OPTIMAL BOUNDARY CONTROL OF NONLINEAR HYPERBOLIC CONSERVATION LAWS WITH SWITCHED BOUNDARY DATA *

SEBASTIAN PFAFF † AND STEFAN ULBRICH ‡

Abstract. We consider the optimal control of initial-boundary value problems for entropy solutions of scalar hyperbolic conservation laws. In particular, we consider initial-boundary value problems where the initial and boundary data switch between different C^1-functions at certain switching points and both, the functions and the switching points, are controlled. We show that the control-to-state mapping is differentiable in a certain generalized sense, which implies Fréchet-differentiability with respect to the control functions and the switching points for the composition with a tracking type functional, even in the presence of shocks. We also present an adjoint-based formula for the gradient of the reduced objective functional.

Key words. optimal control; scalar conservation law; differentiability; adjoint state; shock sensitivity

AMS subject classifications. 49K20; 35L65; 49J50; 35R05

1. Introduction.

In this paper we develop a sensitivity and adjoint calculus for the optimal control of entropy solutions to hyperbolic conservation laws. We consider 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 and the state y is governed by an initial-boundary value problem (IBVP) on $\Omega := (\mathfrak{a}, \mathfrak{b})$, with $-\infty \le \mathfrak{a} < a < b < \mathfrak{b} \le \infty$, for a scalar conservation law

$$\begin{aligned}
y_t + f(y)_x &= g(\cdot, y, u_1), && \text{on } \Omega_T, && && (1.2\text{a})\\
y(0,\cdot) &= u_0, && \text{on } \Omega, && && (1.2\text{b})\\
y(\cdot, a+) &= u_{B,\mathfrak{a}}, && \text{in the sense of (2.2a)} && (\text{if } \mathfrak{a} > -\infty), && (1.2\text{c})\\
y(\cdot, b-) &= u_{B,\mathfrak{b}}, && \text{in the sense of (2.2b)} && (\text{if } \mathfrak{b} < \infty), && (1.2\text{d})
\end{aligned}$$

where $\Omega_T := [0,T] \times \Omega$ and $u = (u_0, u_{B,\mathfrak{a}}, u_{B,\mathfrak{b}}, u_1)$ is the control. The Cauchy problem was already discussed in [31], so that in this paper we mainly focus on the boundary control. Each boundary data u_B consist of a finite collection of functions $u_B^0, \dots, u_B^{n_t} \in C^1([0,T])$ and switching points $0 < t_1 < \cdots < t_{n_t} < T$, combined by

$$u_B(t) = \begin{cases} u_B^0(t), & \text{if } t \in [0, t_1],\\ u_B^j(t), & \text{if } t \in (t_j, t_{j+1}],\ j = 1, \dots, n_t - 1,\\ u_B^{n_t}(t), & \text{if } t \in (t_{n_t}, T]. \end{cases}$$

We want to minimize (1.1) w.r.t. the functions (u_B^j) and the switching points (t_j).

---

*Supported by German Research Foundation (DFG) within the Priority Program 1253 "Optimization with Partial Differential Equations"

†Dept. of Mathematics, TU Darmstadt, pfaff@mathematik.tu-darmstadt.de

‡Dept. of Mathematics, TU Darmstadt, ulbrich@mathematik.tu-darmstadt.de, this author was also supported by DFG within the collaborative research center TRR 154

The present paper forms an essential extension of the results of [31] since several issues arise by including the boundary condition to the model which make the sensitivity calculus quite involved. Furthermore, we present an adjoint based formula for the reduced gradient $\frac{\mathrm{d}}{\mathrm{d}u}J(y(u))$ in the flavor of [32].

It is well known, that hyperbolic conservation laws do not admit unique weak solutions and that one has to consider entropy solutions of (1.2a) instead, in order to maintain the uniqueness, cf. [19] and (2.1). The second important issue is that the Dirichlet-like boundary condition (1.2c)-(1.2d) must not be understood literately but in the BLN-sense (2.2), see [3].

Since entropy solutions develop shocks after finite time even for smooth data, see [6], the sensitivity analysis becomes intricate. Indeed, differentiability of the control-to-state mapping $u \mapsto y(\bar{t}, \cdot; u)$ for the Cauchy problem holds at best with respect to the weak topology of measures.

Even though standard variational calculus fails for these types of problems, the optimal control of conservation laws has been studied intensively in recent years. Several authors investigated the question of the existence of optimal controls for the Cauchy and the initial-boundary value problem, e.g. [1, 2, 29, 30].

Several techniques haven been studied in order to overcome the non-differentiability of the control-to-state mapping, see [4, 7, 8, 9, 12, 30, 31]. The results of this paper are based on [30, 31], where the concept of *shift-differentiability* was introduced. This ansatz also admits an adjoint calculus for the reduced objective function, see also [15, 16, 32]. A crucial tool for our analysis will be the concept of generalized characteristics introduced by Dafermos [13].

To the best of the authors' knowledge, the notion of "switched control" in the context of hyperbolic conservation laws was mentioned for the first time in [18], where also switching in flux function and the source term was considered. The optimal control of convection-reaction equations with switched boundary control was considered in [17].

The results we present in this paper can be used to make the considered infinite dimensional optimal control problem accessible to gradient-based optimization methods and to derive optimality conditions, even in the presence of shocks. This in turn, will give rise to the development of appropriate numerical methods for such problems. Moreover, the result forms the basis for many possible extensions, such as to systems or networks of conservation laws. Since our present result allows for explicit shifting of discontinuities in the boundary data, it is an important step towards the optimal control of networks where the control variables are the switching times between different modes of the node condition. We will demonstrate this possibility by means of the example of a traffic light in a forthcoming paper.

The present paper is organized as follows. In §2 we introduce the considered initial-boundary value problem and collect results on its well-posedness and structural properties of the solution. The main result of our paper will be presented in §3, where we state the shift-differentiability of the control-to-state mappings and the adjoint-based formula for the Fréchet-derivative of the reduced objective function. Those results were already announced in [26]. The proofs of the theorems are postponed to §4.

2. The initial-boundary value problem. In this section we want to discuss the initial-boundary value problem (IBVP).

2.1. Entropic solutions to the IBVP. We are interested in entropy solutions of (1.2), namely solutions that satisfy (1.2a)-(1.2b) in the sense of [19, 3], i.e.

$$(\eta_c(y))_t + (q_c(y))_x - \eta_c'(y) g(\cdot, y, u_1) \leq 0, \qquad \text{in } \mathcal{D}'(\Omega_T), \tag{2.1a}$$

$$\operatorname*{esslim}_{t\to 0+} \|y(t,\cdot) - u_0\|_{1,\Omega\cap(-R,R)} = 0, \qquad \text{for all } R > 0 \tag{2.1b}$$

holds for every (Kružkov-) entropy $\eta_c(\lambda) := |\lambda - c|$, $c \in \mathbb{R}$, and associated entropy flux $q_c(\lambda) := \operatorname{sgn}(\lambda - c)(f(\lambda) - f(c))$.

In order to get a well posed problem, the boundary conditions (1.2c), (1.2d) have to be understood in the sense of [3], that is

$$\min_{k \in I(y(\cdot,\mathfrak{a}+), u_{B,\mathfrak{a}})} \operatorname{sgn}(u_{B,\mathfrak{a}} - y(\cdot,\mathfrak{a}+))(f(y(\cdot,\mathfrak{a}+)) - f(k)) = 0, \quad \text{a.e. on } [0,T], \tag{2.2a}$$

$$\min_{k \in I(y(\cdot,\mathfrak{b}-), u_{B,\mathfrak{b}})} \operatorname{sgn}(y(\cdot,\mathfrak{b}-) - u_{B,\mathfrak{b}})(f(y(\cdot,\mathfrak{b}-)) - f(k)) = 0, \quad \text{a.e. on } [0,T], \tag{2.2b}$$

with $I(\alpha, \beta) := [\min(\alpha,\beta), \max(\alpha,\beta)]$, see also [14, 21, 23, 24].

We should mention that the BLN-condition (2.2) involves the boundary traces $y(\cdot,\mathfrak{a}+)$ and $y(\cdot,\mathfrak{b}-)$. For BV-data the traces exist and, therefore, the condition is well-defined. For L^∞-data one can replace the BLN-condition by the one proposed by Otto [24, 23]. But Vasseur showed [33] that under mild assumptions even for L^∞-entropy solutions there always exist boundary traces in L^∞, that are reached by L^1-convergence. Therefore, the formulation in (2.2) is valid even in the L^∞-setting, see also [11].

Due to notational simplicity in the remaining part of the paper we restrict ourselves to the case of $\Omega = (0, \infty)$. Nevertheless, the presented results can be transferred in a straightforward manner to general intervals $(\mathfrak{a}, \mathfrak{b})$.

2.2. General and structural properties of solutions to the IBVP. In this section we collect some properties of solutions y to the IBVP (1.2). We will work under 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 nonnegative and satisfies $g \in C\left(\Omega_T; C^{0,1}_{\text{loc}}(\mathbb{R}\times\mathbb{R}^m)\right) \cap C^1\left([0,T]; C^1_{\text{loc}}\left(\Omega\times\mathbb{R}\times\mathbb{R}^m\right)\right)$ and for all $M_u > 0$ there exist constants $C_1, C_2 > 0$ such that for all $(t,x,y,u_1) \in [0,T]\times\mathbb{R}^2\times[-M_u, M_u]^m$ holds

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

(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 $U_1 := L^1_{\text{loc}}(\mathbb{R}) \times L^1(0,T) \times L^1_{\text{loc}}([0,T]\times\mathbb{R})^m$.

We will consider the source term g and its control u_1 on the whole real numbers, not only on $\Omega = \mathbb{R}^+$. This is for technical reasons and does not affect the solution which still depends on the restriction of g and u_1 to the spatial domain Ω.

For the major part of this section we will work under the weaker assumption (A1$_{\text{loc}}$) instead of (A1):

(A1$_{\text{loc}}$) (A1) holds with the exception that g has only to be nonnegative in a neighborhood $(-\varepsilon, \varepsilon)$ of the left boundary.

Actually, the results of §3 can also be shown to hold under (A1$_{\text{loc}}$) instead of (A1), but since this leads to some notational and technical inconveniences, we will switch to (A1).

Under the above assumptions, we get the following properties of a solution to (1.2), cf. [3, 11, 24].

Proposition 2.1 (Existence and uniqueness for IBVPs). *Let* ($A1_{\rm loc}$) *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* (1.2) *on* $\Omega = (0,\infty)$. *After a possible modification on a set of measure zero it even holds that* $y \in C([0,T]; L^1_{\rm loc}(\Omega))$. *Moreover, there are constants* $M_y, L_y > 0$ *such that for every* $u, \hat{u} \in U_{\rm ad}$ *and all* $t \in [0,T]$ *the following estimates hold:*

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

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

Remark 2.2. *Under the stronger assumptions* $u_0 \in BV_{\rm loc}(\Omega)$ *and* $u_B \in BV([0,T])$, (1.2) *admits a solution satisfying* $y \in BV([0,T] \times [0,R])$ *for all* $R > 0$ *(cf. [20, 3]).*

The proofs of the main results strongly rely on the theory of generalized characteristics from [13], which will be considered in the remaining part of this section. We will assume that in addition to (A1)-(A2), ($A1_{\rm loc}$)-(A2), respectively, the following assumption holds.

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

Furthermore we will only consider $(u_0, u_1) \in \hat{U}_{\rm ad}$, where

$$\hat{U}_{\rm ad} := \{u \in U_{\rm ad} \ : \ \|u_1\|_{L^\infty(0,T;C^1(\Omega_T)^m)} \le M_u\},$$

$u_0 \in BV_{\rm loc}(\Omega)$ 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 $0 < t_1 < \dots < t_{n_t} < T$ for some $n_t \in \mathbb{N}$.

Using the collected properties, we conclude that for an entropy solution $y \in L^\infty(\Omega_T) \cap C([0,T]; L^1_{\rm loc}(\Omega))$ and all $(t,x) \in (0,T] \times \Omega$ the one-sided limits $y(t,x-)$ and $y(t,x+)$ exist and satisfy $y(t,x-) \ge y(t,x+)$. We chose a pointwise defined representative of $y \in C([0,T]; L^1_{\rm loc}(\Omega))$ and identify $y(t,x)$ with one of the limits $y(t,x-)$ or $y(t,x+)$.

We recall the definition of generalized characteristics in the sense of [13].

Definition 2.3 (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*

$$\dot{\xi}(t) \in [f'(y(t,\xi(t)+)), f'(y(t,\xi(t)-))], \quad a.e. \ on \ [\alpha,\beta]. \tag{2.3}$$

The generalized characteristic is called genuine *if the lower and upper bound in* (2.3) *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

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

Proposition 2.1 yields, that y is bounded in $L^\infty(\Omega_T)$ and hence the maximum speed of a generalized characteristic is bounded, too. Therefore, characteristics either exist for the whole time period $[0, T]$ or leave the spatial domain at some point $(\theta, \xi(\theta)) \in [0, T] \times \partial\Omega$. Moreover it can be shown [13] that (2.3) 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.

In [13] the theory of generalized characteristics is used to exploit structural properties of BV-solutions of conservation laws which are essential for the analysis in the present paper.

