Control and Cybernetics

vol. **48** (2019) No. 2

On the numerical discretization of optimal control problems for conservation laws*

by

Paloma Schäfer Aguilar, Johann Michael Schmitt, Stefan Ulbrich and Michael Moos

Department of Mathematics, Technische Universität Darmstadt,
Dolivostr. 15, 64293 Darmstadt, Germany
aguilar,jschmitt,ulbrich@mathematik.tu-darmstadt.de

Dedicated to Günter Leugering on the occasion of His 65th birthday

Abstract: We analyze the convergence of discretization schemes for the adjoint equation arising in the adjoint-based derivative computation for optimal control problems governed by entropy solutions of conservation laws. The difficulties arise from the fact that the correct adjoint state is the reversible solution of a transport equation with discontinuous coefficient and discontinuous end data. We derive the discrete adjoint scheme for monotone difference schemes in conservation form. It is known that convergence of the discrete adjoint can only be expected if the numerical scheme has viscosity of order $O(h^\alpha)$ with appropriate $0 < \alpha < 1$, which leads to quite viscous shock profiles. We show that by a slight modification of the end data of the discrete adjoint scheme, convergence to the correct reversible solution can be obtained also for numerical schemes with viscosity of order $O(h)$ and with sharp shock resolution. The theoretical findings are confirmed by numerical results.

Keywords: conservation laws, adjoint state, linear transport equation, discontinuous coefficients, finite difference schemes

1. Introduction

We consider optimal control problems for entropy solutions of scalar conservation laws

$$\begin{aligned} &y_t + f(y)_x = 0, \quad (t,x) \in \Omega_T \stackrel{\text{def}}{=} (0,T) \times \mathbb{R}, \\ &y(0,x) = u(x), \quad x \in \mathbb{R}, \end{aligned} \tag{1}$$

where $f \in C^2(\mathbb{R})$ is a strongly convex flux function, i.e.,

$$f \in C^2(\mathbb{R}), \quad f'' \geq m_{f''} > 0 \tag{2}$$

*Submitted: March 2019; Accepted: August 2019

with a contant $m_{f''} > 0$. Here, $u \in L^1(\mathbb{R}) \cap BV(\mathbb{R})$, where $BV(\mathbb{R})$ denotes the space of functions of bounded variation.

The objective function is of the form

$$J(y) \stackrel{\text{def}}{=} \int_{\mathbb{R}} \gamma(x)\, \psi(y(\bar{t}, x), y_d(x))\, dx \tag{3}$$

with a weighting function $\gamma \in C_c^1(\mathbb{R})$, $\psi \in C_{loc}^{1,1}(\mathbb{R})$ and data $y_d \in C^1(I)$. Often, a continuously differentiable regularization term $R(u)$ is added, but we focus only on the differentiability properties and the numerical approximation of the state dependent part (3).

The developments in this paper can also be extended to problems (1) with source terms. This leads to additional technical complications and we prefer to confine our study to conservation laws.

It is well known that, in general, the weak solutions of (1) develop discontinuities after finite time and that uniqueness holds only in the class of entropy solutions. We recall that for given $u \in L^1(\mathbb{R}) \cap BV(\mathbb{R})$ a function $y = y(u) \in L^\infty(\Omega_T)$ is an entropy solution of (1) in the sense of Kružkov (1970) if it satisfies for all convex functions (entropies) $\eta \in C_{loc}^{0,1}(\mathbb{R})$ with corresponding entropy fluxes $q(y) = \int_0^y \eta'(s)\, f'(s)\, ds$ the entropy inequality

$$\eta(y)_t + q(y)_x \leq 0$$

in the sense of distributions and the initial condition in the sense

$$\operatorname*{ess\,lim}_{t \searrow 0} \|y(t, \cdot) - u\|_{1,K} = 0 \quad \forall\, K \subset\subset \mathbb{R}.$$

As we will recall in Section 2, it is known that under a generic nondegeneracy assumption the mapping

$$u \in PC^1(\mathbb{R}; z_1, \ldots, z_N) \mapsto J(y(u)) \tag{4}$$

is Fréchet differentiable, where $PC^1(\mathbb{R}; z_1, \ldots, z_N)$ denotes the space of piecewise C^1 functions with possible discontinuities at $z_1, \ldots, z_N$, see Pfaff and Ulbrich (2015), Ulbrich (2002, 2003). Moreover, the derivative admits the adjoint representation

$$\frac{d}{du} J(y(u)) \cdot \delta u = \int_{\mathbb{R}} p(0, x)\, \delta u(x)\, dx, \tag{5}$$

where p is a reversible solution, according to Definition 1 and Theorem 3 of the adjoint equation, a transport equation with possibly discontinuous coefficient

$$p_t + f'(y) p_x = 0, \quad (t, x) \in \Omega_{\bar{t}}, \tag{6}$$

$$p(\bar{t}, x) = p^{\bar{t}} = \begin{cases} \gamma(x) \psi_y(y(\bar{t}, x), y_d(x)) & \text{if } y(\bar{t}, \cdot) \text{ is continuous at } x \\ \gamma(x) \frac{[\psi(y(\bar{t}, x), y_d(x))]}{[y(\bar{t}, x)]} & \text{if } y(\bar{t}, \cdot) \text{ is discontinuous at } x \end{cases}, \; x \in \mathbb{R}. \tag{7}$$

Here, for a function $v \in BV(\mathbb{R})$, we denote by $[v(x)]$ the jump $[v(x)] := v(x+) - v(x-)$ at x. As we will recall in Section 2, the coefficient $f'(y)$ satisfies a one-sided Lipschitz condition. As a consequence, the solution of (6), (7) is not unique if the state y contains shocks. The correct solution of (6), (7) is the unique reversible solution and can be characterized by a monotonicity criterion, see Definition 1. Equivalently, the reversible solution can be defined along the generalized backward characteristics, see Remark 1.

While the convergence of numerical schemes for the conservation law (1) is very well studied, there exist only few results on the convergence of discretization schemes for the adjoint equation (6), (7), see Bardos and Pironneau (2005), Castro, Palacios and Zuazua (2008), Giles and Ulbrich (2010a,b), Gosse and James (2000), Hajian, Hintermüller and Ulbrich (2019), Homescu and Navon (2003), Ulbrich (2001). However, this is of importance for obtaining convergent approximations for the derivative (5) of the objective functional. In this paper, we will analyze the convergence of numerical schemes for the state equation, the adjoint equation and the resulting discrete approximation of the gradient representation (5). The difficulties result from the fact that the end data $p(\bar{t}, x)$ of (6), (7) are discontinuous at shock locations and have to be propagated in an appropriate fashion by the discrete adjoint scheme to obtain convergence to the correct reversible solution according to Theorem 3. So far, convergence results for the adjoint schemes have only been considered for Lipschitz-continuous end data (see Gosse and James, 2000; Hajian, Hintermüller and Ulbrich, 2019; Ulbrich, 2001) or for schemes with increased numerical viscosity of order $O(h^\beta)$ for $\beta < 1$ (Giles and Ulbrich, 2010a,b).

