,dt \tag{3a}$$

$$\text{subject to} \quad e\in\mathcal{A},\ v\in\mathcal{V},\ t\in(0,T),\ x\in[a^e,b^e] \tag{3b}$$

$$\partial_t \rho^e(x,t) + \partial_x f^e(\rho^e(x,t)) = 0 \tag{3c}$$

$$\partial_t q^e(t) = A^{v,e}(t) \sum_{\bar{e}\in\delta_v^-} f^{\bar{e}}(\rho^{\bar{e}}(x_v^{\bar{e}},t)) - f^e(\rho^e(x_v^e,t)) \tag{3d}$$

$$f^e(\rho^e(x_v^e,t)) = \begin{cases} \min\{A^{v,e}(t)\Big(\sum_{\bar{e}\in\delta_v^-} f^{\bar{e}}(\rho^{\bar{e}}(x_v^{\bar{e}},t))\Big), \mu^e\}; & q^e(t)=0 \\ \mu^e; & q^e(t)>0 \end{cases} \tag{3e}$$

We are concerned with the numerical solution to the previous optimal control problem. For further investigation we apply the following modifications and simplifications: First, in order to avoid the the discontinuous dependence on the queue-length in (3e), we make use of the reformulation presented in [1]. There, equation (3e) has been replaced

$$f^e(\rho^e(x_v^e,t)) = \min\{\mu^e, \frac{q^e(t)}{\varepsilon}\} \quad \text{with } \epsilon \ll 1. \tag{4}$$

See [1] for further remarks. Since adjoint calculus requires the constraints to be differentiable, we replace the function $y \to \min(y/\epsilon, \mu^e)$ in (4) by any smooth approximation $\psi^{e,\delta}(y)$ for the computations following. To be more precise, we assume

there are families of smooth functions $\{\psi^{e,\delta}\}$ such that

$$\lim_{\delta\to 0}\psi^{e,\delta}(y) = \min(y/\epsilon, \mu^e)\ \forall y, \forall e. \tag{5}$$

For notational convenience we drop the superindex δ in the following, since the calculations remain true for all $\delta > 0$. Third, we simplify the notation by introducing functions $h^e(\vec{\rho}, \vec{A}^{v,e})$ for each edge e (resp. $\tilde{e}$) and fixed $v \in \mathcal{V}$ such that $e \in \delta_v^+$ (resp. $\tilde{e} \in \delta_v^+$). We define

$$h^e(\vec{\rho}^e, \vec{A}^v) = A^{v,e}(t)\sum_{\bar{e}\in\delta_v^-} f^{\bar{e}}(\rho^{\bar{e}}), \forall e \in \delta_v^+\backslash\{\tilde{e}\}, \tag{6a}$$

$$h^{\tilde{e}}(\vec{\rho}^e, \vec{A}^v) = \left(1 - \sum_{e\neq\tilde{e}} A^{v,e}(t)\right)\sum_{\bar{e}\in\delta_v^-} f^{\bar{e}}(\rho^{\bar{e}}). \tag{6b}$$

Note that with this definition the assumption $\sum_{e\in\delta_v^+} A^{v,e} = 1$ can be omitted. For example, for an intersection with $\delta_v^- = \{1\}$ and $\delta_v^+ = \{2,3\}$ we have the more explicit form

$$h^2(\vec{\rho}^e, \vec{A}^v) = A^{v,2}(t)f^1(\rho^1),\ h^2(\vec{\rho}^e, \vec{A}^v) = (1 - A^{v,2}(t))f^1(\rho^1). \tag{7}$$

Finally, we summarize the previous modifications and restate the optimal control problem for all $e \in \mathcal{A}$, $v \in \mathcal{V}$, $t \in (0,T)$, $x \in [a^e, b^e]$:

$$\min_{A^{v,e}(t), v\in\mathcal{V}_d}\sum_{e\in\mathcal{A}}\int_0^T\int_{a^e} v^e\rho^e(x,t)\,dx\,dt + \int_0^T q^e(t)\,dt \tag{8a}$$

subject to

$$\partial_t\rho^e(x,t) + v^e\partial_x\rho^e(x,t) = 0,\ \rho^e(x,0) = 0, v^e\rho^e(a,t) = \psi^e(q^e) \tag{8b}$$

$$\partial_t q^e(t) = h^e(\vec{\rho}^e, \vec{A}^v) - \psi^e(q^e),\ q^e(0) = 0. \tag{8c}$$

As technical detail we need to introduce boundary data for those arcs $e \in \mathcal{A}$ which are incoming to the network, i.e., such that $\delta_v^- = \emptyset$. Here, we assume inflow data $\rho_0(t)$ to be given and set $\rho^e(a,t) = \rho_0(t)$ for all $v \in \mathcal{V}$ and $e \in \delta_v^+$ and $\delta_v^- = \emptyset$. From now on we neglect this technical point.

