Adjoint-Based Derivative Computations for the Optimal Control of Discontinuous Solutions of Hyperbolic Conservation Laws

Stefan Ulbrich

Zentrum Mathematik
Technische Universität München
München, Germany

Technical Report, September 2001.

To appear in Systems & Control Letters (2002).

Adjoint-based derivative computations for the optimal control of discontinuous solutions of hyperbolic conservation laws [1]

Stefan Ulbrich

Zentrum Mathematik, Technische Universität München, 80290 München, Germany

Abstract

We propose a rigorous procedure to obtain the adjoint-based gradient representation of cost functionals for the optimal control of discontinuous solutions of conservation laws. Hereby, it is not necessary to introduce adjoint variables for the shock positions. Our approach is based on stability properties of the adjoint equation. We give a complete analysis for the case of convex scalar conservation laws. The adjoint equation is a transport equation with discontinuous coefficients and special reversible solutions must be considered to obtain the correct adjoint-based gradient formula. Reversible solutions of the adjoint transport equation and the required stability properties are analyzed in detail.

Keywords: optimal control, adjoint state, conservation laws, shocks, differentiability, linear transport equations, discontinuous coefficients

1 Introduction

This paper is concerned with the justification of adjoint-based derivative calculations for optimal control problems governed by nonlinear hyperbolic conservation laws with source term. We propose an approach that handles the presence of shock discontinuities in a rigorous way without introducing the shock position as additional state variable. Our analysis is based on the variational calculus for conservation laws developed in Ulbrich [19, 20] and on a detailed study of the stability properties of the adjoint equation – a transport equation with discontinuous coefficient – with respect to its coefficients. Stability properties of the adjoint equation played already a fundamental role for the adjoint-based shock-sensitivity analysis in [19], where we could only sketch the required stability results. The present paper provides a detailed analysis of the adjoint equation. Moreover, we describe a general procedure to obtain also in the case of shocks a rigorous adjoint-based gradient formula for tracking-type functionals.

As a model problem we consider the state equation

$$y_t + f(y)_x = g(t,x,y,u_1), \quad (t,x) \in (0,T) \times \mathbb{R} \stackrel{\text{def}}{=} \Omega_T, \qquad y(0,x) = u_0(x), \quad x \in \mathbb{R}. \tag{1}$$

Hereby, $f : \mathbb{R} \to \mathbb{R}$ is a twice continuously differentiable strictly convex flux function, $u = (u_0, u_1) \in L^\infty(\mathbb{R}) \times L^\infty(\Omega_T)^m$, $m \in \mathbb{N}$, is the control and $g : \Omega_T \times \mathbb{R} \times \mathbb{R}^m$ is a source term. Detailed regularity assumptions on u and g will be given later.

State equations of the form (1) arise, e.g., in model problems for the control of traffic flow [15] or for the optimal design of a duct with flow governed by the quasi-1-D Euler equations [3, 7, 12]. Thus, the scalar inhomogeneous conservation law (1) provides a useful basis for a rigorous analysis of optimal control problems governed by conservation laws.

Email address: `sulbrich@ma.tum.de` (Stefan Ulbrich).

[1] This work was supported by Deutsche Forschungsgemeinschaft under grant UL158-2/1.

It is very well known that even for smooth u and g solutions of the inhomogeneous conservation law (1) develop in general discontinuities (shocks) after a finite time and that entropy solutions provide the physically relevant weak solution, see, e.g., [11, 13]. We recall that $y \in L^\infty(\Omega_T)$ is an entropy solution if for all convex functions $\eta : \mathbb{R} \to \mathbb{R}$ with associated entropy flux q satisfying $q' = \eta' f'$ the entropy inequality

$$\eta(y)_t + q(y)_x \le \eta'(y) g(t, x, y, u_1) \quad \text{in } \mathcal{D}'(\Omega_T)$$

holds and if the initial data are assumed in the sense $\lim_{\tau \to 0+} \frac{1}{\tau} \int_0^\tau \|y(t, \cdot) - u_0\|_{1,loc}\, dt = 0$. We will work with the following regularity and growth assumption on g:

(A1) $g \in L^\infty(\Omega_T; C^{0,1}_{loc}(\mathbb{R} \times \mathbb{R}^m)) \cap L^\infty(0, T; C^1_{loc}(\mathbb{R} \times \mathbb{R} \times \mathbb{R}^m))$ and g is Lipschitz continuous w.r.t. x. Moreover, for all $M_u > 0$ there are $C_1, C_2 > 0$ with

$$g(t, x, y, u_1)\, \mathrm{sgn}(y) \le C_1 + C_2 |y|, \quad \text{for all } (t, x, y, u_1) \in \Omega_T \times \mathbb{R} \times [-M_u, M_u]^m.$$

Then it can be shown, see [18–20] and also [11], that for all $u \in L^\infty(\mathbb{R}) \times L^\infty(\Omega_T)^m$ there exists a unique entropy solution $y \in L^\infty(\Omega_T) \cap C([0, T]; L^1_{loc}(\mathbb{R}))$ of (1) which we denote by $y(\cdot; u)$.

For concreteness, we will consider optimal control problems of the form

$$\min_{u \in \mathcal{U}_{ad}} J(y(u)) + R(u) \quad \text{subject to} \quad y = y(u) \text{ solves (1),} \tag{2}$$

where $\mathcal{U}_{ad}$ is a set of admissible controls, R is a regularization term and J is a general tracking-type functional

$$u \longmapsto J(y(u)) = \int_I \psi(y(\bar{t}, x; u), y_d(x))\, dx \tag{3}$$

with $\bar{t} \in (0, T)$, an interval $I = [A, B]$, $\psi \in C^{1,1}_{loc}(\mathbb{R}^2)$, and data $y_d \in BV(I)$. Then the existence of optimal solutions for (2) is ensured if $\mathcal{U}_{ad}$ is bounded in $L^\infty(\mathbb{R}) \times L^\infty(\Omega_T)^m$ and compact in $L^1_{loc}(\mathbb{R}) \times L^1_{loc}(\Omega_T)^m$ and if the regularization term $R : L^p_{loc}(\mathbb{R}) \times L^p_{loc}(\Omega_T)^m \to \mathbb{R}$ is lower semicontinuous for some $p \in [1, \infty)$, see [18, 20].

In order to justify the application of gradient-based methods for the solution of the control problem (2) it is necessary to obtain differentiability results for the functional (3). Hereby, the presence of shocks poses severe difficulties, since the variation of shock positions enters the variation of (3). This requires a careful study of shock sensitivities, sensitivity equation, and adjoint equation to obtain differentiability results and sensitivity- or adjoint-based derivative formulas for the objective functional. In fact, due to the variation of shocks the control-to-state mapping $u \mapsto y(\bar{t}, \cdot; u)$ is at best differentiable if, e.g., the weak topology of local measures is used on the state space instead of the strong L^1_{loc}-norm (see [2] and [20] for weak differentiability results), but this topology is too weak to imply differentiability results for the objective functional (3). In the recent paper [19], see also [20], we have therefore proposed the concept of shift-differentiability. It is based on nonlinear shift-variations that take into account the shift of shocks in a specific way. This enabled us to obtain in [19, 20] differentiability results for tracking-type functionals J of the form (3). In the present paper we describe a general procedure to derive an adjoint-based gradient representation for J that is rigorous for solutions with shocks and does nevertheless not require to introduce the shock locations as additional state variables, which would lead to inconvenient additional adjoint states for the shock locations. However, the weak differentiability properties of $u \mapsto y(\bar{t}, \cdot; u)$ in the case of shocks are not strong enough to apply the classical adjoint calculus for the functional (3). We note that the linearization of the state equation (1) has necessarily only measure-solutions with singular part on the shock set carrying the shock sensitivities information. Therefore, the formal sensitivity equation

$$\delta y_t + (f'(y)\delta y)_x = g_y\, \delta y + g_{u_1}\, \delta u_1, \quad (t, x) \in \Omega_T, \qquad \delta y(0, x) = \delta u_0(x), \quad x \in \mathbb{R}. \tag{4}$$

must be used with care, since it contains the product $f'(y)\delta y$ of the discontinuous function $f'(y)$ and a measure δy, see [1, 9, 20]. On the other hand, the formal adjoint equation is a transport equation with the discontinuous coefficient $f'(y)$ and admits many solutions, see section 3, which requires the characterization of the ”correct” adjoint state.

Our approach omits these difficulties by considering an ”averaged” variant of the linearization (4) and the associated ”averaged” adjoint equation that take automatically care of the shock sensitivities. Using the variational calculus recently developed in Ulbrich [19, 20] and by extending existence and stability results of Bouchut and James [1] on *reversible solutions* of linear transport equations with discontinuous coefficients we will then take the limit in the "averaged" adjoint equation. This yields a rigorous adjoint-based gradient formula for tracking-type functionals (3) and provides the appropriate interpretation of the adjoint equation itself. Hereby, we will allow the presence of rarefaction waves (for convex flux f generated by upward-jumps in the initial data) which requires some care in the analysis of the adjoint equation.

The results of this paper provide also an analytical framework for the convergence analysis of numerical schemes for adjoint-computations. We will addresse this topic in a forthcoming paper.

The paper is organized as follows. In section 2 we derive an adjoint-based gradient representation for objective funtionals (3). Hereby, we will use existence and stability properties for solutions of an "averaged" adjoint equation that is a transport equation with discontinuous coefficient. In section 3 we give a detailed analysis of this type of transport equations. By extending previous results of Bouchut and James [1] on *reversible* solutions of transport equations we provide in particular the existence and stability properties needed in section 2.

Notations. We use standard notations with the following exceptions: for a set S we denote by $B(S)$ the space of bounded functions $v : S \mapsto \mathbb{R}$ equipped with the sup-norm. For open S we denote by $C^k(S)$ the space of functions with continuous *bounded* derivatives up to order k with norm $\|v\|_{C^k(S)} = \sum_{|\beta|\le k} \|D^\beta v\|_{\infty,S}$. $C^k(S^{cl})$ is the subspace of all $v \in C^k(S)$ such that $D^\beta v$, $|\beta| \le k$, admit a continuous extension to S^{cl}. For $x_0 < x_1 < \dots x_N < x_{N+1}$ and $I = [x_0, x_{N+1}]$ we denote by $PC^k(I; x_1, \dots, x_N)$ the space of piecewise C^k-functions with possible jumps at $x_1, \dots, x_N$, endowed with $\|v\|_{PC^k(I;x_1,\dots,x_N)} = \sum_{i=0}^N \|v\|_{C^k([x_i,x_{i+1}])}$.

2 Differentiability of the reduced objective functional and adjoint-based gradient formula

2.1 Properties of the control-to-state mapping

We have already mentioned that for controls in L^∞ there exists a unique entropy solution $y \in L^\infty(\Omega_T) \cap C([0,T]; L^1_{loc}(\mathbb{R}))$. Moreover precisely, we have the following result, see [11, 14, 16, 18–20].