We will consider monotone finite difference schemes in conservation form for the state equation (1) and the corresponding discrete adjoint scheme for the adjoint equation (6), (7). Variants, where the state is computed by other convergent schemes, ensuring a discrete one-sided Lipschitz condition for the state, are possible. As observed in Giles and Ulbrich (2010a,b) the convergence of the discrete adjoint to the correct adjoint state is in general not ensured. In Giles and Ulbrich (2010a,b) it is shown that a modified Lax-Friedrichs scheme with numerical viscosity of order $O(h^\beta)$ for appropriate $0 < \beta < 1$ yields convergent adjoint approximations. However, the increased numerical viscosity reduces the accuracy of the numerical scheme in smooth regions, as well as the resolution of shocks. In this paper we propose another approach that is inspired by the continuous adjoint equation (6), (7) and does not require an increased numerical viscosity. Numerical results underline the advantages of the approach.

The paper is organized as follows. In Section 2 we recall the known fact on the state equation, the differentiability of objective functionals, the adjoint euqation and an adjoint based derivative represenstation. In Section 3 we derive for monotone difference schemes the corresponding sensitivity scheme and adjoint scheme. We prove convergence of the adjoint scheme for Lipschitz end data to the rversible solution and extend this result subsequently to discontinuous end data as they arise in adjoint based derivative represenstation. This will be achieved by a novel choice of the end data, which we propose in this paper.

In Section 4 we apply the general results to the Engquist-Osher scheme and the modified Lax-Friedrichs scheme and their adjoint schemes. The theoretical findings are illustrated in Section 5 by numerical results.

2. Continuous problem

We summarize known results on entropy solutions of (1), the differentiability properties of the objective function (3), (4) and the adjoint equation (6), (7) to obtain the gradient representation (5).

Proposition 1 *Let* (2) *hold. Then, for any* $u \in L^\infty(\mathbb{R})$ *there exists a unique entropy solution* $y = y(u) \in L^\infty(\Omega_T)$. *After modification on a set of measure zero, one has* $y \in C([0,T]; L^1(-R,R))$ *for all* $R > 0$. *Moreover, let* $u, \hat{u} \in L^\infty(\mathbb{R})$ *be arbitrary and* $M_{f'} = \max_{|s| \le \max(\|u\|_\infty, \|\hat{u}\|_\infty)} |f'(s)|$. *Then*

1. $\|y(t,\cdot;u)\|_\infty \le \|u\|_\infty \quad \forall\, t \in [0,T]$
2. $\|y(t,\cdot;u) - y(t,\cdot;\hat{u})\|_{1,[a,b]} \le \|u - \hat{u}\|_{1,[a-tM_{f'},b+tM_{f'}]} \quad \forall\, t \in [0,T]$
3. *If* $u \in BV(\mathbb{R})$ *and* $u_x \le M_{u'}$ *with* $M_{u'} \in [0,\infty]$ *then* $y(u)$ *satisfies the one-sided Lipschitz condition (OSLC)*

$$y_x(t,\cdot) \le \frac{1}{M_{u'}^{-1} + m_{f''} t} \quad \forall\, t \in (0,T].$$

Moreover,

$$|y(t,\cdot;u)|_{TV,[a,b]} \le |u|_{TV,[a-tM_{f'},b+tM_{f'}]} \quad \forall\, t \in [0,T]$$

Proof See, for example, Brenier and Osher (1988), Málek et al. (1996), and Oleĭnik (1963). □

The differentiability properties of the objective function (3), (4) have been studied in Ulbrich (2002, 2003), see also Pfaff and Ulbrich (2015).

Theorem 1 *Let* (2) *hold, let* $u \in PC^1(\mathbb{R}; z_1, \ldots, z_N)$ *be arbitrary and let* $\bar{t} \in (0,T]$ *be such that* $y(\bar{t},\cdot;u)$ *has on* $supp(\gamma)$ *no shock generation points and finitely many nondegenerate shocks at* $x_1 < x_2 < \ldots < x_K$ *that are all no shock interaction points. Then, the objective function* (3), (4) *is Fréchet differentiable at* u *and the derivative is given by* (5), *where* p *is the reversible solution of* (6), (7), *see Definition 1 and Theorem 3.*

Moreover, $y(\bar{t},\cdot;u)$ *is piecewise* C^1 *and the shock locations* $x_1 < x_2 < \ldots < x_K$ *depend differentially on* u.

Proof See Ulbrich (2002, 2003). □

To introduce reversible solutions we note that (6), (7) has the form

$$\begin{aligned} &p_t + a p_x = 0, \quad (t,x) \in \Omega_{\bar{t}}, \\ &p(\bar{t},x) = p^{\bar{t}}. \end{aligned} \tag{8}$$

If $u \in PC^1(\mathbb{R}; z_1, \ldots, z_N)$ then $a = f'(y)$ satisfies, by Proposition, 1 the OSLC

$$a_x(t,\cdot) = f'(y(t,\cdot))_x \le \left(\max_{|s| \le \|u\|_\infty} f''(s) \right) \frac{1}{1/\|\max(0,u_x)\|_\infty + m_{f''} t}.$$

For simplicity, $u \in PC^1(\mathbb{R}; z_1, \ldots, z_N)$ is assumed in the rest of the paper with $u_x \le M_{u'} < \infty$, i.e. u can only have down-jumps generating shocks. Then there exists $\alpha \in L^1(0,T)$ such that

$$a_x(t,\cdot) \le \alpha(t) \quad \text{for a.a. } t \in [0,\bar{t}]. \tag{9}$$

The case of u having up-jumps generating rarefaction waves can also be handled, see Ulbrich (2001), but this gives rise to some technical complications that are not the focus of this paper.

Reversible solutions for (8) in the case of $p^{\bar{t}} \in C^{0,1}(\mathbb{R})$ have been introduced and analyzed in Bouchut and James (1998) and has been extended to the case of inhomogeneous right hand side and of discontinuous end data in Ulbrich (2002, 2003).

Definition 1 *Consider* (8) *with a satisfying* (9) *and* $p^{\bar{t}} \in C^{0,1}(\mathbb{R})$. *Denote by* $\mathcal{L}$ *the space of Lipschitz continuous solutions of* $p_t + a p_x = 0$. *Then,* $p \in \mathcal{L}$ *is called reversible solution of* (8) *if there exist* $p_1, p_2 \in \mathcal{L}$ *such that* $p = p_1 - p_2$ *and* $(p_1)_x \ge 0$, $(p_2)_x \ge 0$.

The following existence and uniqueness result has been shown in Bouchut and James (1998).

Theorem 2 *Let* $a \in L^\infty(\Omega_{\bar{t}})$ *satisfy the OSLC* (9). *Then, for any* $p^{\bar{t}} \in C^{0,1}(\mathbb{R})$ *there exists a unique reversible solution* $p \in C^{0,1}(\Omega_{\bar{t}}^{cl})$ *of* (8) *with*

$$\begin{aligned} \|p(t,\cdot)\|_{\infty,I_1} &\le \|p^{\bar{t}}\|_{\infty,I_2}, \\ \|p_x(t,\cdot)\|_{\infty,I_1} &\le e^{\int_t^{\bar{t}} \alpha} \|p_x^{\bar{t}}\|_{\infty,I_2}, \end{aligned} \tag{10}$$

where $\Omega_{\bar{t}}^{cl}$ *denotes the closure of* $\Omega_{\bar{t}}$, $I_1 = (x_1, x_2)$ *is arbitrary and* $I_2 = (x_1 - \|a\|_\infty(\bar{t}-t)), x_2 + \|a\|_\infty(\bar{t}-t))$.

