Optimal Control of Nonlinear Hyperbolic Conservation Laws with Switching

Sebastian Pfaff, Stefan Ulbrich and Günter Leugering

Abstract. We consider optimal control problems governed by nonlinear hyperbolic conservation laws at junctions and analyze in particular the Fréchet-differentiability of the reduced objective functional. This is done by showing that the control-to-state mapping of the considered problems satisfies a generalized notion of differentiability. We consider both, the case where the controls are the initial and the boundary data as well as the case where the system is controlled by the switching times of the node condition. We present differentiability results for the considered problems in a quite general setting including an adjoint-based gradient representation of the reduced objective function.

Keywords. optimal control; scalar conservation law; network.

1. Introduction

This paper serves as a final report of the project *Optimal Control of Switched Networks for Nonlinear Hyperbolic Conservation Laws*. In this work we consider optimal control problems for entropy solutions of hyperbolic conservation laws involving objective functionals of the form

$$J(y(u)) := \int_a^b \psi(y(\bar{t}, x; u), y_d(x))\,\mathrm{d}x, \tag{1.1}$$

where $\psi \in C^{1,1}_{\mathrm{loc}}(\mathbb{R}^2)$ and $y_d \in BV([a,b])$ is a desired state. The state y is the entropy solution of either an initial-boundary value problem for a scalar conservation law

$$y_t + f(y)_x = g(\cdot, y, u_1)$$

or of a traffic light problem, as we will call it throughout this paper, where two conservation laws are coupled through a node at which the switching times between red and green phases is controlled. Our motivation is to develop a variational calculus for initial, boundary and switching time control that has the potential to be extended to the optimal control of networks with switching control at nodes.

It is well-known that weak solutions to Cauchy problems of nonlinear hyperbolic conservation laws are in general not unique and that one has to consider entropy solutions, that can be obtained as the vanishing viscosity limit of a parabolic regularization [4, 26]. Even for smooth initial and boundary data entropy solutions can develop discontinuities (shocks) after finite time [7]. This leads to fundamental difficulties for the sensitivity analysis and optimal control theory, since the shock locations depend on the control. Hence, the control-to-state mapping $u \mapsto y(\bar{t}, \cdot; u)$ is at best differentiable with respect to the weak topology of measures and senistivities are necessarily measures with singular part along the shock curves. For networks, where the solutions on two or more intervals are connected by a (possibly controlled) node, the situation gets even more involved.

Motivated by its practical relevance, despite these difficulties the analysis and numerical solution of optimal control problems for hyperbolic conservation laws has become an active research field in recent years.

The existence of optimal controls for the Cauchy problem and the initial-boundary value problem was discussed for example in [1, 2, 35, 36].

The issue of non-differentiability of the solution operator was treated by different authors by introducing generalized notions of differentiability, e.g. [5, 8, 9, 11, 15, 36, 37]. The present work is based on the notion of *shift-differentiability*, that was introduced in [36], where it was also shown to hold for the Cauchy problem. Here the theory of generalized characteristics by Dafermos [17] is a crucial instrument. This approach also includes an adjoint calculus for the reduced objective function, see also [19, 20, 38].

Networks for hyperbolic conservation laws have been considered in various contexts in recent years. Several node conditions have been discussed, most of them are tailored to specific applications, such as traffic modeling [25, 13, 10, 22], gas pipelines [3, 16, 23] or supply chains [21]. The conditions are mostly formulated for Riemann problems and then generalized by wave front tracking. Besides the question of well-posedness also aspects from the optimal control viewpoint have been considered. But these approaches often either consider the linear case or assume the existence of a strong solution. Conservation laws with modal switching have been discussed for the first time in [24], where switching is considered in the fluxes, the boundary condition and the coupling condition at the nodes of a network.

This paper is organized as follows. In section 2 we introduce the two considered problems, the initial-boundary value problem and the traffic light problem. In section 3 we collect results on the well-posedness for these problems and structural properties of the corresponding solutions. The main results will be presented in section 4, where we show the generalized differentiability of the solution operator and the resulting Fréchet-differentiabilty of the reduced objective function.

2. The models

In this paper we focus on two types of problems for a scalar conservation law. The first one is the initial-boundary value problem (IBVP), the second is the traffic light problem (TLP). We will also consider the pure initial value problem (IVP), since on one hand the IVP is helpful to understand the more involved IBVP and on the other hand the traffic light problem is a combination of both. While the IBVP is an important step towards node conditions on networks, the traffic light problem can be seen as a relevant node condition with switching in a traffic network.

2.1. Initial-Boundary Value Problem

The first model problem under consideration is an initial-boundary value problem (IBVP) on an interval $\Omega = (a, b)$, where we explicitly allow for a, b to be $\pm\infty$, respectively. The IBVP is then given by

$$\begin{aligned} y_t + f(y)_x &= g(\cdot, y, u_1), && \text{on } \Omega_T, && (2.1a)\\ y(0,\cdot) &= u_0, && \text{on } \Omega, && (2.1b)\\ y(\cdot, a+) &= u_{B,a}, && \text{in the sense of (2.4a)} \quad (\text{if } a > -\infty), && (2.1c)\\ y(\cdot, b-) &= u_{B,b}, && \text{in the sense of (2.4b)} \quad (\text{if } b < \infty), && (2.1d) \end{aligned}$$

where $\Omega_T := [0, T] \times \Omega$. In order to show existence of a unique solution, following [26, 4], the conservation law (2.1a) has to be understood in sense of an entropic solution, which can be characterized by requiring that for every (Kružkov-) entropy $\eta_c(\lambda) := |\lambda - c|$, $c \in \mathbb{R}$, and associated entropy flux $q_c(\lambda) := \operatorname{sgn}(\lambda - c)(f(\lambda) - f(c))$ the following entropy inequality holds in the sense of distributions

$$(\eta_c(y))_t + (q_c(y))_x \leq \eta_c'(y) g(\cdot, y, u_1) \qquad \text{in } \mathcal{D}'(\Omega_T). \tag{2.2}$$

The initial data in (2.1b) have to be understood in the weak L^1_{loc}-sense, which means that for every $R > 0$

$$\operatorname*{esslim}_{t\to 0+} \|y(t,\cdot) - u_0\|_{1,\Omega\cap(-R,R)} = 0 \tag{2.3}$$

is fulfilled. The Dirichlet-like boundary conditions (2.1c), (2.1d) must not be understood literally. Rather, the solution y of (2.1a)-(2.1d) has to be interpreted as the limit of its parabolic regularization, that is (2.1a)-(2.1d) with the term εy_{xx} added on the right hand side of (2.1a). The boundary condition of the limit solution can then be characterized, as shown in [4], by

$$\min_{k\in I(y(\cdot,a+),u_{B,a})} \operatorname{sgn}(u_{B,a} - y(\cdot, a+))(f(y(\cdot, a+)) - f(k)) = 0, \quad \text{a.e. on } [0, T], \tag{2.4a}$$

$$\min_{k\in I(y(\cdot,b-),u_{B,b})} \operatorname{sgn}(y(\cdot, b-) - u_{B,b})(f(y(\cdot, b-)) - f(k)) = 0, \quad \text{a.e. on } [0, T], \tag{2.4b}$$

with $I(\alpha, \beta) := [\min(\alpha, \beta), \max(\alpha, \beta)]$, see also [18, 28, 30, 31]. In the literature the above formulation of the boundary condition from [4] is sometimes called the *BLN-condition*.

2.2. Traffic Lights

The second topic of this work is a problem that is motivated by a traffic flow problem. This special type of problem is also of great interest because it is a simple example for a network of conservation laws with modular switchings in the node condition. Before we formulate the mathematical problem, we give a short overview on traffic flow modeling by hyperbolic conservation laws.

2.2.1. Macroscopic model for traffic flow. In the mid 1950s, Lighthall and Whitham [29] and Richards [34] proposed a continuum model for heavy traffic. The traffic is described by means of a traffic density ρ and the conservation of cars is ensured by

$$\rho_t + f(\rho)_x = 0, \qquad f(\rho) := \rho v(\rho),$$

where the velocity v of the traffic depends only on the density. This model is widely used and is known as the *LWR*-model. For a detailed overview on traffic flow modeling by partial differential equations we refer to [25]. Usually one assumes that f is a concave function, but since most theoretical results on conservation laws work with convex fluxes, we will make a change of signs and work with a convex flux function. In the following the state y can be interpreted as the negative traffic density $-\rho$. We further assume that the road reaches its maximum density when $y = -1$ and is empty for $y = 0$. The flux $f(y)$ is equal to 0 for these two values and strictly convex in between. In particular, f is negative on $(-1, 0)$.

2.2.2. A traffic light on an open road. We consider a long unidirectional road $I = \mathbb{R}$ that has to be closed for some reason (e.g. because of pedestrian or railway crossings) for some time periods at a specific point $x = 0$, most likely by a traffic light. So the considered time interval $[0, T]$ is split into two different types of phases, namely green $[\sigma_{\mathrm{g}}^{i-1}, \sigma_{\mathrm{r}}^{i})$, $i = 1, \ldots, n_\sigma + 1$ and red phases $[\sigma_{\mathrm{r}}^{i}, \sigma_{\mathrm{g}}^{i})$, $i = 1, \ldots, n_\sigma$ where the incoming traffic at $x = 0$ is or is not allowed to cross respectively. A similar problem was already briefly introduced in [29].