3. Derivation of Optimality Systems for the Optimal Control Problem.

In this section we focus on the issue of numerically solving the optimal control problem. Different approaches are possible. In [6] the optimal control problem has been solved by reformulating it as mixed–integer model. This is possible, if one introduces a coarse grid discretization of (8). In the next subsection we will derive a discrete optimality system for this discretization and – contrary to [6] – solve the latter directly by nonlinear optimization methods. This approach is known as "first discretize then optimize". For numerical comparison between this approach and the one taken in [6] we refer to section 4. Formally, one can also derive the continuous optimality system and discretize the latter. This method is referred to as "first optimize then discretize"; we present the corresponding results in Subsection 3.2. Furthermore, the relation between the approaches "first discretize then optimze" and "first optimize then discretize" will be given in subsection 3.2 for the optimal control problem (8).

3.1. **Optimality System of the Discrete Optimal Control Problem.** First, we consider the discrete optimality system. A coarse grid discretization in space of (8b) is obtained by just a two-point Upwind discretization and (8c) is discretized usind the explicit Euler method. Each arc has length L^e and we introduce a step size Δt such that the CFL condition for each arc and the stiffness restriction of the ordinary differential equation are met. The time steps t_j are numbered by $j = 0, \dots, T$. We use the following abbreviations for all e, j :

$$\rho_j^{e,b} := \rho^e(b^e, t_j),\ \rho_j^{e,a} := \rho^e(a^e, t_j),\ q_j^e := q^e(t_j), A_j^{v,e} := A^{v,e}(t_j) \tag{9}$$

$$h_j^e := h^e(\vec{\rho}^e(x, t_j), \vec{A}^v(t_j)). \tag{10}$$

Due to the boundary condition $v^e \rho^e(a, t) = \psi^e(q^e(t))$ we replace the discrete variable $\rho_j^{e,a}$ by $\psi^e(q_j^e)$ and therefore, $\rho_j^{e,a}$ does not appear explicitly in the discrete optimal control problem below. For the initial data we have

$$\rho_0^{e,b} = \rho_0^{e,a} = q_0^e = 0,\ \forall e. \tag{11}$$

Finally, the discretization of problem (8) reads for $j \geq 1, e \in \mathcal{A}, v \in \mathcal{V}$:

$$\min_{\vec{A}^v,\, v \in \mathcal{V}_d} \sum_{e \in \mathcal{A}} \sum_{j=1}^{T-1} \Delta t \left(\frac{L^e}{2} (\psi^e(q_j^e) + v^e \rho_j^{e,b}) + q_j^e \right) \tag{12a}$$

subject to

$$\rho_{j+1}^{e,b} = \rho_j^{e,b} + \frac{\Delta t}{L^e} (\psi(q_j^e) - v^e \rho_j^{e,b}) \tag{12b}$$

$$q_{j+1}^e = q_j^e + \Delta t (h_j^e - \psi^e(q_j^e)) \tag{12c}$$

For deriving the discrete optimality system we state the precise definition of h^e in the case of the following intersections, see Figure 1. In case A h^2 is independent of

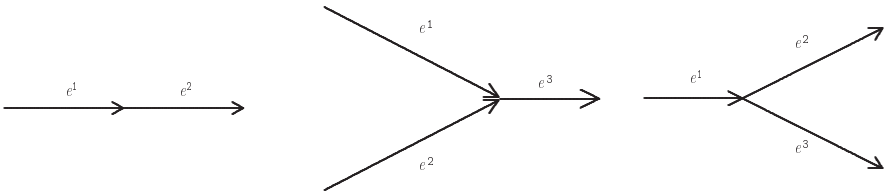

FIGURE 1. Sample intersections labeled as type A,B and C (from left to right)

$\vec{A}^v$ and we have $h^2(\vec{\rho}^e, \vec{A}^v) = v^1 \rho^1(b, t)$. Similarly, in case B we obtain $h^3(\vec{\rho}^e, \vec{A}^v) = v^1 \rho^1(b, t) + v^2 \rho^2(b, t)$. Finally, as already stated, we have in the controlled case C : $h^2(\vec{\rho}^e, \vec{A}^v) = A^{v,2}(t) v^1 \rho^1(b, t),\ h^3(\vec{\rho}^e, \vec{A}^v) = (1 - A^{v,2}(t)) v^1 \rho^1(b, t)$.

Remark 3.1. *The previous system for $\delta = 0$ can be reformulated as mixed–integer problem* [6] *by adding binary variables to reformulate the relation $v^e \rho^e(x_v^e, t) = \min\{\mu^e, q^e(t)/\epsilon\}$: let ζ_j^e be a binary variable, i.e., $\zeta_j^e \in \{0; 1\}$. Then,*

$$v^e \rho_j^{e,a} = \min\{\mu^e, q^e(t)/\epsilon\}$$

is equivalent to

$$\begin{aligned} \mu^e \zeta_j^e &\leq v^e \rho_j^{e,a} \leq \mu^e \\ \frac{q_j^e}{\varepsilon} - M\zeta_j^e &\leq v^e \rho_j^{e,a} \leq \frac{q_j^e}{\varepsilon} \\ \mu^e \zeta_j^e \leq \frac{q_j^e}{\epsilon} &\leq \mu^e(1-\zeta_j^e) + M\zeta_j^e \end{aligned}$$

for $M > 0$ sufficiently large. Indeed, the value ζ_j^e is determined by the relation of μ^e and q_j^e/ϵ. E.g., if $\mu^e > q_j^e/\epsilon$, then we obtain $\zeta_j^e = 1$. Since both formulations are equivalent, we derive the discrete optimality system by considering system (12).