Remark 1 *It can be shown that the unique reversible solution can also be defined along generalized characteristics, see Bouchut and James (1998), Ulbrich (2002,2003). In fact, let for arbitrary* $(t,x) \in \Omega_{\bar{t}}$ *the generalized forward characteristic* $s \in [t,\bar{t}] \mapsto X(s;t,x)$ *be defined by*

$$\frac{d}{ds} X(s;t,x) \in [a(s, X(s;t,x)+), a(s, X(s;t,x)-)].$$

By the OSLC (9), *it can be shown that* $X(\cdot;t,x)$ *is unique. Now, the reversible solution of* (8) *is uniquely defined by*

$$p(s, X(s;t,x)) = p^{\bar{t}}(X(\bar{t};t,x)).$$

Hence, the value $p^{\bar{t}}(z)$ is propagated along all backward characteristics, emanating through $(\bar{t}, z)$, i.e., all $X(\cdot; t, x)$ with $z = X(\bar{t}; t, x)$.

As a consequence, if z is a shock location of $y(\bar{t}, \cdot; u)$, then $p(t, x) = p^{\bar{t}}(z)$ for all (t, x) in the shock funnel confined by the maximal and minimal backward characteristic through $(\bar{t}, z)$.

For discontinuous data $p^{\bar{t}}$ we use the following stability property to define a reversible solution.

THEOREM 3 *Let $a \in L^\infty(\Omega_{\bar{t}})$ satisfy the OSLC (9). Denote by $B(\mathbb{R})$ the Banach space of bounded functions, equipped with the sup-norm, and define*

$$B_{Lip}(\mathbb{R}) \stackrel{\text{def}}{=} \big\{ w \in B(\mathbb{R}) : \quad \exists (w_n) \subset C^{0,1}(\mathbb{R}), (w_n) \text{ bounded in } C(\mathbb{R}) \cap W^{1,1}_{loc}(\mathbb{R}) \text{ such that } w_n \to w \text{ pointwise everywhere} \big\}. \quad (11)$$

Let $p^{\bar{t}} \in B_{Lip}(\mathbb{R})$ and let $(p^{\bar{t}}_n) \subset C^{0,1}(\mathbb{R})$ be any sequence with $(p^{\bar{t}}_n)$ bounded in $C(\mathbb{R}) \cap W^{1,1}_{loc}(\mathbb{R})$ such that $p^{\bar{t}}_n \to p^{\bar{t}}$ pointwise everywhere. Then, the corresponding reversible solution $p_n \in C^{0,1}(\Omega^{cl}_{\bar{t}})$ satisfies

$$p_n \to p \quad \text{in } C([0, \bar{t}]; L^1_{loc}(\mathbb{R})) \text{ and boundedly everywhere on } \Omega^{cl}_{\bar{t}}.$$

Here,

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

satisfies (10) and is independent of the particular sequence $(p^{\bar{t}}_n)$. p is called reversible solution of (8).

PROOF See Ulbrich (2003). □

3. Discrete approximation

3.1. Finite difference schemes for state, sensitivity and adjoint equation

For the discretization of the state equation (1) we consider conservative finite difference schemes. Let $\lambda > 0$ be fixed and set for a grid size $h > 0$

$$\Delta t = \lambda h,\ t_n \stackrel{\text{def}}{=} n\Delta t,\ x_j \stackrel{\text{def}}{=} jh,\ R_j \stackrel{\text{def}}{=} [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}),\ Q^n_j \stackrel{\text{def}}{=} [t_n, t_{n+1}) \times R_j.$$

Given grid values y^n_j at (t_n, x_j), $n \in \mathbb{N}_0, j \in \mathbb{Z}$, we define the difference operators

$$\Delta^+ y^n_j \stackrel{\text{def}}{=} (y^n_{j+1} - y^n_j), \quad \Delta^- y^n_j \stackrel{\text{def}}{=} (y^n_j - y^n_{j-1})$$

and it will be convenient to use this notation also for functions $\phi \in L^1_{loc}(\Omega_T)$ by setting

$$\Delta^+ \phi(t, x) \stackrel{\text{def}}{=} \phi(t, x + h) - \phi(t, x), \quad \Delta^- \phi(t, x) \stackrel{\text{def}}{=} \phi(t, x) - \phi(t, x - h).$$

Moreover, we associate with grid values $(y_j^n)_{j,n}$ and $(v_j)_j$ the piecewise constant functions y_h, y_h^n, and v_h by setting

$$y_h(t,x) = \sum_{n,j} y_j^n \mathbf{1}_{Q_j^n}(t,x), \quad y_h^n(x) = y_h(t_n,x), \quad v_h(x) = \sum_j v_j \mathbf{1}_{R_j}(x),$$

and use the convention $(y_h)_j^n \equiv y_j^n$, $(y_h^n)_j \equiv y_j^n$, $(v_h)_j \equiv v_j$. Finally, given a function $v \in L^1_{loc}(\mathbb{R})$, we obtain a grid function $T_h v$ by the averaging operator

$$T_h v(x) = \frac{1}{h} \int_{R_j} v(\xi)\, d\xi \quad \text{for } x \in R_j.$$

Let N_T such that $T \in [t_{N_T}, t_{N_T+1})$ (analogously, we define $N_{\bar t}$ for $\bar t \in (0,T]$). To discretize the state equation (1) we consider conservative finite difference schemes of the form

$$\begin{aligned} y_j^{n+1} &= y_j^n - \lambda \Delta^- f_{j+\frac12}^{h,n} \stackrel{\text{def}}{=} H(y_{j-K}^n, \dots, y_{j+K}^n), \quad j \in \mathbb{Z},\ n = 0, \dots, N_T - 1, \\ y_j^0 &= u_j, \quad j \in \mathbb{Z}, \end{aligned} \tag{12}$$

where

$$f_{j+\frac12}^{h,n} = f^h(y_{j-K+1}^n, \dots, y_{j+K}^n), \quad \Delta^- f_{j+\frac12}^{h,n} = f_{j+\frac12}^{h,n} - f_{j-\frac12}^{h,n}$$

with a consistent numerical flux f^h, i.e.,

$$f^h \in C^{1,1}_{loc}(\mathbb{R}^{2K}), \quad f^h(y, \dots, y) = f(y) \quad \text{for all } y \in \mathbb{R}. \tag{13}$$

We will sometimes assume that the scheme (12) is monotone, i.e.,

$$H(y_{j-K}^n, \dots, y_{j+K}^n) \quad \text{is nondecreasing in each argument.} \tag{14}$$

The grid function y_h corresponding to y_j^n is an approximation of the entropy solution y. For concreteness, the control $u \in L^\infty(\mathbb{R})$ is approximated by the cell averages

$$u_j = (T_h u)_j. \tag{15}$$

In terms of the associated piecewise constant functions the discrete control-to-state mapping is thus

$$u_h \longmapsto y_h. \tag{16}$$

As discrete approximation of the objective functional (3) we choose, for example

$$u_h \longmapsto J^h(y_h) \stackrel{\text{def}}{=} \int_{\mathbb{R}} \gamma_h(x) \psi(y_h(\bar t, x), y_{d,h}(x))\, dx = \sum_j h\, \gamma_j\, \psi(y_j^{N_{\bar t}}, y_{d,j}), \tag{17}$$

where $\gamma_j = (T_h\gamma)_j$ and $y_{d,j} = (T_h y_d)_j$ with associated grid-functions $\gamma_h, y_{d,h}$.