Proposition 2.4 (Structure of BV-solutions to the IVP). *Let* (A1)-(A3) *hold. Consider an entropy solution $y = y(u)$ of the Cauchy problem* (1.2a)-(1.2b) *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:

(i) *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 almost all $t \in (0, \bar{t})$.*

(ii) *if ξ is genuine, i.e. $y(t, \xi(t)-) = y(t, \xi(t)+)$, $t \in (0, \bar{t})$, then it satisfies*

$$\xi(t) = \zeta(t), \quad t \in [0, \bar{t}], \qquad y(t, \xi(t)) = v(t), \quad t \in (0, \bar{t}),$$
$$u_0(\xi(0)-) \leq v(0) \leq u_0(\xi(0)+), \qquad y(\bar{t}, \xi(\bar{t})-) \geq v(\bar{t}) \geq y(\bar{t}, \xi(\bar{t})+),$$

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

$$\dot{\zeta}(t) = f'(v(t)), \tag{2.4a}$$
$$\dot{v}(t) = g(t, \zeta(t), v(t), u_1(t, \zeta(t))). \tag{2.4b}$$

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

$$(\zeta, v)(\bar{t}) = (\bar{x}, y(\bar{t}, \bar{x}\pm)). \tag{2.4c}$$

The above proposition treats only the pure IVP case. In [25] Perrollaz extends the result to IBVPs. For characteristics that stay inside the spatial domain for the whole considered time interval the assertions of Proposition 2.4 still hold. The following proposition collects the results of [25, §3], where characteristics that enter or leave the domain Ω are discussed. Perrollaz emphasizes that for $\Omega = (0, \infty)$ it is important to have a nonnegative source term in order to prove the third part of Proposition 2.5. This property implies the convexity of genuine characteristics and hence some nondegeneracy of the characteristics near the boundary. A reinspection of the proof shows that the nondegeneracy is preserved if g is nonnegative in a neighborhood of the boundary. For a spatial domain $(-\infty, 0)$ the condition on the source term becomes a local nonpositivity condition.

Proposition 2.5 (Structure of BV-solutions at the boundary). *Let* ($A1_{\text{loc}}$), (A2) *and* (A3) *hold. Consider an entropy solution $y = y(u)$ of the mixed initial-boundary value problem* (1.2) *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 $u_B \in PC^1([0, T]; t_1, \ldots, t_{n_t})$, $f'(u_B) \geq 0$. Then the following holds:*

(i) *Consider* $\theta \in (0,T)$ *with* $f'(y(\theta,0+)) < 0$, *then there exists a genuine backward characteristic* ξ *through* $(\theta,0)$ *with* $\dot{\xi}(\theta) = f'(y(\theta,0+))$.

(ii) *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* (2.4a)-(2.4b) *associated to* ξ *by Proposition 2.4 on every* $[\tilde{t},\bar{t}] \subset (\theta,\bar{t}]$. *Then with* $v(\theta) := \lim_{t\searrow\theta} v(t)$ *it holds*

$$u_B(\theta+) \le v(\theta) \le u_B(\theta-).$$

(iii) *Let* ξ *be a genuine 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*

$$f'(v(\theta)) \le 0 \quad \text{and} \quad f(v(\theta)) \ge f(u_B(\theta-)),$$

where $v(\theta) := \lim_{t\nearrow\theta} v(t)$.

The following Lemma on the differentiability of the solution operator of the characteristic equation (2.4) is a consequence of a result on ordinary differential equations (cf. [30, Proposition 3.4.5, Lemma 3.4.6] or [27, §5.6]). Together with Propositions 2.4 and 2.5 this Lemma can be used to shed light on the local differentiability of a solution y to the I(B)VP, as we will see in §4.2. The derivatives occurring in Lemma 2.6 can be expressed by means of the solution $(\delta\zeta,\delta v)(\cdot;\theta,z,w,u_1;\delta\theta,\delta z,\delta w,\delta u_1)$ of the *linearized characteristic equation*:

$$\dot{\delta\zeta}(t) = f''(v(t))\delta v(t), \tag{2.5a}$$

$$\dot{\delta v}(t) = g_x(:)\delta\zeta(t) + g_w(:)\delta v(t) + g_{u_1}(:)u_{1x}(\dot{:})\delta\zeta(t) + g_{u_1}(:)\delta u_1(\dot{:}), \tag{2.5b}$$

$$(\delta\zeta,\delta v)(\theta) = (\delta z - f'(w)\delta\theta, \delta w - g(\theta,z,w,u_1(\theta,z))\delta\theta), \tag{2.5c}$$

where $(:) = (t,\zeta(t),v(t),u_1^i(\dot{:}))$ and $(\dot{:}) = (t,\zeta(t))$.

Lemma 2.6. *Let* (A1$_{\text{loc}}$) *and* (A3) *hold and denote for* $(\theta,z,w,u_1) \in [0,T]\times\mathbb{R}^2\times C([0,T];C^1(\mathbb{R})^m)$ *by* $(\zeta,v)(\cdot;\theta,z,w,u_1)$ *the solution of* (2.4a)-(2.4b) *for initial data*

$$(\zeta,v)(\theta;\theta,z,w,u_1) = (z,w).$$

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

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

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. *The derivative is given in terms of the solution of the linearized characteristic equation* (2.5) *by*

$$d_{(\theta,z,w,u_1)}(\zeta,w)\cdot(\delta\theta,\delta z,\delta w,\delta u_1) = (\delta\zeta,\delta v)(\cdot;\theta,z,w,u_1;\delta\theta,\delta z,\delta w,\delta u_1).$$

Finally, for any closed $S \subset [0,T] \times \mathbb{R}$, *any fixed* $(\bar{\theta}, \bar{z}) \in [0,T] \times \mathbb{R}$ *and any bounded intervals* $\mathcal{T} \subseteq [0,T]$, $\hat{I}$ *the mappings*

$$(\theta, u_B, u_1) \in C(S;\mathcal{T}) \times C^1(\mathcal{T}) \times C(0,T;C^1(\mathbb{R})^m) \longmapsto (\zeta, v)(\cdot_t, \theta, \bar{z}, u_B(\theta), u_1) \in C(S)^2$$

$$(z, u_0, u_1) \in C(S;\hat{I}) \times C^1(\hat{I}) \times C(0,T;C^1(\mathbb{R})^m) \longmapsto (\zeta, v)(\cdot_t, \bar{\theta}, z, u_0(z), u_1) \in C(S)^2$$

are continuously Fréchet-differentiable.

From this point we will assume (A1) to hold instead of ($\mathrm{A1_{loc}}$). In the case of IBVPs, for given $\bar{t} \in (0,T)$, the point

$$\theta^{\Delta} := \sup\{t \in [0,\bar{t}] \ : \ f'(y(t,0+;u)) < 0\} \tag{2.6}$$

and the maximal backward characteristic ξ^{Δ} through $(\theta^{\Delta}, 0)$, ensured by Proposition 2.5 to exist if $\theta^{\Delta} \in (0,T)$, is of special interest. We collect some properties of it.

LEMMA 2.7. *Let the assumption of Proposition 2.5 hold with the modification that also* (A1) *is satisfied and let* $f'(u_B) \geq \alpha > 0$. *Consider* $\theta^{\Delta} \in (0,T)$ *from* (2.6), *then the following holds:*

(i) *The characteristic* ξ^{Δ} *exists on the whole interval* $[0,\theta^{\Delta}]$ *and meets* $\{t = 0\}$ *in a point* z^{Δ}.

(ii) *There is no genuine backward characteristic* ξ *through* $(\bar{t}, \bar{x})$ *with* $\bar{t} > \theta^{\Delta}$ *and* $\bar{x} \in \Omega$ *that intersects the domain*

$$D^- := \{(t,x) \in \Omega_{\bar{t}} \ : \ t \in (0,\theta^{\Delta}), \ x \in [0, \xi^{\Delta}(t)]\}. \tag{2.7}$$

(iii) *The point* $(\theta^{\Delta}, 0)$ *is a shock generating point, i.e.* $\theta^{\Delta} = t_{j^{\Delta}}$ *for some* $j^{\Delta} \in \{0, \ldots, n_t\}$ *and* $u_B(\theta^{\Delta}+) > u_B(\theta^{\Delta}-)$.

Proof. The first assertion is clear, since by the sign condition on g the genuine characteristic ξ^{Δ} is convex and, therefore, has negative speed.

The second statement is a consequence of the first and the fact that genuine characteristics may only intersect each other at their endpoints.

Assume that the last assertion does not hold. Then there is a (maximal) genuine forward characteristic ξ, that stays genuine up to some time $\tilde{t} > \theta^{\Delta}$. Moreover, there is a monotone decreasing sequence $(x_k)_{k \in \mathbb{N}}$ with limit $\xi(\tilde{t})$. Since ξ is genuine, $y(\tilde{t}, x_k)$ converges to $y(\tilde{t}, \xi(\tilde{t}))$. Denote by $(\tilde{\zeta}, \tilde{v}) := (\zeta, v)(\cdot; \tilde{t}, \xi(\tilde{t}), y(\tilde{t}, \xi(\tilde{t})), u_1)$ the backward solution of (2.4) associated with ξ and by $(\zeta_k, v_k) := (\zeta, v)(\cdot; \tilde{t}, x_k, y(\tilde{t}, x_k+), u_1)$ the respective solutions associated with the maximal backward characteristics ξ_k through $(\tilde{t}, x_k)$. Lemma 2.6 yields that $(\zeta_k, v_k)(\theta^{\Delta}) \searrow (0, \tilde{v}(\theta^{\Delta}))$, but since ξ was chosen maximal, $\zeta_k(\theta^{\Delta}) > 0$ holds for all $k \in \mathbb{N}$. The convergence of v_k yields $\dot{\zeta}_k(\theta^{\Delta}) = f'(v_k(\theta^{\Delta})) \to f'(\tilde{v}(\theta^{\Delta})) = \dot{\xi}(\theta^{\Delta}) \geq \alpha > 0$. By the continuity of $\dot{\zeta}_k$ this ensures the existence of an $\hat{t} < \theta^{\Delta}$, such that for k sufficiently large $\dot{\zeta}_k(t) > 0$ for all $t \in [\hat{t}, \theta^{\Delta}]$, which in turn implies $\zeta_k(\hat{t}) < \zeta_k(\theta^{\Delta})$. If k is large enough to ensure $\zeta_k(\theta^{\Delta}) < \xi^{\Delta}(\hat{t})$, this implies $\zeta_k(\hat{t}) < \xi^{\Delta}(\hat{t})$. This means ζ_k intersects the domain D^-, which is a contradiction to the second statement, since ζ_k coincides with the genuine characteristic ξ_k through $(\tilde{t}, x_k) \in (\theta^{\Delta}, T] \times \Omega$. □

REMARK 2.8. *In the setting of* ($\mathrm{A1_{loc}}$) *there may be multiple transition points* $\theta_1^{\Delta}, \ldots, \theta_{n_\triangle}^{\Delta}$ *between inflow and outflow parts of the boundary. But one can show, that those points cannot be arbitrarily close to each other. The set* D^- *of points that do not lie on a generalized backward characteristic through a point* $(\bar{t}, \bar{x})$, *needs no longer be connected.*

3. Shift-differentiability. In this section we give the main result of this paper, that is the shift-differentiability of the control-to-state mapping for (1.2). It is well-known, that entropy solutions develop shocks after a finite time, even for smooth data. This leads to complications for the treatment of optimal control problems concerning such discontinuous solutions. Since the shock positions depend on the control, classical notions of differentiability of the control-to-state mapping $u \mapsto y(u)$ do only hold in very weak topologies, that are not strong enough to imply the Fréchet-differentiability of the reduced objective functional. The following example illustrates the situation by means of a Riemann problem.

EXAMPLE 1. *Consider the parametrized Cauchy problem*

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

Then a representative of the entropy solution is given by

$$y^\varepsilon(t,x) = \varepsilon + \operatorname{sgn}(\varepsilon t - x).$$

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, differentiability does only hold in the weak topology of the measure space* $\mathcal{M}([a,b])$.

3.1. Definitions and preliminary work. To overcome the above mentioned lack of differentiability, in [8] and [30, 31] the authors introduced a variational calculus, that addresses the reason of the non-differentiability, that is the shift of shock positions in the solution resulting from a variation of the control. Instead of only considering additive variations (e.g. in L^1), the so-called shift-variations also allow for horizontal shifts of discontinuities. We recall the definitions of the notions of shift-variations and shift-differentiability.

DEFINITION 3.1 (Shift-variations, shift-differentiability).

(i) *Let* $a < b$ *and* $v \in BV([a,b])$. *For* $a < x_1 < x_2 < \cdots < x_{n_x} < b$ *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*

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

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

(ii) *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* $a < x_1 < x_2 < \cdots < x_{n_x} < b$ *and* $D_s v(\bar u) \in \mathcal{L}(U, L^r([a,b])\times\mathbb{R}^{n_x})$ *for some* $r \in (1,\infty]$, *such that for* $\delta u \in U$, $(\delta v,\delta x) := D_s v(\bar u)\cdot\delta u$ *holds*

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

This variational concept is indeed strong enough to directly imply the Fréchet-differentiability of tracking type functionals as in (1.1) (see [30, Lemma 3.2.3]) as long as y_d and $y(\bar t,\cdot)$ do not share discontinuities on $[a,b]$. The derivative is given by

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

with

$$\bar{\psi}_y(x) := \int_0^1 \psi_y\left(y(\bar{t}, x+; u)\right) + \tau[y(\bar{t}, x; u)], y_d(x+) + \tau[y_d(x)]\right) \,\mathrm{d}\tau. \tag{3.1}$$

The proof of Theorem 3.3 and the formula for the gradient of the reduced objective function in Theorem 3.6 are based on an appropriately defined adjoint state. Formally the adjoint equation reads as follows

$$p_t + f'(y)p_x = -g_y(\cdot, y, u_1)p, \qquad \text{on } \Omega_{\bar{t}}, \tag{3.2a}$$

$$p(\bar{t}, \cdot) = p^{\bar{t}}, \qquad \text{on } \Omega. \tag{3.2b}$$

As already discussed in [30] for the Cauchy problem, the classical adjoint calculus is not applicable in the present context. The equation (3.2) is a linear transport equation with discontinuous coefficients, since y may contain shocks, which makes the analysis more involved. Nevertheless, for $\Omega = \mathbb{R}$, $g \equiv 0$ and Lipschitz continuous end data $p^{\bar{t}}$, Bouchut and James [5] give a definition of a *reversible solution* for (3.2), which satisfies a crucial duality relation.