Now it is straightforward to derive the discrete optimality system for (12). We denote the Lagrange multipliers for the discretized partial differential equation by λ_j^e and for the discretized ordinary differential equation by p_j^e. The discrete Lagrangian is given by

$$L(\vec{\rho}_j^e, \vec{q}_j^e, \vec{A}_j^v, \vec{\lambda}_j^e, \vec{p}_j^e) = \sum_{e\in\mathcal{A}} \sum_{j=1}^{T-1} \Delta t \left(\frac{L^e}{2}(\psi^e(q_j^e) + v^e \rho_j^{e,b}) + q_j^e \right) - \tag{13a}$$

$$\sum_{e\in\mathcal{A}} \sum_{j=1}^{T} \Delta t L^e \lambda_j^e \left(\frac{\rho_{j+1}^{e,b} - \rho_j^{e,b}}{\Delta t} - \frac{\psi(q_j^e) - v^e \rho_j^{e,b}}{L^e} \right) - \tag{13b}$$

$$\sum_{e\in\mathcal{A}} \sum_{j=1}^{T} \Delta t\, p_j^e \left(\frac{q_{j+1}^e - q_j^e}{\Delta t} - (h_j^e - \psi(q_j^e)) \right), \tag{13c}$$

if we set $\lambda_T^e = p_T^e = 0$. Assuming sufficient constraint qualifications the first–order optimality system is given by equations (12c) and (12b) and the following additional equations for $j \leq T, e \in \mathcal{A}$ and $v \in \mathcal{V}$:

$$\lambda_{j-1}^e = \Delta t \frac{v^e}{2} + \lambda_j^e - \frac{\Delta t}{L^e}\left(\phi_j^e - v^e \lambda_j^e\right), \tag{14a}$$

$$\phi_j^e := \sum_{\bar{e}\in\delta_v^+ \text{ s.t. } e\in\delta_v^-} p_j^{\bar{e}} \frac{\partial}{\partial \rho^e} h_j^{\bar{e}}, \tag{14b}$$

$$p_{j-1}^e = \Delta t(1 + \frac{L^e}{2}(\psi^e)'(q_j^e)) + p_j^e - \Delta t\left(p_j^e - \lambda_j^e\right)(\psi^e)'(q_j^e), \tag{14c}$$

$$0 = \sum_{e\in\delta_v^+} p_j^e \frac{\partial}{\partial A^{v,\bar{e}}} h_j^e \tag{14d}$$

The summation in the definition of the function ϕ^e is understood in the following way: For a fixed intersection $v \in \mathcal{V}$ such that $e \in \delta_v^-$ we sum over all $\bar{e} \in \delta_v^+$. Hence, the function ϕ^e depends on the type of intersection and for clearity we state its explicit form for the cases $A - C$ introduced above: In case A we have $\phi_j^2 = 0$ and $\phi_j^1 = p_j^1 v^1$. In case B we obtain $\phi_j^1 = p_j^3 v^3$ and $\phi_j^2 = p_j^3 v^3$. Finally, for the interesting case C we find $e = 1$ which implies $\phi_j^1 = A^{v,2} p_j^2 v^2 + (1 - A^{v,2}) p_j^3 v^3$. Furthermore, we obtain with the previous definitions for $\bar{e} \neq \tilde{e}$:

$$\sum_{e\in\delta_v^+} p_j^e\, \partial_{A^{v,\bar{e}}} h_j^e = \left(p_j^{\bar{e}} - p_j^{\tilde{e}}\right) \sum_{e\in\delta_v^-} v^e \rho_j^e. \tag{15}$$

Summarizing, the optimality system to (12) is given by (12c,12b) and (14). Changing the objective functions only affects the first term on the right hand side in formulas (14a) and (14c). In section 4 we will present results on the solution to the

optimality system of (discretized) mixed partial and ordinary differential equations for different objective functions.