Proposition 1. (Existence, stability, and Oleinik's entropy condition)
Let (A1) *hold. Then for all $u = (u_0, u_1) \in L^\infty(\mathbb{R}) \times L^\infty(\Omega_T)^m \stackrel{\text{def}}{=} \mathcal{U}_\infty$ there exists a unique entropy solution $y = y(\cdot; u) \in L^\infty(\Omega_T)$ of* (1)*. Moreover, $y \in C([0,T]; L^1_{loc}(\mathbb{R}))$ after modification on a set of measure zero.*

Let $M_u > 0$ and $\mathcal{U}_{ad} \subset \{u \in \mathcal{U}_\infty : \|u_0\|_\infty \le M_u, \|u_1\|_\infty \le M_u\}$. Then there are $M_y > 0$ and $L_y > 0$ such that for all $u, \hat{u} \in \mathcal{U}_{ad}$ and all $t \in [0,T]$ the stability estimates hold

(i) $\|y(t,\cdot;u)\|_\infty \le M_y$,

(ii) $\|y(t,\cdot;u) - \hat{y}(t,\cdot;\hat{u})\|_{1,[a,b]} \le L_y(\|u_0 - \hat{u}_0\|_{1,I_t} + \|u_1 - \hat{u}_1\|_{1,[0,t]\times I_t})$,

where $a < b$ are arbitrary and $I_t \stackrel{\text{def}}{=} [a - tM_{f'}, b + tM_{f'}]$, $M_{f'} \stackrel{\text{def}}{=} \max_{|y|\le M_y} |f'(y)|$.

Let in addition $f'' \ge m_{f''} > 0$ hold and set $\hat{\mathcal{U}}_{ad} = \{u \in \mathcal{U}_{ad} : \|u_1\|_{L^\infty(0,T;C^1(\Omega_T)^m)} \le M_u\}$. Then there exists a constant $M > 0$ such that for all $u \in \hat{\mathcal{U}}_{ad}$ and all $t \in (0,T]$ Oleinik's entropy condition

$$y_x(t,\cdot;u) \le \frac{1}{(1 - e^{-cMt})M^{-1} + e^{-cMt}(\sup\{Lip^+(u_0), M\})^{-1}} \tag{5}$$

holds in the sense of distributions, where $c \stackrel{\text{def}}{=} m_{f''}$, $Lip^+(u_0) \stackrel{\text{def}}{=} \text{ess sup}_{x\ne z} \left(\frac{u_0(x)-u_0(z)}{x-z}\right)_+$. Hereby, $M > 0$ depends on the Lipschitz constants of $g(t,x,y,u_1(t,x))$ w.r.t. x, y and $g \equiv 0$ allows the limit $M \to 0$. In particular, $y(t,\cdot) \in BV_{loc}(\mathbb{R})$ for all $t \in (0,T]$ and $y \in BV([\sigma,T] \times [-R,R])$ for all $\sigma, R > 0$.

Since ψ is locally Lipschitz-continuous, it is obvious by (i) and (ii) of Proposition 1 that the mapping $u \in (\mathcal{U}_{ad} \subset L^1_{loc}) \mapsto J(y(u))$ is Lipschitz-continuous for J given by (3) and bounded $\mathcal{U}_{ad} \subset L^\infty$. However, the mapping $u \mapsto y(\bar{t},\cdot;u) \in L^1_{loc}(\mathbb{R})$ is in general not differentiable if $y(\bar{t},\cdot;u)$ contains a shock, since the variation of the shock position allows at best differentiability results for, e.g., $u \mapsto y(\bar{t},\cdot;u) \in \mathcal{M}_{loc}(\mathbb{R}) -$ weak*, where $\mathcal{M}_{loc}(\mathbb{R}) -$ weak* denotes the space of locally bounded Borel measures on $\mathbb{R}$ equipped with the usual weak*-topology. We refer to [19] for a simple example that illustrates this fact. On the other hand, an application of the chain rule to obtain the differentiability of the functional J in (3) would require the differentiability of $u \mapsto y(\bar{t}, \cdot; u) \in L^p_{loc}(\mathbb{R})$ for some $p > 1$, which does not hold in the case of shocks.

2.2 *Differentiability of the objective functional*

The previous considerations show that the differentiability of the objective functional can only be obtained if the sensitivity of shocks and their contribution to the variation of the objective functional are studied in detail. To this purpose, we have proposed in [19, 20] the concept of shift-differentiability that is based on a first-order approximation in L^1_{loc} of the actual variation $y(\bar{t}, \cdot; u + \delta u) - y(\bar{t}, \cdot; u)$ by a nonlinear shift-variation. The usefulness of shift-differentiability lies in the fact that it implies the Fréchet-differentiability of tracking-type functionals J in (3). More precisely, let $\mathcal{U}$ be a control space, fix some $u \in \mathcal{U}$ and assume that $y(\bar{t}, \cdot; u)$ is piecewise C^1 on a neighborhood of $I = [A, B]$ with shocks at $A < x_1 < \ldots < x_K < B$. Similar to Fréchet-differentiability we call $u \in \mathcal{U} \mapsto y(\bar{t}, \cdot; u) \in L^1(I)$ *shift-differentiable* at u if there exists a bounded linear operator

$$D_s y(\bar{t}, \cdot; u) : \quad \delta u \in \mathcal{U} \longmapsto (\delta y^{\bar{t}}, \delta x_1, \ldots, \delta x_K) \in L^r(I) \times \mathbb{R}^K, \quad r > 1, \tag{6}$$

such that with $(\delta y^{\bar{t}}, \delta x) = D_s y(\bar{t}, \cdot; u) \cdot \delta u$

$$\left\| y(\bar{t}, \cdot; u + \delta u) - y(\bar{t}, \cdot; u) - \delta y^{\bar{t}} - \Sigma^{(x_i)}_{y(\bar{t}, \cdot; u)}(\delta x) \right\|_{1,I} = o(\|\delta u\|_{\mathcal{U}})$$

holds, where the *shift-correction* $\Sigma^{(x_i)}_{y(\bar{t}, \cdot; u)}(\delta x)$ is defined by

$$\Sigma^{(x_k)}_{y(\bar{t}, \cdot; u)}(\delta x_1, \ldots, \delta x_K)(x) \stackrel{\text{def}}{=} \sum_{k=1}^{K} [y(\bar{t}, x_k; u)] \operatorname{sgn}(\delta x_k) \mathbf{1}_{I(x_k, x_k + \delta x_k)}(x), \quad x \in I.$$

Hereby, $[y(\bar{t}, x_k; u)] \stackrel{\text{def}}{=} y(\bar{t}, x_k -; u) - y(\bar{t}, x_k +; u)$ denotes the jump across x_k and $I(x_k, x_k + \delta x_k)$ is the interval enclosed between the minimum and maximum of $x_k, x_k + \delta x_k$.

Remark 2. $u \mapsto y(\bar{t}, \cdot; u)$ is for example shift-differentiable if the shock locations of $y(\bar{t}, \cdot; u)$ vary smoothly and connect smoothly varying states.

Once the shift-differentiability of $u \in \mathcal{U} \mapsto y(\bar{t}, \cdot; u) \in L^1(I)$ is shown, it follows the Fréchet-differentiability of tracking-type functionals J in (3) as long as y_d is continuous at $x_1, \ldots, x_K$, see [19, 20]. In fact, the Fréchet-derivative of (3) is given by

$$d_u J(y(u)) \cdot \delta u = (\psi_y(y(\bar{t}, \cdot; u), y_d), \delta y^{\bar{t}})_{2,I} + \sum_{k=1}^{K} \bar{\psi}_y(x_k)[y(\bar{t}, x_k; u)]\, \delta x_k, \tag{7}$$

where $\bar{\psi}_y(x)$ is the everywhere defined "mean value" representative of $\psi_y(y(\bar{t}, \cdot; u), y_d)$

$$\bar{\psi}_y(x) \stackrel{\text{def}}{=} \int_0^1 \psi_y(\tau y(\bar{t}, x+; u) + (1-\tau) y(\bar{t}, x-; u), \tau y_d(x+) + (1-\tau) y_d(x-))\, d\tau. \tag{8}$$

Remark 3. Of course, $\psi_y(y(\bar{t}, \cdot; u), y_d)$ in the first term of (7) can be replaced by $\bar{\psi}_y$. □

We have the following result [19, 20].

Theorem 4. (Shift-differentiability of entropy solutions, differentiability of objective functionals)
Let (A1) *hold, let $f'' \geq m_{f''} > 0$ and let g be affine linear w.r.t. y. Consider for arbitrary $z_1 < \ldots < z_N$ the control space $\mathcal{U} = PC^1(\mathbb{R}; z_1, \ldots, z_N) \times L^\infty(0, T; C^1(\mathbb{R})^m)$.*

Then for $I = [A, B]$, $\bar{t} \in (0, T]$, the mapping $u \in \mathcal{U} \longmapsto y(\bar{t}, \cdot; u) \in L^1(I)$ with $y(\cdot; u)$ denoting the entropy solution of (1) *is shift-differentiable at any $u \in \mathcal{U}$ such that $y(\bar{t}, \cdot; u)$ has on I no shock generation points and finitely many nondegenerate shocks $A < x_1 < \ldots < x_K < B$ that are all no shock interaction points. Moreover, the objective functional* (3) *is Fréchet-differentiable at u with derivative* (7) *if y_d is continuous at $x_1, \ldots, x_K$.*

Proof. The proof can be obtained by a careful application of Dafermos' theory of generalized characteristics [5] and can be found in [19, 20]. □

2.3 Adjoint-based gradient representation

We will now use the differentiability result of Theorem 4 to convert (7) to a more convenient adjoint-based gradient representation. In view of numerical approximations and the wish to handle complicated shock structures it is hereby highly desirable that the resulting adjoint- or sensitivity-based derivative formulas *do not* require the introduction of additional adjoint states or sensitivties for the shock position. Rather, one would like to find derivative formulas that are based on appropriately defined measure-solutions of the sensitivity equation (4) or the solution of the corresponding adjoint equation, respectively.