In [30, 32] it was shown that the reversible solution of (3.2) is exactly the solution along the generalized characteristics of y. Using this characterization, the notion could be extended to more general source terms g and discontinuous end data. In the case of a bounded domain, here $\Omega = (0, \infty)$, the above 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) value is prescribed, cf. Lemma 2.7. The following definition treats this issue by setting the adjoint state to 0 on D^-.

DEFINITION 3.2 (Adjoint state). *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 (3.2) for $\Omega = (0,\infty)$ is characterized by the following requirements:*

(i) *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

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

(ii) *For every $(t, x) \in D^-$ there holds $p(t, x) = 0$ with D^- from (2.7).*

3.2. Shift-differentiability of solutions to the IBVP. We now state our main results for the initial-boundary value problem on $\Omega = (0, \infty)$, the results for general intervals are quite similar. We show the shift-differentiability of the control-to-state mapping, from this we derive the Fréchet-differentiability of the reduced objective function and finally give a formula for its gradient.

As one can see from the formulation of the boundary condition (2.2a), the solution remains unchanged if one replaces u_B by $\max(u_B, f'^{-1}(0))$. This motivates to only consider boundary data satisfying $u_B \geq f'^{-1}(0)$ and thus the introduction of the space

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

for given $0 < t_1 < t_2 < \cdots < t_{n_t}$.

Let us fix a current control $u = (u_0, u_B, u_1)$ where $u_B \in U_B^\alpha$ for some small $\alpha > 0$, $u_0 \in PC^1(\Omega; x_1, \ldots, x_{n_x})$ for some $0 < x_1 < x_2 < \cdots < x_{n_x}$ and $u_1 \in C([0,T]; C^1(\mathbb{R})^m)$. We analyze the shift-differentiability of $\delta u \mapsto y(\bar{t}, \cdot; u + \delta u)$ w.r.t. the perturbation 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 shock creating discontinuities, but not of rarefaction centers. For this purpose we define

$$\begin{aligned} \mathbf{S}_{(x_i)} &:= \{ s \in \mathbb{R}^{n_x} \ : \ u_0(x_i-) < u_0(x_i+) \ \Rightarrow s_i = 0, \ i = 1, \ldots, n_x \}, \\ \mathbf{S}_{(t_j)} &:= \{ s \in \mathbb{R}^{n_t} \ : \ u_B(t_j-) > u_B(t_j+) \ \Rightarrow s_j = 0, \ j = 1, \ldots, n_t \} \end{aligned}$$

and consider variations in

$$\begin{aligned} W := {} & PC^1(\Omega; x_1, \ldots, x_{n_x}) \times \mathbf{S}_{(x_i)} \\ & \times PC^1([0,T]; t_1, \ldots, t_{n_t}) \times \mathbf{S}_{(t_j)} \times C([0,T]; C^1(\mathbb{R})^m). \end{aligned} \tag{3.3}$$

Under the nondegeneracy condition on the shocks given at the end of §4.1, we get the following result.

Theorem 3.3 (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 $0 < x_1 < x_2 < \cdots < x_{n_x}$, $0 < t_1 < t_2 < \cdots < t_{n_t}$ $u_0 \in PC^1(\Omega; x_1, \ldots, x_{n_x})$, $u_B \in U_B^\alpha$ 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 $y = y(u) \in L^\infty(\Omega_T) \cap C([0,T]; L^1_{\mathrm{loc}}(\Omega))$ the entropy solution of the initial-boundary value problem* (1.2) *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 nondegenerated according to Definition 4.1. Further assume that the pointwise defined boundary trace $y(\cdot, 0+; u)$ satisfies*

$$\operatorname*{essinf}_{t \, : \, u_B(t) \neq y(t, 0+; u)} |f(u_B(t)) - f(y(t, 0+; u))| > 0$$

and that the transition point θ^Δ from (2.6) *is nondegenerated according to Definition 4.12.*

For W from (3.3) *we consider the mapping*

$$\begin{aligned} & (\delta u_0, \delta x, \delta u_B, \delta t, \delta u_1) \in W \longmapsto \\ & \qquad y\left(\bar{t}, \cdot; u_0 + S_{u_0}^{(x_i)}(\delta u_0, \delta x), u_B + S_{u_B}^{(t_j)}(\delta u_B, \delta t), u_1 + \delta u_1\right) \in L^1(a,b). \end{aligned} \tag{3.4}$$

If $(x_i), (t_j)$ are genuine discontinuities of u_0 and u_B, i.e. $u_0(x_i-) \neq u_0(x_i+)$ and $u_B(t_j-) \neq u_B(t_j+)$, respectively, the mapping (3.4) *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 3.4. *If u_0 or u_B are continuous at some x_i or t_j, respectively, similarly to the second assertion of [30, Theorem 3.3.2], the shift-differentiability of* (3.4) *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 of Theorem 3.3 is presented in §4. This requires a proper analysis of the solution y in small neighborhoods of different types of generalized backward characteristics.

The following corollary is a simple consequence of the above theorem and [30, Lemma 3.2.3].

Corollary 3.5 (Fréchet-differentiability of the reduced objective). *Let the assumptions of Theorem 3.3 hold and consider J as defined 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.*

In the following theorem we give a representation of the gradient of the reduced objective function based on the appropriate notion of an adjoint state from Definition 3.2.

Theorem 3.6 (Formula for the reduced gradient). *Let the assumptions of Corollary 3.5 hold and let the terminal data $p^{\bar{t}}$ in* (3.2) *be given by $\bar{\psi}_y$ defined in* (3.1). *Then there exists an adjoint state p according to Definition 3.2 as the reversible solution of the adjoint equation* (3.2)*, 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).

Using the notation of Lemma 2.7, the derivative of the reduced functional $\delta u \in W \mapsto \hat{J}(\delta u) = J(y(u+\delta u))$ is given by

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

The proof will be given at the end of §4.

Remark 3.7. *One can show, that Theorem 3.3 and 3.6 hold also true, if assumption* (A1) *is replaced by* ($\mathrm{A1}_{\mathrm{loc}}$).

4. Proofs of the main results. The proof of Theorem 3.3 follows the strategy of [31]: We start by classifying various types of continuity points $\bar{x} \in [a,b]$ of $y(\bar{t},\cdot)$. Afterwards, we show the Fréchet-differentiability of $y(\bar{t},\cdot)$ w.r.t. the control in neighborhoods of those points. This is done by combining the results of Propositions 2.4 and 2.5 and Lemma 2.6. The shock points of $y(\bar{t},\cdot)$ will be classified in a similar fashion. The proof of differentiability of the shock positions is obtained by an adjoint argument based on the notion of an adjoint state according to Definition 3.2.

For the whole section we will work in the setting of Theorem 3.3.

4.1. Classification. For continuity points $\bar{x}$ of $y(\bar{t},\cdot;)$ we denote the unique backward characteristic by $\bar{\xi}$. By Proposition 2.4 the characteristic $\bar{\xi}$ coincides with the solution $\zeta(\cdot;\bar{t},\bar{x},y(\bar{t},\bar{x}),u_1)$ of (2.4). We emphasize that $\bar{\xi}$ may not "touch" the boundary at $\{x=0\}$ and return to Ω, since $f'(u_B) \geq \alpha > 0$, but either leaves the domain at some time $\bar{\theta}$ or stays inside and ends in a point $\bar{z}$ at $t=0$. The case for continuity points has already been categorized and analyzed in [30, 31]. We therefore only briefly recall the classifications and will use the differentiability results of [30,

§3.3, §4]. We denote by $\bar{w} := v(0;\bar{t},\bar{x},y(\bar{t},\bar{x}),u_1)$ the value of v-part of the solution of (2.4) associated with $\bar{\xi}$ by Proposition 2.4.

Case C**:** *$\bar{z} \neq x_l$ for $l = 1,\dots,n_x$.*
There exists an interval J with $\bar{z} \in J$ and $u_I|_J \in C^1(J)$ and

$$\frac{d}{dz}\zeta(t;0,z,u_I(z),u_1)|_{z=\bar{z}} \geq 0, \quad t \in [0,\bar{t}]. \tag{4.1}$$

We say that $\bar{x}$ is of class C^c if even

$$\frac{d}{dz}\zeta(t;0,z,u_I(z),u_1)|_{z=\bar{z}} \geq \beta > 0, \quad t \in [0,\bar{t}]. \tag{4.2}$$

As shown in [31] (4.2) holds if $(\bar{t},\bar{x})$ is no shock generation point.

Case CB**:** *$\bar{z} = x_l$ for some $l \in \{1,\dots,n_x\}$ and $u_I(x_l-) = u_I(x_l+)$.*
By the same arguments the one-sided derivatives satisfy (4.1). If even the one-sided version of (4.2) holds, we call $\bar{x}$ of case CB^c.

Case R**:** *$\bar{z} = x_l$ for some $l \in \{1,\dots,n_x\}$ and $\bar{w} \in (u_I(x_l-),u_I(x_l+))$.*
In this case we have

$$\frac{d}{dw}\zeta(t;0,\bar{z},w,u_1)|_{w=\bar{w}} \geq 0, \quad t \in [0,\bar{t}]. \tag{4.3}$$

The R^c-Case is characterized by the stronger inequality

$$\frac{d}{dw}\zeta(t;0,\bar{z},w,u_1)|_{w=\bar{w}} \geq \beta t > 0, \quad t \in (0,\bar{t}]. \tag{4.4}$$

which, again is ensured by the requirement that no shock is generated at $(\bar{t},\bar{x})$.

Case RB**:** *$\bar{z} = x_l$ for some $l \in \{1,\dots,n_x\}$, $u_I(x_l-) < u_I(x_l+)$ and $\bar{w} \in \{u_I(x_l+),u_I(x_l-)\}$.*
The point $(\bar{t},\bar{x})$ lies on the left or right boundary of a rarefaction wave. The one-sided derivatives satisfy (4.1) and (4.3), respectively. If even (4.2) and (4.4) are satisfied, $\bar{x}$ is of class RB^c.

From now on we will only consider continuity points whose backward characteristics leave the spatial domain at some time $\bar{\theta} \in [0,\bar{t})$. Again, we denote by $\bar{w} := v(\bar{\theta};\bar{t},\bar{x},y(\bar{t},\bar{x}),u_1)$ the value of the corresponding v-part of the solution of the characteristic equation at the characteristic's end point. We distinguish the following cases.

Case C_B**:** *$\bar{\theta} \neq t_l$ for $l = 1,\dots,n_t$.*
There exists an interval $\mathcal{T} \subset (0,\bar{t})$ with $\bar{\theta} \in \mathcal{T}$ and $u_B|_{\mathcal{T}} \in C^1(\mathcal{T})$ and

$$\frac{d}{d\theta}\zeta(t;\theta,0,u_B(\theta),u_1)|_{\theta=\bar{\theta}} \leq 0, \quad \bar{\theta} \leq t \leq \bar{t}. \tag{4.5}$$

We say that $\bar{x}$ is of class C_B^c if even

$$\frac{d}{d\theta}\zeta(t;\theta,0,u_B(\theta),u_1)|_{\theta=\bar{\theta}} \leq -\beta < 0, \quad \bar{\theta} \leq t \leq \bar{t}. \tag{4.6}$$

Similar to the C^c-case one can show that (4.6) holds if $(\bar{t},\bar{x})$ is no shock generation point.

Case CB_B**:** *$\bar{\theta} = t_l$ for some $l \in \{1,\dots,n_t\}$ and $u_B(t_l-) = u_B(t_l+)$.*
By the same arguments the one-sided derivatives satisfy (4.5). If even the one-sided version of (4.6) holds, we call $\bar{x}$ of case CB_B^c.

Case R_B**:** *$\bar\theta = t_l$ for some $l \in \{1,\dots,n_t\}$ and $\bar w \in (u_B(t_l+), u_B(t_l-))$.*
By the same arguments as for the R-Case there holds

$$\frac{d}{dw}\zeta(t;\bar\theta,0,w,u_1)|_{w=\bar w} \geq 0, \quad \bar\theta \leq t \leq \bar t. \tag{4.7}$$

The R_B^c-Case is characterized by the stronger inequality

$$\frac{d}{dw}\zeta(t;\bar\theta,0,w,u_1)|_{w=\bar w} \geq \beta(t-\bar\theta) > 0, \quad \bar\theta < t \leq \bar t, \tag{4.8}$$

which, again is ensured by the requirement that no shock is generated at $(\bar t,\bar x)$.

Case RB_B**:** *$\bar\theta = t_l$ for some $l \in \{1,\dots,n_t\}$, $u_B(t_l-) > u_B(t_l+)$ and $\bar w \in \{u_B(t_l+), u_B(t_l-)\}$.*
The point $(\bar t,\bar x)$ lies on the left or right boundary of a rarefaction wave. The one-sided derivatives satisfy (4.5) and (4.7), respectively. If even (4.6) and (4.8) are satisfied, $\bar x$ is of class RB_B^c.

Case C_0**:** *$\bar\xi(0) = 0$ and $u_B(0+) = u_0(0+)$.*
The right-hand derivative $\frac{d}{d\theta}\zeta(t;\theta,0,u_B(\theta),u_1)|_{\theta=0+}$ meets the C_B-condition (4.5) and $\frac{d}{dz}\zeta(t;0,z,u_0(z),u_1)|_{z=0}$ satisfies (4.1). If even (4.6) and (4.2) are satisfied, we call $\bar x$ of Case C_0^c.

Case R_0**:** *$\bar\xi(0) = 0$ and $\bar w \in (u_B(0+), u_0(0+))$.*
That is a R_B-point, satisfying (4.7) at $\bar\theta = 0$. Analogously, we call $\bar x$ of class R_0^c if (4.8) is met.

Case $RB_{B,0}$**:** *$\bar\xi(0) = 0$, $u_B(0+) < u_0(0+)$, and $\bar w = u_B(0+)$).*
That is a RB_B-point with $\bar\theta = 0$ at the left boundary of a rarefaction wave with center at $(0,0)$ with the one-sided derivatives satisfying (4.5) and (4.7).