The assumptions ensure that the discrete control-to-state mapping (16) and, consequently, also the discrete objective functional (17) is continuously differentiable. Obviously, we have

$$d_{u_h} y_h \cdot \delta u_h = \mu_h, \tag{18}$$

where μ_h is the discrete sensitivity and the corresponding grid values μ_j^n solve the discrete sensitivity equation obtained by linearizing the scheme (12)

$$\begin{aligned} \mu_j^{n+1} &= \mu_j^n - \lambda \sum_{k=1-K}^{K} \Delta^-(f^{h,n}_{y_k, j+\frac{1}{2}} \mu^n_{j+k}), \\ \mu_j^0 &= \delta u_j, \end{aligned} \tag{19}$$

where $f^{h,n}_{y_k, j+\frac{1}{2}} = f^h_{y_k}(y_{j+1-K}, \dots, y_{j+K})$ and $f^h_{y_k}$, $k = 1-K, \dots, K$, denotes the partial derivative of $f^h(y_{1-K}, \dots, y_K)$ with respect to the $(k+K)$-th argument y_k.

If we set

$$a^n_{j+\frac{1}{2},k} = f^{h,n}_{y_k, j+\frac{1}{2}} \tag{20}$$

then the discrete sensitivity equation (19) reads

$$\begin{aligned} \mu_j^{n+1} &= \mu_j^n - \lambda \sum_{k=1-K}^{K} \Delta^-(a^n_{j+\frac{1}{2},k} \mu^n_{j+k}) \\ \mu_j^0 &= \delta u_j. \end{aligned} \tag{21}$$

Using (18), it becomes obvious that the action of the derivative of the discrete objective functional (17) is given by

$$\begin{aligned} &\frac{d}{du_h} J^h(y_h(u_h)) \cdot \delta u_h = \frac{d}{dy_h} J^h(y_h) \cdot \mu_h \\ &\quad = \int_{\mathbb{R}} \gamma_h(x) \psi_y(y_h(\bar{t}, x), y_{d,h}(x))\, \mu_h(\bar{t}, x)\, dx = \sum_{x_j \in I} h \gamma_j \psi_y(y_j^{N_{\bar{t}}}, y_{d,j})\, \mu_j^{N_{\bar{t}}} \end{aligned} \tag{22}$$

with the sensitivities μ_j^n according to (19) (or equivalently (21)) and associated grid function μ_h.

To derive the discrete adjoint scheme for (12), we introduce the discrete Lagrangian

$$\begin{aligned} &L(y_h, u_h, p_h) \\ &= J^h(y_h(u_h)) - h \sum_j \left(p_j^0 (y_j^0 - u_j) \sum_{n=0}^{N_{\bar{t}}-1} p_j^{n+1} (y_j^{n+1} - y_j^n + \lambda \Delta^- f^{h,n}_{j+\frac{1}{2}}) \right). \end{aligned}$$

Then, by standard adjoint calculus we obtain

$$\frac{d}{du_h} J^h(y_h(u_h)) \cdot \delta u_h = L_{u_h}(y_h, u_h, p_h) \cdot \delta u_h = h \sum_j p_j^0 \delta u_j, \tag{23}$$

where p_h solves the discrete adjoint equation

$$L_{y_h}(y_h, u_h, p_h) = 0,$$

which is equivalent to

$$L_{y_j^n}(y_h, u_h, p_h) = 0 \quad \forall\, j \in \mathbb{Z},\ n = 0, \ldots, N_{\bar{t}}.$$

This yields by using the linearization (21) the discrete adjoint scheme

$$p_j^n = p_j^{n+1} + \lambda \sum_{k=1-K}^{K} a_{j-k+\frac{1}{2},k} \Delta^+ p_{j-k}^{n+1}, \quad j \in \mathbb{Z},\ n = 0, \ldots, N_{\bar{t}} - 1, \tag{24}$$

$$p_j^{N_{\bar{t}}} = \gamma_j \psi_y(y_j^{N_{\bar{t}}}, y_{d,j}), \quad j \in \mathbb{Z}. \tag{25}$$

It is well known that monotone finite difference schemes converge to the unique entropy solution.

THEOREM 4 *Consider a scheme* (12)–(13) *that is monotone, see* (14). *Then for any* $u, \hat{u} \in L^\infty \cap L^1(\mathbb{R})$ *the corresponding grid function* y_h *satisfies*

1. $\|y_h(t, \cdot; u)\|_\infty \le \|u_h\|_\infty \le \|u\|_\infty \quad \forall\, t \in [0, T]$
2. $\|y_h(t, \cdot; u_h) - y(t, \cdot; \hat{u}_h)\|_1 \le \|u_h - \hat{u}_h\|_1 \le \|u - \hat{u}\|_1 \quad \forall\, t \in [0, T]$
3. *If* $u \in BV(\mathbb{R})$ *then*

$$|y_h(t, \cdot; u_h)|_{TV} \le |u_h|_{TV} \le |u|_{TV} \quad \forall\, t \in [0, T]$$

4. $y_h \to y$ *in* $L^\infty(0, T; L^1_{loc}(\mathbb{R}))$ *as* $h \searrow 0$, *where* $y = y(u)$ *is the entropy solution of* (1).
5. *There exists a constant* $C(t) > 0$ *such that*

$$\|y^h(t, \cdot; u_h) - y(t, \cdot; u)\|_1 \le C(t)\, |u|_{TV} h^{1/2} \quad \forall\, t \in [0, T],\ 0 < h \le h_0.$$

PROOF See, for example, Crandall and Majda (1980), point 5 is demonstrated in Kuznetsov (1976). □

REMARK 2 *For piecewise smooth solutions there exist improved versions of point 5., see, for example, Teng and Zhang (1997).*

Analogously to the OSLC (9) for entropy solutions, many standard schemes, such as the (modified) Lax-Friedrichs scheme and Engquist-Osher scheme satisfy for initial data $u^h = T_h u$, $u \in BV(\mathbb{R})$, $u_x \le M_{u'}$ a discrete OSLC of the form

$$\frac{\Delta^+ y_j^n}{h} \le \frac{1}{M_{u'}^{-1} + \beta n \Delta t} \quad \forall\, j \in \mathbb{Z},\ n = 0, \ldots, N_T - 1, \tag{26}$$

with a constant $\beta > 0$, see Nessyahu and Tadmor (1992) and point 4 of Theorem 4. Using an interpolation inequality between the one-sided Lipschitz norm and the L^1-norm, one can show the following.

Theorem 5 *Let the assumptions of Theorem 4 hold and let* (12)–(13) *satisfy the discrete OSLC* (26)*. Then, for any* $t > 0$ *and* $x \in \mathbb{R}$ *there exists a constant* $C(t) > 0$ *such that*

$$|y(t,x) - y_h(t,x)| \leq C(t) \left(1 + \max_{|\xi - x| \leq h^{1/3}} |y_x(t,\xi)|\right) h^{1/3}$$

Proof See Nessyahu and Tadmor (1992). □

3.2. Convergence of the adjoint scheme for Lipschitz continuous end data

We study now the convergence properties of the discrete adjoint scheme (24). Instead of the end condition (25), we consider first the case of

$$p_j^{N_{\bar{t}}} = p_j^{\bar{t}} \stackrel{\text{def}}{=} (T_h p^{\bar{t}})_j \tag{27}$$

for $p^{\bar{t}} \in C^{0,1}(\mathbb{R})$.