3.2. **Optimality system of the continuous optimal control problem.** In this subsection, we turn our attention to the continuous optimality system for (8); we will show that the optimality system (12c), (12b) and (14) from subsection 3.1 is a valid discretization of the former. For the derivation of the continuous optimality system to (8) the Lagrangian reads

$$L(\vec{\rho}^e, \vec{A}^v, \vec{q}^e, \vec{\Lambda}^e, \vec{P}^e) = \sum_{e\in\mathcal{A}} \int_0^T \int_{a^e}^{b^e} v^e \rho^e dx dt + \int_0^T q^e dt - \tag{16a}$$

$$\sum_{e\in\mathcal{A}} \int_0^T \int_{a^e}^{b^e} \Lambda^e \partial_t \rho^e + \Lambda^e v^e \partial_x \rho^e dx dt - \tag{16b}$$

$$\sum_{e\in\mathcal{A}} \int_0^T P^e \left(\partial_t q^e - h^e(\vec{\rho}^e, \vec{A}^v) + \psi^e(q^e) \right) dt \tag{16c}$$

In this setup the adjoint variables are denoted as $\Lambda^e(x,t)$ and $P^e(x,t)$; we use captial letters to highlight their difference from the previously introduced quantities λ_j^e and p_j^e. The relation between these variables is discussed below. We formally obtain the continuous optimality system for all $t, x \in [a^e, b^e], e \in \mathcal{A}$ as

$$\partial_t \rho^e + v^e \partial_x \rho^e = 0,\ \rho^e(x,0) = 0,\ v^e \rho^e(a,t) = \psi^e(q^e), \tag{17a}$$

$$\partial_t q^e = h^e(\vec{\rho}^e, \vec{A}^v) - \psi^e(q^e),\ q^e(0) = 0, \tag{17b}$$

$$-\partial_t \Lambda^e - v^e \partial_x \Lambda^e = v^e,\ \Lambda^e(x,T) = 0, \tag{17c}$$

$$v^e \Lambda^e(b,t) = \sum_{\bar{e}\in\delta_v^+ \text{ s.t. } e\in\delta_v^-} P^{\bar{e}}(t)\ \frac{\partial}{\partial \rho^{\bar{e}}} h^{\bar{e}}(\vec{\rho}^e, \vec{A}^v), \tag{17d}$$

$$-\partial_t P^e = 1 - (P^e - \Lambda^e(a,t))\,(\psi^e)'(q^e),\ P^e(T) = 0, \tag{17e}$$

$$\sum_{e\in\delta_v^+} P^e\ \frac{\partial}{\partial A^{v,\bar{e}}} h^e(\vec{\rho}^e, \vec{A}^v) = 0. \tag{17f}$$

Recall that in the limit case $\delta = 0$, we have by definition $\psi^e(y) \to \min\{y/\epsilon, \mu^e\}$. Therefore, we obtain

$$(\psi^e)'(q^e) \to \frac{1}{\epsilon} H(\mu^e - q^e/\epsilon),\ \delta \to 0, \tag{18}$$

where $H(x)$ is the Heaviside function. Hence, in the limit the dynamics of the adjoint queue P^e is governed by a discontinuous right–hand side.

Finally, we show that in fact (14), (12b), (12c) is a suitable discretization of (17). We proceed by reformulating the discrete optimal control problem in the introduced variables defined by

$$\Lambda_j^{e,a} := \lambda_j^e - \frac{L^e}{2},\ \ P_j^e := p_j^e, \tag{19}$$

Then, (14),(12c),(12b) read

$$\frac{\rho_{j+1}^{e,b}-\rho_j^{e,b}}{\Delta t} = -\frac{v^e}{L^e}(\rho_j^{e,b}-\rho_j^{e,a}),\ \rho_0^e=0,\ v^e\rho_j^{e,a}=\psi^e(q_j^e), \tag{20a}$$

$$\frac{q_{j+1}^e-q_j^e}{\Delta t} = h_j^e-\psi^e(q_j^e),\ q_0^e=0 \tag{20b}$$

$$\frac{\Lambda_{j-1}^{e,a}-\Lambda_j^{e,a}}{\Delta t} = v^e-\frac{v^e}{L^e}\left(\Lambda_j^{e,b}-\Lambda_j^{e,a}\right),\ \Lambda_T^{e,a}=0, \tag{20c}$$

$$v^e\Lambda_j^{e,b} = \sum_{\bar e\in\delta_v^+}\ \sum_{\text{s.t. } e\in\delta_v^-} p_j^{\bar e}\frac{\partial}{\partial\rho^e}h_j^{\bar e}, \tag{20d}$$

$$\frac{P_{j-1}^e-P_j^e}{\Delta t} = 1-\left(P_j^e-\Lambda_j^{e,a}\right)(\psi^e)'(q_j^e), \tag{20e}$$

$$0 = \sum_{e\in\delta_v^+} P_j^e\,\frac{\partial}{\partial A^{v,\bar e}}h_j^e \tag{20f}$$

Obviously, (20) is an Upwind and explicit Euler discretization of (17). Note that the discrete Lagrangian multiplier λ_j^e and the discretized Lagrange multiplier $\Lambda_j^{e,a}$ satisfy

$$\Lambda_j^{e,a}=\lambda_j^e+O(L^e) \tag{21}$$

and L^e is in fact the discretization stepwidth in space. Therefore, if we formally let $L^e,\Delta t\to 0$ for $L^e/\Delta t$ fixed, we see that $\lambda^e\to\Lambda^e$ and furthermore, the discrete Lagrangian tends to the continuous Lagrangian.