Case RB_0**:** *$\bar\xi(0) = 0$, $u_B(0+) < u_0(0+)$, and $\bar w = u_0(0+)$).*
That is a R_B-point from [31] with $\bar z = 0$ at the right boundary of a rarefaction wave with center at $(0,0)$ with the one-sided derivatives satisfying (4.1) and (4.7).

Although we have to classify cases C_0, R_0, $RB_{0,B}$ and RB_0, that describe the different possible situation for backward characteristics ending at the boundary at time $\bar\theta = 0$, they do not really need special treatment in the analysis of the problem since they are only special cases of other classes.

The shock points are categorized by the classes of their minimal and maximal characteristics $\bar\xi_l$ and $\bar\xi_r$. Using these classifications we are now able to define the nondegeneracycondition for shock points $\bar x$.

Definition 4.1 (Nondegeneracy of shock points). *A point $\bar x$ of discontinuity of $y(\bar t,\cdot;u)$ is called nondegenerated, if it is no shock interaction point and is of class X_lX_r with $X_l, X_r \in \{C^c, C_B^c, R^c, R_B^c, R_0^c\}$.*

4.2. Differentiability at continuity points. Let $\bar x$ be a continuity point of $y(\bar t,\cdot)$ of Class C_B^c, i.e. (4.6) is satisfied. Then by continuity we can find $\theta_r < \bar\theta < \theta_l$ and $\kappa > 0$, such that (after a possible reduction of $\mathcal{T}$ and β) holds

$$\frac{d}{d\theta}\zeta(t;\theta,0,u_B(\theta),u_1) \leq -\beta < 0, \quad \text{for all } (t,\theta) \in \mathbf{T}_{\bar t}, \tag{4.9}$$

where for every $s > 0$ the set $\mathbf{T}_s$ is defined by

$$\mathbf{T}_s := \{(t,\theta) \in [0,s] \times \mathcal{T} \;:\; t \geq \theta\} \quad \text{with} \quad \mathcal{T} = (\theta_r - \kappa, \theta_l + \kappa).$$

Lemma 4.2. *Let $u_B \in PC^1([0,T];t_1,\dots,t_{n_t})$, $u_1 \in C^1(\Omega_T)^m$ and let (4.9) hold for some $\beta,\kappa > 0$. Then the following holds:*

(i) *There exist* $\tau > 0$ *and a neighborhood* $V \subset C^1(\mathcal{T}) \times C^1(\Omega_T)^m$ *of* $(u_B|_{\mathcal{T}}, u_1)$ *such that*

$$\frac{d}{d\theta}\zeta(t;\theta,0,\hat{u}_B(\theta),\hat{u}_1) \le -\frac{\beta}{2} < 0, \qquad \forall (t,\theta) \in \mathbf{T}_{\bar{t}+\tau}, \quad \forall (\hat{u}_B,\hat{u}_1) \in V.$$

(ii) *Consider* $\hat{u} \in V$ *and a point* $(t,x) \in S = S(\tau)$, *where*

$$S(\tau) := \{(t,x) \in [\theta_r, \bar{t}+\tau] \times \mathbb{R} \;:\; x \in [\xi_l(\max(t,\theta_l)), \xi_r(t)]\}$$

and $\xi_{l/r}(t) := \zeta(t;\theta_{l/r},0,u_B(\theta_{l/r}),u_1)$, $t \in [\theta_{l/r}, \bar{t}+\tau]$. *Then the equation*

$$x = \zeta(t;\theta,0,\hat{u}_B(\theta),\hat{u}_1)$$

is uniquely solvable w.r.t. θ *on* $\mathcal{T}$ *from* (4.9) *with solution* $\theta = \Theta(t,x,\hat{u}_B,\hat{u}_1)$. *Moreover, let* $Y_B(t,x,\hat{u}_B,\hat{u}_1)$ *be defined by*

$$Y_B(t,x,\hat{u}_B,\hat{u}_1) := v(t;\Theta(t,x,\hat{u}_B,\hat{u}_1),0,\hat{u}_B(\Theta(t,x,\hat{u}_B,\hat{u}_1)),\hat{u}_1).$$

Then

(iii) $\Theta(\cdot,\hat{u}_B,\hat{u}_1), Y_B(\cdot,\hat{u}_B,\hat{u}_1) \in C^{0,1}(S)$.

(iv) *The mapping*

$$(x,\hat{u}_B,\hat{u}_1) \in (\max(\xi_l(t),0),\xi_r(t)) \times V \longmapsto (\Theta,Y_B)(t,x,\hat{u}_B,\hat{u}_1), \qquad t \in [\theta_r, \bar{t}+\tau)$$

is continuously Fréchet-differentiable with derivatives

$$\begin{aligned} &d_{(x,u_B,u_1)}\Theta(t,x,\hat{u}_B,\hat{u}_1)\cdot(\delta x,\delta u_B,\delta u_1) \\ &\qquad = \frac{\delta x - \delta\zeta(t;\theta,0,\hat{u}_B(\theta),\hat{u}_1;0,0,\delta u_B(\theta),\delta u_1)}{\delta\zeta(t;\theta,0,\hat{u}_B(\theta),\hat{u}_1;1,0,\hat{u}_B'(\theta),0)} \\ &d_{(x,u_B,u_1)}Y_B(t,x,\hat{u}_B,\hat{u}_1)\cdot(\delta x,\delta u_B,\delta u_1) \\ &= \delta v(t;\theta,0,\hat{u}_B(\theta),\hat{u}_1;1,0,\hat{u}_B'(\theta),0)\cdot d_{(x,u_B,u_1)}\Theta(t,x,\hat{u}_B,\hat{u}_1)\cdot(\delta x,\delta u_B,\delta u_1) \\ &\qquad + \delta v(t;\theta,0,\hat{u}_B(\theta),\hat{u}_1;0,0,\delta u_B(\theta),\delta u_1). \end{aligned}$$

(v) *The mapping*

$$(\hat{u}_B,\hat{u}_1) \in V \longmapsto (\Theta,Y_B)(\cdot,\hat{u}_B,\hat{u}_1) \in C(S)^2$$

is continuously Fréchet-differentiable with derivative

$$d_{(u_B,u_1)}(\Theta,Y_B)(\cdot,\hat{u}_B,\hat{u}_1)\cdot(\delta u_B,\delta u_1) = d_{(x,u_B,u_1)}(\Theta,Y_B)(\cdot,\hat{u}_B,\hat{u}_1)\cdot(0,\delta u_B,\delta u_1).$$

Proof. The proof is completely analogous to the one of [31, Lemma 4.1]. We use the stability results of Lemma 2.6 and the properties of genuine characteristics obtained by Propositions 2.4 and 2.5. ☐

LEMMA 4.3. *Let* $u = (u_0,u_B,u_1) \in PC^1(\Omega;x_1,\ldots,x_{n_x}) \times PC^1([0,T];t_1,\ldots,t_{n_t}) \times C^1(\Omega_T)^m$ *and let* $\bar{x}$ *be a* C_B^c*-point. Then the following statements are true:*

(i) *There exists a maximal open interval* $I \ni \bar{x}$, *such that* $\{\bar{t}\} \times I$ *contains no point of the shock set and that all backward characteristics intersect* $x = 0$ *in a point* $\theta \neq t_l$.

(ii) $y(\bar{t},\cdot;u)$ *is continuously differentiable on* I.

(iii) *Let* $\hat{I} := (x_l,x_r)$ *be an interval with* $x_l,x_r \in I$. *Denote by* $\xi_{l/r}$ *the genuine backward characteristics through* $(\bar{t},x_{l/r})$ *with endpoints* $\theta_{l/r}$ *at* $x = 0$. *Then there exist* $\kappa,\beta > 0$, *such that* (4.9) *is satisfied.*

(iv) *Let $M_\infty > 0$, then there exist $R > 0$, $\nu > 0$, such that after the possible reduction of τ and V from Lemma 4.2*

$$y(t,x;\hat{u}) = Y_B(t,x,\hat{u}_B,\hat{u}_1) \quad \forall (t,x) \in S,\ \forall \hat{u} \in \hat{V}$$

holds, where

$$\begin{aligned}\hat{V} := \{(\hat{u}_0,\hat{u}_B,\hat{u}_1) \in BV(\Omega) \times BV(0,T) \times C^1(\Omega_T)^m \ :\\ (\hat{u}_B|_{\mathcal{T}}, u_1) \in V,\ \|\hat{u}_B - u_B\|_{\infty,[0,T]\setminus\mathcal{T}} < M_\infty,\ \|\hat{u}_B - u_B\|_{1,[0,T]\setminus\mathcal{T}} < \nu,\\ \|\hat{u}_0 - u_0\|_\infty < M_\infty,\ \|\hat{u}_0 - u_0\|_{1,(0,R)} < \nu\}.\end{aligned}$$

Proof. The proof is analogous to the one of [31, Lemma 4.4]. We, therefore, only sketch the main idea.

(i)-(ii) We use Lemma 4.2 and the fact that $y(\bar{t},\cdot;u)$ has bounded total variation and hence is continuous outside a countable set. Since genuine characteristics cannot escape the stripe S from Lemma 4.2, $y(\bar{t},\cdot;u)$ coincides with $Y(\bar{t},\cdot,u)$ in a neighborhood of $\bar{x}$.

(iii) This statement follows by taking a finite covering of the compact interval $[x_l, x_r]$.

(iv) By the first part of the proof we know, that for the current control u the functions $y(\cdot;u)$ and $Y_B(\cdot,u)$ coincide on S. By reducing V and τ we can construct a slightly wider stripe $\tilde{S}$ with the same properties. By the stability of the solutions of (2.4) obtained by Lemma 2.6 and the L^1-stability from Proposition 2.1 one can show that in between the two stripes there are continuity points of $y(\cdot;\hat{u})$ whose backward characteristics end in $\mathcal{T}$ for all $\hat{u} \in \hat{V}$. Thus, the same holds for all continuity points in the smaller stripe S and so the assertion follows as in the first part of the proof. ☐

Let $\bar{x}$ be a R_B^c-point of $y(\bar{t},\cdot)$, i.e. (4.8) is satisfied. Then by continuity we can find $w_l < \bar{w} < w_r$ and $\kappa > 0$, such that (after a possible reduction of β)

$$\frac{d}{dw}\zeta(t;\bar{\theta},0,w,u_1)|_{w=\bar{w}} \geq \beta(t-\bar{\theta}) > 0, \quad \text{for all } (t,w) \in (\bar{\theta},\bar{t}] \times J_w \tag{4.10}$$

holds, where $J_w := (w_l - \kappa, w_r + \kappa)$.

LEMMA 4.4. *Let $u_B \in PC^1([0,T];t_1,\ldots,t_{n_t})$, $u_1 \in C^1(\Omega_T)^m$ and let (4.10) hold for some $\beta,\kappa > 0$. Then the following holds:*

(i) *There exist $\tau > 0$ and a neighborhood $V_1 \subset C^1(\Omega_T)^m$ of u_1 such that*

$$\frac{d}{dw}\zeta(t;\bar{\theta},0,w,\hat{u}_1) \geq \frac{\beta}{2}(t-\bar{\theta}) > 0, \qquad \forall (t,w) \in (\bar{\theta},\bar{t}+\tau] \times J_w, \quad \forall \hat{u}_1 \in V_1.$$

(ii) *Consider $\hat{u}_1 \in V_1$ and a point $(t,x) \in S = S(\tau)$, where*

$$S(\tau) := \big\{(t,x) \ :\ t \in (\bar{\theta},\bar{t}+\tau],\ x \in [\xi_l(t),\xi_r(t)]\big\}$$

and $\xi_{l/r}(t) := \zeta(t;\bar{\theta},w_{l/r},u_1)$, $t \in (\bar{\theta},\bar{t}+\tau]$. Then the equation

$$x = \zeta(t;\bar{\theta},0,w,\hat{u}_1)$$

is uniquely solvable on J_w with solution $w = W(t,x,\hat{u}_1)$.
Moreover, let $Y_R(t,x,\hat{u}_1)$ be defined by

$$Y_R(t,x,\hat{u}_1) := v(t;\bar{\theta},0,W(t,x,\hat{u}_1),\hat{u}_1).$$

Then

(iii) $W(\cdot,\hat u_1), Y_R(\cdot,\hat u_1) \in C^{0,1}(S \cap \{t \geq s\})$ *for all* $s \in (\bar\theta, \bar t)$.
(iv) *The mapping*

$$(x,\hat u_1) \in (\xi_l(t),\xi_r(t)) \times V_1 \longmapsto (W,Y_R)(t,x,\hat u_1), \qquad t \in (s, \bar t + \tau)$$

is continuously Fréchet-differentiable with derivatives

$$d_{(x,u_1)} W(t,x,\hat u_1)\cdot(\delta x,\delta u_1) = \frac{\delta x - \delta\zeta(t;\bar\theta,0,1,\hat u_1;0,0,0,\delta u_1)}{\delta\zeta(t;\bar\theta,0,w,u_1;0,0,1,0)}, \tag{4.11}$$

$$\begin{aligned} d_{(x,u_1)} Y_R(t,x,\hat u_1)\cdot(\delta x,\delta u_1) &= \delta v(t;\bar\theta,0,w,\hat u_1;0,0,0,\delta u_1) \\ &+ \delta v(t;\bar\theta,0,w,\hat u_1;0,0,1,0)\cdot d_{(x,u_1)}W(t,x,\hat u_1)\cdot(\delta x,\delta u_1). \end{aligned} \tag{4.12}$$

(v) *The mapping*

$$\hat u_1 \in V_1 \longmapsto (W,Y_R)(\cdot,\hat u_1) \in C(S \cap \{t \geq s\})^2$$

is continuously Fréchet-differentiable with derivative

$$d_{u_1}(W,Y_R)(\cdot,\hat u_1)\cdot\delta u_1 = d_{(x,u_1)}(W,Y_R)(\cdot,\hat u_1)\cdot(0,\delta u_1).$$

(vi) *There exists* $C > 0$ *such that*

$$|d_{u_1}(W,Y_R)(\cdot,\hat u_1)\cdot\delta u_1| \leq C(t-\bar\theta)\,\|\delta u_1\|_{C^1(\Omega_T)^m} \quad \text{in } S.$$

Thus, the operator $d_{u_1}(W,Y_R)(\cdot,\hat u_1)$ *can continuously be extended to* $(\bar\theta,0)$ *by* 0.