The analysis in this subsection is similar to that in Gosse and James (2000), where only the case of Lipschitz end data is considered. We will provide all estimates that are necessary to extend the convergence analysis later to the case of discontinuous end data, which is not considered in Gosse and James (2000).

To carry out the convergence analysis it will be convenient to associate with the grid values $a^n_{j+1/2,k}$ the functions

$$a_{k,h}(t,x) \stackrel{\text{def}}{=} \sum_{j,n} a^n_{j+\frac{1}{2},k} \mathbf{1}_{Q_j^n}(t,x), \qquad a^n_{k,h}(x) \stackrel{\text{def}}{=} a_{k,h}(t_n,x). \tag{28}$$

Moreover, we introduce for a grid function y_h and an interval I the discrete Lipschitz semi-norm

$$|y_h(t,\cdot)|_{Lip_h(I)} \stackrel{\text{def}}{=} \sup_{x \in I} \frac{|y_h(t,x+h) - y_h(t,x)|}{h}$$

and recall that for $t \in [t_n, t_{n+1})$ there holds

$$|y_h(t,\cdot)|_{TV,I} = |y_h^n|_{TV,I} = \sum_{x_{j+\frac{1}{2}} \in I} |\Delta^+ y_j^n|, \tag{29}$$

as long as $I \cap \partial I$ does not contain some $x_{j+1/2}$.

To ensure consistency with the continuous problem, we will need the following properties of the coefficients $a^n_{j,k}$.

Assumption 1 *There are constants* $M_a, h_0 > 0$ *such that for all* $h = \Delta t/\lambda \leq h_0$

$$\|a_{k,h}\|_\infty \leq M_a, \quad -K < k \leq K, \tag{30}$$

and

$$a_h \stackrel{\text{def}}{=} \sum_{k=1-K}^{K} a_{k,h} \to a \quad \textit{in } L^1_{loc}(\Omega_t^{cl}) \textit{ as } h \to 0. \tag{31}$$

Moreover, there exist a function $\alpha \in L^1(0,T)$ *and some* $h_0 > 0$ *such that for all* $h = \Delta t/\lambda \le h_0$ *the discrete OSLC holds*

$$\sum_{k=1-K}^{K} \Delta^+ a^n_{j-k+\frac{1}{2},k} \le \frac{h}{\Delta t} \int_{t_n}^{t_{n+1}} \alpha(s)\, ds \quad \forall\, j \in \mathbb{Z},\ n = 0, \dots, N_{\bar{t}} - 1. \tag{32}$$

In Assumption 2 further on in this section, we state the properties of the numerical flux function f^h that ensure Assumption 1.

We start by deriving a priori estimates for the adjoint scheme (24), (27). In order to derive the L^∞-stability we note that (24) can be written in the form

$$p^n_j = \sum_{k=-K}^{K} B^n_{j,k} p^{n+1}_{j-k}, \tag{33}$$

where with the Kronecker-symbol $\delta_{0,k}$ there is

$$\begin{aligned} &B^n_{j,k} = \delta_{0,k} + \lambda(a^n_{j-k-\frac{1}{2},k+1} - a^n_{j-k+\frac{1}{2},k}), \quad -K < k < K, \\ &B^n_{j,-K} = \lambda a^n_{j+K-\frac{1}{2},1-K}, \quad B^n_{j,K} = -\lambda a^n_{j-K+\frac{1}{2},K}. \end{aligned} \tag{34}$$

It is easy to check that

$$\sum_{k=-K}^{K} B^n_{j,k} = 1. \tag{35}$$

To derive bounds for the total variation of the associated grid function p_h we note that the difference of (24) for $j+1$ and j can be written as

$$\Delta^+ p^n_j = \sum_{k=-K}^{K} C^n_{j,k} \Delta^+ p^{n+1}_{j-k}, \tag{36}$$

where

$$\begin{aligned} &C^n_{j,k} = \delta_{0,k} + \lambda(a^n_{j-k+\frac{1}{2},k+1} - a^n_{j-k+\frac{1}{2},k}), \quad -K < k < K, \\ &C^n_{j,-K} = \lambda a^n_{j+K+\frac{1}{2},1-K}, \quad C^n_{j,K} = -\lambda a^n_{j-K+\frac{1}{2},K}. \end{aligned} \tag{37}$$

We observe that

$$\sum_{k=-K}^{K} C^n_{j+k,k} = 1 \tag{38}$$

and the following relation between $B^n_{j,k}$ and $C^n_{j,k}$ holds

$$\begin{aligned} C^n_{j,k} &= B^n_{j,k} + \lambda \Delta^- a^n_{j-k+\frac{1}{2},k+1}, \quad -K \le k < K \\ C^n_{j,K} &= B^n_{j,K}. \end{aligned} \tag{39}$$

Lemma 1 *If the coefficients $B^n_{j,k}$ in (3.2) satisfy*

$$B^n_{j,k} \ge 0, \quad -K \le k \le K, \quad \text{for all } \ j \in \mathbb{Z}, \ \ 0 \le n \le N_\tau - 1, \tag{40}$$

then the solution of the adjoint scheme (24), (27) satisfies

$$|p^n_j| = \|p_h\|_{\infty, Q^n_j} \le \|p^{\bar{t}}_h\|_{\infty, I^n_j},$$

where $I^n_j \overset{\text{def}}{=} [x_j - K(N_{\bar{t}} - n)h, x_j + K(N_{\bar{t}} - n)h]$. In particular, we have

$$\|p_h\|_{\infty, \Omega^{cl}_{t_{N_{\bar{t}}}}} \le \|p^{\bar{t}}_h\|_\infty \le \|p^{\bar{t}}\|_\infty. \tag{41}$$

Proof This follows directly from (33), (35) and (40). □

We will need the following discrete Gronwall inequality.

Proposition 2 *Let $b_n, M_n \ge 0$, $n \in \mathbb{N}_0$, with*

$$M_{n+1} \le (1 + \Delta t\, b_n)\, M_n, \quad n \ge 0.$$

Then

$$M_{n+1} \le M_0 \, \exp\left(\sum_{n'=0}^{n} \Delta t\, b_{n'} \right).$$

Proof Denote by M, b the piecewise constant functions with $M(t) = M_n$, $b(t) = b_n$ for $t \in [t_n, t_{n+1})$, $t_n = n\Delta t$. Then, for $t \in [t_n, t_{n+1}]$ there holds

$$M(t) \le M(t_n) + \int_{t_n}^{t} b(s) M(s)\, ds$$

and summing over n gives

$$M(t) \le M_0 + \int_0^t b(s) M(s)\, ds \quad \forall\, t \ge 0.$$

Now, the classical Gronwall lemma yields

$$M(t) \le M_0 e^{\int_0^t b(s)\, ds}$$

and inserting $t = t_{n+1}$ concludes the proof. □

LEMMA 2 *Assume that the discrete OSLC* (32) *holds. If the coefficients* $B^n_{j,k}$, $C^n_{j,k}$ *in* (3.2) *and* (3.2) *satisfy*

$$B^n_{j,k}, C^n_{j,k} \geq 0, \quad -K \leq k \leq K, \quad \forall\, j \in \mathbb{Z}, \quad 0 \leq n \leq N_{\bar{t}} - 1,$$

then the solution of the adjoint scheme (24), (27) *satisfies*

$$\frac{|\Delta^+ p^n_j|}{h} \leq |p^{\bar{t}}_h|_{Lip_h(I^n_j)}\, e^{\int_{t_n}^{t_{N_{\bar{t}}}} \alpha(s)\, ds},$$

where

$$I^n_j \stackrel{\text{def}}{=} [x_j - K(N_{\bar{t}} - n)h, x_j + K(N_{\bar{t}} - n)h].$$