For the sequel we assume $\sigma = (\sigma_{\mathrm{g}}^0, \sigma_{\mathrm{r}}^1, \sigma_{\mathrm{g}}^1, \ldots, \sigma_{\mathrm{g}}^{n_\sigma}, \sigma_{\mathrm{r}}^{n_\sigma+1}) \in \Sigma$, where

$$\Sigma := \left\{ \nu \in \mathbb{R}^{2(n_\sigma+1)} \;:\; 0 = \nu_1 < \nu_2 < \cdots < \nu_{2n_\sigma+1} < \nu_{2n_\sigma+2} = T \right\}. \tag{2.5}$$

for the sake of simplicity. The presented analysis can also be carried over to the case where the first and/or the final phase is a red phase.

During the i-th green phase a solution y of such a traffic light problem (TLP) is determined by solving a Cauchy problem on $\Omega_{\mathrm{g},i} := [\sigma_{\mathrm{g}}^{i-1}, \sigma_{\mathrm{r}}^{i}] \times \mathbb{R}$ with initial data

$$u_0 = y(\sigma_{\mathrm{g}}^{i-1}-, \cdot), \quad i = 2, \ldots, n_\sigma + 1.$$

Here, $y(\sigma_{\mathrm{g}}^{i-1}-, \cdot)$ is the final state of the previous red phase.

For the i-th red phase the solution y consists of two parts, namely y_1 and y_2, its restriction to the incoming and outgoing part $I_1 := (-\infty, 0)$

and $I_2 := (0,\infty)$ of the road. The restriction y_1 is the solution of an initial-boundary value problem on $\Omega^1_{\mathrm{r},i} := [\sigma^i_{\mathrm{r}}, \sigma^i_{\mathrm{g}}] \times I_1$ with initial value $y(\sigma^i_{\mathrm{r}}-, \cdot)$ and boundary data $u_{B,0} \equiv -1$. Similarly, y_2 solves an IBVP on $\Omega^2_{\mathrm{r},i} := [\sigma^i_{\mathrm{r}}, \sigma^i_{\mathrm{g}}] \times I_2$ with $u_{B,0} \equiv 0$. For the first green phase, i.e. the first IVP, the initial data are given by some function u_I. The traffic light problem can then formulated in the following way:

$$
\begin{aligned}
y_t + f(y)_x &= g(\cdot, y, u_1), && \text{on } \Omega_{\mathrm{g},i+1}, && i = 0, \dots, n_\sigma, && j = 1,2, && \text{(2.6a)}\\
y_t + f(y)_x &= g(\cdot, y, u_1), && \text{on } \Omega^j_{\mathrm{r},i}, && i = 1, \dots, n_\sigma, && j = 1,2, && \text{(2.6b)}\\
y(0, \cdot) &= u_I, && \text{on } I, && && && \text{(2.6c)}\\
y(\sigma^i_g, \cdot)\big|_{I_j} &= y_j(\sigma^i_g-, \cdot), && \text{on } I_j, && i = 1, \dots, n_\sigma, && j = 1,2, && \text{(2.6d)}\\
y_j(\sigma^i_r, \cdot) &= y(\sigma^i_r-, \cdot)\big|_{I_j}, && \text{on } I_j, && i = 1, \dots, n_\sigma, && j = 1,2, && \text{(2.6e)}\\
y_1(\cdot, 0-) &= -1, && \text{on } [\sigma^i_r, \sigma^i_g], && i = 1, \dots, n_\sigma, && && \text{(2.6f)}\\
y_2(\cdot, 0+) &= 0, && \text{on } [\sigma^i_r, \sigma^i_g], && i = 1, \dots, n_\sigma. && && \text{(2.6g)}
\end{aligned}
$$

The conservation laws (2.6a), (2.6b) model the conservation of cars. The source term g can be seen as additional traffic that enter or leave the road from minor roads or parking lots, that are not modeled in detail. The boundary conditions that model the red lights (red light conditions) (2.6f), (2.6g) guarantee, that during these periods no cars enter or leave the two roads over the artificial boundary, since, as stated in 2.2.1, the flux $f(y)$ is equal to zero for $y \in \{-1, 0\}$. Moreover, even if formally one has to interpret these boundary conditions in the BLN-sense, we will see that under mild assumptions they may be considered literally. We will discuss these conditions more detailed in section 3.2. The continuity conditions between the phases (2.6d), (2.6e) describe the transition from one phase into another.

3. Properties of entropy solutions

In this section we collect important properties of the solutions to (2.1) and (2.6).

3.1. General and structural properties of solutions to IBVPs

First we consider the initial value problem (2.1a)-(2.1b) for $\Omega = \mathbb{R}$. We make the following assumptions:

(A1) The flux function satisfies $f \in C^2(\mathbb{R})$ and there exists $m_{f''} > 0$ such that $f'' \geq m_{f''}$. The source term satisfies $g \in L^\infty\left(\Omega_T; C^{0,1}_{\mathrm{loc}}(\mathbb{R} \times \mathbb{R}^m)\right) \cap L^\infty\left(0, T; C^1_{\mathrm{loc}}(\mathbb{R} \times \mathbb{R} \times \mathbb{R}^m)\right)$ and for all $M_u > 0$ there exist constants $C_1, C_2 > 0$ such that for all $(t, x, y, u_1) \in \Omega_T \times \mathbb{R} \times [-M_u, M_u]^m$ it holds that

$$g(t, x, y, u_1)\mathrm{sgn}(y) \leq C_1 + C_2|y|.$$

(A2) The set of admissible controls U_{ad} is bounded in $U_\infty := L^\infty(\mathbb{R}) \times L^\infty(\Omega_T)^m$ by some constant M_u and closed in $U_1 := L^1_{\mathrm{loc}}(\mathbb{R}) \times L^1_{\mathrm{loc}}(\Omega_T)^m$.

We recall Proposition 1 from [38], that covers some of the most important properties of the solution to the IVP.

Proposition 3.1 (Existence and Uniqueness for Cauchy problems). *Let (A1) and (A2) hold. Then for every* $u = (u_0, u_1) \in U_\infty$ *there exists a unique entropy solution* $y = y(u) \in L^\infty(\Omega_T)$ *of* (2.1a)-(2.1b) *on* $\Omega = \mathbb{R}$. *After a possible modification on a set of measure zero it even holds* $y \in C([0,T]; L^1_{\mathrm{loc}}(\mathbb{R}))$. *There are constants* $M_y, L_y > 0$ *such that for every* $u, \hat{u} \in U_{\mathrm{ad}}$ *and all* $t \in [0,T]$ *the following estimates hold:*

$$\begin{aligned} &\|y(t,\cdot;u)\|_\infty \leq M_y, \\ &\|y(t,\cdot;u) - y(t,\cdot;\hat{u})\|_{1,[a,b]} \leq L_y(\|u_0 - \hat{u}_0\|_{1,I_t} + \|u_1 - \hat{u}_1\|_{1,[0,t]\times I_t}), \end{aligned}$$

where $a < b$ *and* $I_t := [a - tM_{f'}, b + tM_{f'}]$, $M_{f'} := \max_{|y|\leq M_y} |f'(y)|$.

Set $\hat{U}_{\mathrm{ad}} := \{u \in U_{\mathrm{ad}} \; : \; \|u_1\|_{L^\infty(0,T;C^1(\Omega_T)^m)} \leq M_u\}$. *Then there is a constant* $M > 0$ *such that for all* $u \in \hat{U}_{\mathrm{ad}}$ *and all* $t \in (0,T]$ *Oleinik's entropy condition*

$$y_x(t,\cdot;u) \leq \left((1 - e^{-m_{f''}Mt})M^{-1} + e^{-m_{f''}Mt}(C_{u,M})^{-1})\right)^{-1}$$

holds with $C_{u,M} := \max\left\{M, \operatorname{esssup}_{x\neq z} \frac{u_0(x)-u_0(z)}{x-z}\right\}$. *In particular* $y(t,\cdot) \in BV_{\mathrm{loc}}(\mathbb{R})$ *for all* $t \in (0,T]$ *and* $y \in BV([s,T] \times [-R,R])$ *for all* $s, R > 0$.

For the case of an initial-boundary value problem we have a similar result. We restrict ourselves to the case of $\Omega = (0,\infty)$. The first thing to mention here is the fact that the BLN-condition (2.4a) involves the boundary trace $y(\cdot, 0+)$. When Bardos, le Roux and Nédélec stated this formulation they only considered the case where the solution has bounded total variation, see Remark 3.3. In order to also allow for L^∞-data in [31, 30] Otto proposed another characterization of the boundary condition that is equivalent to the one in [4] if the boundary trace exists. But Vasseur showed [39] that under mild assumptions even for L^∞-entropy solutions there always exist boundary traces. Therefore the formulation in (2.4a) (and (2.4b)) is valid even in the L^∞-setting, see also [14].