We have already observed that the differentiability properties of $u \mapsto y(\cdot; u)$ do not allow an application of the classical adjoint calculus to obtain a gradient representation for J. Therefore, we proceed more carefully. Let with the notations of Theorem 4 the mapping $u \longmapsto y(\bar{t}; \cdot; u)$ be shift-differentiable at $u \in \mathcal{U}$. Now let $\delta u \in \mathcal{U}$ be arbitrary and set $\hat{u} \stackrel{\text{def}}{=} u + \delta u$, $\hat{y} \stackrel{\text{def}}{=} y(\cdot; u + \delta u)$, $y = y(\cdot; u)$, and $\Delta y \stackrel{\text{def}}{=} \hat{y} - y$. Using that $y, \hat{y} \in L^\infty(\mathbb{R}) \cap C([0,T]; L^1_{loc}(\mathbb{R}))$ are weak solutions of (1) we obtain by subtracting (1) for $\hat{y}$ and y that

$$\Delta y_t + (\bar{f}'(\hat{y}, y)\Delta y)_x = g(\cdot, \hat{y}, \hat{u}_1) - g(\cdot, y, u_1) \quad \text{in } \mathcal{D}'(\Omega_T), \quad \Delta y(0, \cdot) = \delta u_0$$

with the averaged coefficient

$$\bar{f}'(\hat{y}, y) \stackrel{\text{def}}{=} \int_0^1 f'(\tau\hat{y} + (1-\tau)y)\, d\tau.$$

Since $y, \hat{y} \in C([0,T]; L^1_{loc}(\mathbb{R})) \cap L^\infty(\mathbb{R})$, this yields for any test function satisfying

$$\hat{p} \in L^\infty(\Omega_{\bar{t}}) \cap C([0,\bar{t}]; L^2(\mathbb{R})) \cap C^{0,1}([\sigma, \bar{t}] \times \mathbb{R}) \text{ for all } \sigma \in (0, \bar{t}), \quad \text{supp}_x(\hat{p}) \text{ bounded} \tag{9}$$

the identity

$$\begin{aligned}(\hat{p}(\bar{t}, \cdot), \Delta y(\bar{t}, \cdot))_2 = (\hat{p}(0, \cdot), \delta u_0)_2 + (\hat{p}_t + \bar{f}'(\hat{y}, y)\hat{p}_x + g_y(\cdot, y, \hat{u}_1)\hat{p}, \Delta y)_{2,\Omega_{\bar{t}}} \\ + (\hat{p}, g(\cdot, y, \hat{u}_1) - g(\cdot, y, u_1))_{2,\Omega_{\bar{t}}},\end{aligned} \tag{10}$$

where we have used that g is by assumption affine linear w.r.t. y. As we will show in section 3, it is possible to construct for given $p^{\bar{t}} \in C^{0,1}_c(\mathbb{R})$ a *reversible* solution $\hat{p}$ for the "averaged" adjoint equation

$$\hat{p}_t + \bar{f}'(\hat{y}, y)\hat{p}_x = -g_y(\cdot, y, \hat{u}_1)\hat{p}, \quad (t,x) \in \Omega_{\bar{t}}, \qquad \hat{p}(\bar{t}, x) = p^{\bar{t}}(x), \quad x \in \mathbb{R} \tag{11}$$

that has the regularity (9), see Corollary 15, as well as the following stability property, see Theorem 17: For all $\sigma \in (0, \bar{t})$, $r \in [1, \infty)$ and any ε-neighborhood F_ε of the set of rarefaction centers $F = \{z \in \mathbb{R} : [u_0(z)] < 0\}$ we have

$$\left.\begin{aligned}&\hat{p} \to p \quad \text{in } C([0,\bar{t}]; L^r(\mathbb{R})) \cap C([\sigma, \bar{t}] \times \mathbb{R}) \cap C([0,\bar{t}] \times (\mathbb{R} \setminus F_\varepsilon)), \\ &\|\hat{p}\|_\infty \le M_p, \quad \operatorname*{supp}_x(\hat{p}) \text{ uniformly bounded}\end{aligned}\right\} \quad \text{as } \|\delta u\|_{\mathcal{U}} \to 0 \tag{12}$$

with a constant $M_p > 0$. Hereby, p satisfies again (9) and is the reversible solution of

$$p_t + \bar{f}'(y)p_x = -g_y(\cdot, y, u_1)p, \quad (t,x) \in \Omega_{\bar{t}}, \qquad p(\bar{t}, x) = p^{\bar{t}}(x), \quad x \in \mathbb{R}. \tag{13}$$

Our aim is to take the limit in (10) for regular end data $p^{\bar{t}}$. Then we let $p^{\bar{t}}$ converge to $\bar{\psi}|_I$, cf. (7), (8).

Clearly, the shift-differentiability result of Theorem 4 holds also for a slightly larger interval $\bar{I} = [A - \rho, B + \rho]$. We thus obtain with $(\delta y, \delta x) = D_s y(\bar{t}, \cdot; u) \cdot \delta u$

$$(p^{\bar{t}}, \Delta y(\bar{t}, \cdot))_{2,\bar{I}} = (p^{\bar{t}}, \delta y^{\bar{t}})_{2,\bar{I}} + \sum_{k=1}^K \int_{x_k}^{x_k + \delta x_k} p^{\bar{t}}(x)\, dx\, [y(\bar{t}, x_k; u)] + o(\|p^{\bar{t}}\|_\infty \|\delta u\|_{\mathcal{U}}). \tag{14}$$

Using (A1), (11), (12) and (14) we deduce from (10) that

$$\begin{aligned}(p^{\bar{t}}, \delta y^{\bar{t}})_{2,\bar{I}} + \sum_{k=1}^K p^{\bar{t}}(x_k)[y(\bar{t}, x_k)]\delta x_k = (p(0, \cdot), \delta u_0)_2 + (g_{u_1}^T(\cdot, y, u_1)p, \delta u_1)_{2,\Omega_{\bar{t}}} + o(M_p\|\delta u\|_{\mathcal{U}}) \\ + O\left(\|p^{\bar{t}}\|_{1,\mathbb{R}\setminus\bar{I}} + \textstyle\sum_k \|p^{\bar{t}} - p^{\bar{t}}(x_k)\|_{1,I(x_k, x_k + \delta x_k)}\right).\end{aligned} \tag{15}$$

On the other hand, J in (3) has by Theorem 4 the derivative (7). Comparing (7) and (15), we would like to choose the discontinuous end data $p^{\bar t} = \bar\psi_y|_I$, which requires an extension of the admissible end data in (13). We will show in section 3, Corollary 15, that for end data

$$p^{\bar t} \in B_{Lip}(\mathbb{R}) \stackrel{\text{def}}{=} \left\{ p^{\bar t} \in B(\mathbb{R}) \ : \ \begin{array}{l} p^{\bar t} \text{ is the pointwise everywhere limit of a sequence} \\ (p_n^{\bar t}) \text{ in } C^{0,1}(\mathbb{R}),\ (p_n^{\bar t}) \text{ bounded in } C(\mathbb{R}) \cap W^{1,1}_{loc}(\mathbb{R}) \end{array} \right\}$$

there exists a reversible solution

$$p \in B(\Omega_{\bar t}) \cap C^{0,1}([0,\bar t]; L^1_{loc}(\mathbb{R})) \cap B([0,\bar t]; BV_{loc}(\mathbb{R})) \cap BV_{loc}(\Omega_{\bar t}^{cl})$$

of (13) such that for any such sequence $(p_n^{\bar t})$ in $C^{0,1}(\mathbb{R})$ and corresponding solution sequence (p_n) of (13)

$$p_n \to p \ \text{ boundedly everywhere on } ((0,\bar t] \times \mathbb{R}) \cup ([0,\bar t] \times (\mathbb{R} \setminus F)) \text{ and in } C([0,\bar t]; L^1_{loc}(\mathbb{R})). \tag{16}$$

Since $y(\bar t, \cdot), y_d \in BV(I)$, it is easy to see that $p^{\bar t} = \bar\psi_y|_I \in B_{Lip}(\mathbb{R})$ and without restriction we can choose the approximating sequence $(p_n^{\bar t})$ such that $p_n^{\bar t} \in C_c^{0,1}(\bar I)$, $p_n^{\bar t}|_{I(x_k, x_k+\delta x_k)} \equiv \bar\psi_y(x_k)$.

With $r > 1$ from (6) we have $|(p_n^{\bar t} - \bar\psi_y, \delta y^{\bar t})_{2,\bar I}| \le \|p_n^{\bar t} - \bar\psi_y|_I\|_{r',\bar I} \|\delta y^{\bar t}\|_{r,I} = O(\|p_n^{\bar t} - \bar\psi_y|_I\|_{r',\bar I} \|\delta u\|_{\mathcal{U}})$, $1/r' + 1/r = 1$. Since (15) holds for $p_n^{\bar t}$ and corresponding p_n we thus obtain with (7)

$$\begin{aligned} d_u J(y(u)) \cdot \delta u &= (p(0,\cdot), \delta u_0)_2 + (g_{u_1}^T(\cdot, y, u_1) p, \delta u_1)_{2,\Omega_{\bar t}} \\ &\quad + O\left((\|p_n^{\bar t} - \bar\psi_y|_I\|_{r',\bar I} + \|p_n - p\|_{C([0,\bar t]; L^1(\mathbb{R}))}) \|\delta u\|_{\mathcal{U}} \right) + o(M_{p_n} \|\delta u\|_{\mathcal{U}}). \end{aligned}$$

Now we have $\|p_n^{\bar t} - \bar\psi_y|_I\|_{r',\bar I} \to 0$ by the Lebesgue convergence theorem, $\text{supp}_x(p_n)$ are uniformly bounded and thus $\|p_n - p\|_{C([0,\bar t]; L^1(\mathbb{R}))} \to 0$ by (16) for $n \to \infty$. Therefore, the last two terms are $o(\|\delta u\|_{\mathcal{U}})$ and we conclude that

$$d_u J(y(u)) \cdot \delta u = (p(0,\cdot), \delta u_0)_2 + (g_{u_1}^T(\cdot, y, u_1) p, \delta u_1)_{2,\Omega_{\bar t}}, \tag{17}$$

where p is the reversible solution of (13) for data $p^{\bar t} = \bar\psi_y$ given by (8). Although we have still to introduce appropriate *reversible* solutions of (11) and (13) having the stability properties (12) and (16), which will be done in the next section, we already state the following result.

Theorem 5. *Let the assumptions of Theorem 4 hold. Then the gradient respresentation of J in* (3) *with respect to the scalar product of $L^2(\mathbb{R}) \times L^2(\Omega_{\bar t})^m$ is given by* (17)*, where p is the reversible solution of* (13) *for data $p^{\bar t} = \bar\psi_y$ given by* (8)*.*

Remark 6. We emphasize that the described approach for the derivation of the adjoint-based gradient representation can also be applied to systems of conservation laws. It is rigorous as soon as $u \mapsto y(\bar t, \cdot; u)$ is shift-differentiable and the averaged adjoint equation is stable with respect to coefficients and end data. □

Before we turn to the study of transport equations of the form (11) and (13) in the next section we collect some properties of the coefficients. For convenience, we set

$$\hat a \stackrel{\text{def}}{=} \bar f'(\hat y, y), \quad \hat b \stackrel{\text{def}}{=} g_y(\cdot, y, \hat u_1), \quad a \stackrel{\text{def}}{=} f'(y), \quad b \stackrel{\text{def}}{=} g_y(\cdot, y, u_1).$$

By the assumptions of Theorem 4 (A1) holds and g_y does not depend on y. It is now not difficult to see that

$$\hat b, b \in L^\infty(0,T; C^{0,1}(\mathbb{R})) \quad \text{and} \quad \hat b \to b \ \text{ in } L^\infty(0,T; C(\mathbb{R})) \text{ as } \|\delta u\|_{\mathcal{U}} \to 0. \tag{18}$$

Moreover, we have by Proposition 1, (i)–(ii), that

$$\hat a \to a \ \text{ in } L^1_{loc}(\Omega_T) \text{ and in } L^\infty(\Omega_T)\text{-weak}^* \text{ as } \|\delta u\|_{\mathcal{U}} \to 0. \tag{19}$$

Finally, $f'' > 0$ and Oleinik's condition (5) yield $C > 0$ and for any $\varepsilon > 0$ a constant $M_\varepsilon > 0$ with

$$\hat a_x(t,\cdot), a_x(t,\cdot) \le \frac{C}{t}, \quad \hat a_x(t,\cdot)|_{\mathbb{R} \setminus F_\varepsilon^{cl}}, a_x(t,\cdot)|_{\mathbb{R} \setminus F_\varepsilon^{cl}} \le M_\varepsilon \ \text{ for } \|\delta u\|_{\mathcal{U}} \le 1 \tag{20}$$

in the sense of distributions, where F_ε denotes again the ε-neighborhood of $F = \{z \in \mathbb{R} : [u_0(z)] < 0\}$. Hereby and throughout we use the following convenient convention.

$$\text{For } F = \emptyset \text{ we set } F_\varepsilon = \emptyset \text{ for all } \varepsilon > 0. \tag{21}$$

The second estimate in (20) follows from (5) by using the Lip^+-boundedness of $u_0, \hat{u}_0$ on $\mathbb{R} \setminus F$ and the finite propagation speed of $\hat{y}, y$.