Proof. The proof of first five assertions is completely analogous to the one of Lemma 4.2. For the last statement we apply Gronwall's Lemma to (2.5b) and obtain a constant $C > 0$ such that

$$|\delta v(t;\bar\theta,0,w,\hat u_1;0,0,0,\delta u_1)| \leq C(t-\bar\theta)\,\|\delta u_1\|_{C^1(\Omega_T)^m}, \quad t \in (\bar\theta, \bar t + \tau).$$

The first line of (2.5a) then implies

$$|\delta\zeta(t;\bar\theta,0,w,\hat u_1;0,0,0,\delta u_1)| \leq C(t-\bar\theta)^2\,\|\delta u_1\|_{C^1(\Omega_T)^m}.$$

In addition $\delta\zeta(t;\bar\theta,0,w,\hat u_1;0,0,1,0) \geq \frac{\beta}{2}(t-\bar\theta)$ holds by the first statement. The last assertion follows now from (4.11)-(4.12). ∎

Lemma 4.5. *Let* $u = (u_0,u_B,u_1) \in PC^1(\Omega;x_1,\dots,x_{n_x}) \times PC^1([0,T];t_1,\dots,t_{n_t}) \times C^1(\Omega_T)^m$ *and let* $\bar x$ *be a* R_B^c*-point. Then the following statements are true:*

(i) *There exists a maximal open interval* $I \ni \bar x$*, such that* $\{\bar t\} \times I$ *contains no point of the shock set and that all backward characteristics intersect* $x = 0$ *in* $\bar\theta$.

(ii) $y(\bar t,\cdot;u)$ *is continuously differentiable on* I.

(iii) *Let* $\hat I := (x_l,x_r)$ *be an interval with* $x_l,x_r \in I$. *Denote by* $\xi_{l/r}$ *the genuine backward characteristics through* $(\bar t, x_{l/r})$ *and set* $w_{l/r} := v(\bar\theta;\bar t,x_{l/r},y(\bar t,x_{l/r};u),u_1)$. *Then there exist* $\kappa,\beta > 0$*, such that* (4.10) *is satisfied.*

(iv) *Let* $\mathcal{T} \subset (0,T)$ *be an arbitrary neighborhood of* $\bar\theta$. *Let further* $M_\infty > 0$ *and* $s \in (\bar\theta,\bar t)$*, then there exist* $R > 0$, $\nu > 0$*, such that (for possibly smaller* τ *and* V_1*)*

$$y(t,x;\hat u) = Y_R(t,x,\hat u_1) \quad \forall (t,x) \in S \cap \{t \geq s\},\ \forall \hat u \in \hat V$$

holds, where

$$\begin{aligned} \hat V := \{(\hat u_0,\hat u_B,\hat u_1) \in BV(\Omega)\times BV(0,T)\times C^1(\Omega_T)^m \ :\ u_1 \in V_1, \\ \|\hat u_B - u_B\|_{\infty,[0,T]\setminus\mathcal{T}} < M_\infty,\ \|\hat u_B - u_B\|_{1,[0,T]\setminus\mathcal{T}} + \|\hat u_B - u_B\|_{\infty,\mathcal{T}} < \nu, \\ \|\hat u_0 - u_0\|_\infty < M_\infty,\ \|\hat u_0 - u_0\|_{1,(0,R)} < \nu\}. \end{aligned}$$

Proof. The proof is completely analogous to the one of [31, Lemma 4.8], we proceed as in the proof of Lemma 4.3. □

Remark 4.6. *Lemma 4.5 is also valid for $\bar\theta = 0$. Only the last statement requires a slight modification: Let $\mathcal{T} = [0, \varepsilon_{\mathcal{T}})$, $J = [0, \varepsilon_J)$ for some arbitrary values $0 < \varepsilon_{\mathcal{T}} < T$, $\varepsilon_J > 0$. Let further $M_\infty > 0$ and $s \in (0, \bar t)$, then there exist $R > 0$, $\nu > 0$, such that (for possibly smaller τ and V_1)*

$$y(t, x; \hat u) = Y_R(t, x, \hat u_1) \quad \forall (t, x) \in S \cap \{t \geq s\},\ \forall \hat u \in \hat V^0$$

holds, where

$$\begin{aligned}\hat V^0 := \{(\hat u_0, \hat u_B, \hat u_1) \in BV(\Omega) \times BV(0,T) \times C^1(\Omega_T)^m \ :\ u_1 \in V_1,\\ \|\hat u_B - u_B\|_{\infty,[0,T]\setminus\mathcal{T}} < M_\infty,\ \|\hat u_B - u_B\|_{1,[0,T]\setminus\mathcal{T}} + \|\hat u_B - u_B\|_{\infty,\mathcal{T}} < \nu,\\ \|\hat u_0 - u_0\|_{\infty,\Omega\setminus J} < M_\infty,\ \|\hat u_0 - u_0\|_{1,(0,R)} + \|\hat u_0 - u_0\|_{\infty,J} < \nu\}.\end{aligned}$$

Remark 4.7.

(i) *The results from [31] concerning the differentiability in a neighborhood of continuity points whose backward characteristic end in a C^1-part of u_0 (Case C^c) or a rarefaction center (Case R^c) at $(0, \bar z)$ are also valid in the considered (IBVP-) setting. Only the neighborhood $\hat V$ from the last statement of [31, Lemma 4.4 and 4.8] has to be modified to*

$$\begin{aligned}\hat V^{C^c} := \{(\hat u_0, \hat u_B, \hat u_1) \in BV(\Omega) \times BV(0,T) \times C^1(\Omega_T)^m \ :\ (\hat u_0|_J, u_1) \in V,\\ \|\hat u_0 - u_0\|_{\infty,\Omega\setminus J} < M_\infty,\ \|\hat u_0 - u_0\|_{1,(0,R)\setminus J} < \nu,\\ \|\hat u_B - u_B\|_{\infty,[0,T]} < M_\infty,\ \|\hat u_B - u_B\|_{1,[0,T]} < \nu\},\\ \hat V^{R^c} := \{(\hat u_0, \hat u_B, \hat u_1) \in BV(\Omega) \times BV(0,T) \times C^1(\Omega_T)^m \ :\ u_1 \in V_1,\\ \|\hat u_0 - u_0\|_{\infty,\Omega\setminus J} < M_\infty,\ \|\hat u_0 - u_0\|_{1,(0,R)\setminus J} + \|\hat u_0 - u_0\|_{\infty,J} < \nu,\\ \|\hat u_B - u_B\|_{\infty,[0,T]} < M_\infty,\ \|\hat u_B - u_B\|_{1,[0,T]} < \nu\},\end{aligned}$$

with V, J and V_1, J from [31, Lemma 4.1, Lemma 4.5], respectively.

(ii) *For continuity points $\bar x$ whose backward characteristic $\bar\xi$ end in a continuity point $\bar z$ of u_0 with $\bar z = x_i$ for a $i \in \{1, \ldots, n_x\}$ (Case CB^c) [31, Lemma 8.1] yields that* (3.4) *is Fréchet differentiable from $B^W_\rho(0)$ to $L^r(a, b)$ for all $r \in [1, \infty)$ and $[a, b] \subset \hat I$ by glueing together the local solutions $Y_\pm$ from [31, Lemma 4.1] along $\bar\xi$. The same can be done for CB^c_B- and C^c_0-points.*

(iii) *For continuity points $\bar x$ whose backward characteristic $\bar\xi$ end at the boundary of a rarefaction at $t = 0$ (Case RB^c) [31, Lemma 9.1] yields that* (3.4) *is continuously Fréchet differentiable from $B^W_\rho(0)$ to $L^r(a, b)$ for all $r \in [1, \infty)$ and $[a, b] \subset \hat I$ by glueing together the local solutions Y, Y_r from [31, Lemma 4.1, Lemma 4.5] along $\bar\xi$. The same can be done for RB^c_B- and RB^c_0-points.*

4.3. Differentiability at shock points. In this section we investigate the stability of the shock position w.r.t. small perturbations in the control.

We start by studying the stability of the shock positions under variations in the boundary (and initial) data, as well as in the source term.

Lemma 4.8 (Stability of the shock position). *Let $u = (u_0, u_B, u_1)$, where $u_0 \in PC^1(\Omega; x_1, \ldots, x_{n_x})$, $u_B \in PC^1([0,T]; t_1, \ldots, t_{n_t})$ and $u_1 \in C^1(\Omega_T)^m$. Furthermore, let $\bar x$ be a $C^c_B C^c_B$-point, i.e. a point of discontinuity of $y(\bar t, \cdot; u)$ on a shock curve η*

with minimal/maximal characteristics $\xi_{l/r}$ *that end at the boundary* $\{x=0\}$ *at time* $\bar\theta^{l/r}$ *and both satisfy the* C_B^c*-condition* (4.9) *for some* $\beta,\kappa>0$ *and* $\theta_l^{l/r}\geq\bar\theta^{l/r}\geq\theta_r^{l/r}$. *Denote by* $Y_{B,l/r},\Theta_{l/r},V_{l/r},S_{l/r},\mathcal{T}_{l/r}$ *the respective objects obtained by applying Lemma 4.2 to* $\xi_{l/r}$.

After a possible reduction of $V_{l/r}$ *and* $\tau>0$, *there exists a neighborhood* $I:=(x_l,x_r)$ *of* $\bar x$ *such that the following holds:*

(i) $y(\cdot;u)$ *is locally given by*

$$y(t,x;u)=\begin{cases}Y_{B,l}(t,x,u_B,u_1) & \text{if } (t,x)\in S_l\cap\{x<\eta(t)\},\\ Y_{B,r}(t,x,u_B,u_1) & \text{if } (t,x)\in S_r\cap\{x>\eta(t)\},\end{cases}.$$

(ii) *Let* $M_\infty>0$, *then there exist* $R>0,\nu>0$ *such that for*

$$\begin{aligned}\hat V:=\{&(\hat u_0,\hat u_B,\hat u_1)\in BV(\Omega)\times BV(0,T)\times C^1(\Omega_T)^m\ :\ (\hat u_B|_{\mathcal{T}_{l/r}},\hat u_1)\in V_{l/r},\\ &\|\hat u_0\|_\infty\leq M_\infty,\ \|\hat u_0-u_0\|_{1,[0,R]}\leq\nu,\ \|\hat u_B\|_\infty\leq M_\infty,\ \|\hat u_B-u_B\|_1\leq\nu\}\end{aligned}$$

equipped with the seminorm

$$\|\hat u\|_{\mathcal{V}}:=\|\hat u_B\|_{C^1(\mathcal{T}_l\cap\mathcal{T}_r)}+\|\hat u_B\|_{1,[\bar\theta^r,\bar\theta^l]}+\|\hat u_1\|_{C^1(\Omega_T)}$$

there is a Lipschitz continuous function

$$x_s:\hat u\in(\hat V,\|\cdot\|_{\mathcal{V}})\quad\longmapsto\quad x_s(\hat u)$$

with $x_s(u)=\bar x$, *such that for all* $\hat u\in\hat V$ *holds*

$$y(\bar t,x;\hat u)=\begin{cases}Y_{B,l}(\bar t,x,\hat u_B|_{\mathcal{T}_l},\hat u_1) & \text{if } x\in(x_l,x_s(\hat u)),\\ Y_{B,r}(\bar t,x,\hat u_B|_{\mathcal{T}_r},\hat u_1) & \text{if } x\in(x_s(\hat u),x_r).\end{cases}$$

Proof. The first assertion can be proven in a similar fashion as the second statement in Lemma 4.3 by using the backward stability of genuine backward characteristics according to Lemma 2.6.

A reinspection of the proof of [31, Lemma 5.1] shows, that the class (C^c) of the extreme characteristics is not explicitly used, but only the results on the local solutions $Y_\pm$ on the stripes $S_\pm$ obtained in [31, Lemma 4.4]. Therefore, we analogously use $Y_{B,l/r}$ on the stripes $S_{l/r}$ from Lemma 4.3 and follow the procedure of [31]. ∎

Remark 4.9. *The stability result of Lemma 4.8 can be proven in the same way for all shocks* X_lX_r *with* $X_{l/r}\in\{C^c,R^c,C_B^c,R_B^c,R_0^c\}$.

Lemma 4.10 (Differentiability of the shock position). *Let the assumptions of Lemma 4.8 hold. Consider a sufficiently small neighborhood* $\hat W\subset W$ *of* 0 *with* W *as defined in* (3.3). *Furthermore, let* g *be affine linear w.r.t.* y. *Denote by*

$$D:=\left\{(t,x)\in[\bar\theta^r,\bar t]\times\mathbb{R}\ :\ x\in[\xi_l(\max(t,\bar\theta^l)),\xi_r(t)]\right\}$$

the area confined by the extreme characteristics $\xi_{l/r}$. *Then the mapping*

$$\begin{aligned}&\delta w=(\delta w_0,\delta x,\delta w_B,\delta t,\delta u_1)\in\hat W\\ &\qquad\longmapsto\quad x_s(u_0+S_{u_0}^{(x_i)}(\delta w_0,\delta x),u_B+S_{u_B}^{(t_j)}(\delta w_B,\delta t),u_1+\delta u_1)\end{aligned}\tag{4.13}$$

is continuously Fréchet-differentiable and the derivative is given by

$$d_u x_s(u)\cdot\delta w = (p(\cdot,0+), f'(u_B)\delta w_B))_{2,[\bar\theta^r,\bar\theta^l]} + \sum_{\substack{j\in\{1,\dots,n_t\},\\ t_j\in[\bar\theta^r,\bar\theta^l]\\ \delta t_j\neq 0}} p(t_j,0+)[f(u_B(t_j))]\delta t_j + (pg_{u_1}(\cdot;y(\cdot;u),u_1),\delta u_1)_{2,D}, \quad (4.14)$$

where p is the adjoint state according to Definition 3.2 of (3.2) *for constant end data* $p^{\bar t} = 1/[y(\bar t, x_s(u);u)]$.