In particular, we have for all $0 \leq n \leq N_{\bar{t}} - 1$ *and* $t \in [t_n, t_{n+1})$

$$|p_h(t,\cdot)|_{Lip_h(\mathbb{R})} \leq |p^{\bar{t}}_h|_{Lip_h(\mathbb{R})}\, e^{\int_{t_n}^{t_{N_{\bar{t}}}} \alpha(s)\, ds} \leq \|p^{\bar{t}}_x\|_\infty\, e^{\int_{t_n}^{t_{N_{\bar{t}}}} \alpha(s)\, ds}.$$

PROOF We use the abbreviation $N = N_{\bar{t}}$. By (36) and the nonnegativity of the coefficients $C^n_{j,k}$, we conclude that

$$|\Delta^+ p^n_j| \leq \sum_{k=-K}^{K} C^n_{j,k} |\Delta^+ p^{n+1}_{j-k}| \leq \left(\sum_{k=-K}^{K} C^n_{j,k} \right) \sup_{|j'-j| \leq K} |\Delta^+ p^{n+1}_{j'}|.$$

By inserting (39) and using (35), we derive

$$|\Delta^+ p^n_j| \leq \left(1 + \sum_{k=-K}^{K-1} \frac{\Delta t}{h} \Delta^- a^n_{j-k+\frac{1}{2},k+1} \right) \sup_{|j'-j| \leq K} |\Delta^+ p^{n+1}_{j'}|.$$

Since $\Delta^- a^n_{j-k+\frac{1}{2},k+1} = \Delta^+ a^n_{j-(k+1)+\frac{1}{2},k+1}$, we have

$$\sum_{k=-K}^{K-1} \Delta^- a^n_{j-k+\frac{1}{2},k+1} = \sum_{k=1-K}^{K} \Delta^+ a^n_{j-k+\frac{1}{2},k}.$$

Thus, we conclude, by (32), that

$$|\Delta^+ p^n_j| \leq \left(1 + \int_{t_n}^{t_{n+1}} \alpha(s)\, ds \right) \sup_{|j'-j| \leq K} |\Delta^+ p^{n+1}_{j'}|.$$

As in the proof of Lemma 1, we fix some $(t_{n'}, x)$, $0 \leq n' \leq N-1$, and set

$$I^n \stackrel{\text{def}}{=} [x - K(n-n')h, x + K(n-n')h].$$

After dividing by Δt the last estimate gives for all $n = n', \ldots, N-1$

$$|p^n_h|_{Lip_h(I^n)} \leq \left(1 + \Delta t \left(\frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \alpha(s)\, ds \right) \right) |p^{n+1}_h|_{Lip_h(I^{n+1})}.$$

Thus, the discrete Gronwall lemma in Proposition 2 yields for $n = n', \ldots, N-1$

$$|p_h(t_n,\cdot)|_{Lip_h(I^n)} \leq \left(|p_h^N|_{Lip_h(I^N)}\right) \cdot e^{\int_{t_n}^{t_N} \alpha(s)\,ds}.$$

From this the assertions of the Lemma follow immediately. □

The next Lemma estimates the discrete Lipschitz constant of p_h with respect to t in terms of the discrete Lipschitz constant with respect to x.

Lemma 3 *The solution of the adjoint scheme* (24) *satisfies*

$$\frac{|p_j^{n+1} - p_j^n|}{\Delta t} \leq \sup_{-K<k\leq K} |a^n_{j-k+\frac{1}{2},k}| \sum_{k=1-K}^{K} \frac{|\Delta^+ p_{j-k}^{n+1}|}{h}. \tag{42}$$

If, in addition, (30) *holds then we have in particular*

$$\frac{|p_j^{n+1} - p_j^n|}{\Delta t} \leq 2KM_a |p_h(t_{n+1},\cdot)|_{Lip_h(\mathbb{R})}.$$

Proof From (24) we see that

$$|p_j^{n+1} - p_j^n| \leq \lambda \sum_{k=1-K}^{K} |a^n_{j-k+\frac{1}{2},k}||\Delta^+ p_{j-k}^{n+1}|.$$

Now the lemma is obvious. □

Lemma 4 *Let for the coefficients* $C^n_{j,k}$ *in* (3.2) *hold*

$$C^n_{j,k} \geq 0, \quad -K \leq k \leq K, \quad \textit{for all} \;\; j \in \mathbb{Z}, \;\; 0 \leq n \leq N_{\bar{t}} - 1.$$

Then, for any $n = 0, \ldots, N_{\bar{t}} - 1$ *and any open interval* $I = (z_1, z_2)$ *the solution of the adjoint scheme* (24), (27) *satisfies*

$$|p_h(t^n,\cdot)|_{TV,I} \leq |p_h^{\bar{t}}|_{TV,I^n} \leq |p^{\bar{t}}|_{TV,I^n+[-h,h]},$$

where

$$I^n = (z_1 - K(N_{\bar{t}} - n)h, z_2 + K(N_{\bar{t}} - n)h).$$

Moreover, if, in addition, (30) *holds then one has for any* $0 \leq n' < n \leq N_{\bar{t}}$

$$\|p_h(t^n,\cdot) - p_h(t^{n'},\cdot)\|_{1,I} \leq (t^n - t^{n'})2KM_a \|p_h\|_{L^\infty(t^{n'},t^n;BV(\hat{I}))}, \tag{43}$$

where $\hat{I} \stackrel{\text{def}}{=} (z_1 - Kh, z_2 + Kh)$.

Proof We use the abbreviation $N = N_{\bar{t}}$. By (36) and the nonnegativity of the coefficients $C^n_{j,k}$, we obtain, as before

$$|\Delta^+ p_j^n| \leq \sum_{k=-K}^{K} C^n_{j,k} |\Delta^+ p_{j-k}^{n+1}|.$$

Let an open interval $I = (z_1, z_2)$ be given, fix some $t_{n'}$, $0 \le n' \le N-1$, and set

$$I^n \stackrel{\text{def}}{=} (z_1 - K(n-n')h, z_2 + K(n-n')h).$$

Let $n \in \{n', \dots, N-1\}$ be arbitrary. Summing the last inequality for all j with $x_{j+1/2} \in I^n$ yields, by (29)

$$|p_h^n|_{TV,I^n} \le \sum_{x_{j+\frac{1}{2}} \in I^n} \sum_{k=-K}^{K} C_{j,k}^n |\Delta^+ p_{j-k}^{n+1}|.$$

Using the nonnegativity of $C_{j,k}^n$ together with (38), we obtain the estimate

$$|p_h^n|_{TV,I^n} \le \sum_{x_{j+\frac{1}{2}} \in I^n} \sum_{k=-K}^{K} C_{j,k}^n |\Delta^+ p_{j-k}^{n+1}| = \sum_{k=-K}^{K} \sum_{x_{j+k+\frac{1}{2}} \in I^n} C_{j+k,k}^n |\Delta^+ p_j^{n+1}|$$

$$\le \sum_{x_{j+\frac{1}{2}} \in I^{n+1}} \left(\sum_{k=-K}^{K} C_{j+k,k}^n \right) |\Delta^+ p_j^{n+1}| = \sum_{x_{j+\frac{1}{2}} \in I^{n+1}} |\Delta^+ p_j^{n+1}| = |p_h^{n+1}|_{TV,I^{n+1}}.$$