It remains to analyze the linear transport equations (11), (13) and to justify the stability properties (12) and (16). This will be carried out in the next section.

3 The adjoint equation: Linear transport equations with discontinuous coefficient

We consider the backward problem for a transport equation of the form

$$p_t + ap_x = -bp + c, \quad (t, x) \in \Omega_\tau \stackrel{\text{def}}{=} (0, \tau) \times \mathbb{R} \qquad p(\tau, \cdot) = p^\tau, \tag{22}$$

where $\tau \in (0, T]$, $b, c \in L^\infty(0, T; C^{0,1}(\mathbb{R}))$ and $a \in L^\infty(\Omega_T)$ satisfies the one-sided Lipschitz condition (OSLC)

$$a_x(t, \cdot) \le \alpha(t), \quad \alpha \in L^1(0, T), \tag{23}$$

or at least the weakened one-sided Lipschitz condition

$$a_x(t, \cdot) \le \alpha(t), \quad \alpha \in L^1(\sigma, T), \quad \text{for all } \sigma \in (0, T).$$

As observed earlier, the latter case is appropriate if we want to consider the adjoint equation for solutions with rarefaction waves. The adjoint equation (13) and the averaged adjoint equation (11) are of the form (22). In view of (20) we use the more flexible assumption

$$a_x(t, \cdot) \le \alpha(t), \quad \alpha \in L^1(\sigma, T), \quad \text{for all } \sigma \in (0, T), \qquad a_x(t, \cdot)|_{\mathbb{R}\setminus E} \le \tilde{\alpha}(t), \quad \tilde{\alpha} \in L^1(0, T) \tag{24}$$

with a closed set $E \subset \mathbb{R}$ and ε-neighborhood E_ε. This contains (23) if we choose $E = \emptyset$. Clearly, (20) implies (24) for all sets $E = F_\varepsilon^{cl}$, $\varepsilon > 0$, with the convention (21).

Transport equations (22) have been studied in a different context by several authors under the strong OSLC (23). Conway [4] shows that under the strong OSLC (23) for Lipschitz continuous b, c and any $p^\tau \in Lip(\mathbb{R})$ there exists a Lipschitz continuous solution to (22) which is not necessarily unique. For the non-uniqueness of Lipschitz solutions we refer to the simple sgn-example given by Conway [4] with $a(t, x) = -\operatorname{sgn}(x)$ and $b, c \equiv 0$. Similar results were obtained in the context of uniqueness proofs for (1) in [10, 14] and of error estimates for approximate solutions of (1) in [17]. To ensure uniqueness and stability of solutions, Bouchut and James introduce in the recent paper [1] for a satisfying the strong OSLC (23), $b, c \equiv 0$ and data $p^\tau \in Lip(\mathbb{R})$ special *reversible* solutions that are unique and stable with respect to a. The reversible solutions of [1] are not directly extendible to the general case $b, c \not\equiv 0$ but will nevertheless form the basis of our approach, since they provide a generalized backward flow associated with a that will allow us to define reversible solutions by using the characteristic equation. The results of this section extend and augment the existing results in the following directions:

- We work under the weakened OSLC (24). This is essential to handle the adjoint equation (13) in the case of rarefaction waves.
- We admit discontinuous end data. We have already seen that this is essential to obtain the adjoint-based gradient representation (17) for tracking type functionals (3).
- We cover the nonhomogeneous case $b, c \not\equiv 0$ which is necessary to handle adjoint equations for conservation laws with source term, where $b \not\equiv 0$, as well as adjoint equations for cost functionals with distributed observation, where $c \not\equiv 0$.
- We will derive precise regularity results of reversible solutions, also for the case of the weakened OSLC (24). This yields in particular quite precise regularity results for the L^2-gradient (17) of (3).

To motivate our definition of reversible solutions of (22) it is instructive to consider the case of smooth coefficients and data. If a, b, c, p^τ are smooth then it is well known that (22) admits unique classical solutions given by the characteristic equation. In fact, let $D_b = \{(s, t) \in \mathbb{R}^2 \,:\, 0 \le t \le s \le T\}$ and define the characteristic backward flow $X : D_b \times \mathbb{R} \longrightarrow \mathbb{R}$ by requiring that for all $(t, x) \in \Omega_T$

$$X(t; t, x) = x, \quad \frac{d}{ds} X(s; t, x) = a(s, X(s; t, x)), \;\; s \in [t, T]. \tag{25}$$

Then the solution of (22) is given by the characteristic equation

$$\left.\begin{aligned} p(\tau, X(\tau;t,x)) &= p^\tau(X(\tau;t,x)) \\ \frac{d}{ds}p(s, X(s;t,x)) &= (-bp+c)(s, X(s;t,x)), \quad s \in (t,\tau), \end{aligned}\right\} \; (t,x) \in \Omega_T. \tag{26}$$

Our definition of reversible solutions p for the case of discontinuous a can be motivated as follows: Obviously, the backward flow X satisfies the composition formula

$$X(s;t,X(t;\sigma,z)) = X(s;\sigma,z) \quad \text{for all} \quad 0 \le \sigma \le t \le s \le T, \;\; z \in \mathbb{R}. \tag{27}$$

and with $\sigma = 0$ and $x = X(t;0,z)$ we see that (26) is equivalent to

$$\begin{aligned} p(\tau, X(\tau;0,z)) &= p^\tau(X(\tau;0,z)), \;\; z \in \mathbb{R} \\ \frac{d}{dt}p(t, X(t;0,z)) &= (-bp+c)(t, X(t;0,z)), \;\; t \in (0,\tau), \end{aligned} \tag{28}$$

Moreover, since by (27) holds $X(s;t,X(t;0,z)) = X(s;0,z)$ and $X(s;s,x) = x$, we see by (28) that for any $s \in (0,T]$ the function $X(s;\cdot,\cdot)$ is the unique classical solution to

$$X_t(s;\cdot) + aX_x(s;\cdot) = 0, \;\; (t,x) \in (0,s) \times \mathbb{R}, \qquad X(s;s,x) = x, \;\; x \in \mathbb{R}. \tag{29}$$

Thus, $X(s;\cdot,\cdot)$ solves a homogeneous transport equation with coefficient a and can be defined as a reversible solution in the sense of Bouchut and James [1] also for the case, where $a \in L^\infty(\Omega_T)$ satisfies the OSLC (23). This yields a definition of the generalized backward flow X that is stable with respect to perturbations – e.g., smoothing – of a and is thus consistent with the smooth case (it turns out that $X(\cdot;t,x)$ is nothing else but the Filippov-solution of the ODE (25)). This justifies to use (28) for the definition of p and the stability of X will ensure the stability of p with respect to the coefficient a.

3.1 Reversible solutions

We recall several results from Bouchut and James [1] that are the starting point of our approach. Denote by $\mathcal{L}_h$ the space of Lipschitz-continuous solutions to

$$p_t + ap_x = 0, \quad (t,x) \in \Omega_\tau \tag{30}$$

We have seen that solutions with prescribed end data p^τ are in general not unique. To obtain uniqueness Bouchut and James define in [1] for the homogeneous case $b, c \equiv 0$ reversible solutions as follows:

Definition 7. (Reversible solution for $b, c \equiv 0$, [1])
$p \in \mathcal{L}_h$ is called *reversible solution* of (30) if there exist $p_1, p_2 \in \mathcal{L}_h$ such that $(p_1)_x \ge 0$, $(p_2)_x \ge 0$ and $p = p_1 - p_2$. □

We have the following result of [1].

Theorem 8. (Existence and uniqueness of reversible solution for $b, c \equiv 0$, [1])
Let $a \in L^\infty(\Omega_T)$ satisfy the OSLC (23)*. Then for any $p^\tau \in Lip(\mathbb{R})$ there exists a unique reversible solution $p \in C^{0,1}_{loc}(\Omega^{cl}_\tau)$ of* (30) *with $p(\tau,\cdot) = p^\tau$. Moreover,*

$$\|p(t,\cdot)\|_{\infty,I} \le \|p^\tau\|_{\infty,J}, \quad \|p_x(t,\cdot)\|_{\infty,I} \le e^{\int_t^\tau \alpha} \|p^\tau_x\|_{\infty,J}$$

with $I = (x_1, x_2)$, $J = (x_1 - \|a\|_\infty(\tau - t), x_2 + \|a\|_\infty(\tau - t))$.