We make the following assumptions:

(A1') The flux function satisfies $f \in C^2(\mathbb{R})$ and there exists $m_{f''} > 0$ such that $f'' \geq m_{f''}$. The source term is non-negative and satisfies $g \in C\left(\Omega_T; C^{0,1}_{\mathrm{loc}}(\mathbb{R} \times \mathbb{R}^m)\right) \cap C^1\left([0,T]; C^1_{\mathrm{loc}}(\Omega \times \mathbb{R} \times \mathbb{R}^m)\right)$ and for all $M_u > 0$ there exist constants $C_1, C_2 > 0$ such that for all $(t,x,y,u_1) \in \Omega_T \times \mathbb{R} \times [-M_u, M_u]^m$ holds:

$$g(t,x,y,u_1)\mathrm{sgn}(y) \leq C_1 + C_2|y|.$$

(A2') The set of admissible controls U_{ad} is bounded in $U_\infty := L^\infty(\mathbb{R}) \times L^\infty(0,T) \times L^\infty([0,T] \times \mathbb{R})^m$ by some constant M_u and closed in $U_1 := L^1_{\mathrm{loc}}(\mathbb{R}) \times L^1(0,T) \times L^1_{\mathrm{loc}}([0,T] \times \mathbb{R})^m$.

For technical reasons we consider the source term and the corresponding control u_1 not only for the considered spatial domain. Of course the solution depends only on its restriction to Ω_T.

Under the above assumptions we get the following properties of a solution to (2.1), c.f. [4, 14, 31].

Proposition 3.2 (Existence and Uniqueness for IBVPs). *Let (A1') and (A2') hold. Then for every $u = (u_0, u_B, u_1) \in U_\infty$ there exists a unique entropy solution $y = y(u) \in L^\infty(\Omega_T)$ of (2.1) on $\Omega = (0, \infty)$. After a possible modification on a set of measure zero it even holds that $y \in C([0,T]; L^1_{\text{loc}}(\Omega))$. Moreover, there are constants $M_y, L_y > 0$ such that for every $u, \hat{u} \in U_{\text{ad}}$ and all $t \in [0,T]$ the following estimates hold:*

$$\begin{aligned} \|y(t,\cdot;u)\|_\infty &\leq M_y, \\ \|y(t,\cdot;u) - y(t,\cdot;\hat{u})\|_{1,[a,b]} &\leq L_y(\|u_0 - \hat{u}_0\|_{1,I_t} \\ &\qquad + \|u_B - \hat{u}_B\|_{1,[0,t]} + \|u_1 - \hat{u}_1\|_{1,[0,t]\times I_t}), \end{aligned}$$

where $a < b$ and $I_t := [a - tM_{f'}, b + tM_{f'}] \cap \Omega$, $M_{f'} := \max_{|y|\leq M_y} |f'(y)|$.

Remark 3.3. Under the stronger assumptions $u_0 \in BV_{\text{loc}}(\Omega)$ and $u_B \in BV([0,T])$, (2.1) admits a solution satisfying $y \in BV([0,T] \times [0,R])$ for all $R > 0$ (c.f. [27, 4]).

The basic idea behind the proof of the main result of this work is the theory of generalized characteristics from [17], which will be considered in the remaining part of this section. We will assume that in addition to (A1)-(A2), (A1')-(A2') respectively, the following assumption holds.

(A3) g is globally Lipschitz w.r.t. x and y.

Furthermore we will only consider $(u_0, u_1) \in \hat{U}_{\text{ad}}$ (see Proposition 3.1), $u_0 \in BV_{\text{loc}}(\Omega)$ and boundary data $u_B \in PC^1([0,T]; t_1, \dots, t_{n_t})$, that is a piecewise continuously differentiable function with possible kinks or discontinuities at $0 < t_1 < \cdots < t_{n_t}$ for some $n_t \in \mathbb{N}$.

Using the properties collected in Propositions 3.1 and 3.2, we conclude that $y \in L^\infty(\Omega_T) \cap C([0,T]; L^1_{\text{loc}}(\Omega))$ has the following properties: For all $(t,x) \in (0,T] \times \Omega$ the one-sided limits $y(t,x-)$ and $y(t,x+)$ exist and satisfy $y(t,x-) \geq y(t,x+)$. It will be convenient to work with a pointwise defined representative of $y \in C([0,T]; L^1_{\text{loc}}(\Omega))$ where $y(t,x)$ is identified with one of the limits $y(t,x-)$ or $y(t,x+)$.

We now recall the definition of a generalized characteristic in the sense of Dafermos from [17].

Definition 3.4 (Generalized characteristics). A Lipschitz curve

$$[\alpha, \beta] \subset [0,T] \to \Omega_T, \ t \mapsto (t, \xi(t))$$

is called a *generalized characteristic* on $[a,b]$ if

$$\dot{\xi}(t) \in [f'(y(t,\xi(t)+)), f'(y(t,\xi(t)-))], \quad \text{a.e. on } [\alpha,\beta]. \tag{3.1}$$

The generalized characteristic is called *genuine* if the lower and upper bound in (3.1) coincide for almost all $t \in [\alpha, \beta]$.

In the following we will also call ξ a (generalized) characteristic instead of $t \mapsto (t, \xi(t))$. It will also be useful to introduce notions of *extreme* or *maximal/minimal characteristics* $\xi_\pm$, that satisfy

$$\dot{\xi}_\pm(t) = f'(y(t, \xi(t)\pm)).$$

Since by Proposition 3.1 or 3.2 y is (essentially) bounded on Ω_T, an a priori bound on the speed of all generalized characteristic is known. Therefore, characteristics do not escape and they either exist for the whole time period $[0, T]$ or (in the bounded case) leave the spatial domain at some point $(\theta, \xi(\theta)) \in [0, T] \times \partial\Omega$. Moreover it can be shown [17] that (3.1) can be restricted to

$$\dot{\xi}(t) = \begin{cases} f'(y(t, \xi(t))) & \text{if } f'(y(t, \xi(t)+)) = f'(y(t, \xi(t)-)) \\ \frac{[f(y(t,\xi(t)))]}{[y(t,\xi(t))]} & \text{if } f'(y(t, \xi(t)+)) \neq f'(y(t, \xi(t)-)) \end{cases}, \quad \text{a.e. on } [\alpha, \beta],$$

where for $\varphi \in BV(\mathbb{R})$ the expression

$$[\varphi(x)] := \varphi(x-) - \varphi(x+)$$

denotes the height of the jump of φ across x.

Based on the notion of generalized characteristics in [17] Dafermos exploits structural properties of BV-solutions that are essential for the analysis in the present paper.

Proposition 3.5 (Structure of BV-Solutions). *Let (A1)-(A3) hold. Consider an entropy solution $y = y(u)$ of the Cauchy problem* (2.1a)-(2.1b) *on $\Omega = \mathbb{R}$ for controls $u = (u_0, u_1) \in \hat{U}_{\text{ad}}$, $u_0 \in BV_{\text{loc}}(\mathbb{R})$.*

For $(\bar{t}, \bar{x}) \in \Omega_T$ fixed denote by ξ a backward characteristic on $[0, \bar{t}]$ through $(\bar{t}, \bar{x})$. Then ξ has the following properties:

1. *if ξ is an extreme backward characteristic, i.e. $\xi = \xi_\pm$, then ξ is genuine, i.e. $y(t, \xi_\pm(t)-) = y(t, \xi_\pm(t)+)$ for all $t \in (0, \bar{t})$.*
2. *if ξ is genuine, i.e. $y(t, \xi(t)-) = y(t, \xi(t)+)$, $t \in (0, \bar{t})$, then it satisfies*

$$\xi(t) = \zeta(t), \quad t \in [0, \bar{t}], \qquad y(t, \xi(t)) = v(t), \quad t \in (0, \bar{t}), \tag{3.2a}$$

$$u_0(\xi(0)-) \leq v(0) \leq u_0(\xi(0)+), \quad y(\bar{t}, \xi(\bar{t})-) \geq v(\bar{t}) \geq y(\bar{t}, \xi(\bar{t})+), \tag{3.2b}$$

where (ζ, v) is a solution of the characteristic equation

$$\dot{\zeta}(t) = f'(v(t)), \tag{3.3a}$$

$$\dot{v}(t) = g(t, \zeta(t), v(t), u_1(t, \zeta(t))). \tag{3.3b}$$

For extreme characteristics $\xi_\pm$ the initial values are given by

$$(\zeta, v)(\bar{t}) = (\bar{x}, y(\bar{t}, \bar{x}\pm)). \tag{3.3c}$$

Although this classical result by Dafermos is widely-known and gives important information about the inner structure of entropy solutions, the earliest extension to be found in the literature to a bounded spatial domain is in a work of Perrollaz in [32] published in 2013. Here the situation for characteristics ξ that stay inside the spatial domain for the whole considered time interval is exactly the same as in Proposition 3.5. But there are two cases that require special consideration. On the one hand there are backward characteristics through some $(\bar{t}, \bar{x}) \in \Omega_T$ that leave the spatial domain at some time, say $\theta \in (0, \bar{t})$, on the other hand one has to consider characteristics that enter the spatial domain at some time θ.