4. **Numerical results.** In fact, there are two different approaches for solving the optimal control problem given in (12b), (12c) and (14). On the one hand, we use a steepest descent method for a suitable cost functional. We consecutively solve the equations of state (12b) and (12c) for a given initial control $\vec{A}_0^v\equiv 0$ and the adjoint equations (14a)–(14c) which in turn are needed to evaluate the gradient (15). Using the Armijo–Goldstein rule for the choice of the stepsizes we update the controls $\vec{A}_0^v$ and iterate the described procedure.

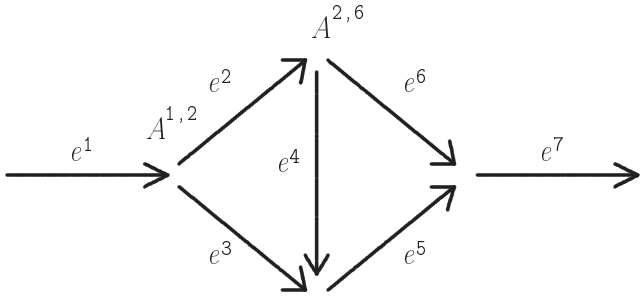

FIGURE 2. Sample network with controls $A_j^{1,2}$ and $A_j^{2,6}$

On the other hand, we reformulate (12b), (12c) and (14) as a mixed-integer programming (MIP) model. The continuous optimal control probelm (3) is discretized

using a two point Upwind scheme for the PDE and a explicit Euler method for the ODE (see [6] for some details). The proposed MIP model is stated as follows:

$$\min \sum_{e\in\mathcal{A}} \sum_{j=1}^{T-1} \Delta t \left(\frac{L^e}{2}(v^e \rho_j^{e,a} + v^e \rho_j^{e,b}) + q_j^e \right) \tag{22a}$$

subject to

$$\begin{aligned}
\rho_{j+1}^{e,b} &= \rho_j^{e,b} + \frac{\Delta t}{L^e}(v^e \rho_j^{e,a} - v^e \rho_j^{e,b}) && (22b)\\
\sum_{e\in\delta_v^-} v^e \rho_j^{e,b} &= \sum_{e\in\delta_v^+} h_j^e && (22c)\\
q_{j+1}^e &= q_j^e + \Delta t(h_j^e - v^e \rho_j^{e,a}) && (22d)\\
\mu^e \zeta_j^e &\le v^e \rho_j^{e,a} \le \mu^e && (22e)\\
\frac{q_j^e}{\varepsilon} - M\zeta_j^e &\le v^e \rho_j^{e,a} \le \frac{q_j^e}{\varepsilon} && (22f)\\
\rho_j^{e,a}, \rho_j^{e,b}, h_j^e, q_j^e &\ge 0, && (22g)\\
\zeta_j^e &\in \{0,1\}, && (22h)
\end{aligned}$$

with $e \in \mathcal{A}$, $j = 1, \ldots, T$, and M a sufficiently large constant. The essential difference to (12b) is to rewrite the nonlinearity in (4) by introducing binary variables ζ_j^e. This leads finally to a mixed-integer problem and not just a linear programming (LP) model. For solving the mixed–integer problem the standard optimization software solver ILOG CPLEX [12] is used.

4.1. **Gradient computations.** At first we compare the gradient of the cost functional obtained by finite differences to the gradient obtained by the adjoint equations for a suitable network. We use the network depicted in Figure 2 for this test since it has only two variable controls $A_j^{1,2}$ and $A_j^{2,6}$ at time j (recall that $A_j^{1,3} = 1 - A_j^{1,2}$ and $A_j^{2,5} = 1 - A_j^{2,6}$ due to the coupling conditions). We discretize the control–space $[0,1] \times [0,1]$ using 16 points in both the $A_j^{1,2}$ and $A_j^{2,6}$ component.

We set the time–horizon $T = 4$, use $NT = 200$ time–intervals and set $\epsilon = 1$. We use a one–sided forward difference scheme to compare the gradient at time–interval j, $j = 1, \ldots, NT$:

$$\partial_{A_j^{v,e}} J(\vec{A}^v) := \frac{J(\vec{A}^v + \delta) - J(\vec{A}^v)}{\delta} \tag{23}$$

where $\delta = 0.001$. For the cost–functional we chose the nonlinear function

$$J(\vec{A}^v) := \left(\sum_{e\in\mathcal{A}} \sum_j \Delta t \left(\frac{L^e}{2}(\psi(q_j^e) + v^e \rho_j^{e,b}) + q_j^e \right) \right)^2 . \tag{24}$$

Further, we set $L^3 = L^6 = 10$ and $L^e = 1$ for $e \in \mathcal{A}\backslash\{e^3, e^6\}$. The processing rates are $\mu^e = 1$, $\forall e$. This implies that the lowest functional value should be attained for $A_j^{1,2} = 1$ and $A_j^{2,6} = 0$ for all j as confirmed in Figure 3. The inflow–profile on e^1 is chosen as

$$f^{in}(t) = \begin{cases} 0.852 & t \le 2 \\ 0 & t > 2 \end{cases}$$

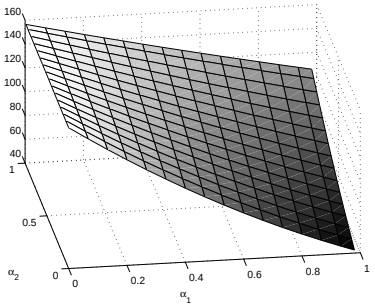

FIGURE 3. Plot of the cost functional (24) corresponding to Figure 2.