This convenient characterization of reversible solutions is not extendible to the case $b \ne 0$ or $c \ne 0$. However, we will use the following generalized backward flow for a introduced in [1] together with the characteristic equation (28):

Definition 9. (Generalized backward flow)
Let $D_b = \{(s,t) \in \mathbb{R}^2 \,:\, 0 \le t \le s \le T\}$ and let $a \in L^\infty(\Omega_T)$ satisfy the OSLC (23). Then the generalized backward flow $X \in Lip(D_b \times \mathbb{R})$ associated with a is defined by the requirement that $X(s;\cdot,\cdot)$ is for any $s \in (0,T]$ the unique reversible solution to

$$X_t(s;\cdot,\cdot) + aX_x(s;\cdot,\cdot) = 0, \;\; (t,x) \in (0,s) \times \mathbb{R}, \qquad X(s;s,x) = x, \;\; x \in \mathbb{R}.$$

Moreover, we set $X(0;0,x) = x$. $\square$

One can show [1] that $X(s;t,x)$ satisfies

$$\|X_s\|_{\infty,\overset{\circ}{D}_b\times\mathbb{R}} \le \|a\|_\infty, \quad \|X_t\|_{\infty,\mathbb{R}\times\overset{\circ}{D}_b} \le \|a\|_\infty e^{\int_0^T \alpha}, \quad \|X_x(s;t,\cdot)\|_\infty \le e^{\int_t^s \alpha} \text{ for all } (s,t) \in D_b. \quad (31)$$

Moreover, $X_x \ge 0$, $x \longmapsto X(s;t,x)$ is surjective for all $(s,t) \in D_b$ and

$$X_{sx}(s;t,x) \le \alpha(s) X_x(s;t,x) \quad \text{for a.a. } s \in (0,T) \quad (32)$$

on $\overset{\circ}{D}_b \times \mathbb{R}$, see [1]. Thus, for arbitrary $z_1 < z_2$ and $0 \le t \le \sigma \le s \le T$ we obtain

$$0 \le X(s;t,z_2) - X(s;t,z_1) \le X(\sigma;t,z_2) - X(\sigma;t,z_1) + \int_\sigma^s \alpha(r)(X(r;t,z_2) - X(r;t,z_1))\,dr,$$

and hence the Gronwall lemma yields

$$0 \le X(s;t,z_2) - X(s;t,z_1) \le (X(\sigma;t,z_2) - X(\sigma;t,z_1)) e^{\int_\sigma^s \alpha} \quad \text{for all } \ t \le \sigma \le s \le T. \quad (33)$$

Moreover, it is shown in [1] that the composition formula (27) holds and that for any $(t,x) \in \Omega_T$

$$X_s(s;t,x) \in [a(s, X(s;t,x)+), a(s, X(s;t,x)-)] \quad \text{for a.a. } s \in (0,T). \quad (34)$$

We recall that the one-sided Lipschitz condition (23) or (24) implies that $a(t,\cdot) \in BV_{loc}(\mathbb{R})$ for a.a. t and thus the left- and right-sided limits in (34) exist.

Remark 10. (34) shows that $X(\cdot;t,x)$ is a solution of (25) in the sense of Filippov [6]. On the other hand, it is shown in [6] that the strong OSLC (23) implies existence and uniqueness of Filippov-solutions for (25). Thus, the generalized backward flow X of Definition 9 coincides with the unique flow obtained by solving (25) in the sense of Filippov.

In the case of the adjoint equation (22) we have $a = f'(y)$ with the solution y of (1). Thus, $X(\cdot;t,x)$ is nothing else but the generalized forward characteristic through (t,x) in the sense of Dafermos [5]. $\square$

For our analysis the following stability result of [1] for the generalized backward flow X, which extends a classical stability result of Filippov-solutions [6], will be important.

Theorem 11. (Stability of the generalized backward flow, [1, Thm. 4.1.15])
Let (a_n) be a bounded sequence in $L^\infty(\Omega_T)$ with $a_n \to a$ in $L^\infty(\Omega_T)$-weak and let for a bounded sequence (α_n) in $L^1(0,T)$ and $\alpha \in L^1(0,T)$ the OSLCs hold*

$$(a_n)_x(t,\cdot) \le \alpha_n(t), \quad a_x(t,\cdot) \le \alpha(t) \quad \textit{for a.a. } t \in (0,T).$$

Denote by X_n and X the generalized backward flows associated with a_n and a according to Definition 9, respectively. Then it holds $X_n \to X$ in $C(D_b \times [-R,R])$ for all $R > 0$.

Finally, it is shown in [1] that for $b,c \equiv 0$ the reversible solution of (22) is given by $p(t,x) = p^\tau(X(\tau;t,x))$ and is thus the broad solution according to (28) defined along the generalized characteristics. This motivates our following definition of reversible solutions for (22) in the general case $b,c \not\equiv 0$.

Definition 12. (Reversible solution)
Denote by $B(\mathbb{R})$ the Banach space of bounded functions equipped with the $\sup$-norm and let

$$p^{\bar{t}} \in B_{Lip}(\mathbb{R}) \overset{\text{def}}{=} \left\{ w \in B(\mathbb{R}) \ : \ \begin{array}{l} w \text{ is the pointwise everywhere limit of a sequence} \\ (w_n) \text{ in } C^{0,1}(\mathbb{R}),\ (w_n) \text{ bounded in } C(\mathbb{R}) \cap W^{1,1}_{loc}(\mathbb{R}) \end{array} \right\}$$

Let $a \in L^\infty(\Omega_T)$, $a_x(t,\cdot) \le \alpha(t)$, $\alpha \in L^1(0,T)$ and $b,c \in L^\infty(0,T;C^{0,1}(\mathbb{R}))$. Then a reversible solution of (22) is defined as follows. For any $z \in \mathbb{R}$ define $p(t, X(t;0,z))$ as solution of

$$p(\tau, X(\tau;0,z)) = p^\tau(X(\tau;0,z)), \quad \frac{d}{dt} p(t, X(t;0,z)) = (-bp + c)(t, X(t;0,z)) \text{ for a.a. } t \in (0,\tau). \quad (35)$$

If merely $\alpha \in L^1(\sigma,T)$ for all $\sigma > 0$ holds then we define p first on the domains $(\sigma,\tau) \times \mathbb{R}$ and then on Ω_τ by exhaustion. $\square$

In order to handle the weakened OSLC (24) the following observation is important.

Remark 13. The composition formula (27) yields the identity $X(t;s,x) = X(t;0,z)$ with $x = X(s;0,z)$ for $s \le t \le T$. Therefore, as in the classical case (35) implies that p satisfies for all $0 \le s < \tau$ and $x \in \mathbb{R}$

$$p(\tau, X(\tau;s,x)) = p^\tau(X(\tau;s,x)), \quad \frac{d}{dt}p(t,X(t;s,x)) = (-bp+c)(t,X(t;s,x)) \text{ for a.a. } t \in [s,\tau]. \quad (36)$$

Thus, for $0 \le \sigma < s \le \tau$ the reversible solution of (22) on $(\sigma,\tau)\times\mathbb{R}$ is an extension of the reversible solution of (22) on $(s,\tau)\times\mathbb{R}$. This justifies the construction of p by exhaustion in the case of the weakened OSLC (24). Moreover, we see from (31) that $p(s,x)$ depends only on the values of a, b and c in the triangle $\{(t,z)\in\Omega_\tau \,:\, t\in[s,\tau], z\in[x-\|a\|_\infty(t-s), x+\|a\|_\infty(t-s)]\}$. □

3.2 Existence and uniqueness of reversible solutions

We are now in the position to show the following existence and uniqueness result.

Theorem 14. (Existence, uniqueness and regularity of reversible solutions under strong OSLC) *Let $a \in L^\infty(\Omega_T)$ satisfy the OSLC* (23). *Let $b, c \in L^\infty(0,T;C^{0,1}(\mathbb{R}))$. Then the following holds:*
For all $p^\tau \in C^{0,1}(\mathbb{R})$ there exists a unique reversible solution p of (22). *Moreover, $p \in C^{0,1}(\Omega_\tau^{cl})$ and p solves* (22) *almost everywhere on Ω_τ. Furthermore, $\|p\|_{p\in C^{0,1}(\Omega_\tau^{cl})}$ has independently of τ a bound depending on $\|b\|_{L^\infty(0,T;C^{0,1}(\mathbb{R}))}$, $\|c\|_{L^\infty(0,T;C^{0,1}(\mathbb{R}))}$, $\|p^\tau\|_{C^{0,1}(\mathbb{R})}$, $\|a\|_\infty$, and $\|\alpha\|_1$. Finally, for all $t\in[0,\tau]$, $z_1<z_2$ and $0\le s<\hat{s}\le\tau$ with*

$$I = [z_1,z_2], \quad J = [z_1 - \|a\|_\infty(\tau-t), z_2+\|a\|_\infty(\tau-t)], \quad I_s^{\hat{s}} = [s,\hat{s}]\times I, \quad J_t^\tau = [t,\tau]\times J$$

the following estimates hold:

$$\|p(t)\|_{B(I)} \le (\|p^\tau\|_{B(J)} + \|c\|_{L^1(0,\tau;B(J))})e^{\|b\|_{L^1(0,\tau;B(J))}}, \quad (37)$$

$$\|p_x(t)\|_{1,I} \le (\|p_x^\tau\|_{1,J} + \|b_x\|_{1,J_t^\tau}\|p\|_{\infty,J_t^\tau} + \|c_x\|_{1,J_t^\tau})e^{\|b\|_{L^1(t,\tau;B(J))}}, \quad (38)$$

$$\|p_t\|_{1,I_s^{\hat{s}}} \le (\hat{s}-s)(\|bp-c\|_{L^\infty(s,\hat{s};L^1(I))} + \|a\|_{\infty,I_s^{\hat{s}}}\|p_x\|_{L^\infty(s,\hat{s};L^1(I))}). \quad (39)$$

In particular, one has with constants C, M depending on $\|p^\tau\|_{W^{1,1}(J)}$, $\|a\|_{\infty,[0,\tau]\times J}$, $\|b\|_{L^\infty(0,\tau;W^{1,1}(J))}$, and $\|c\|_{L^\infty(0,\tau;W^{1,1}(J))}$, but not depending on α

$$\|p\|_{W^{1,1}((0,\tau)\times I)} + \|p\|_{B((0,\tau];W^{1,1}(I))} \le M, \quad \|p(\hat{s})-p(s)\|_{1,I} \le C\,(\hat{s}-s). \quad (40)$$

Before we prove this theorem, we state the following corollary that gives an existence and uniqueness result under the weakened OSLC (24) and covers also the case of discontinuous end data.

Corollary 15. (Existence and uniqueness under weakened OSLC and for discontinuous end data) *Let the assumptions of Theorem* 14 *hold with the relaxation that only the weakened OSLC* (24) *is satisfied for a closed set $E\subset\mathbb{R}$. Then the following holds:*

(i) *For any $p^\tau \in C^{0,1}(\mathbb{R})$ there exists a unique reversible solution p of* (22). *Moreover, for all $\sigma\in(0,\tau)$ and $\varepsilon>0$ we have with the convention* (21)

$$p \in B(\Omega_\tau)\cap C^{0,1}([\sigma,\tau]\times\mathbb{R})\cap C^{0,1}([0,\tau]\times(\mathbb{R}\setminus E_\varepsilon))\cap C^{0,1}([0,\tau];L^1_{loc}(\mathbb{R}))\cap B([0,\tau];BV_{loc}(\mathbb{R})),$$

(37)–(40) *hold for all $t\in(0,\tau]$ and p satisfies* (36) *for all $s\in(0,\tau)$. Finally, $\|p\|_{C^{0,1}([0,\tau]\times(\mathbb{R}\setminus E_\varepsilon))}$ has independently of τ a bound depending on ε, $\|b\|_{L^\infty(0,T;C^{0,1}(\mathbb{R}))}$, $\|c\|_{L^\infty(0,T;C^{0,1}(\mathbb{R}))}$, $\|p^\tau\|_{C^{0,1}(\mathbb{R})}$, $\|a\|_\infty$, and $\|\tilde{\alpha}\|_1$.*

(ii) *For end data $p^\tau\in B_{Lip}(\mathbb{R})$ there exists a unique reversible solution $p\in B(\Omega_\tau)$ according to Definition 12. Moreover, p satisfies the pointwise bound* (37) *and*

$$p\in B(\Omega_\tau)\cap C^{0,1}([0,\tau];L^1_{loc}(\mathbb{R}))\cap B([0,\tau];BV_{loc}(\mathbb{R}))\cap BV_{loc}(\Omega_\tau^{cl}).$$

Let (p_n^τ) be any sequence in $C^{0,1}(\mathbb{R})$ that is bounded in $C(\mathbb{R})\cap W^{1,1}_{loc}(\mathbb{R})$ and converges pointwise everywhere to p^τ. Then the corresponding reversible solutions p_n of (22) *according to* (ii) *satisfy*

$$p_n \to p \text{ boundedly everywhere on } ((0,\tau]\times\mathbb{R})\cup([0,\tau]\times(\mathbb{R}\setminus E)) \text{ and in } C([0,\tau];L^1_{loc}(\mathbb{R})).$$

We prove now Theorem 14 and subsequently Corollary 15.