It turns out that for $\Omega = (0, \infty)$ the non-negativity condition on g is crucial for the second and third part of Proposition 3.6, since it avoids some degeneracy of the characteristics near the boundary. For a spatial domain $(-\infty, 0)$ the condition on the source term becomes a non-positivity condition and consequently this leads to the requirement that g has to vanish if one considers general intervals (a, b) of finite length. But as mentioned, this property is only important near the boundary and can therefore be weakened to a local condition.

The following proposition collects the results of section 3 in [32].

Proposition 3.6. *Let (A1'), (A2') and (A3) hold. Consider an entropy solution $y = y(u)$ of the mixed initial-boundary value problem* (2.1) *on $\Omega = (0, \infty)$ for controls $u = (u_0, u_B, u_1) \in U_{\text{ad}}$ with $(u_0, u_1) \in \hat{U}_{\text{ad}}$, $u_0 \in BV_{\text{loc}}(\mathbb{R})$ and $u_B \in PC^1([0, T]; t_1, \ldots, t_{n_t})$. Then the following holds:*

1. *Consider $\theta \in (0, T)$ with $f'(y(\theta, 0+)) < 0$, then there exists a genuine backward characteristic ξ through $(\theta, 0)$ with $\dot{\xi}(\theta) = f'(y(\theta, 0+))$.*
2. *Let ξ be a genuine characteristic through $(\bar{t}, \bar{x}) \in \Omega_T$ satisfying $\xi(t) \in \Omega$ for $t \in (\theta, \bar{t}] \subset [0, T]$ and $\lim_{t \searrow \theta} \xi(t) = 0$. Denote by (ζ, v) the solution of the characteristic equation* (3.3a)-(3.3b) *associated to ξ by Proposition* 3.5 *on every $[\tilde{t}, \bar{t}] \subset (\theta, \bar{t}]$. Then with $v(\theta) := \lim_{t \searrow \theta} v(t)$ it holds*
$$u_B(\theta+) \leq v(\theta) \leq u_B(\theta-). \tag{3.4}$$
3. *Let ξ be a forward characteristic in $[0, \tilde{t}] \times \Omega$ for every $\tilde{t} \in (0, \theta)$ and (ζ, v) be the associated solution of the characteristic equation. If now $\lim_{t \nearrow \theta} \xi(t) = 0$ then*
$$f'(\bar{v}) \leq 0 \quad \text{and} \quad f(\bar{v}) \geq f(u_B(\theta-)), \tag{3.5}$$
where $\bar{v} := \lim_{t \nearrow \theta} v(t)$.

This connection between the genuine characteristics and the characteristic equation is very useful, since by the following lemma, that is a consequence of a result on ordinary differential equations (c.f. Proposition 3.4.5 and Lemma 3.4.6 in [36] or chapter 5.6 in [33]) this yields some important information on the local differentiability properties of a solution y of the I(B)VP.

Lemma 3.7. *Let (A1') and (A3) hold and denote for $(\theta, z, w, u_1) \in [0, T] \times \mathbb{R}^2 \times C^1([0, T] \times \mathbb{R}^m)$ by $(\zeta, v)(\cdot, \theta, z, w, u_1)$ the solution of* (3.3a)-(3.3b) *for inital data*
$$(\zeta, v)(\theta) = (z, w).$$

Let $M_w, M_u > 0$ be given and set

$$\mathcal{B}_i := [0,T] \times \mathbb{R}^2 \times L^2(0,T; C^i(\mathbb{R})^m), \quad i = 0, 1,$$
$$\bar{\mathcal{B}} := \left\{ (\theta, z, w, u_1) \in \mathcal{B}_1 \ : \ |w| < M_w, \quad \|u_1\|_{C^1([0,T]\times\mathbb{R}^m)} < M_u \right\}.$$

Then the mapping

$$(\theta, z, w, u_1) \in (\bar{\mathcal{B}}, \|\cdot\|_{\mathcal{B}_i}) \longmapsto (\zeta, v)(\cdot, \theta, z, w, u_1) \in C([0,T])^2$$

is Lipschitz continuous for $i = 0$ and continuously Fréchet-differentiable for $i = 1$ and on $\bar{\mathcal{B}}$ the right hand side is uniformly Lipschitz w.r.t. t.

Lemma 3.7 is a direct generalization of the first assertion of Lemma 3.4.6 in [36] to the case where the dependence on the time θ where the initial datum is specified, is considered, too. The remaining statements of Lemma 3.4.6 can also be carried over to this generalized case.

3.2. General and structural properties of solutions to traffic light problems

In this section we analyze the structure of solutions to traffic light problems. We consider $u_I \in BV_{\mathrm{loc}}(\mathbb{R})$ and u_1 bounded in $C^1([0,T] \times \mathbb{R})^m$. Since a solution of a TLP is a concatenation of solutions to IVPs and IBVPs on a finite number of time slabs, the existence, uniqueness and stability properties can easily be transferred to such solutions.

We add the following requirements to our setting.

(A4) g is non-positive on $(-\infty, 0)$, non-negative on $(0, \infty)$ and vanishes on $(-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$. In addition g is chosen such that $-1 \le y \le 0$ is guaranteed. Furthermore, let $U_{\mathrm{ad}} \subset \{(u_0, u_1) \in U_\infty \ : \ -1 \le u_0 \le 0\}$ and let $\Sigma_{\mathrm{ad}} \subset \Sigma$ be a closed set in $[0,T]$, with Σ defined in (2.5).

Remark 3.8. The condition on g in (A4) holds clearly for the choice $g \equiv 0$.

Corollary 3.9 (Existence and Uniqueness for traffic light problems). *Let (A1'), (A2') and (A4) hold. Then for every $u = (u_0, u_1) \in U_\infty$ and $\sigma \in \Sigma_{\mathrm{ad}}$ there exists a unique entropy solution $y = y(u, \sigma) \in L^\infty(\Omega_T)$ of (2.1) on $\Omega = (0, \infty)$. After a possible modification on a set of measure zero it even holds $y \in C([0,T]; L^1_{\mathrm{loc}}(\Omega))$.*

Moreover for every $t \in [0,T]$ and $a < b$ we have the following stability estimates:

1. *For fixed $u \in U_{\mathrm{ad}}$ there is $L_\Sigma > 0$ such that for all $\tilde{\sigma}, \hat{\sigma} \in \Sigma_{\mathrm{ad}}$ holds*
$$\|y(t, \cdot; u, \tilde{\sigma}) - y(t, \cdot; u, \hat{\sigma})\|_{1,(a,b)} \le L_\Sigma \|\tilde{\sigma} - \hat{\sigma}\|.$$
2. *For fixed $\sigma \in \Sigma_{\mathrm{ad}}$ there is $L_U > 0$ such that for all $\tilde{u}, \hat{u} \in U_{\mathrm{ad}}$ holds*
$$\|y(t, \cdot; \tilde{u}, \sigma) - y(t, \cdot; \hat{u}, \sigma)\|_{1,(a,b)} \le L_U \left(\|\tilde{u}_0 - \hat{u}_0\|_{1,I_t} + \|\tilde{u}_1 - \hat{u}_1\|_{1,[0,t]\times I_t} \right),$$
where $I_t := [a - tM_{f'}, b + tM_{f'}] \cap \Omega$, $M_{f'} := \max_{|y| \le M_y} |f'(y)|$.

One can easily verify that the structural features for solutions to IBVPs provided by Propositions 3.5 and 3.6 also hold for genuine backward characteristics ξ that correspond to the solution $y = (y_1, y_2)$ of a TLP as long as they do not touch the switching points $(\sigma, 0)$.

We now discuss what happens to the solution $y = (y_1, y_2)$ during a red phase and at the beginning of the green phase. The following considerations are illustrated in Figure 1.

First we consider the initial-boundary value problem for y_1 during a red phase $[\sigma_{\mathrm{r}}^i, \sigma_{\mathrm{g}}^i]$. Here especially the situation on the boundary is of interest. We recall, that the boundary data $u_{B,0}$ are chosen to be equal to -1 and that by assumption (A4) $-1 \leq y = (y_1, y_2) \leq 0$ holds. Therefore the BLN-boundary condition (2.4b) becomes

$$\min_{k \in [-1, y(\cdot,0-)]} \operatorname{sgn}(y(\cdot,0-)+1)(f(y(\cdot,0-)) - f(k)) = 0.$$

Since k may be chosen equal to $y(\cdot,0-)$, the condition is equivalent to

$$\operatorname{sgn}(y(t,0-)+1)(f(y(t,0-)) - f(k)) \geq 0, \quad \forall k \in [-1, y(t,0-)].$$