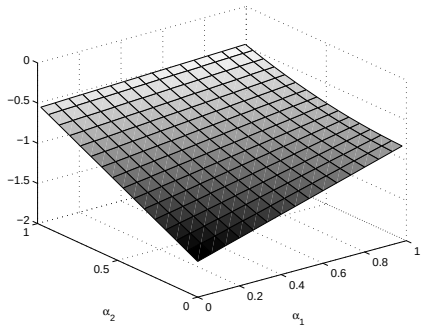

FIGURE 4. First component of the gradient computed by the adjoint scheme at $j = 50$

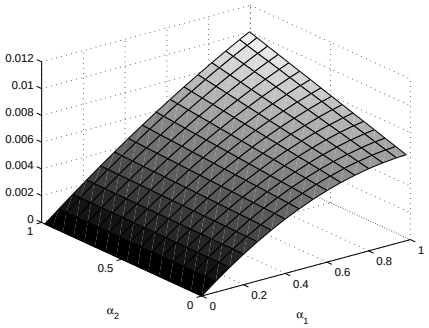

FIGURE 5. Second component of the gradient computed by the adjoint scheme at $j = 200$

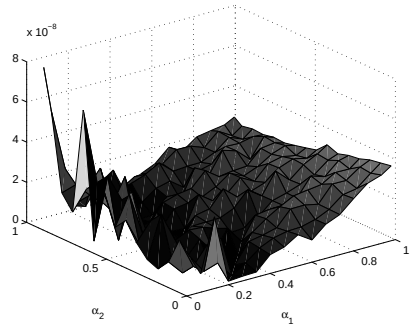

FIGURE 6. Relative error in partial w.r.t. $A^{2,6}_{200}$ at $j = 200$

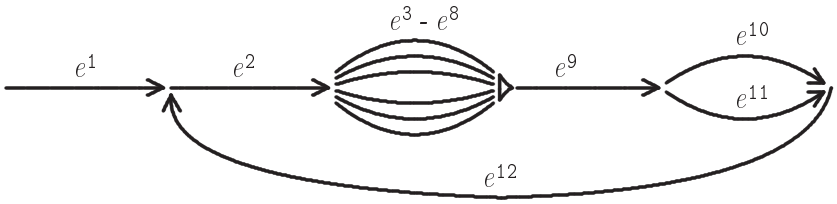

FIGURE 7. Sample network

With this inflow profile the gradient w.r.t. $A^{1,2}_{50}$ is nonzero and is depicted in Figure 4. The controls $A^{1,2}_j$ with $j > 2$ can be chosen arbitrarily since the inflow is zero and hence the gradient w.r.t. these controls needs to vanish. However, since

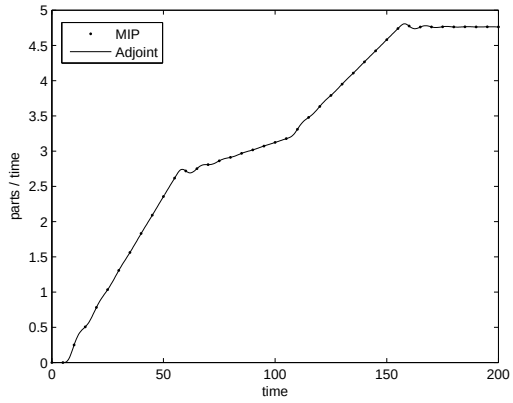

FIGURE 8. Optimal output for processor 12.

in this particular setup queue 2 is nonempty at time $j = 200$, the gradient w.r.t. $A^{2,6}_{200}$ does not vanish, cf. Figure 5. The relative error in the second component is of order $1e^{-8}$ and can be found in Figure 6.
Finally, we mention that we have conducted extensive test with different objective–functionals and varying the parameters $\epsilon \in [0.01, 1]$, $\delta \in \{1e^{-2}, 1e^{-3}, 1e^{-4}, 1e^{-5}\}$ and $NT \in [20, 400]$; we never encountered a relative error in the gradient larger than $1e^{-6}$.