Proof of Theorem 14. We show first that p is well defined. Let $(t,x) \in \Omega_\tau$ be arbitrary. Since $z \longmapsto X(t;0,z)$ is surjective, there is $z \in \mathbb{R}$ with $x = X(t;0,z)$. Now (35) defines the values of p on the curve $(s, X(s;0,z))$, $t \le s \le \tau$. If z is not unique then we get for all $\tilde{z}$ with $x = X(t;0,\tilde{z})$ by (33) that $X(s;0,\tilde{z}) = X(s;0,z)$ for all $s \in [t,T]$. Thus, the definition does not depend on the choice of z.

By (31) we see that $X(s;0,z) \in J$ for all $x \in I$ and $s \in [t,T]$. Thus, (35) gives for all z with $x = X(t;0,z) \in I$ and $s \in [t,\tau]$

$$|p(s, X(s;0,z))| \le \|p^\tau\|_{B(J)} + \|c\|_{L^1(t,\tau;B(J))} + \int_s^\tau \|b(r)\|_{B(J)} |p(r, X(r;0,z))|\, dr.$$

Now the Gronwall lemma yields (37). Hence, there exists a unique $p \in B(\Omega_\tau)$ satisfying (35).

We show that p is Lipschitz-continuous. Let $z_1 < z_2$ be arbitrary. Then

$$\Delta p(t) \stackrel{\text{def}}{=} p(t, X(t;0,z_2)) - p(t, X(t;0,z_1))$$

satisfies $\Delta p(\tau) = p^\tau(X(\tau;0,z_2)) - p^\tau(X(\tau;0,z_1))$ and

$$\begin{aligned}\frac{d}{dt}\Delta p(t) &= c(t, X(t;0,z_2)) - c(t, X(t;0,z_1)) - b(t, X(t;0,z_2))\Delta p(t) \\ &\quad - (b(t, X(t;0,z_2)) - b(t, X(t;0,z_1)))p(t, X(t;0,z_1)).\end{aligned}$$

Thus, setting $I(t) = [X(t;0,z_1), X(t;0,z_2)]$ we get for all $t \in [0,\tau]$

$$\begin{aligned}|\Delta p(t)| \le &\int_t^\tau \|b(s)\|_{B(I(s))} |\Delta p(s)|\, ds + |p^\tau(X(\tau;0,z_2)) - p^\tau(X(\tau;0,z_1))| \\ &+ \int_t^\tau \|p(s)\|_{B(I(s))} |b(s, X(s;0,z_2)) - b(s, X(s;0,z_1))|\, ds \\ &+ \int_t^\tau |c(s, X(s;0,z_2)) - c(s, X(s;0,z_1))|\, ds.\end{aligned} \tag{41}$$

Hence, we have by (33) with $\Delta X(t) \stackrel{\text{def}}{=} X(t;0,z_2) - X(t;0,z_1)$

$$|\Delta p(t)| \le \Big(\|p_x^\tau\|_\infty + \|b\|_{L^1(0,\tau;C^{0,1})}\|p\|_{B(\Omega_\tau)} + \|c\|_{L^1(0,\tau;C^{0,1})}\Big) e^{\int_t^\tau \alpha} \Delta X(t) + \int_t^\tau \|b(s)\|_{B(I(s))}|\Delta p(s)|\, ds$$

and by Gronwall for all $t \in [0,\tau]$

$$|\Delta p(t)| \le \Delta X(t)\big(\|p_x^\tau\|_\infty + \|b\|_{L^1(0,\tau;C^{0,1})}\|p\|_{B(\Omega_\tau)} + \|c\|_{L^1(0,\tau;C^{0,1})}\big) e^{\|b\|_{L^1(0,\tau;B(\mathbb{R}))} + \int_\tau^t \alpha}.$$

This yields a uniform bound for $\|p_x\|_\infty$. Finally, let $z \in \mathbb{R}$ and $t, \hat{t} \in [0,\tau]$, $t < \hat{t}$, be arbitrary then by (31)

$$\begin{aligned}|p(\hat{t}, X(t;0,z)) - p(t, X(t;0,z))| &\le |p(\hat{t}, X(\hat{t};0,z)) - p(t, X(t;0,z))| \\ &\quad + |p(\hat{t}, X(\hat{t};0,z)) - p(\hat{t}, X(t;0,z))| \\ &\le (\|b\|_{B(\Omega_\tau)}\|p\|_{B(\Omega_\tau)} + \|c\|_{B(\Omega_\tau)} + \|a\|_\infty \|p_x\|_{B(\Omega_\tau)})|\hat{t} - t|.\end{aligned} \tag{42}$$

Hence, $p \in C^{0,1}(\Omega_\tau)$ and one easily checks that $\|p\|_{C^{0,1}(\Omega_\tau^{cl})}$ is bounded by a constant only depending on the asserted quantities.

Finally, p solves (22) a.e. in Ω_T. In fact, for a.a. $(t,x) = (t, X(t;0,z))$ in Ω_T the Lipschitz-function p is differentiable. Moreover, since $a(t,\cdot) \in BV_{loc}(\mathbb{R})$ for a.a. t by the one-sided Lipschitz condition, we have $a(t,x-) = a(t,x+)$ for a.a. $(t,x) \in \Omega_T$ and thus from (34) that

$$X_s(t;0,z) = a(t, X(t;0,z))$$

for a.a. $(t,x) = (t, X(t;0,z))$. Now the chain rule yields with (35) that (22) is satisfied for all these (t,x).

Since for $w \in C^{0,1}(I(t))$ holds $\|w\|_{TV(I(t))} = \|w_x\|_{1,I(t)}$, see for example [8], we find for arbitrary $\varepsilon > 0$ points $x_0 < x_1 \ldots < x_N$ in $I(t)$ such that

$$\|p_x(t)\|_{1,I(t)} \le \sum_{i=1}^N |p(t,x_i) - p(t,x_{i-1})| + \varepsilon.$$

Summing (41) for the pairs x_{i-1}, x_i instead of z_1, z_2 yields thus for all $0 \le t \le \tau$

$$\|p_x(t)\|_{1,I(t)} \le \varepsilon + \int_t^\tau \|b(s)\|_{\infty,I(s)} \|p_x(s)\|_{1,I(s)}\, ds + \|p_x^\tau\|_{1,I(\tau)} + \int_t^\tau (\|p(s)\|_{\infty,I(s)} \|b_x(s)\|_{1,I(s)} + \|c_x(s)\|_{1,I(s)})\, ds.$$

Since $\varepsilon > 0$ was arbitrary, the same holds for $\varepsilon = 0$. With $I = I(t)$ and $J = [X(t;0,z_1) - \|a\|_\infty(\tau - t), X(t;0,z_2) + \|a\|_\infty(\tau - t)]$ holds $I(s) \subset J$ for $t \le s \le \tau$ by (31). Now (38) follows by Gronwall.

Since p solves (22) a.e. in Ω_τ, we get for arbitrary $0 \le s < \hat{s} \le \tau$ and $I_s^{\hat{s}} = [s, \hat{s}] \times I$

$$\|p_t\|_{1,I_s^{\hat{s}}} \le \|bp - c\|_{1,I_s^{\hat{s}}} + \|a\|_{\infty,I_s^{\hat{s}}} \|p_x\|_{1,I_s^{\hat{s}}} \le (\hat{s} - s)(\|bp - c\|_{L^\infty(s,\hat{s};L^1(I))} + \|a\|_{\infty,I_s^{\hat{s}}} \|p_x\|_{L^\infty(s,\hat{s};L^1(I))}).$$

This is exactly (39). Now (40) follows directly from (37)–(39). Hereby, we use that $W^{1,1}(J) \hookrightarrow C(J)$. ☐

Proof of Corollary 15. (i): The assertions follow by applying Theorem 14 on $(\sigma,\tau)\times\mathbb{R}$ and letting $\sigma \to 0+$. In fact, for any $\sigma \in (0,\tau)$, the reversible solution is given by (36) for $s = \sigma$ and satisfies (37)–(40) on $(\sigma,\tau)\times\mathbb{R}$. Moreover, p satisfies (36) by Remark 13 also for all $s \in [\sigma,\tau)$. Therefore, the reversible solution obtained for some $\sigma > 0$ is an extension for all reversible solutions corresponding to larger σ which justifies the definition of p by exhaustion of Ω_τ.

Now (40) yields $p \in C^{0,1}((0,\tau]; L^1_{loc}(\mathbb{R})) \cap B((0,\tau]; W^{1,1}_{loc}(\mathbb{R}))$, and we may therefore extend p uniquely to $C^{0,1}([0,\tau]; L^1_{loc}(\mathbb{R}))$. Clearly, p satisfies (37)–(40) for $t \in (0,\tau]$. Since $p(t,\cdot)$ is for $t > 0$ by (40) uniformly bounded in $W^{1,1}_{loc}(\mathbb{R}) \hookrightarrow BV_{loc}(\mathbb{R})$ and $p(t,\cdot) \to p(0,\cdot)$ in $L^1_{loc}(\mathbb{R})$ for $t \to 0+$, we obtain $p(0,\cdot) \in BV_{loc}(\mathbb{R})$ by the lower semicontinuity of $\|\cdot\|_{BV}$ under L^1-convergence [8]. Together with $p \in B((0,\tau]; W^{1,1}_{loc}(\mathbb{R}))$ we conclude that $p \in B([0,\tau]; BV_{loc}(\mathbb{R}))$.