Hereby, we have used, besides (29), that for any $k = -K, \dots, K$ there holds

$$\{j \,:\, x_{j+k+1/2} \in I^n\} \subset \{j \,:\, x_{j+1/2} \in I^{n+1}\}.$$

This proves the first assertion. Now let, in addition, (30) hold. Using Lemma 3, the weighted sum of (42) for $\{j \,:\, R_j \cap I \ne \emptyset\} = \{j \,:\, x_{j+1/2} \in (z_1, z_2 + h)\}$ with weights $\Lambda_1(R_j \cap I)$ $(= h$ if $R_j \subset I$, Λ_1 is the Lebesgue measure on $\mathbb{R})$ yields, by (29)

$$\frac{\|p_h^{n+1} - p_h^n\|_{1,I}}{\Delta t} \le M_a \sum_{k=1-K}^{K} \sum_{x_{j+\frac{1}{2}} \in (z_1, z_2+h)} |\Delta^+ p_{j-k}^{n+1}| \le 2KM_a |p_h^{n+1}|_{TV,\hat{I}}.$$

Summing over n and applying the triangle inequality on the left hand side yields (43). The proof is complete. □

THEOREM 6 *Let $a \in L^\infty(\Omega_{\bar{t}})$, $p^{\bar{t}} \in C^{0,1}(\mathbb{R})$ and assume that* (30), (31), *and* (32) *hold (then a satisfies automatically the OSLC* (9)*). Moreover, let the coefficients $B_{j,k}^n, C_{j,k}^n$ in* (3.2) *and* (3.2) *satisfy*

$$B_{j,k}^n, C_{j,k}^n \ge 0, \quad -K \le k \le K, \quad \text{for all } \; j \in \mathbb{Z}, \quad 0 \le n \le N_{\bar{t}} - 1.$$

Then, the solution of the adjoint scheme (24)–(27) *converges locally uniformly to the unique reversible solution $p \in C^{0,1}(\Omega_{\bar{t}}^{cl})$ of* (8)*, i.e.,*

$$p_h \to p \quad \text{in } B([0,\bar{t}] \times [-R,R]) \text{ for all } R > 0 \text{ as } h = \Delta t/\lambda \to 0.$$

Remark 3 *Note that Theorem 6 does not require that the state y_h be generated by the scheme, to which the adjoint scheme belongs, it is only important that y_h ensures* (30)*,* (31)*, and* (32)*. Hence, also an optimize-then-discretize approach is covered.*

We show first the following auxiliary result.

Lemma 5 *Under the assumptions of Theorem 6 any sequence $h_i \to 0$ contains a subsequence $h_i' \to 0$ such that the corresponding solutions $p_{h_i'}$ of the adjoint scheme* (24)*,* (27) *satisfy with $I = [-R, R]$ for all $R > 0$*

$$p_{h_i'} \to p \quad \text{in } B([0,\bar{t}] \times I), \tag{44}$$

where

$$p \in C^{0,1}([0,\bar{t}] \times \mathbb{R}) \tag{45}$$

is a solution of (8)*.*

Proof By Lemmas 1, 2 we find a constant $M_p > 0$ such that for $h \le h_0$ there holds

$$\|p_h\|_{B(\Omega_{\bar{t}})} \le M_p. \tag{46}$$

Moreover, by Lemma 2 there exists a constant L_x with

$$|p_h|_{B([0,\bar{t}];Lip_{p_h}(\mathbb{R}))} \le L_x. \tag{47}$$

Using (47), Lemma 3 yields for all $t \in [\sigma, \bar{t})$ the discrete Lipschitz estimate in time

$$\sup_{t' \in (t,\bar{t}]} \frac{\|p_h(t',\cdot) - p_h(t,\cdot)\|_{B(\mathbb{R})}}{|t'-t| + \Delta t} \le 2KM_aL_x \stackrel{\text{def}}{=} L_t. \tag{48}$$

We show next, by an Arzela-Ascoli type argument, that any sequence $h_i \to 0$ contains a subsequence (h_i') such that with $I \stackrel{\text{def}}{=} [-R, R]$ for all $R > 0$ (44) holds, where $p \in C^{0,1}(\Omega_{\bar{t}})$.

In fact, we choose a countable dense subset (z_l) of $\Omega_{\bar{t}}^{cl}$ and may by (46) select a diagonal subsequence (h_i') such that $p_{h_i'}(z_l)$ converges for all z_l.

Now, let $R > 0$ be arbitrary but fixed and set $D_R \stackrel{\text{def}}{=} [0,\bar{t}] \times [-R, R]$. For every $\delta > 0$ we then find an L_δ, such that the δ-balls $(B_\delta(z_l))_{1 \le l \le L_\delta}$ cover the compact set D_R. Then, we find N_δ such that

$$\sup_{1 \le l \le L_\delta} |p_{h_i'}(z_l) - p_{h_j'}(z_l)| \le \delta \quad \text{for all } i, j \ge N_\delta$$

and may choose N_δ without restriction such that $(\Delta t')_i, (h')_i < \delta$ for $i \ge N_\delta$. Then, for any $(t,x) \in D_R$ there is $1 \le l \le L_\delta$ with $(t,x) \in D_R \cap B_\delta(z_l)$ and (47)–(48) yield

$$|p_{h_i'}(t,x) - p_{h_j'}(t,x)| \le \delta + (L_x + L_t)2\delta \quad \forall\, i, j \ge N_\delta.$$

This shows that

$$p_{h_i'} \to p \quad \text{in } B(D_R). \tag{49}$$

Moreover, since the choice of the dense set (z_l) and the diagonal sequence (h_i') does not depend on R, this holds on D_R for all $R > 0$. The regularity properties (45) of the limit p are by the convergence (49) inherited from the properties (46), (47) and (48) of p_h.

In the next step we show that p solves (8). Clearly, p satisfies the end condition, since by (27) and (44) we have for all $I = [-R, R]$, $R > 0$,

$$\|p(\bar{t}, \cdot) - p^{\bar{t}}\|_{B(I)} \le \lim_{i\to\infty} \|(p - p_{h_i})(\bar{t}, \cdot)\|_{B(I)} + \|p_{h_i}(\bar{t}, \cdot) - p^{\bar{t}}\|_{B(I)} = 0.$$

In the following it will be convenient to recall for any $\phi \in B(\Omega_{\bar{t}})$ the notation

$$\Delta^+\phi(t,x) \stackrel{\text{def}}{=} \phi(t, x+h) - \phi(t,x), \quad \Delta^-\phi(t,x) \stackrel{\text{def}}{=} \phi(t,x) - \phi(t, x-h).$$

Using (28), the adjoint scheme (24) can be written down as

$$\frac{p_h(t,x) - p_h(t-\Delta t, x)}{\Delta t} + \sum_{k=1-K}^{K} a_{k,h}(t-\Delta t, x-kh)\frac{\Delta^+ p_h(t, x-kh)}{h} = 0.$$