4.2. **Quality of solutions of discrete adjoint calculus compared with the mixed-integer model.** As a next step, we compare results computed by the adjoint approach and the mixed-integer programming (MIP) model presented in (22). We show that this kind of discretization induces same results for the cost functional as the discrete adjoint approach by focussing on the optimal control problem of routing of goods through a network.

e	1	2	3	4	5	6	7	8	9	10	11	12
μ^e	100	8	10	0.5	0.5	10	0.5	2	20	3.5	2.5	8

TABLE 1. Processing rates μ^e

NT	Adjoint	MIP
200	7.31	5.52
400	26.10	17.06
800	45.10	68.09
2000	124.58	592.61

TABLE 2. CPU times in sec for sample network Figure 7

In the following, we consider the network in Figure 7. It consists of 11 processors and queues and we have the six free controls $A^{2,3}(t), A^{2,4}(t), A^{2,5}(t), A^{2,6}(t), A^{2,7}(t)$ and $A^{9,10}(t)$.

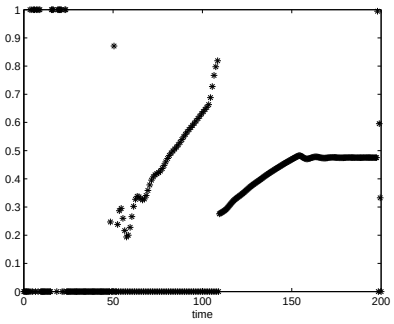

FIGURE 9. Plot of the distribution rate $A_j^{9,10}$ computed by the MIP.

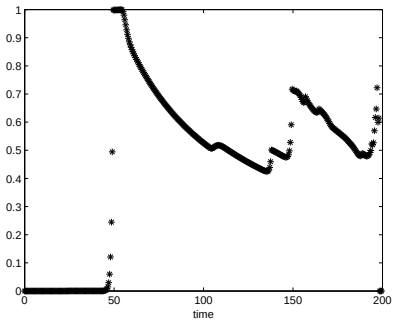

FIGURE 10. Plot of the distribution rate $A_j^{9,10}$ computed by the adjoint approach.

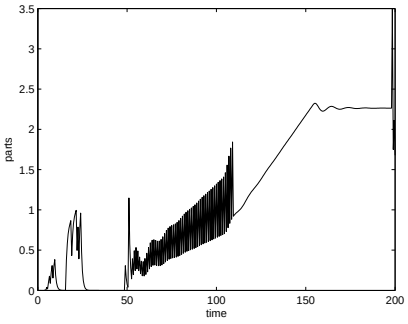

FIGURE 11. Optimal queue length q_j^{10} computed by the MIP.

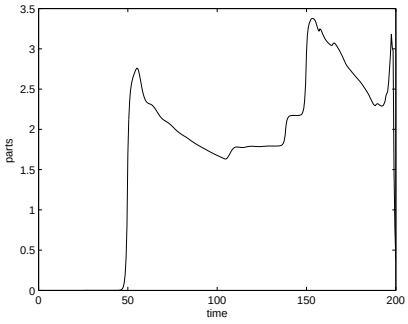

FIGURE 12. Optimal queue length q_j^{10} computed by the adjoint approach.

The artificial arc 1 is used to prescribe an inflow profile which is given by

$$f(t) = \begin{cases} 0.5 & 0 \leq t \leq \frac{T}{4} \\ 0.1 & \frac{T}{4} < t \leq \frac{T}{2} \\ 0.3 & \frac{T}{2} < t \leq \frac{3}{4}T \\ 0 & \frac{3}{4}T < t \leq T \end{cases} \tag{25}$$

Our goal is to maximize the output of processor 12 on a given time–interval $[0, T]$. We use an equidistant time–discretization with NT time–intervals and choose the following reduced cost functional

$$J(\vec{A}^v) = \sum_{j=2}^{NT+1} -\frac{v^{12}\rho_j^{12,b}}{j}. \tag{26}$$

In the example below we define $T = 200$, $NT = 400$, $\epsilon = 1$ and set $L^e = v^e = 1$ for all edges except for $e = 2$; here we use $L^2 = 1$ and $v^2 = 2$. The corresponding processing rates are given in Table 1.

In Figures 8 – 12 we present results for the optimal routing problem by pointing out similarities and differences between the adjoint and discrete approach. The computation of the adjoint approach takes $37.781s$ using 32 iterations and for the MIP $16.60s$ using 16 iterations. In the adjoint approach, we terminate the iteration if the relative error of two consecutive iterates is less than $tol := 1e^{-6}$ - consistent with the default accuracy in ILOG CPLEX [12]. Both approaches yield an optimal functional value of $J^*(\vec{A}^v) = -6.49$.

In Figure 8 we plot the optimal outflow profile computed by the two approaches. We observe that for this particular example the curves coincide. However, the computed optimal controls and time evolution of the queues differ considerably. In Figures 9 and 10 we plot the optimal control feeding parts into queue 10. Furthermore, we present the evolution over time for the queue 10 in Figures 11 and 12 for the MIP and the adjoint approach, respectively, and the maximum queue–length in Figures 11 and 12.

Since the optimal functional values coincide we see that we do not have a unique minimizer to our optimal control problem.