Proof. Throughout the proof we will use the following abbreviations:

$$\begin{aligned}\delta u_0 :=& S^{(x_i)}_{u_0}(\delta w_0,\delta x), \quad \delta u :=(\delta u_0,\delta u_B,\delta u_1),\\ \delta u_B :=& S^{(t_j)}_{u_B}(\delta w_B,\delta t), \quad\ \ \tilde u := u+\delta u.\end{aligned} \quad (4.15)$$

By the obvious relations

$$\begin{aligned}\|\delta u_B\|_{1,[0,T]} \le& T\,\|\delta w_B\|_\infty + 2\,\|u_B\|_\infty\,\|\delta t\|_1 \le C\,\|\delta w\|_W\,,\\ \|\delta u_0\|_{1,[0,R]} \le& R\,\|\delta w_0\|_\infty + 2\,\|u_0\|_\infty\,\|\delta x\|_1 \le C\,\|\delta w\|_W\,,\end{aligned}$$

it is easy to see, that for $\hat W$ sufficiently small there holds $\tilde u \in \hat V$.

By $y := y(\cdot;u)$, $\tilde y := y(\cdot;\tilde u)$ we denote the respective solutions of (1.2) and by $\Delta y := \tilde y - y$ their difference. As in [31, §7] one of the key points of the proof is the fact that for $\varepsilon > 0$ sufficiently small and $\hat x_{l/r} := x_s(u) \mp \varepsilon \in (x_l, x_r)$ (from Lemma 4.8) the following equality holds:

$$\int_{\hat x_l}^{\hat x_r} \Delta y(\bar t, x)\,\mathrm{d}x = (x_s(\tilde u) - x_s(u))[y(\bar t, x_s(u))] + O((\varepsilon + \|\delta w\|_W)\,\|\delta w\|_W). \quad (4.16)$$

The above equation is obtained as in [31] with the only modification that we replace $\tilde y, y$ by $Y_{B,l/r}(\cdot;\tilde u_B,\tilde u_1)$ or $Y_{B,l/r}(\cdot;u_B,u_1)$, respectively.

The rest of this proof will be concerned with the derivation of an adjoint-based formula for the left hand side of (4.16). We avoid the introduction of a detailed analysis for linear transport equations with discontinuous coefficients on bounded domains, instead we show how the considered equation can be modified so that the results of [31, §6] can be used. For $(t,x)\in(0,\bar t]\times\mathbb{R}^+$ we define

$$\begin{aligned} a(t,x) &:= f'(y(t,x)), & \tilde a(t,x) &:= \int_0^1 f'(y(t,x)+\lambda\Delta y(t,x))\,\mathrm{d}\lambda,\\ b(t,x) &:= g_y(t,x,y(t,x),u_1), & \tilde b(t,x) &:= g_y(t,x,\tilde y(t,x),\tilde u_1).\end{aligned}$$

Using the above abbreviations and the assumption that g is affine linear w.r.t y, we deduce that the difference of $\tilde y$ and y is a weak solution of

$$\partial_t\Delta y + \partial_x(\tilde a\Delta y) = \tilde b\Delta y + g(\cdot,y,\tilde u_1) - g(\cdot,y,u_1). \quad (4.17)$$

We extend the functions $a,\tilde a,b,\tilde b$ to $[0,\bar t]\times\mathbb{R}$ by setting $(\tilde a,a)(t,x) := (M_{f'},M_{f'})$ if $x<0$ and $(\tilde b,b)(t,x) := (\tilde b,b)(t,\max(0,x)+)$, where $M_{f'}$ is an a priori bound on f'.

Denote by $t_{r_1} < \cdots < t_{r_{\bar K}}$ the increasing sequence of all rarefaction centers at the boundary in $\{t_1,\dots,t_{n_t}\}$ and set $t_{r_0} := 0$, $t_{r_{\bar K+1}} := \bar t$. On $(t_{r_{\bar K}},\bar t]$ the adjoint state p

according to Definition 3.2 can be interpreted as the restriction of the reversible solution to the extended adjoint equation (3.2) on $(t_{r_{\tilde K}}, \bar t] \times \mathbb{R}$ in the sense of [31, Definition 6.5] for the same end data. The same holds for all proximate time slabs $[t_{r_{s-1}}, t_{r_s}]$, $s = 1, \dots, \tilde K$, where the respective end data is chosen as $p^{t_{r_s}} := p(t_{r_s}+, \max(x, 0))$.

We define $\tilde p$ as the solution on $(0, \bar t] \times \mathbb{R}^+$ of the averaged adjoint equation

$$\partial_t \tilde p + \tilde a \partial_x \tilde p = -\tilde b \tilde p, \quad \tilde p(\bar t, \cdot) = p^{\bar t} \equiv \frac{1}{[y(\bar t, x_s(u); u)]} \tag{4.18}$$

in a similar fashion: On $(t_{r_{\tilde K}}, \bar t]$ the function $\tilde p$ is the reversible solution to the extended averaged adjoint equation (4.18) on $(t_{r_{\tilde K}}, \bar t] \times \mathbb{R}$ (acc. [31, Definition 6.5]) and on every time slab $(t_{r_{s-1}}, t_{r_s}]$, $s = 1, \dots, \tilde K$, $\tilde p$ is the restriction on $(t_{r_{s-1}}, t_{r_s}] \times \mathbb{R}^+$ of the reversible solution of

$$\partial_t \tilde p + \tilde a \partial_x \tilde p = -\tilde b \tilde p, \quad \tilde p(t_{r_s}, \cdot) = p^{t_{r_s}}$$

on $(t_{r_{s-1}}, t_{r_s}] \times \mathbb{R}$.

By Lemma 4.8 the genuine backward characteristics $\hat\zeta_{l/r}$ of y through $(\bar t, \hat x_{l/r})$ end in $\mathcal{T}_{l/r} \times \{0\}$ and for all $t \le \bar t$ the relation $\hat\zeta_l(t) \le \xi_l(t) \le \xi_r(t) \le \hat\zeta_r(t)$ holds where the respective curves exist. Furthermore, Lemma 4.3 yields that for $\|\delta w\|_W$ sufficiently small $\tilde y$ coincides with $Y_{B,l/r}(\cdot, \tilde u_B, \tilde u_1)$ near $\hat\zeta_{l/r}$. Denote by $\hat\theta_{l/r}$ the respective endpoints of $\hat\zeta_{l/r}$. For $\delta > 0$ define the sets

$$\Omega_\delta^s := \Big((t_{r_s}, t_{r_{s+1}}] \times \mathbb{R}^+ \Big) \setminus \Big((t_{r_s}, t_{r_s} + \delta] \times (0, \delta] \Big), \qquad \Omega_\delta := \bigcup_{s=0}^{\tilde K} \Omega_\delta^s$$

and

$$A_\delta := \Big\{ (t, x) \in \Omega_\delta \ : \ t \in [\hat\theta_r, \bar t], \ \gamma(t) \le x \le \hat\zeta_r(t) \Big\}, \qquad A_\delta^s := A_\delta \cap \Omega_\delta^s,$$

where $\gamma(t) := \hat\zeta_l(\max(t, \hat\theta_l))$ is the extension of $\hat\zeta_l$ to $[\hat\theta_r, \bar t]$. By construction [31, Theorem 6.10] yields that $\tilde p \in C^{0,1}(A_\delta^s) \cap C([t_{r_s}, t_{r_{s+1}}]; L^2_{\mathrm{loc}}(\mathbb{R}^+))$, $s = 0, \dots, \tilde K$. We multiply (4.17) by $\tilde p$ and apply integration by parts on every A_δ^s.

At this point we highly recommend the reader to consider Fig. 4.1 before continuing reading the proof. There, the domains of integration, especially the excluded parts around rarefaction centers, are illustrated. Integration by parts yields

$$\begin{aligned}
&\int_{\hat x_l}^{\hat x_r} \tilde p(\bar t, x) \Delta y(\bar t, x) \,\mathrm{d}x \\
&\quad = \int_{[\hat\theta_r, \hat\theta_l] \setminus \bigcup_{s=1}^{\tilde K} [t_{r_s}, t_{r_s} + \delta]} \tilde p(t, 0+)(f(\tilde u_B) - f(u_B))(t) \,\mathrm{d}t \\
&\qquad + \iint_{A_\delta} \Delta y (\partial_t \tilde p + \tilde a \partial_x \tilde p + \tilde b \tilde p) \,\mathrm{d}x \,\mathrm{d}t \\
&\qquad + \sum_{\substack{s \in \{1, \dots, \tilde K\} \\ t_{r_s} \ge \hat\theta_r}} \int_{\gamma(t_{r_s})}^{\hat\zeta_r(t_{r_s})} (p^{t_{r_s}}(x) - \tilde p(t_{r_s}+, x)) \Delta y(t_{r_s}, x) \,\mathrm{d}x \\
&\qquad + \iint_{A_\delta} \tilde p (g(\cdot, y, \tilde u_1) - g(\cdot, y, u_1)) \,\mathrm{d}x \,\mathrm{d}t
\end{aligned}$$

$$
\begin{aligned}
&+\int_{\hat\theta_l}^{\bar t} \tilde p(t,\hat\zeta_l(t))(-f'(y)\Delta y + f(\tilde y) - f(y))(t,\hat\zeta_l(t))\,\mathrm{d}t\\
&+\int_{\hat\theta_r}^{\bar t} \tilde p(t,\hat\zeta_r(t))(f'(y)\Delta y - f(\tilde y) + f(y))(t,\hat\zeta_r(t))\,\mathrm{d}t\\
&+\sum_{\substack{s\in\{1,\dots,\tilde K\}\\ t_{r_s}\in[\hat\theta_r,\hat\theta_l]}} \int_{t_{r_s}}^{t_{r_s}+\delta} \tilde p(t,\delta)(\Delta y\tilde a)(t,\delta+)\,\mathrm{d}t\\
&+\sum_{\substack{s\in\{1,\dots,\tilde K\}\\ t_{r_s}\in[\hat\theta_r,\hat\theta_l]}} \int_0^{\delta} \Big(\tilde p(t_{r_s}+,x)\Delta y(t_{r_s},x) - \tilde p(t_{r_s}+\delta,x)\Delta y(t_{r_s}+\delta,x)\Big)\,\mathrm{d}x. \qquad (4.19)
\end{aligned}
$$

In the second line, we can replace the boundary trace by the boundary data by Proposition 2.5 and Lemma 4.3. Since by [31, Lemma 6.10] $\tilde p$ solves (4.18) almost everywhere on A_δ, the second integral on the right hand side vanishes. The functions $\tilde p$, $\tilde a$, $g(\cdot,y,\tilde u)$ and $g(\cdot,y,u)$ are essentially bounded on Ω_T and admit the respective traces for all $\delta\ge 0$. Moreover, $u_B\in PC^1([0,T];t_1,\dots,t_{n_t})$ and thus we can pass to the limit $\delta\to 0$ in (4.19) and obtain

$$
\begin{aligned}
&\int_{\hat x_l}^{\hat x_r} \tilde p(\bar t,x)\Delta y(\bar t,x)\,\mathrm{d}x\\
&\quad=\int_{\hat\theta_r}^{\hat\theta_l} \tilde p(t,0+)(f(\tilde u_B)-f(u_B))(t)\,\mathrm{d}t\\
&\qquad+\sum_{\substack{s\in\{1,\dots,\tilde K\}\\ t_{r_s}\ge\hat\theta_r}} \int_{\gamma(t_{r_s})}^{\hat\zeta_r(t_{r_s})} (p^{t_{r_s}}(x)-\tilde p(t_{r_s}+,x))\Delta y(t_{r_s},x)\,\mathrm{d}x\\
&\qquad+\int_{\hat\theta_r}^{\bar t}\int_{\gamma(t)}^{\hat\zeta_r(t)} \tilde p(g(\cdot,y,\tilde u_1)-g(\cdot,y,u_1))\,\mathrm{d}x\,\mathrm{d}t\\
&\qquad+\int_{\hat\theta_l}^{\bar t} \tilde p(t,\hat\zeta_l(t))(-f'(y)\Delta y + f(\tilde y) - f(y))(t,\hat\zeta_l(t))\,\mathrm{d}t\\
&\qquad+\int_{\hat\theta_r}^{\bar t} \tilde p(t,\hat\zeta_r(t))(f'(y)\Delta y - f(\tilde y) + f(y))(t,\hat\zeta_r(t))\,\mathrm{d}t\\
&\quad=:I_1+\Sigma I_2+I_3+I_4+I_5. \qquad (4.20)
\end{aligned}
$$

For the integrals along the confining characteristics I_4 and I_5 the boundedness of $\tilde p$ for all $\delta w\in\hat W$ and the Lipschitz continuity of $u\mapsto Y_{l/r}(\cdot;u)$ ensured by Lemma 4.2 yield an estimate

$$I_{4/5}=O(\|\delta w\|_W^2).$$

Using similar bounds as for $I_{4/5}$ and the backward stability of genuine backward characteristics one obtains

$$\left|I_3-\iint_D p g_{u_1}(t,x,\hat y,\hat u_1)\delta u_1(t,x)\,\mathrm{d}x\,\mathrm{d}t\right| \le C(\varepsilon+\|\tilde p-p\|_{1,D})\,\|\delta w\|_W + o(\|\delta w\|_W).$$