Here the first factor is strictly positive whenever $y(\cdot,0-) \neq -1$ and for $k = -1$, the second factor is negative if $f(y(\cdot,0-)) \neq 0$. Hence we can deduce that there are only two possibilities for the boundary trace, namely $y(t,0-) \in \{0,-1\}$ for almost all $t \in [\sigma_{\mathrm{r}}^i, \sigma_{\mathrm{g}}^i]$. If $y(\theta,0-) = 0$ for some $\theta \in (\sigma_{\mathrm{r}}^i, \sigma_{\mathrm{g}}^i)$, then the existence of a backward characteristic ξ satisfying $\dot{\xi}(\theta) = f'(0) < 0$ can be deduced from Proposition 3.6. Since by the sign condition on the source term, all genuine characteristics on $\Omega_{\mathrm{r},i}^1$ are concave and since two genuine characteristics may not intersect each other, this implies that $y(t,0-) = 0$ must hold for all $t \in (\sigma_{\mathrm{r}}^i, \theta]$. Conversely speaking, this means, that if $y(\tilde{\theta},0-) = -1$ for some $\tilde{\theta} \in (\sigma_{\mathrm{r}}^i, \sigma_{\mathrm{g}}^i)$, then $y(t,0-) = -1$ holds for all $t \in [\tilde{\theta}, \sigma_{\mathrm{g}}^i]$. Consequently, if the initial data of the IBVP on $\Omega_{\sigma_{\mathrm{g}}^i}^1$ are bounded away from 0 in a small neighborhood of the right boundary at $x = 0$, $y(\cdot,0-) = -1$ holds during the whole time slab. We will assume this property for the sequel. In this case a generalized characteristic η emanates from $(\sigma_{\mathrm{r}}^i, 0)$ having strictly negative speed at least for a small time period $(\sigma_{\mathrm{r}}^i, \tilde{\tau})$, see Figure 1. (More precisely, η is either a shock or a characteristic traveling with speed $f'(-1)$.) After that period, it keeps traveling with non-positive speed at least up to $t = \sigma_{\mathrm{g}}^i$. The solution y_1 is constantly equal to -1 on the nonempty set $\{(t,x) \in [\sigma_{\mathrm{r}}^i, \sigma_{\mathrm{g}}^i] \times [-\varepsilon, 0) \ : \ \eta(t) < x\}$ with ε from assumption (A4). The situation for y_2 is completely analogous. If the initial data of the IBVP on $\Omega_{\mathrm{r},i}^2$ are bounded away from -1 in a small neighborhood of the boundary at $x = 0$, $y(\cdot,0+) = 0$ holds during the whole time slab and we conclude that $y_2 = 0$ on a set $\{(t,x) \in [\sigma_{\mathrm{r}}^i, \sigma_{\mathrm{g}}^i] \times (0, \varepsilon] \ : \ \bar{\eta}(t) > x\}$. Therefore, for every $t \in (\sigma_{\mathrm{r}}^i, \sigma_{\mathrm{r}}^i)$ the solution $y(t,\cdot)$ is known at least in a small neighborhood of $x = 0$, compare the area filled with characteristics in Figure 1. We now examine the solution y on the subsequent green phase $[\sigma_{\mathrm{g}}^i, \sigma_{\mathrm{r}}^{i+1}]$. Here the situation at $x = 0$ is again of special interest. By the previous considerations we know that there is a $\delta > 0$ such that with u_0 being the initial

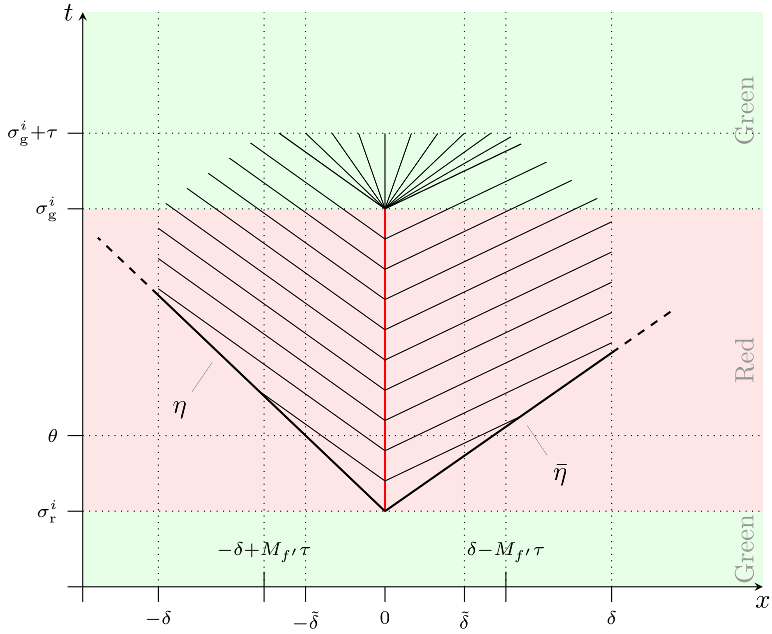

FIGURE 1. Characteristics in a neighborhood of a red phase.

data of the considered Cauchy problem on $\Omega_{{\rm g},i+1}$, $u_0(x) = \frac{1}{2}(\operatorname{sgn}(x) - 1)$ for all $x \in (-\delta, \delta)$. Together with the finite propagation speed this implies that locally y is the solution of a Riemann problem producing a rarefaction wave.

We subsume the previous considerations in the following lemma. The assertions and occuring quantities are also illustrated in Figure 1.

Lemma 3.10. *Let (A1'), (A2') and (A4) hold and let $u_I \in BV_{\rm loc}(\mathbb{R})$, $u_1 \in C^1([0,T] \times \mathbb{R})^m$. Consider for $i = 1, \dots, n_\sigma$ the i-th red phase of the traffic light problem* (2.6)*. Assume that the final state of the i-th green phase $y(\sigma_{\rm r}^i, \cdot)$ is bounded away from 0 on $(-\tilde{\varepsilon}, 0]$ and bounded away from -1 on $[0, \tilde{\varepsilon})$ for some $\tilde{\varepsilon} > 0$. Then the solution of* (2.6) *satisfies the following equations.*

1. *For every $\theta \in (\sigma_{\rm r}^i, \sigma_{\rm g}^i)$ there exists $\tilde{\delta} > 0$ such that there holds*

$$y_1(t,x) = -1, \qquad (t,x) \in (\theta, \sigma_{\rm g}^i) \times (-\tilde{\delta}, 0),$$
$$y_2(t,x) = 0, \qquad (t,x) \in (\theta, \sigma_{\rm g}^i) \times (0, \tilde{\delta}).$$

2. *There exists $\delta > 0$ such that for all $0 < \tau < \frac{\delta}{2M_{f'}}$ there holds*

$$y(\sigma_{\rm g}^i + \tau, x) = \begin{cases} f'^{-1}\left(\frac{x}{\tau}\right), & \text{if } x \in [f'(-1)\tau, f'(0)\tau], \\ 0, & \text{if } x \in (f'(0)\tau, \delta - M_{f'}\tau), \\ -1, & \text{if } x \in (-\delta + M_{f'}\tau, f'(-1)\tau), \end{cases}$$

with $M_{f'}$ from Corollary 3.9.

4. Shift-Differentiability

In this section we present the main results of this paper, namely the shift-differentiable dependence of the control-to-state mapping for the considered problems (2.1) and (2.6).

4.1. Motivation and preliminary work

One of the main difficulties that arise when one considers optimal control problems concerning entropy solutions of hyperbolic conservation laws is, that the control-to-state mapping $u \mapsto y(u)$ is generally not differentiable in a sense, that is strong enough in order to simply deduce Fréchet-differentiability of the reduced objective functional. This issue of non-differentiability is caused by the presence of shocks in the entropic solution, even for smooth (e.g. C^∞) data. However we illustrate the situation by means of an example where the data are discontinuous, namely a Riemann problem.

Example. Consider the parametrized Cauchy problem

$$\begin{aligned} y_t^\varepsilon + \left(\tfrac{1}{2}\left(y^\varepsilon\right)^2\right)_x &= 0 && \text{on } [0,T]\times\mathbb{R} \\ y^\varepsilon(0,\cdot) &= \varepsilon - \operatorname{sgn} && \text{on } \mathbb{R}. \end{aligned}$$

Then the entropy solution is almost everywhere given by

$$y^\varepsilon(t,x) = \begin{cases} \varepsilon + 1, & \text{if } x \le \varepsilon t, \\ \varepsilon - 1, & \text{if } x > \varepsilon t. \end{cases}$$

Furthermore, consider the mapping $S : \mathbb{R} \to L^1([a,b]),\ \varepsilon \mapsto y^\varepsilon(\bar{t},\cdot)$. Clearly S is not differentiable in 0, since the obvious candidate for the derivative, $1 + 2\bar{t}\delta_0$, where δ_0 denotes the Dirac measure at $x = 0$, does not belong to $\mathcal{L}(\mathbb{R}, L^1([a,b]))$. In fact, differentiabilty does only hold in the weak topology of the measure space $\mathcal{M}([a,b])$.

In order to still achieve a differentiability result for the reduced objective, a non-standard variational calculus was introduced in [9] and [36, 37]. The so called shift-variations mimic the observed behavior of the solution in the neighborhood of discontinuities. Shift-variations consist of an additive part (in L^1) and a second part that allows for horizontal shifts of discontinuities. We recall the definitions of the notions of shift-variations and shift-differentiability.

Definition 4.1 (Shift-variations, shift-differenbtiability). 1. Let $a < b$ and $v \in BV([a,b])$. For $a < x_1 < x_2 < \cdots < x_N < b$ we associate with $(\delta v, \delta x)$ the *shift-variation* $S_v^{(x_i)}(\delta v, \delta x) \in L^1([a,b])$ of v by

$$S_v^{(x_i)}(\delta v, \delta x)(x) := \delta v(x) \cdot \sum_{i=1}^{n} [v(x_i)] \operatorname{sgn}(\delta x_i) \mathbb{1}_{I(x_i, x_i + \delta x_i)}(x),$$

where $[v(x_i)] := v(x_i-) - v(x_i+)$ and $I(\alpha,\beta) := [\min(\alpha,\beta), \max(\alpha,\beta)]$.