4.3. **Computational times.** The numerical results conclude with a comparison of computational times of the adjoint–based approach and the mixed-integer formulation. Our computations are performed on the network given in Figure 7 with default parameters $v^e = L^e = 1, e = 2, \ldots, 12$, $\epsilon = 1$ and time horizon $T = 200$. To obtain a stable discretization both models have to satisfy the following restriction:

$$\Delta t \leq \min\{\epsilon; \frac{L^e}{v^e} : e \in \mathcal{A}\}. \tag{27}$$

Resulting from (27) the parameter NT describes the number of time intervals. We increase NT by varying the ratio of L^1/v^1. The MIP is solved using the interior point method implemented in ILOG CPLEX [12].

As Table 2 indicates the MIP is superior if one wants to use up to approximately 600 time-steps (corresponding to $\Delta t \in [0.\overline{3}, 1]$). As NT increases the adjoint approach becomes more attractive. For values of $\Delta t < 0.\overline{3}$ it computes an optimal solution faster than the MIP. At present the MIP fails to compute a solution for $\Delta t \leq 0.05$ since the system becomes too large and the preprocessing procedure produces infeasible solutions.

5. **Summary.** In this work we have extended certain optimization techniques to supply networks. We derive a continuous and discrete optimality system and show that the latter can be interpreted as an upwind and explicit Euler discretization of the former. The results are compared to the ones obtained by a mixed–integer formulation. For the testcases under consideration the optimal solutions (i.e., the optimal values of the objective function) of the adjoint approach and the MIP formulation introduced in [6] coincide. However, the optimal controls differ qualitatively since they are not unique. The usage of the adjoint method as presented here is limited. With the MIP more complex and praxis–relevant questions can be modeled and solved quite easily; processor shutdown due to maintenance or min–up/min–down times are just two examples. It is desirable to develop an adjoint–based approach capable of treating these aspects.

REFERENCES

[1] D. Armbruster, C. de Beer, M. Freitag, T. Jagalski and C. Ringhofer, *Autonomous control of production networks using a pheromone approach*, Submitted, Physica, 2005.

[2] D. Armbruster, P. Degond and C. Ringhofer. *A model for the dynamics of large queuing networks and supply chains*, SIAM J. Applied Mathematics, Vol. 66 (2006), 896–920.

[3] D. Armbruster, P. Degond and C. Ringhofer. *Kinetic and fluid models for supply chains supporting policy attributes*, Submitted, Transp. Theory and Stat. Phys., 2004.

[4] D. Armbruster, D. Marthaler and C. Ringhofer. *Kinetic and fluid model hierarchies for supply chains*, SIAM J. on Multiscale Modeling, Vol. 2(1)(2004) 43-61.

[5] C. D'Apice, R. Manzo. *A fluid dynamic model for supply chains*, NHM, Vol. 1 (2006) 379-398.

[6] 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*, Submitted, 2006.

[7] S. Göttlich, M. Herty and A. Klar. *Network models for supply chains*, Comm. Math. Sci., Vol. 3(2005), 545–559.

[8] S. Göttlich, M. Herty and A. Klar. *Modelling and optimization of supply chains on complex networks*, Comm. Math. Sci., Vol. 4(2006), 315–330.

[9] M. Herty, M. Gugat, A. Klar and G. Leugering. *Optimal control for traffic flow networks*, Journal of Optimization Theory and Applications, Vol. 126(2005),589–615.

[10] M. Hinze and K. Kunisch. *Second order methods for optimal control of time-dependent fluid flow*, Siam J. Control and Optim., Vol. 40 (2001), 925–946.

[11] M. Hinze and R. Pinnau. *An optimal control approach to semiconductor design* M3AS, Vol. 12(2002), 89–107.

[12] ILOG CPLEX Division, 889 Alder Avenue, Suite 200, Incline Village, NV 89451, USA. Information available at URL http://www.cplex.com.

[13] C. Kelly. "Iterative Methods for Optimization," SIAM Frontiers in Applied Mathematics, 1999

[14] S. G. Nash, and A. Sofer. "Linear and Nonlinear Programming," The McGraw–Hill Companies, New York, St. Louis, San Francisco, 1996.

[15] P. Spellucci. "Numerische Verfahren der nichtlinearen Optimierung," Birkhäuser Verlag, Basel, Boston, Berlin, 1993.

[16] S. Ulbrich. *A sensitivity and adjoint calculus for discontinuous solutions of hyperbolic conservation laws with source terms*, SIAM Journal on Control and Optimization, Vol. 41(2002), 740–797.

[17] S. Ulbrich. *Adjoint-based derivative computations for the optimal control of discontinuous solutions of hyperbolic conservation laws*, Systems & Control Letters, Vol. 48(2003), 309–324.

Received for publication September 2006.

E-mail address: goettlich@mathematik.uni-kl.de
E-mail address: herty@rhrk.uni-kl.de
E-mail address: kirchner@mathematik.uni-kl.de
E-mail address: klar@itwm.fhg.de