For ΣI_2, that appears because of the lack of continuity of $\tilde{p}$ in time, we have

$$
\begin{aligned}
|\Sigma I_2| \leq & \sum_{\substack{s\in\{1,\dots,\tilde{K}\}\\ t_{r_s}>\hat{\theta}_l}} \|(p-\tilde{p})(t_{r_s}+,\cdot)\|_{\infty,[\hat{\zeta}_l(t_{r_s}),\hat{\zeta}_r(t_{r_s})]} \|\Delta y(t_{r_s},\cdot)\|_{1,[\hat{\zeta}_l(t_{r_s}),\hat{\zeta}_r(t_{r_s})]} \\
& + \sum_{\substack{s\in\{1,\dots,\tilde{K}\}\\ t_{r_s}\in[\hat{\theta}_r,\hat{\theta}_l]}} \|(p-\tilde{p})(t_{r_s}+,\cdot)\|_{1,(0,\tilde{\delta})} \|y(t_{r_s},\cdot;\tilde{u})-y(t_{r_s},\cdot;u)\|_{\infty,(0,\tilde{\delta})} \\
& + \sum_{\substack{s\in\{1,\dots,\tilde{K}\}\\ t_{r_s}\in[\hat{\theta}_r,\hat{\theta}_l]}} \|(p-\tilde{p})(t_{r_s}+,\cdot)\|_{\infty,(\tilde{\delta},\hat{\zeta}_r(t_{r_s}))} \|\Delta y(t_{r_s},\cdot))\|_{1,[\delta,\hat{\zeta}_r(t_{r_s})]}\,. \qquad (4.21)
\end{aligned}
$$

We show, that $o(\|\delta w\|_W)$ by using the results on the convergence of $\tilde{p}$ from [30, Theorem 3.7.10]: Since all rarefaction centers are located at $x=0$, we have $\tilde{p}\to p$ uniformly in the first and the last line of (4.21). Together with the L^1_{loc}-stability of the state, ensured by Proposition 2.1, this implies that these two terms are $o(\|\delta w\|_W)$. In the neighborhood of rarefaction centers $\tilde{p}\to p$ holds at least in $C([\hat{\theta}_r,\bar{t}];L^1_{\mathrm{loc}}(\mathbb{R}^+))$ in the second line. The careful reinspection of the proofs of Lemmas 4.3 and 4.5 shows, that in the second line $y(\cdot;\tilde{u})$ also locally coincides with $Y_B(\cdot,\tilde{u})$ from Lemma 4.2 and, therefore, we conclude that this term is $o(\|\delta w\|_W)$ and thus

$$\Sigma I_2 = o(\|\delta w\|_W).$$

Let $t_{j'}\in[\hat{\theta}^r,\hat{\theta}^l]$, $j'\in\{1,\dots,n_t\}$ be a shock generating discontinuity of u_B. We assume w.l.o.g. that $t_{j'}$ is the only discontinuity of this type on $[\hat{\theta}^r,\hat{\theta}^l]$. Multiple discontinuities will only cause that in the following computation one has to deal with the finite sum of certain integrals, instead of a single one. But of course, the same arguments hold also in that case.

For the first integral I_1 we consider the two different types of variations, δw_B and δt, separately. Using the notation $\mathcal{T}:=I(t_{j'},t_{j'}+\delta t_{j'})$ and the triangle inequality, we obtain

$$
\begin{aligned}
&\left| I_1 - \int_{\bar{\theta}^r}^{\bar{\theta}^l} pf'(u_B)\delta w_B \,\mathrm{d}t - p(t_{j'},0+)[f(u_B(t_{j'}))]\delta t_{j'} \right| \\
&\quad \leq \left| \int_{\hat{\theta}_r}^{\hat{\theta}_l} \tilde{p}(\cdot,0+)(f(u_B+\delta w_B)-f(u_B))\,\mathrm{d}t - \int_{\bar{\theta}^r}^{\bar{\theta}^l} pf'(u_B)\delta w_B\,\mathrm{d}t \right| \\
&\qquad + \left| \int_{\mathcal{T}} \tilde{p}(\cdot,0+)(f(u_B+\delta w_B+\operatorname{sgn}(\delta t_{j'})[u_B(t_{j'})]) - f(u_B+\delta w_B))\,\mathrm{d}t \right. \\
&\qquad\qquad \left. - p(t_{j'},0+)[f(u_B(t_{j'}))]\delta t_{j'} \right| \\
&\quad \leq \Big(L_\Theta\varepsilon \|\tilde{p}(\cdot,0)\|_\infty + \|(\tilde{p}-p)(\cdot,0)\|_{1,[\bar{\theta}_r,\bar{\theta}_l]\setminus\bigcup_{s=1}^{\tilde{K}}[t_{r_s},t_{r_s}+\delta]} \Big) M_{f'} \|\delta w_B\|_\infty \\
&\qquad + \sum_{s=1}^{\tilde{K}} \left| \int_{t_{r_s}}^{t_{r_s}+\delta} (\tilde{p}-p)(\cdot,0)(f(u_B+\delta w_B)-f(u_B))\,\mathrm{d}t \right| \\
&\qquad + \|p(\cdot,0)\|_{1,[\hat{\theta}_r,\hat{\theta}_l]} M_{f''} \|\delta w_B\|_\infty^2 + \|(\tilde{p}-p)(\cdot,0)\|_{\infty,\mathcal{T}} M_{f'} |[u_B(t_{j'})]|\cdot|\delta t_{j'}|
\end{aligned}
$$

$$+ \left| \int_{\mathcal{T}} p(\cdot, 0+) \left(f(u_B + \delta w_B + \operatorname{sgn}(\delta t_{j'})[u_B(t_{j'})]) - f(u_B + \delta w_B) \right) \right.$$
$$\left. - \operatorname{sgn}(\delta t_{j'}) p(t_{j'}, 0+)[f(u_B(t_{j'}))] \,\mathrm{d}t \right|.$$

The convergence $\tilde{p} \to p$ in $C(A_\delta^s)$ for all $s \in \{0, \ldots, \tilde{K}\}$ implies $\|(\tilde{p} - p)(\cdot, 0)\|_{\infty, \mathcal{T}} \to 0$ if $\mathcal{T}$ and δ are sufficiently small. The same holds for $\|(\tilde{p} - p)(\cdot, 0)\|_{1, [\hat{\theta}_r, \hat{\theta}_l] \setminus \bigcup_{s=1}^{\tilde{K}} [t_{r_s}, t_{r_s}+\delta]}$. For the remaining intervals $[t_{r_s}, t_{r_s} + \delta]$ we consider the extensions of p and $\tilde{p}$ on $(t_{r_s}, t_{r_{s+1}}] \times \mathbb{R}$. We apply the divergence theorem to the vector field

$$(t, x) \quad \longmapsto \quad \begin{pmatrix} \tilde{p} - p \\ M_{f'}(\tilde{p} - p) \end{pmatrix} (t, x) \cdot (f(u_B + \delta w_B) - f(u_B)) \left(t - \frac{x}{M_{f'}} \right)$$

on the triangle

$$D_\Delta := \{(t, x) \in [t_{r_s}, t_{r_s} + \delta] \times (-\infty, 0] \; : \; x \geq -M_{f'}(t_{r_s} - t + \delta))\}$$

and obtain

$$\left| \int_{t_{r_s}}^{t_{r_s}+\delta} (\tilde{p} - p)(\cdot, 0)(f(u_B + \delta w_B) - f(u_B)) \,\mathrm{d}t \right|$$
$$\leq \|\delta w_B\|_\infty \left(\|(\tilde{p} - p)(t_{r_s}, \cdot)\|_{1, [-M_{f'}\delta, 0]} + \|b(\tilde{p} - p)\|_{1, D_\Delta} + \left\| \tilde{p}(\tilde{b} - b) \right\|_{1, D_\Delta} \right)$$
$$= o(\|\delta w\|_W).$$

We now consider the differentiability of the shock position w.r.t. $\delta t_{j'}$. We make several estimates in the following order: We consider the dependence on δw_B, replace $p(t, 0)$ by $p(t_{j'}, 0)$ using the Lipschitz continuity of p on A_δ and finally show that even the remaining term is of order $O(\|\delta w\|_W^2)$ by overestimating the remainder of the Taylor expansion:

$$\left| \int_{\mathcal{T}} p(\cdot, 0+) \left(f(u_B + \delta w_B + \operatorname{sgn}(\delta t_{j'})[u_B(t_{j'})]) - f(u_B + \delta w_B) \right) \right.$$
$$\left. - \operatorname{sgn}(\delta t_{j'}) p(t_{j'}, 0+)[f(u_B(t_{j'}))] \,\mathrm{d}t \right|$$
$$\leq 2|\delta t_{j'}| \, \|p(\cdot, 0+)\|_{\infty, \mathcal{T}} \, M_{f'} \, \|\delta w_B\|_\infty$$
$$+ \left| \int_{\mathcal{T}} (p(t, 0+) - p(t_{j'}, 0+)) \left(f(u_B + \operatorname{sgn}(\delta t_{j'})[u_B(t_{j'})]) - f(u_B) \right) \,\mathrm{d}t \right|$$
$$+ \left| p(t_{j'}, 0+) \right.$$
$$\left. \cdot \int_{\mathcal{T}} f(u_B + \operatorname{sgn}(\delta t_{j'})[u_B(t_{j'})]) - f(u_B) - \operatorname{sgn}(\delta t_{j'})[f(u_B(t_{j'}))] \,\mathrm{d}t \right|$$
$$\leq 2|\delta t_{j'}| \, \|p(\cdot, 0+)\|_{\infty, \mathcal{T}} \, M_{f'} \, \|\delta w_B\|_\infty + |\delta t_{j'}|^2 L_p 2 M_f$$
$$+ \; |p(t_{j'}, 0+)| \cdot \left| \int_{\mathcal{T}} \int_t^{t_{j'}} f'(u_B(s + 0 \cdot \delta t_{j'})) u_B'(s + 0 \cdot \delta t_{j'}) \,\mathrm{d}s \,\mathrm{d}t \right.$$
$$+ \int_{\mathcal{T}} \int_{t_{j'}}^t f'(u_B(t_{j'} - 0 \cdot \delta t_{j'}) + u_B(s + 0 \cdot \delta t_{j'}) - u_B(t_{j'} + 0 \cdot \delta t_{j'}))$$
$$\left. \cdot u_B'(s + 0 \cdot \delta t_{j'}) \,\mathrm{d}s \,\mathrm{d}t \right|$$

$$\leq 2|\delta t_{j'}|\,\|p(\cdot,0+)\|_{\infty,\mathcal{T}}\, M_{f'}\,\|\delta w_B\|_\infty + |\delta t_{j'}|^2 L_p 2M_f \\ + 2M_{f'}|p(t_{j'},0+)|\cdot|\delta t_{j'}|^2\,\|u_B\|_{PC^1([0,T];t_1,\dots,t_{n_t})} = O(\|\delta w\|_W^2). \tag{4.22}$$

The use of u'_B in the above computations is valid, since for sufficiently small $\delta t_{j'}$ the interval $\mathcal{T}$ has no point of discontinuity t_j in its interior.

This shows, that (4.13) is Fréchet-differentiable and that the derivative is given by (4.14). The continuity of the derivative follows again from the stability of the solution of the adjoint state w.r.t. perturbation of the coefficients of (3.2). □

Remark 4.11. *By the same arguments as above, one can show that the assertions of Lemma 4.10 also hold for any shock X_lX_r with $X_{l/r} \in \{C_B^c, R_B^c, R_0^c\}$. Shocks of class X_lX_r with $X_{l/r} \in \{C^c, R^c, R_0^c\}$ have already been considered in [31, Lemma 5.2, Corollary 5.3].*

A quite similar formulation of the above Lemma holds if the shock funnel contains the area D^- from (2.7) if the following nondegeneracy condition on the transition point θ^Δ holds in addition to the properties already collected in Lemma 2.7.

Definition 4.12 (Nondegeneracy of θ^Δ). *In the setting of Theorem 3.3 the* transition point $\theta^\Delta = t_{j^\Delta}$*, cf. Lemma 2.7, from* (2.6) *is called* nondegenerated*, if the following conditions are satisfied:*

(i) *One can construct a stripe S around ξ^Δ such that one can define a local solution Y on S as in [31, Lemma 4.1 or Lemma 4.5] with the same properties and that for any $\delta w \in B_\rho^W(0)$ with $\rho > 0$ sufficiently small and δu from* (4.15)*, $y(\cdot;u+\delta u)$ coincides with Y on $S \cap \{t \leq \theta^\Delta + \delta t_{j^\Delta}\}$.*

(ii) *$f(u_B(\theta^\Delta+)) > f(Y(\theta^\Delta, 0, u))$.*

Lemma 4.13. *Let $u = (u_0, u_B, u_1)$, where $u_0 \in PC^1(\Omega; x_1, \dots, x_{n_x})$, $u_B \in PC^1([0,T]; t_1, \dots, t_{n_t})$ and $u_1 \in C^1(\Omega_T)^m$. Furthermore, let $\bar{x}$ be a $C_B^cC^c$-point, i.e. a point of discontinuity of $y(\bar{t},\cdot;u)$ on a shock curve η with minimal characteristic ξ_l that ends at the boundary $\{x=0\}$ at time $\bar{\theta}$ and satisfies the C_B^c-condition* (4.9) *and with maximal characteristic ξ_r that ends at time $\{t=0\}$ at the point $\bar{z}$ and satisfies the C^c-condition,*

$$J = (z_l - \kappa, z_r + \kappa), \quad \text{and} \quad \frac{d}{dz}\zeta(t;0,z,u_0(z),u_1) \geq \beta > 0, \quad \text{for all } t \in [0,\bar{t}],\ z \in J,$$

for some $\beta, \kappa > 0$ and $\theta_l \geq \bar{\theta} \geq \theta_r$, $z_l \leq \bar{z} \leq z_r$. Denote by $Y_B/Y, \Theta/J, V_{l/r}, S_{l/r}$ the respective objects obtained by applying Lemma 4.2 and its analogue [31, Lemma 4.1] to $\xi_{l/r}$.

Assume in addition that the transition point $\theta^\Delta = t_{j^\Delta}$ is nondegenerated according to Definition 4.12.