2. Let U be a real Banach space and $D \subset U$ open. Consider a locally bounded mapping $D \to L^\infty(\mathbb{R})$, $u \mapsto v(u)$. For $\bar{u} \in U$ with $v(\bar{u}) \in BV([a,b])$, we call v *shift-differentiable at* $\bar{u}$ if there exist $a < x_1 < x_2 < \cdots < x_N < b$ and $D_s v(\bar{u}) \in \mathcal{L}(U, L^r([a,b]) \times \mathbb{R}^N)$ for some $r \in (1, \infty]$, such that for $\delta u \in U$, $(\delta v, \delta x) := D_s v(\bar{u}) \cdot \delta u$

$$\left\| v(u+\delta u) - v(u) - S_v^{(x_i)}(\delta v, \delta x) \right\|_{1,[a,b]} = o(\|\delta u\|_U).$$

The utility of this variational concept lies in the feature that it implies the Fréchet-differentiability of tracking type functionals as in (1.1) (see Lemma 3.2.3 in [36]) as long as y_d and $y(\bar{t}, \cdot)$ do not share discontinuities on $[a,b]$. The derivative is given by

$$d_u J(y(u)) \cdot \delta u = (\psi_y(y(\bar{t}, \cdot; u), y_d), \delta y)_{2,[a,b]} + \sum_{i=1}^{N} \bar{\psi}_y(x_i)[y(\bar{t}, \cdot; u)]\delta x_i,$$

with

$$\bar{\psi}_y(x) := \int_0^1 \psi_y\left(y(\bar{t}, x+; u)) + \tau[y(\bar{t}, x; u)], y_d(x+) + \tau[y_d(x)]\right) \, \mathrm{d}\tau.$$

4.2. Shift-Differentiability of solutions to IBVPs and traffic light problems

We now state the main results. First we consider the differentiability of the solution operator for the initial-boundary value problem. We restrict ourselves to the case $\Omega = (0, \infty)$, where the result for general intervals is similar. A reinspection of the formulation of the boundary condition (2.4a) motivates to only consider boundary data with $u_B \geq f'^{-1}(0)$, since both the choices u_B and $\max(u_B, f'^{-1}(0))$ as boundary data will yield the same solution. Therefore it is useful to define the space

$$U_B^\alpha := \{\varphi \in PC^1([0,T]; t_1 \ldots, t_K) \;:\; f'(\varphi) \geq \alpha\} \tag{4.1}$$

for given $0 < t_1 < t_2 < \cdots < t_K$. Consider $u = (u_0, u_B, u_1)$ where $u_B \in U_B^\alpha$ for some small $\alpha > 0$, $u_0 \in PC^1(\Omega; x_1, \ldots, x_N)$ for some $0 < x_1 < x_2 < \cdots < x_N$ and $u_1 \in C([0,T]; C^1(\mathbb{R})^m)$. We want to investigate the shift-differentiable dependence of $\delta u \mapsto y(\bar{t}, \cdot; u + \delta u)$ on δu. In addition to usual variations in the controls, we additionally consider some shift-variations of the initial and the boundary data. This means that we consider explicit shifts of discontinuities that create shocks, but no rarefactions. For this purpose we define

$$\mathbf{S}_{(x_i)} := \{s \in \mathbb{R}^N \;:\; u_0(x_i-) < u_0(x_i+) \;\Rightarrow s_i = 0, \; i = 1, \ldots, N\},$$
$$\mathbf{S}_{(t_j)} := \{s \in \mathbb{R}^K \;:\; u_B(t_j-) > u_B(t_j+) \;\Rightarrow s_j = 0, \; j = 1, \ldots, K\}$$

and consider variations in

$$\begin{aligned} W := {} & PC^1(\Omega; x_1, \ldots, x_N) \times \mathbf{S}_{(x_i)} \\ & \times PC^1([0,T]; t_1, \ldots, t_K) \times \mathbf{S}_{(t_j)} \times C([0,T]; C^1(\mathbb{R})^m). \end{aligned} \tag{4.2}$$

Under a nondegeneracy condition on the shocks (see Definition 3.6.1 in [36]) we get the following result.

Theorem 4.2 (Shift-Differentiability for IBVPs). *Let (A1') and (A3) hold and let in addition g be affine linear w.r.t. y. Let $\Omega = (0, \infty)$ and $0 < x_1 < x_2 < \cdots < x_N$, $0 < t_1 < t_2 < \cdots < t_K$ $u_0 \in PC^1(\Omega; x_1, \ldots, x_N)$, $u_B \in U_B^\alpha$ for some $\alpha > 0$ and $u_1 \in C([0,T]; C^1(\mathbb{R})^m)$. For $u = (u_0, u_B, u_1)$ denote by $y = y(u) \in L^\infty(\Omega_T) \cap C([0,T]; L^1_{\text{loc}}(\Omega))$ the entropy solution of the initial-boundary value problem* (2.1) *on Ω_T. Let $0 < a < b$ and $\bar{t} \in (0,T)$ such that on $[a,b]$ $y(\bar{t}, \cdot; u)$ has no shock generation points and only a finite number of shocks at $a < \bar{x}_1 < \cdots < \bar{x}_{\bar{N}} < b$, that all are neither degenerated nor shock interaction points. Further assume that for almost all $t \in [0,T]$ the boundary trace $y(\cdot, 0+; u) \in L^\infty(0,T)$ satisfies $u_B(t) \neq y(t, 0+; u) \;\Rightarrow\; f(u_B(t)) \neq f(y(t, 0+; u))$.*

For W from (4.2) *we consider the mapping*

$$
\begin{aligned}
&(\delta u_0, \delta x, \delta u_B, \delta t, \delta u_1) \in W \longmapsto \\
&y\left(\bar{t}, \cdot; u_0 + S_{u_0}^{(x_i)}(\delta u_0, \delta x), u_B + S_{u_B}^{(t_j)}(\delta u_B, \delta t), u_1 + \delta u_1\right) \in L^1(a,b). \quad (4.3)
\end{aligned}
$$

If $(x_i), (t_j)$ are real discontinuities of u_0, u_B, i.e. $u_0(x_i-) \neq u_0(x_i+)$ and $u_B(t_j-) \neq u_B(t_j+)$, respectively, then the mapping (4.3) *is continuously shift-differentiable on a sufficiently small neighborhood $B_\rho^W(0) := \{\delta_u \in W \; : \; \|\delta u\|_W \leq \rho\}$. The shift-derivative satisfies $T_s(0) = D_s y(\bar{t}, \cdot; u) \in \mathcal{L}(W, PC([a,b]; \bar{x}_1, \ldots, \bar{x}_{\bar{N}}) \times \mathbb{R}^{\bar{N}})$.*

Remark 4.3. If u_0 or u_B are continuous at some x_i or t_j, respectively, similarly to the second assertion of Theorem 3.3.2 in [36], the shift-differentiability of (4.3) in 0 is preserved. The shift-derivative satisfies $T_s(0) \in \mathcal{L}(W, PC([a,b]; \bar{x}_1, \ldots, \bar{x}_{\bar{N}}, \tilde{x}_1, \ldots, \tilde{x}_{\tilde{N}}) \times \mathbb{R}^{\bar{N}})$, where the set of discontinuities of $y(u)$ is augmented by continuity points $\tilde{x}_k$ that are starting points of genuine backward characteristics that end in a (pseudo-) discontinuity x_i or t_j.

The proof can be obtained by a very careful extension of the proof of Theorem 3.3.2 in [36]. This requires a proper analysis of the solution y in small neighborhoods of different types of generalized backward characteristics. A detailed proof will be presented in a forthcoming paper.

The following corollary is a simple consequence of the above theorem and Lemma 3.2.3 in [36].

Corollary 4.4. *Let the assumptions of Theorem* 4.2 *hold and consider J defined as in* (1.1). *If y_d is continuous in a small neighborhood of $\{\bar{x}_1, \ldots, \bar{x}_{\bar{N}}\}$, then the reduced objective functional $\delta u \in W \mapsto J(y(u + \delta u))$ is continuously Fréchet-differentiable on $B_\rho^W(0)$ for $\rho > 0$ small enough.*

An adjoint-based formula for the gradient of the considered mapping will be presented in Theorem 4.8.

For the traffic light problem we have a very similar result.