Now let E be a closed set such that $a_x(t,\cdot)|_{\mathbb{R}\setminus E} \le \tilde{\alpha}(t)$ with some $\tilde{\alpha} \in L^1(0,T)$, cf. (24). Let E_ε be an arbitrary ε-neighborhood of E and set $\hat{a} = a(1-\varphi)$ with the function

$$\varphi(t,x) = \mathbf{1}_{[0,\varepsilon/(1+4\|a\|_\infty)]}(t) \max\{0, 1 - 4\,\mathrm{dist}(x, E_{\varepsilon/4})/\varepsilon\}. \tag{43}$$

Then φ is Lipschitz with respect to x and $\mathrm{supp}\,\varphi \subset [0,\varepsilon/(1+4\|a\|_\infty)] \times E_{\varepsilon/2}$. Since a satisfies also the weak OSLC (24), we have obviously $\hat{a}_x \le \hat{\alpha}$ for some $\hat{\alpha} \in L^1(0,T)$. Denote by $\hat{p}$ the reversible solution for a replaced by $\hat{a}$. Then $\hat{p} \in C^{0,1}(\Omega_\tau^{cl})$ by Theorem 14. Now the values of $p|_{[0,\tau]\times(\mathbb{R}\setminus E_\varepsilon)}$ and $\hat{p}|_{[0,\tau]\times(\mathbb{R}\setminus E_\varepsilon)}$ depend by Remark 13 only on the values of a and $\hat{a}$ outside the support of φ and there a coincides with $\hat{a}$. Hence, we have $p|_{[0,\tau]\times(\mathbb{R}\setminus E_\varepsilon)} = \hat{p}|_{[0,\tau]\times(\mathbb{R}\setminus E_\varepsilon)}$ which yields $p \in C^{0,1}([0,\tau]\times(\mathbb{R}\setminus E_\varepsilon))$. The bound for the norm now follows from Theorem 14. If $E = \emptyset$ then the strong OSLC (23) holds and we can clearly choose $E = E_\varepsilon = \emptyset$ by Theorem 14.

(ii): Exactly as at the beginning of the proof of Theorem 14 we obtain that the reversible solution according to Definition 12 exists, is unique and satisfies the bound (37) for $t > 0$. In particular, we have $p \in B(\Omega_\tau)$.

Now let (p_n^τ) be a sequence in $C^{0,1}(\mathbb{R})$, bounded in $C(\mathbb{R})\cap W^{1,1}_{loc}(\mathbb{R})$, that converges pointwise everywhere to p^τ and denote the corresponding reversible solutions of (22) by p_n. Then (p_n) is bounded in $B(\Omega_\tau)$ by (37). Now $p_n - p$ satisfies (36) with $c = 0$ and $p_n^\tau - p^\tau$ instead of p^τ for all $s > 0$. Applying the Gronwall lemma yields for all $(t,x) \in \Omega_\tau$ and with z such that $x = X(t;s,z)$ similarly as in (37) for all $t \in (s,\tau]$

$$|(p_n - p)(t,x)| \le |(p_n^\tau - p^\tau)(X(\tau;s,z))| e^{\|b(\cdot,X(\cdot;s,z))\|_{\infty,[s,\tau]}(\tau - t)}.$$

This shows that $p_n \to p$ everywhere on $(0,\tau]\times\mathbb{R}$, and boundedly everywhere, since (p_n) is bounded in $B((0,\tau]\times\mathbb{R})$ by (37).

Moreover, (p_n) is bounded in $C^{0,1}([0,\tau]; L^1_{loc}(\mathbb{R})) \cap B([0,\tau]; BV_{loc}(\mathbb{R})) \cap W^{1,1}_{loc}(\Omega_\tau^{cl})$ by (40) in Theorem 14 and the arguments in the proof of (i), since (p_n^τ) is bounded in $W^{1,1}_{loc}(\mathbb{R})$. Now $p_n \to p$ boundedly everywhere on $(0,\tau]\times\mathbb{R}$ implies that $p_n(t,\cdot) \to p(t,\cdot)$ in $L^1_{loc}(\mathbb{R})$ boundedly for all $t \in (0,\tau]$. Therefore, we have also $p \in C^{0,1}([0,\tau]; L^1_{loc}(\mathbb{R}))$. Moreover, we conclude by an Arzela-Ascoli argument that $p_n \to p$ in $C([0,\tau]; L^1_{loc}(\mathbb{R}))$: we have $p_n(t,\cdot) \to p(t,\cdot)$ in $L^1_{loc}(\mathbb{R})$ for all t in a dense countable subset of $[0,\tau]$ and get now uniform convergence by using the uniform Lipschitz-continuity. In particular, the boundedness of (p_n) in $B([0,\tau]; BV_{loc}(\mathbb{R})) \cap W^{1,1}_{loc}(\Omega_\tau^{cl})$ yields $p \in B([0,\tau]; BV_{loc}(\mathbb{R})) \cap BV_{loc}(\Omega_\tau^{cl})$ by the lower semicontinuity of the *BV*-norm under L^1-convergence.

Now let E be a closed set such that $a_x(t,\cdot)|_{\mathbb{R}\setminus E} \le \tilde{\alpha}(t)$ with some $\tilde{\alpha} \in L^1(0,T)$, cf. (24). If $E = \emptyset$ then the strong OSLC (23) holds and thus (36) is satisfied also for $s = 0$ by Theorem 14. Hence, $p_n \to p$ converges even boundedly everywhere on $[0,\tau] \times \mathbb{R}$.

If $E \neq \emptyset$ we have to show that $p_n(0,\cdot) \to p(0,\cdot)$ boundedly everywhere on the open set $\mathbb{R} \setminus E$. We know by (i) that $p_n \in C([0,\tau] \times (\mathbb{R} \setminus E))$ and (p_n) is bounded in $B((0,\tau] \times \mathbb{R})$. Thus, (p_n) is bounded in $B([0,\tau] \times (\mathbb{R} \setminus E))$ and it remains to show that $p_n(0,x) \to p(0,x)$ for all $x \in \mathbb{R} \setminus E$. Let any such x be given. Then for $\varepsilon > 0$ small enough we have $x \notin E_\varepsilon$. With φ as in (43) we set again $\hat{a} = a(1-\varphi)$ and denote by $\hat{p}_n$, $\hat{p}$ the reversible solutions for $\hat{a}$ instead of a. As in the proof of (i) $\hat{a}$ satisfies the strong OSLC $\hat{a}_x \le \hat{\alpha} \in L^1(0,T)$ and therefore $\hat{p}_n \to \hat{p}$ boundedly everywhere on $[0,\tau] \times \mathbb{R}$ as we have already shown. But as in the proof of (i) we conclude with Remark 13 that $\hat{p}_n(0,x) = p_n(0,x)$ and $\hat{p}(0,x) = p(0,x)$. □

3.3 Stability of reversible solutions

We study now the stability properties of reversible solutions with respect to the coefficients a, b, c and the end data. Our aim is to prove in particular the stability property (12) if the coefficients converge in the sense (18), (19). The following stability result extends a similar result of [1] for the case $b = c \equiv 0$ to the case $b, c \neq 0$. A further extension to the weakened OSLC (24) will follow in Theorem 17.

Theorem 16. (Stability of reversible solutions under strong OSLC)
Let the following assumptions hold:

(a) (a_n) *is a bounded sequence in* $L^\infty(\Omega_T)$ *with* $a_n \to a$ *in* $L^\infty(\Omega_T)$*-weak* and for a bounded sequence* (α_n) *in* $L^1(0,T)$ *and* $\alpha \in L^1(0,T)$ *the OSLCs hold*

$$(a_n)_x(t,\cdot) \le \alpha_n(t), \quad a_x(t,\cdot) \le \alpha(t) \quad \textit{for a.a. } t \in (0,T),$$

(b) $(b_n), (c_n)$ *are sequences in* $L^\infty(0,T;C^{0,1}(\mathbb{R}))$*, bounded in* $L^1(0,T;C(\mathbb{R}))$*, with* $b_n \to b$*,* $c_n \to c$ *in* $L^1(0,T;C_{loc}(\mathbb{R}))$*, where* $b, c \in L^\infty(0,T;C^{0,1}(\mathbb{R}))$*,*

(c) (p_n^τ) *is a sequence in* $C^{0,1}(\mathbb{R})$*, bounded in* $C(\mathbb{R})$*, with* $p_n^\tau \to p^\tau$ *in* $C_{loc}(\mathbb{R})$*, where* $p^\tau \in C^{0,1}(\mathbb{R})$*.*

Then the reversible solutions p_n *of*

$$(p_n)_t + a_n(p_n)_x = -b_n p_n + c_n, \quad p_n(\tau,\cdot) = p_n^\tau \tag{44}$$

satisfy

$$p_n \to p \quad \textit{in } C([0,\tau] \times [-R,R])$$

for all $R > 0$*, where* p *is the reversible solution of* (22)*.*

Proof. Denote by X and X_n the backward flows according to Definition 9 for a and a_n, respectively. By Theorem 11 it holds $X_n \to X$ in $C(D_b \times [-R,R])$ for any $R > 0$. By the definition of reversible solutions we have for all $z \in \mathbb{R}$

$$p_n(\tau, X_n(\tau;0,z)) = p_n^\tau(X_n(\tau;0,z)), \qquad \frac{d}{dt} p_n(t, X_n(t;0,z)) = (-b_n p_n + c_n)(t, X_n(t;0,z)).$$

For the reversible solution p of (22) holds (35). Fix some $R > 0$ and consider an arbitrary $(t,x) \in [0,T] \times [-R,R]$. There exist $z, z_n \in \mathbb{R}$ with $x = X(t;0,z) = X_n(t;0,z_n)$ and we have $X(s;0,z), X_n(s;0,z_n) \in [R - M_a\tau, R + M_a\tau] \overset{\text{def}}{=} J$ according to (31) for all $s \in [t,\tau]$ with an upper bound M_a for $\|a_n\|_\infty$ and $\|a\|_\infty$. Since $X(s;t,x) = X(s;0,z)$ and $X_n(s;t,x) = X_n(s;0,z_n)$ by (27), we have for a.a. $s \in (t,\tau)$

$$p_n(\tau, X_n(\tau;t,x)) = p_n^\tau(X_n(\tau;t,x)), \quad \frac{d}{ds} p_n(s, X_n(s;t,x)) = (-b_n p_n + c_n)(s, X_n(s;t,x)),$$
$$p(\tau, X(\tau;t,x)) = p^\tau(X(\tau;t,x)), \quad \frac{d}{ds} p(s, X(s;t,x)) = (-bp + c)(s, X(s;t,x)).$$

Therefore, the difference $\Delta p_n(s) \overset{\text{def}}{=} p(s, X(s;t,x)) - p_n(s, X_n(s;t,x))$ satisfies

$$|\Delta p_n(\tau)| = |p^\tau(X(\tau;t,x)) - p_n^\tau(X_n(\tau;t,x))| \le \|p^\tau - p_n^\tau\|_{C(J)} + \|p_x^\tau\|_{\infty,J} \|X - X_n\|_{C(D_b \times J)} \tag{45}$$

and for a.a. $s \in (t,\tau)$

$$\frac{d}{ds}\Delta p_n(s) = (b_n(s, X_n(s;t,x)) - b(s, X(s;t,x)))p(s, X(s;t,x)) \\ - b_n(s, X_n(s;t,x))\Delta p_n(s) + c(s, X(s;t,x)) - c_n(s, X_n(s;t,x)).$$

Thus, we get with $J_\tau \stackrel{\text{def}}{=} [0,\tau] \times J$

$$\begin{aligned} |\Delta p_n(s)| &\le \|p\|_{C(J_\tau)}\big(\|b_n - b\|_{L^1(0,\tau;C(J))} + \|b_x\|_{L^1(0,\tau;L^\infty(J))}\|X - X_n\|_{C(D_b\times J)}\big) + \|c_n - c\|_{L^1(0,T;C(J))} \\ &\quad + \|c_x\|_{L^1(0,\tau;L^\infty(J))}\|X - X_n\|_{C(D_b\times J)} + |\Delta p_n(\tau)| + \int_s^\tau \|b_n(r,\cdot)\|_{C(J)}|\Delta p_n(r)|\,dr \\ &\stackrel{\text{def}}{=} \eta_n + \int_s^\tau \|b_n(r,\cdot)\|_{C(J)}|\Delta p_n(r)|\,dr. \end{aligned}$$

Obviously, this inequality holds for all $(t,x) \in [0,\tau] \times [-R,R]$ and η_n does not depend on $(t,x) \in [0,\tau] \times [-R,R]$. Now the Gronwall lemma yields

$$|p(t,x) - p_n(t,x)| = |\Delta p_n(t)| \le \eta_n\, e^{\|b_n\|_{L^1(0,\tau;C(J))}}$$

for all $(t,x) \in [0,\tau] \times [-R,R]$. By assumptions (a)–(c), (37), and (45) we see that $\|b_n\|_{L^1(0,\tau;C(J))}$ is uniformly bounded and $\eta_n \to 0$. This shows that $\lim_{n\to\infty} \|p - p_n\|_{C([0,T]\times[-R,R])} = 0$. □

We have already observed that only the weakened OSLC (24) is satisfied if the initial data have an up-jump (which generates a rarefaction wave), cf. (20). In this case we have the following variant of Theorem 16.