Consider a sufficiently small neighborhood $\hat{W} \subset W$ of 0 with W as defined in (3.3)*. Furthermore, let g be affine linear w.r.t. y. Denote by*

$$D := \left\{(t,x) \in [0,\bar{t}] \times \mathbb{R} \ :\ x \in [\xi_l(\max(t,\bar{\theta})) + \xi^\Delta(\min(t,\theta^\Delta)), \xi_r(t)]\right\}$$

the area confined by the extreme characteristics $\xi_{l/r}$ and ξ^Δ with θ^Δ and ξ^Δ from Lemma 2.7. Then the mapping (4.13) *is continuously Fréchet-differentiable and the*

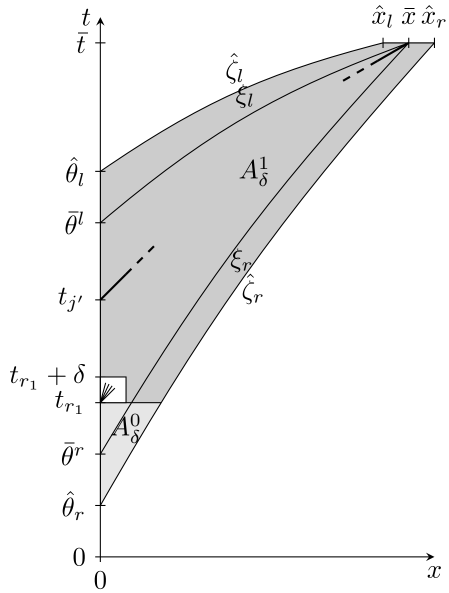

FIG. 4.1. *Proof of Lemma 4.10*

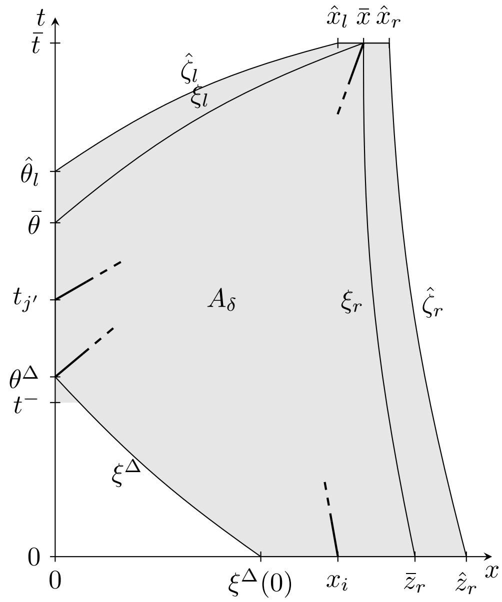

FIG. 4.2. *Proof of Lemma 4.13*

derivative is given by

$$
\begin{aligned}
d_u x_s(u)\cdot\delta w &= (p(\cdot,0+), f'(u_B)\delta w_B))_{2,[\theta^\Delta,\bar\theta]} + (p(0,\cdot),\delta w_0))_{2,[\xi^\Delta(0),\bar z]} \\
&\quad + \sum_{\substack{j\in\{j^\Delta+1,\dots,n_t\},\\ t_j\le\bar\theta,\,\delta t_j\neq 0}} p(t_j,0+)[f(u_B(t_j))]\delta t_j + \sum_{\substack{i\in\{1,\dots,n_x\},\\ x_i\in[\xi^\Delta(0),\bar z],\\ \delta x_i\neq 0}} p(0,x_i)[u_0(x_i)]\delta x_i \\
&\quad + p(\theta^\Delta,0+)(f(y(\theta^\Delta-,0+;u)) - f(u_B(\theta^\Delta+)))\delta t_{j^\Delta} \\
&\quad + (pg_{u_1}(\cdot;y(\cdot;u),u_1),\delta u_1)_{2,D}, \qquad (4.23)
\end{aligned}
$$

where p is the adjoint state according to Definition 3.2 of (3.2) *for constant end data* $p^{\bar t} = 1/[y(\bar t, x_s(u);u)]$.

Proof. The proof is very similar to the one of Lemma 4.10. In (4.19) the area A_δ, where integration by parts is applied, must be slightly modified, see the shadowed area in Fig. 4.2.

This leads to some additional terms in (4.20):

$$
\begin{aligned}
I_{6,1} &:= \int_{t_-}^{t_{j^\Delta}} \tilde p(t,0+)(f(\tilde u_B(t)) - f(Y(t,0,u)))\,\mathrm{d}t, \\
I_{6,2} &:= \int_{t_{j^\Delta}}^{t_+} \tilde p(t,0+)(f(Y(t,0,\tilde u)) - f(u_B(t)))\,\mathrm{d}t, \\
I_7 &:= \int_0^{\xi^\Delta(t_-)} \tilde p(t_-,x)(Y(t_-,x,\tilde u) - Y(t_-,x,u))\,\mathrm{d}x, \\
I_8 &:= \int_0^{t_-} \tilde p(t,\xi^\Delta(t))(-f'(y)\Delta y + f(\tilde y) - f(y))(t,\xi^\Delta(t))\,\mathrm{d}t,
\end{aligned}
$$

where $t_- := \min(t_{j^\Delta}, t_{j^\Delta} + \delta t_{j^\Delta})$ and $t_+ := \max(t_{j^\Delta}, t_{j^\Delta} + \delta t_{j^\Delta})$.

One of the integrals $I_{6,1}$ and $I_{6,2}$ vanishes, the other one converges in $O(\|\delta w\|_W^2)$ to the shock sensitivity (the second last term in (4.23)), similar to (4.22).

The stability of Y and ξ^Δ yield $I_7 = o(\|\delta w\|_W)$ and $I_8 = O(\|\delta w\|_W^2)$ (cf. $I_{4/5}$). ◻

Remark 4.14. *By the same arguments as above, one can show that the assertions of Lemma 4.13 also hold for any shock $X_l X_r$ with $X_l \in \{C_B^c, R_B^c, R_0^c\}$, $X_r \in \{C^c, R^c, R_0^c\}$.*

Proof of Theorem 3.3. The previous considerations of this section can now be combined as in the proof of [31, Theorem 10.1] yielding the asserted result of Theorem 3.3. ◻

Proof of Theorem 3.6. The proof is mainly a combination of the one of Lemma 4.10 and [32, Theorem 5]. Basically, the adjoint calculus from Lemma 4.10 is used on the whole domain in order to find a first order approximation of $\int_\Omega \bar{\psi}_y \Delta y \, \mathrm{d}x$ instead of the lefthand side of (4.16). ◻

5. Conclusion and future work. We have presented a result on the differentiability of reduced objective functions for the optimal control of hyperbolic conservation laws on bounded domains. In our setting explicit shifts of discontinuities in the boundary data are possible. We have shown that the control-to-state mapping for the considered problem is shift-differentiable by using the stability of solutions to the characteristic equation and an appropriate adjoint calculus. We have shown that even derivatives w.r.t. the positions of the shock generating discontinuities exist and can be computed via the adjoint state. Once we have shown this, the result for the reduced objective function is a consequence of [30, Lemma 3.2.3]. We have also derived an adjoint-based formula for the gradient of the reduced objective, which is an important step towards the accessibility of such problems by gradient based optimization methods.

While on the one hand the investigated IBVP is a classical problem that has not been discussed in context of the presented ansatz in the literature so far, on the other hand the ability of taking derivatives w.r.t. the position of discontinuities provides the possibility of extending our approach to networks with switched node conditions. Indeed, in a forthcoming paper we will investigate such situations by means of a traffic light on a road, where the traffic is modeled by hyperbolic conservation laws using the LWR-model, cf. [22, 28], and the control variables are the switching times between green and red phases, where the traffic is or is not allowed to cross a certain point, respectively. Here the explicit shift of rarefaction centers has to be studied in addition to the results of the present paper. This is done by using the fact that the solution near rarefaction centers, i.e. green switching times, is thoroughly known.

In the future we want to reconsider the I(B)VP case and try to treat the shift of rarefaction centers for IBVPs in the setting of the present paper by extending the technique we used for the green traffic light.

We plan to extend the considered approach under suitable assumptions to systems of conservation laws where one has multiple conserved quantities, as in case of the Euler equation.

The presented sensitivity and adjoint calculus forms the basis for the numerical treatment of the considered optimal control problems. For the Cauchy problem there exist several works on the convergence of optimal solutions of discretized optimal control problems, e.g. [9, 29], and the convergence of sensitivities, adjoints and reduced gradients, see [15, 16, 30, 31, 32] and also [9] for an alternating descent method.

Our current investigations focus on the extension of those convergence results to the problem considered in this paper, where we follow the approach in [10] to approximate the boundary conditions (2.2). A particular issue will be the appropriate discrete approximation of variations for the shift of discontinuities in the boundary condition. Here, we will consider and compare two different approaches. In the first one, we consider the variation of the times step sizes between switching times, while for the latter we want to use 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, cf. [26]. The second author was in parts supported by the DFG within the collaborative research center TRR 154 "Mathematical Modeling, Simulation and Optimization Using the Example of Gas Networks". Moreover, we gratefully acknowledge discussions with G. Leugering and T. I. Seidman.

REFERENCES

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

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

[3] Claude Bardos, Alain Y. le Roux, and Jean-Claude Nédélec, *First order quasilinear equations with boundary conditions*, Comm. Partial Differential Equations, 4 (1979), pp. 1017–1034.

[4] Stefano Bianchini, *On the shift differentiability of the flow generated by a hyperbolic system of conservation laws*, Discrete Contin. Dyn. Syst., 6 (2000), pp. 329–350.

[5] Fransois Bouchut and Fransois James, *One-dimensional transport equations with discontinuous coefficients*, Nonlinear Anal., 32 (1998), pp. 891–933.

[6] Alberto Bressan, *Hyperbolic systems of conservation laws*, vol. 20 of Oxford Lecture Ser. Math. Appl., Oxford University Press, Oxford, 2000. The one-dimensional Cauchy problem.

[7] Alberto Bressan and Graziano Guerra, *Shift-differentiability of the flow generated by a conservation law*, Discrete Contin. Dyn. Syst., 3 (1997), pp. 35–58.

[8] Alberto Bressan and Andrea Marson, *A variational calculus for discontinuous solutions of systems of conservation laws*, Comm. Partial Differential Equations, 20 (1995), pp. 1491–1552.

[9] Carlos Castro, Francisco Palacios, and Enrique 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 (2008), pp. 369–416.

[10] Bernardo Cockburn, Frédéric Coquel, and Philippe G. LeFloch, *Convergence of the finite volume method for multidimensional conservation laws*, SIAM J. Numer. Anal., 32 (1995), pp. 687–705.

[11] Giuseppe M. Coclite, Kenneth H. Karlsen, and Young-Sam Kwon, *Initial-boundary value problems for conservation laws with source terms and the Degasperis-Procesi equation*, J. Funct. Anal., 257 (2009), pp. 3823–3857.

[12] Rinaldo M. Colombo and Alessandro Groli, *On the optimization of the initial boundary value problem for a conservation law*, J. Math. Anal. Appl., 291 (2004), pp. 82–99.

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

[14] Fransois Dubois and Philippe LeFloch, *Boundary conditions for nonlinear hyperbolic systems of conservation laws*, J. Differential Equations, 71 (1988), pp. 93–122.

[15] Mike Giles and Stefan 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 (2010), pp. 882–904.

[16] ———, *Convergence of linearized and adjoint approximations for discontinuous solutions of conservation laws. Part 2: Adjoint approximations and extensions*, SIAM J. Numer. Anal., 48 (2010), pp. 905–921.

[17] Falk M. Hante and Günter Leugering, *Optimal boundary control of convection-reaction*

transport systems with binary control functions, in Hybrid Systems: Computation and Control, Rupak Majumdar and Paulo Tabuada, eds., vol. 5469 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2009, pp. 209–222.

[18] Falk M. Hante, Günter Leugering, and Thomas I. Seidman, *Modeling and analysis of modal switching in networked transport systems*, Appl. Math. Optim., 59 (2009), pp. 275–292.

[19] Stanislav N. Kružkov, *First order quasilinear equations in several independent variables*, Math. USSR Sb., 10 (1970), pp. 217–243.

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

[21] Philippe LeFloch, *Explicit formula for scalar nonlinear conservation laws with boundary condition*, Math. Methods Appl. Sci., 10 (1988), pp. 265–287.

[22] Michael J. Lighthill and Gerald B. Whitham, *On kinematic waves. II. A theory of traffic flow on long crowded roads*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 229 (1955), pp. 317–345.

[23] Josef Málek, Jindrich Nečas, Mirko Rokyta, and Michael Růžička, *Weak and measure-valued solutions to evolutionary PDEs*, vol. 13 of Applied Mathematics and Mathematical Computation, Chapman & Hall, London, 1996.

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

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

[26] Sebastian Pfaff, Stefan Ulbrich, and Günter Leugering, *Optimal control of nonlinear hyperbolic conservation laws with switching*, in Trends in PDE Constrained Optimization, vol. 165 of Internat. Ser. Numer. Math., Springer International Publishing, 2014, pp. 109–131.

[27] Elijah Polak, *Optimization*, vol. 124 of Appl. Math. Sci., Springer-Verlag, New York, 1997.

[28] Paul I. Richards, *Shock waves on the highway*, Oper. Res., 4 (1956), pp. 42–51.

[29] Stefan 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), vol. 133 of Internat. Ser. Numer. Math., Birkhäuser, Basel, 1999, pp. 287–299.

[30] ———, *Optimal control of nonlinear hyperbolic conservation laws with source terms*, Habilitation, Zentrum Mathematik, Technische Universität München, Germany, 2001.

[31] ———, *A sensitivity and adjoint calculus for discontinuous solutions of hyperbolic conservation laws with source terms*, SIAM J. Control Optim., 41 (2002), pp. 740–797.

[32] ———, *Adjoint-based derivative computations for the optimal control of discontinuous solutions of hyperbolic conservation laws*, Systems Control Lett., 48 (2003), pp. 313–328. Optimization and control of distributed systems.

[33] Alexis Vasseur, *Strong traces for solutions of multidimensional scalar conservation laws*, Arch. Ration. Mech. Anal., 160 (2001), pp. 181–193.