Theorem 4.5 (Shift-Differentiability for traffic light problems). *Let (A1') and (A4) hold and let in addition g be affine linear w.r.t. y. Let $x_1 < x_2 < \cdots < x_N$, $\sigma = (\sigma_{\text{g}}^0, \sigma_{\text{r}}^1, \sigma_{\text{g}}^1, \ldots, \sigma_{\text{g}}^{n_\sigma}, \sigma_{\text{r}}^{n_\sigma+1}) \in \Sigma_{\text{ad}}$, $PC^1(\mathbb{R}; x_1, \ldots, x_N)$ and*

$u_1 \in C([0,T];C^1(\mathbb{R})^m)$. For $\sigma \in \Sigma_{\mathrm{ad}}$ denote by $y = y(\sigma) \in L^\infty(\Omega_T) \cap C([0,T];L^1_{\mathrm{loc}}(\Omega))$ the solution of the traffic light problem (2.6). Let $a < b$ and $\bar{t} \in (\sigma_{\mathrm{g}}^{n_\sigma}, \sigma_{\mathrm{r}}^{n_\sigma+1})$ such that on $[a,b]$ $y(\bar{t},\cdot;\sigma)$ has no shock generation points and only a finite number of shocks at $a < \bar{x}_1 < \cdots < \bar{x}_{\bar{N}} < b$, that all are neither degenerated nor shock interaction points. Furthermore assume that for almost all $t \in [\sigma_{\mathrm{g}}^i, \sigma_{\mathrm{r}}^i]$, $i = 1,\dots,n_\sigma$ the boundary traces $(y(\cdot,0-;\sigma), y(\cdot,0+;\sigma)) \in L^\infty(0,T)^2$ are equal to $(-1,0)$.

Finally let $\Sigma_0 := \{\nu \in \mathbb{R}^{2(n_\sigma+1)} \ : \ \nu_1 = \nu_{2(n_\sigma+1)} = 0\}$ then the mapping

$$\delta\sigma \in \Sigma_0 \longmapsto y\left(\bar{t},\cdot;\sigma+\delta\sigma\right) \in L^1(a,b)$$

is continuously shift-differentiable on a sufficiently small neighborhood $B_\rho^\Sigma(0) := \{\delta\sigma \in \Sigma_0 \ : \ \|\delta\sigma\| \le \rho\}$. The shift-derivative satisfies $T_s(0) = D_s y(\bar{t},\cdot;\sigma) \in \mathcal{L}(\Sigma_0, PC([a,b];\bar{x}_1,\cdot,\bar{x}_{\bar{N}}) \times \mathbb{R}^{\bar{N}})$.

It is important to emphasize that in comparison to the result for the initial (-boundary) value problem, also green switching times, i.e. rarefaction centers, may explicitly be shifted. This is because the solution in a neighborhood of such points is thoroughly known for TLPs, see Lemma 3.10, whereas the structure for general rarefaction waves may be more delicate.

As for the IBVP, one can deduce from Lemma 3.2.3 in [36] the total differentiability for reduced objective functionals.

Corollary 4.6. *Let the assumptions of Theorem 4.5 hold and consider J defined as in (1.1). If y_d is continuous in a small neighborhood of $\{\bar{x}_1,\dots,\bar{x}_{\bar{N}}\}$, then the reduced functional $\delta\sigma \in \Sigma_0 \mapsto J(y(\sigma+\delta\sigma))$ is continuously differentiable on $B_\rho^\Sigma(0)$ for $\rho > 0$ small enough.*

One also may consider the optimal control of the traffic light problem for fixed switching times where the source term and the initial data is controlled. Here one can obtain similar results as for the initial (-boundary) value problem without any traffic lights.

4.3. Adjoint Equation

The sensitivity of the shock position, that is needed in order to obtain the shift-differentiabilty result of Theorem 4.2, is based on an adjoint-argument. As already discussed in [36] for the Cauchy problem, the classical adjoint calculus is not applicable for problems concerning discontinuous solutions of hyperbolic equations. Nevertheless one can define an adjoint state as a solution of the following equation

$$p_t + f'(y)p_x = -g_y(\cdot,y,u_1)p, \qquad \text{on } \Omega_{\bar{t}}, \tag{4.4a}$$

$$p(\bar{t},\cdot) = p^{\bar{t}}, \qquad \text{on } \Omega. \tag{4.4b}$$

The adjoint equation (4.4) is a linear transport equation with discontinuous coefficients, since y may contain shocks. In [6] Bouchut and James showed that for $\Omega = \mathbb{R}$, $g \equiv 0$ and Lipschitz continuous end data $p^{\bar{t}}$ equation (4.4) does not admit a unique solution within the space of Lipschitz continuous functions. Nevertheless they give a definition and a characterization of a *reversible solution* for (4.4), which satisfies a crucial duality relation. In [36, 38] this notion was extended to more general source terms g and discontinuous end data. In this case the reversible solution p can be characterized as the solution along generalized characteristics of the state y. For the IBVP on $\Omega = (0,\infty)$ we have to deal with the fact that this definition might lead to an underdetermined problem, since not all characteristics on $\Omega_{\bar{t}}$ intersect the line $\{\bar{t}\} \times \Omega$, where the initial (or terminal) condition acts. One can show by the theory of generalized characteristics that the set D of points that lie on a genuine characteristic that does not reach the line $\{\bar{t}\} \times \Omega$ is a connected set that lies in the lower left corner of the space-time cylinder $\Omega_{\bar{t}}$.

Definition 4.7. Let $p^{\bar{t}}$ be a bounded function that is the pointwise everywhere limit of a sequence (w_n) in $C^{0,1}(0,\infty)$, with (w_n) bounded in $C(0,\infty) \cap W^{1,1}_{\mathrm{loc}}(0,\infty)$. The adjoint state p associated to (4.4) for $\Omega = (0,\infty)$ is characterized by the following requirements:

1. For every generalized characteristic ξ of y through $(\bar{t},\bar{x}) \in \Omega_T$

$$t \mapsto p^{\xi}(t) = p(t,\xi(t))$$

   is the solution of the ordinary differential equation

$$\begin{aligned} \dot{p}^{\xi}(t) &= -g_y(t,\xi(t),y(t,\xi(t)),u_1(t,\xi(t)))p^{\xi}(t), \qquad t \in (0,\bar{t}] \;:\; \xi(t) > 0, \\ p^{\xi}(\bar{t}) &= p^{\bar{t}}(\bar{x}). \end{aligned}$$

2. For every $(t,x) \in D$ there holds $p(t,x) = 0$, where

$$D := \{(t,x) \in \Omega_{\bar{t}} \;:\; t \in [0,\tau],\; x \le \tilde{\xi}(t)\}.$$

   Here $\tilde{\xi}$ denotes the maximal backward characteristic through $(\tau,0)$, where $\tau := \operatorname{esssup}\{t \in [0,\bar{t}] \;:\; f'(y(t,0+)) < 0\}$.

Using the above definition of an adjoint state, we are now able to formulate a representation of the gradient of the reduced objective function.

Theorem 4.8. *Let the assumptions of Corollary* 4.4 *hold and let the terminal data in* (4.4) *be given by*

$$p^{\bar{t}}(t,x) := \gamma(x) \int_0^1 \psi_y(y(\bar{t},x+) + \tau[y(\bar{t},x)],x)\, d\tau.$$

Then there exists an adjoint state p according to Definition 4.7*, satisfying*

$$p \in B((0,\bar{t}) \times (0,\infty)) \cap BV_{\mathrm{loc}}([0,\bar{t}] \times [0,\infty)),$$

where $B((0,\bar{t}) \times (0,\infty))$ denotes the space of measurable bounded functions (defined pointwise everywhere).

The derivative of the reduced functional $\delta u \in W \mapsto \hat{J}(\delta u) = J(y(u+\delta u))$ for $\rho > 0$ small enough is given by

$$
\begin{aligned}
\hat{J}'(0) \cdot \delta u = (p, g_{u_1}(\cdot, y, u_1)\delta u_1)_{2,(0,\bar{t})\times\mathbb{R}^+} \\
+ (p(0,\cdot), \delta u_0)_{2,\mathbb{R}^+} + (p(\cdot,0), f'(u_B)\delta u_B)_{2,(0,\bar{t})} \\
+ \sum_{i=1}^{N} p(0,x_i)[u_0(x_i)]\delta x_i + \sum_{j=1}^{K} p(t_j,0)[f(u_B(t_j))]\delta t_j.
\end{aligned}
$$

5. Conclusion and Outlook

We have presented a generalized differentiability result for an initial-boundary value problem for a nonlinear hyperbolic conservation law on an interval by using the theory of generalized characteristics. This property implies the Fréchet-differentiability of the reduced objective functional, for which we also presented an adjoint-based gradient representation. The result is an important step to make such problems accessible to gradient based optimization algorithms. Furthermore we have discussed the dependence of the state on the switching times of a traffic light on a single road. Also in this case we were able to show shift-differentiability by similar arguments. The considered problem for the traffic light can also be seen as a network problem involving one node and two edges and can be in a straight-forward manor extended to the case of multiple incoming and outgoing roads that are connected by a similar modular node condition that time dependently connects some pairs of incoming and outgoing roads and closes others. If one chooses the sequence of modes in such a way, that no road is open for two or more consecutive time phases, the same arguments as for the traffic light problem can be used. Questions for future research will be whether one may drop the latest assumption. Moreover we will have to investigate the case when the boundary data of the red light condition (2.6f), (2.6g) is not assumed by the boundary trace, which means that the traffic light turns red, when either the incoming road is empty near the traffic light or the outgoing road has already reached its maximum capacity. This becomes more important, if one considers multiple traffic lights in a row. Another interesting modification of the traffic problem is the case where the flux functions on the two sides of the junction are not necessarily the same. Moreover, it will be of interest how the shift-differentiability concept applies to networks of three edges, that are connected by more common node conditions, as those from [10] and [13].