Theorem 17. (Stability of reversible solutions under weakened OSLC)
Let the assumptions (a)–(c) *of Theorem 16 hold with the following modifications:*

(a) *In* (a) *the sequence* (α_n) *is bounded in* $L^1(\sigma,T)$ *and* $\alpha \in L^1(\sigma,T)$ *merely for all fixed* $\sigma > 0$.

(b) *In* (b) *the sequences* $(b_n), (c_n)$ *are in addition bounded in* $L^\infty(0,T;W^{1,1}_{loc}(\mathbb{R}))$.

(c) *In* (c) *the sequence* (p_n^τ) *is in addition bounded in* $W^{1,1}_{loc}(\mathbb{R})$.

Then the reversible solutions p_n *of* (44) *are uniformly bounded in* $B(\Omega_\tau) \cap W^{1,1}((0,\tau) \times (-R,R)) \cap C^{0,1/r}([0,\tau];L^r(-R,R))$ *for all* $r \in [1,\infty)$ *and* $R > 0$. *Moreover,*

$$p_n \to p \quad \text{in } C([\sigma,\tau] \times [-R,R]) \cap C([0,\tau];L^r(-R,R))$$

for all $\sigma > 0$, $R > 0$ *and* $r \in [1,\infty)$, *where* p *is the reversible solution of* (22).

If in addition E *is a closed set such that* $a_x(t,\cdot)|_{\mathbb{R}\setminus E} \le \tilde\alpha(t)$, $\tilde\alpha \in L^1(0,T)$, $(a_n)_x(t,\cdot)|_{\mathbb{R}\setminus E} \le \tilde\alpha_n(t)$, $(\tilde\alpha_n)$ *bounded in* $L^1(0,T)$, *then one has moreover*

$$p_n \to p \quad \text{in } C([0,\tau] \times ([-R,R] \setminus E_\varepsilon))$$

for any ε*-neighborhood* E_ε *of* E *and all* $R > 0$.

Proof. The reversible solutions p_n and p are by definition reversible solutions on any domain $[\sigma,\tau] \times \mathbb{R}$, $\sigma > 0$, see Remark 13, and there the strong OSLC condition of Theorem 16, (a) holds. Therefore, Theorem 16 is applicable on these domains and yields that $p_n \to p$ in $C([\sigma,\tau] \times [-R,R])$ for all $R > 0$. The uniform boundedness of p_n in $B(\Omega_\tau) \cap W^{1,1}((0,\tau) \times (-R,R)) \cap C^{0,1}([0,\tau];L^1(-R,R))$ is a direct consequence of Corollary 15 and the boundedness assumptions on b_n, c_n, and p_n^τ. Setting $I = [-R,R]$ and using the interpolation inequality $\|\cdot\|_{r,I} \le \|\cdot\|_{1,I}^{1/r}\|\cdot\|_{\infty,I}^{1-1/r}$ for $r \in [1,\infty)$, we conclude that (p_n) is also bounded in $C^{0,1/r}([0,\tau];L^r(-R,R))$. Now for $\sigma > 0$ and all $t \in [0,\sigma]$ we have with $I = [-R,R]$

$$\|p_n(t) - p(t)\|_{1,I} \le \|p_n(\sigma) - p(\sigma)\|_{1,I} + (\sigma - t)\|p_n - p\|_{C^{0,1}([0,\tau];L^1(I))} \le 2R\|p_n(\sigma) - p(\sigma)\|_{C(I)} + C\sigma$$

where C is a uniform bound for $\|p_n - p\|_{C^{0,1}([0,\tau];L^1(I))}$. This shows that

$$\|p_n - p\|_{C([0,\tau];L^1(-R,R))} \le 2R\|p_n - p\|_{C([\sigma,\tau]\times[-R,R])} + C\sigma.$$

We have already shown that the first term on the right-hand side converges to zero for any fixed $\sigma > 0$. Thus, given any $\varepsilon > 0$, the right hand side is $\le \varepsilon$ by choosing $\sigma = \varepsilon/(2C)$ for all n sufficiently large. This shows that $p_n \to p$ in $C([0,\tau]; L^1(-R,R))$. By interpolation with the uniform L^∞-bound for p_n and p, the same holds in $C([0,\tau]; L^r(-R,R))$ for all $r \in [1,\infty)$.

Now let E be a closed set such that $a_x(t,\cdot)|_{\mathbb{R}\setminus E} \le \tilde{\alpha}(t)$, $\alpha \in L^1(0,T)$, $(a_n)_x(t,\cdot)|_{\mathbb{R}\setminus E} \le \tilde{\alpha}_n(t)$, $(\tilde{\alpha}_n)$ bounded in $L^1(0,T)$. Moreover, let $\varepsilon > 0$ be arbitrary and E_ε the ε-neighborhood of E. By (a) we find M_a with $\|a_n\|_\infty, \|a\|_\infty \le M_a$. As used previously, we set $\hat{a} = a(1-\varphi)$ and $\hat{a}_n = a_n(1-\varphi)$ with the function

$$\varphi(t,x) = \mathbf{1}_{[0,\varepsilon/(1+4M_a)]}(t) \max\{0, 1 - 4\,\mathrm{dist}(x, E_{\varepsilon/4})/\varepsilon\}.$$

Using in addition the bounds $a_x \le \alpha$ and $(a_n)_x \le \alpha_n$, there exist obviously $\hat{\alpha} \in L^1(0,T)$ and a bounded sequence $(\hat{\alpha}_n)$ in $L^1(0,T)$ with $\hat{a}_x \le \hat{\alpha}$ and $(\hat{a}_n)_x \le \hat{\alpha}_n$. Now Theorem 16 yields $\hat{p}_n \to \hat{p}$ in $C([0,\tau] \times [-R,R])$ for all $R > 0$. Moreover, Remark 13 yields as before that $\hat{p}_n|_{[0,\tau]\times([-R,R]\setminus E_\varepsilon)} = p_n|_{[0,\tau]\times([-R,R]\setminus E_\varepsilon)}$ and $\hat{p}|_{[0,\tau]\times([-R,R]\setminus E_\varepsilon)} = p|_{[0,\tau]\times([-R,R]\setminus E_\varepsilon)}$, since the propagation speed of the backward flow is bounded by M_a. Hence, we have shown that $p_n \to p$ in $C([0,\tau] \times ([-R,R] \setminus E_\varepsilon))$. This completes the proof $\square$

References

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

[2] F. Bouchut, F. James, Differentiability with respect to initial data for a scalar conservation law, in: Hyperbolic problems: theory, numerics, applications, Vol. I (Zürich, 1998), Birkhäuser, Basel, 1999, pp. 113–118.

[3] E. M. Cliff, M. Heinkenschloss, A. R. Shenoy, An optimal control problem for flows with discontinuities, J. Optim. Theory Appl. 94 (1997) 273–309.

[4] E. D. Conway, Generalized solutions of linear differential equations with discontinuous coefficients and the uniqueness question for multidimensional quasilinear conservation laws, J. Math. Anal. Appl. 18 (1967) 238–251.

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

[6] A. F. Filippov, Differential equations with discontinuous right-hand side, Amer. Math. Soc. Transl. Ser. 2 42 (1964) 199–231.

[7] P. D. Frank, G. R. Shubin, A comparison of optimization-based approaches for a model computational aerodynamics design problem, J. Comput. Phys. 98 (1992) 74–89.

[8] E. Giusti, Minimal surfaces and functions of bounded variation, Birkhäuser Verlag, Basel, 1984.

[9] E. Godlewski, P.-A. Raviart, The linearized stability of solutions of nonlinear hyperbolic systems of conservation laws. A general numerical approach, Math. Comput. Simulation 50 (1999) 77–95, modelling ’98 (Prague).

[10] D. Hoff, The sharp form of Oleĭnik’s entropy condition in several space variables, Trans. Amer. Math. Soc. 276 (1983) 707–714.

[11] S. N. Kružkov, First order quasilinear equations in several independent variables, Math. USSR-Sb. 10 (1970) 217–243.

[12] T.-P. Liu, Nonlinear resonance for quasilinear hyperbolic equation, J. Math. Phys. 28 (1987) 2593–2602.

[13] J. Málek, J. Nečas, M. Rokyta, M. Růžička, Weak and measure-valued solutions to evolutionary PDEs, Chapman & Hall, London, 1996.

[14] O. A. Oleĭnik, Discontinuous solutions of non-linear differential equations, Amer. Math. Soc. Transl. Ser. 2 26 (1963) 95–172, translation from Usp. Mat. Nauk, 12 (1957), pp. 3-73.

[15] M. Slemrod, Existence of optimal controls for control systems governed by nonlinear partial differential equations, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 1 (1974) 229–246.

[16] J. Smoller, Shock waves and reaction-diffusion equations, 2nd ed., Springer-Verlag, New York, 1994.

[17] E. Tadmor, Local error estimates for discontinuous solutions of nonlinear hyperbolic equations, SIAM J. Numer. Anal. 28 (1991) 891–906.

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

[19] S. Ulbrich, A sensitivity and adjoint calculus for discontinuous solutions of hyperbolic conservation laws with source terms, Technical Report TR00-10, Department of Computational and Applied Mathematics, Rice University, Houston, TX, accepted for publication (in revised form) in SIAM J. Control Optim. (2000).

[20] S. Ulbrich, Optimal control of nonlinear hyperbolic conservation laws with source terms, Habilitationsschrift, Zentrum Mathematik, Technische Universität München (2001).