Finally, our results form the basis for the convergence analysis of numerical approximations of the considered optimal control problems. So far, there exist several results in the context of initial value problems with initial control and sometimes also with control in the source term. The convergence of optimal solutions of dicretized optimal control problems was considered e.g. in [11, 35]. The convergence of sensitivities, adjoints and reduced gradients was analyzed in [19, 20, 36, 37, 38], see also [11] for an alternating descent method. We are currently investigating the extension of these results to the case of the initial-boundary value problem with boundary control and to the traffic light problem. Here, we follow the approach in [12] for the discrete approximation of the boundary condition, where the convergence to the unique entropy solution of the initial-boundary value problem according to [4] is shown. A particular issue will be the appropriate discrete approximation of shift variations for boundary controls. We plan to consider the variation of the times step sizes between switching times as well as disretization techniques with fixed time steps.

Acknowledgment

The authors gratefully acknowledge the support of the German Research Foundation (DFG) within the Priority Program 1253 "Optimization with Partial Differential Equations" under grant UL158/8-1. Moreover, we gratefully acknowledge discussions with T. I. Seidman.

References

[1] F. Ancona and G. M. Coclite. On the attainable set for Temple class systems with boundary controls. *SIAM J. Control Optim.*, 43(6):2166–2190 (electronic), 2005.

[2] F. Ancona and A. Marson. On the attainable set for scalar nonlinear conservation laws with boundary control. *SIAM J. Control Optim.*, 36(1):290–312 (electronic), 1998.

[3] M. K. Banda, M. Herty, and A. Klar. Coupling conditions for gas networks governed by the isothermal Euler equations. *Netw. Heterog. Media*, 1(2):295–314, 2006.

[4] C. Bardos, A. Y. le Roux, and J.-C. Nédélec. First order quasilinear equations with boundary conditions. *Comm. Partial Differential Equations*, 4(9):1017–1034, 1979.

[5] S. Bianchini. On the shift differentiability of the flow generated by a hyperbolic system of conservation laws. *Discrete Contin. Dynam. Systems*, 6(2):329–350, 2000.

[6] F. Bouchut and F. James. One-dimensional transport equations with discontinuous coefficients. *Nonlinear Anal.*, 32(7):891–933, 1998.

[7] A. Bressan. *Hyperbolic systems of conservation laws*, volume 20 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2000. The one-dimensional Cauchy problem.

[8] A. Bressan and G. Guerra. Shift-differentiability of the flow generated by a conservation law. *Discrete Contin. Dynam. Systems*, 3(1):35–58, 1997.

[9] A. Bressan and A. Marson. A variational calculus for discontinuous solutions of systems of conservation laws. *Comm. Partial Differential Equations*, 20(9-10):1491–1552, 1995.

[10] G. Bretti, R. Natalini, and B. Piccoli. Numerical approximations of a traffic flow model on networks. *Netw. Heterog. Media*, 1(1):57–84, 2006.

[11] C. Castro, F. Palacios, and E. Zuazua. An alternating descent method for the optimal control of the inviscid Burgers equation in the presence of shocks. *Math. Models Methods Appl. Sci.*, 18(3):369–416, 2008.

[12] B. Cockburn, F. Coquel, and P. G. LeFloch. Convergence of the finite volume method for multidimensional conservation laws. *SIAM J. Numer. Anal.*, 32(3):687–705, 1995.

[13] G. M. Coclite, M. Garavello, and B. Piccoli. Traffic flow on a road network. *SIAM J. Math. Anal.*, 36(6):1862–1886, 2005.

[14] G. M. Coclite, K. H. Karlsen, and Y.-S. Kwon. Initial-boundary value problems for conservation laws with source terms and the Degasperis-Procesi equation. *J. Funct. Anal.*, 257(12):3823–3857, 2009.

[15] R. M. Colombo and A. Groli. On the optimization of the initial boundary value problem for a conservation law. *J. Math. Anal. Appl.*, 291(1):82–99, 2004.

[16] R. M. Colombo, G. Guerra, M. Herty, and V. Schleper. Optimal control in networks of pipes and canals. *SIAM J. Control Optim.*, 48(3):2032–2050, 2009.

[17] C. M. Dafermos. Generalized characteristics and the structure of solutions of hyperbolic conservation laws. *Indiana Univ. Math. J.*, 26(6):1097–1119, 1977.

[18] F. Dubois and P. LeFloch. Boundary conditions for nonlinear hyperbolic systems of conservation laws. *J. Differential Equations*, 71(1):93–122, 1988.

[19] M. Giles and S. Ulbrich. Convergence of linearized and adjoint approximations for discontinuous solutions of conservation laws. Part 1: Linearized approximations and linearized output functionals. *SIAM J. Numer. Anal.*, 48(3):882–904, 2010.

[20] M. Giles and S. Ulbrich. Convergence of linearized and adjoint approximations for discontinuous solutions of conservation laws. Part 2: Adjoint approximations and extensions. *SIAM J. Numer. Anal.*, 48(3):905–921, 2010.

[21] S. Göttlich, M. Herty, and A. Klar. Modelling and optimization of supply chains on complex networks. *Commun. Math. Sci.*, 4(2):315–330, 2006.

[22] M. Gugat, M. Herty, A. Klar, and G. Leugering. Optimal control for traffic flow networks. *J. Optim. Theory Appl.*, 126(3):589–616, 2005.

[23] M. Gugat, M. Herty, and V. Schleper. Flow control in gas networks: exact controllability to a given demand. *Math. Methods Appl. Sci.*, 34(7):745–757, 2011.

[24] F. M. Hante, G. Leugering, and T. I. Seidman. Modeling and analysis of modal switching in networked transport systems. *Appl. Math. Optim.*, 59(2):275–292, 2009.

[25] H. Holden and N. H. Risebro. A mathematical model of traffic flow on a network of unidirectional roads. *SIAM J. Math. Anal.*, 26(4):999–1017, 1995.

[26] S. N. Kružkov. Quasilinear parabolic equations and systems with two independent variables. *Trudy Sem. Petrovsk.*, (5):217–272, 1979.

[27] A. Y. le Roux. Étude du problème mixte pour une équation quasi-linéaire du premier ordre. *C. R. Acad. Sci. Paris Sér. A-B*, 285(5):A351–A354, 1977.

[28] P. LeFloch. Explicit formula for scalar nonlinear conservation laws with boundary condition. *Math. Methods Appl. Sci.*, 10(3):265–287, 1988.

[29] M. J. Lighthill and G. B. Whitham. On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proc. Roy. Soc. London. Ser. A.*, 229:317–345, 1955.

[30] J. Málek, J. Nečas, M. Rokyta, and M. Růžička. *Weak and measure-valued solutions to evolutionary PDEs*, volume 13 of *Applied Mathematics and Mathematical Computation*. Chapman & Hall, London, 1996.

[31] F. Otto. Initial-boundary value problem for a scalar conservation law. *C. R. Acad. Sci. Paris Sér. I Math.*, 322(8):729–734, 1996.

[32] V. Perrollaz. Asymptotic stabilization of entropy solutions to scalar conservation laws through a stationary feedback law. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 30(5):879–915, 2013.

[33] E. Polak. *Optimization*, volume 124 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1997. Algorithms and consistent approximations.

[34] P. I. Richards. Shock waves on the highway. *Operations Res.*, 4:42–51, 1956.

[35] S. Ulbrich. On the existence and approximation of solutions for the optimal control of nonlinear hyperbolic conservation laws. In *Optimal control of partial differential equations (Chemnitz, 1998)*, volume 133 of *Internat. Ser. Numer. Math.*, pages 287–299. Birkhäuser, Basel, 1999.

[36] S. Ulbrich. Optimal control of nonlinear hyperbolic conservation laws with source terms. Habilitation, Zentrum Mathematik, Technische Universität München, Germany, 2001.

[37] S. Ulbrich. A sensitivity and adjoint calculus for discontinuous solutions of hyperbolic conservation laws with source terms. *SIAM J. Control Optim.*, 41(3):740–797 (electronic), 2002.

[38] S. Ulbrich. Adjoint-based derivative computations for the optimal control of discontinuous solutions of hyperbolic conservation laws. *Systems Control Lett.*, 48(3-4):313–328, 2003. Optimization and control of distributed systems.

[39] A. Vasseur. Strong traces for solutions of multidimensional scalar conservation laws. *Arch. Ration. Mech. Anal.*, 160(3):181–193, 2001.

Sebastian Pfaff
Technische Universität Darmstadt
Department of Mathematics
Dolivostr. 15, 64293 Darmstadt, Germany
e-mail: `pfaff@mathematik.tu-darmstadt.de`

Stefan Ulbrich
Technische Universität Darmstadt
Department of Mathematics
Dolivostr. 15, 64293 Darmstadt, Germany
e-mail: `ulbrich@mathematik.tu-darmstadt.de`

Günter Leugering
Universität Erlangen-Nürnberg
Department of Mathematics
Cauerstr. 11, 91058 Erlangen, Germany
e-mail: `leugering@math.fau.de`