For convenience, we write in the sequel h instead of h_i. It is obvious that

$$\frac{p_h(t,x) - p_h(t-\Delta t, x)}{\Delta t} \to p_t \quad \text{in } \mathcal{D}'(\Omega_{\bar{t}}) \text{ as } h \to 0,$$

since for all $\phi \in \mathcal{D}(\Omega_{\bar{t}})$ and $h \le \operatorname{dist}(\operatorname{supp}\phi, \partial\Omega_{\bar{t}})$ there holds

$$\begin{aligned}\int_{\Omega_{\bar{t}}} \phi \frac{p_h(t,x) - p_h(t-\Delta t, x)}{\Delta t}\, dx\, dt &= \int_{\Omega_{\bar{t}}} \frac{\phi(t,x) - \phi(t+\Delta t, x)}{\Delta t}\, p_h\, dx\, dt \\ &\to -\int_{\Omega_{\bar{t}}} \phi_t\, p\, dx\, dt \quad \text{as } h \to 0.\end{aligned}$$

Hereby, we have used (44) and the fact that the difference quotient of ϕ converges boundedly everywhere to ϕ_t. Moreover, we have for all $\phi \in \mathcal{D}(\Omega_{\bar{t}})$ and with a_h defined in (31)

$$\begin{aligned}&\int_{\Omega_{\bar{t}}} \phi \sum_{k=1-K}^{K} a_{k,h}(t-\Delta t, x-kh)\frac{\Delta^+ p_h(t,x-kh)}{h}\, dx\, dt \\ &= \int_{\Omega_{\bar{t}}} \sum_{k=1-K}^{K} \phi(t, x+kh) a_{k,h}(t-\Delta t, x)\, \frac{p_h(t,x+h) - p_h(t,x)}{h}\, dx\, dt \\ &= \int_{\Omega_{\bar{t}}} \phi(t,x)\, a_h(t-\Delta t, x)\frac{p_h(t,x+h) - p_h(t,x)}{h}\, dx\, dt \\ &\quad + \int_{\Omega_{\bar{t}}} \sum_{k=1-K}^{K} (\phi(t,x+kh) - \phi(t,x)) a_{k,h}(t-\Delta t, x)\frac{\Delta^+ p_h(t,x)}{h}\, dx\, dt \\ &=: I_1 + I_2.\end{aligned}$$

Now choose $R > 0$ large enough such that $\operatorname{supp}\phi \subset [0,\bar{t}] \times [-R,R] \stackrel{\text{def}}{=} D_R$. I_2 tends to zero, since (30) and (47) yield

$$|I_2| \leq M_a L_x \sum_{k=1-K}^{K} \|\phi(t,x+kh) - \phi(t,x)\|_{1,\Omega_{\bar{t}}} \to 0 \quad \text{as } h \to 0.$$

To analyze the first term, we note that, as above

$$\frac{p_h(t,x+h) - p_h(t,x)}{h} \to p_x \quad \text{in } \mathcal{D}'(\Omega_{\bar{t}})$$

and also in $L^\infty(D_R)$-weak*, since its absolute value is on D_R bounded by L_x. On the other hand, (31) yields

$$\phi a_h(\cdot - \Delta t, \cdot) = \phi a + \phi\big((a(\cdot - \Delta t, \cdot) - a) + (a_h - a)(\cdot - \Delta t, \cdot)\big) \to \phi\, a \quad \text{in } L^1(\Omega_{\bar{t}})$$

as $h \to 0$ and has support in D_R. Therefore, we obtain

$$\lim_{h_i \to 0} I_1 + I_2 = \int_{\Omega_{\bar{t}}} \phi\, a\, p_x \, dx\, dt.$$

This shows that the limit p of $p_{h_i'}$ satisfies (8) in the sense of distributions, where p_t and p_x are the distributional derivatives. Using the local Lipschitz-continuity of p on $\Omega_{\bar{t}}$, p is by Rademacher's theorem almost everywhere differentiable in the classical sense with partial derivatives p_t and p_x. Thus, p is a classical solution of (8) and has the regularity (45). The lemma is proven. □

As a quite immediate consequence we can now prove Theorem 6, since reversible solutions can be characterized by the monotonicity property of Definition 1.

Proof (of Theorem 6) It is obvious, by (31) and (32), that the limit coefficient a satisfies the OSLC (9). Thus, (8) has, by Theorem 2, a unique reversible solution $p \in C^{0,1}(\Omega_{\bar{t}})$. We know from Bouchut and James (1998), see Definition 1, that a Lipschitz continuous solution p of (8) on $\Omega_{\bar{t}}$ is the unique reversible solution, if and only if there exist Lipschitz solutions p_1, p_2 of $p_t + a p_x = 0$ with

$$(p_1)_x, (p_2)_x \geq 0, \quad p = p_1 - p_2 \quad \text{on } \Omega_{\bar{t}}. \tag{50}$$

Now, given the sequence of end data $p_h^{\bar{t}}$, we easily find by collecting only up jumps or down jumps, respectively, monotone increasing end data $(p_l)_h^{\bar{t}}$, $l = 1, 2$, with discrete Lipschitz constant $L_{\bar{t}}$, such that $p_h^{\bar{t}} = (p_1)_h^{\bar{t}} - (p_2)_h^{\bar{t}}$. Hence, we find a sequence $h_i \to 0$ with $(p_l)_{h_i}^{\bar{t}} \to p_l^{\bar{t}} \in Lip(\mathbb{R})$ in $B_{loc}(\mathbb{R})$ as $i \to \infty$, where $(p_l^{\bar{t}})_x \geq 0$, $l = 1, 2$, and $p^{\bar{t}} = p_1^{\bar{t}} - p_2^{\bar{t}}$. Using Lemma 5, we can choose a subsequence h_i' such that for $i \to \infty$ the corresponding solutions $p_{h_i'}$ and $(p_l)_{h_i'}$ of (24)–(25) converge in $B([0,\bar{t}] \times [-R,R])$ for all $R > 0$ to Lipschitz solutions p, p_l of (8) for end data $p^{\bar{t}}$, $p_l^{\bar{t}}$, respectively. In particular, we have $p = p_1 - p_2$.

It remains to show that $(p_l)_x \geq 0$, $l = 1, 2$. But (36) yields, together with $C^n_{j,k} \geq 0$, the monotonicity of the adjoint scheme (24). Thus, the monotonicity properties of the end data $(p_l)^{\bar{t}}_h$, $l = 1, 2$, are preserved. Hence, $p = p_1 - p_2$ is the unique reversible solution of (8).

Therefore, we have shown that any sequence $h_i \to 0$ contains a subsequence h'_i with $p_{h'_i} \to p$ in the sense of (44), where p is the unique reversible solution of (8). Thus, $p_h \to p$ in the sense of (44) holds generally for $h \to 0$ by a subsequence-subsequence-argument. The proof is complete. □

We recall that the coefficient in the adjoint scheme (24) is defined by

$$a^n_{j+\frac{1}{2},k} = f^{h,n}_{y_k, j+\frac{1}{2}}. \tag{20}$$

Moreover, the adjoint equation (6), (7) can be written in the form (8) with the coefficient

$$a = f'(y). \tag{51}$$

In order to apply the convergence results of the previous section we have to verify Assumption 1 as well as

$$B^n_{j,k}, C^n_{j,k} \geq 0, \quad -K \leq k \leq K, \quad \text{for all } \; j \in \mathbb{Z}, \;\; 0 \leq n \leq N_{\bar{t}} - 1. \tag{52}$$

For convenience, we recall that by (3.2), (20)

$$\begin{aligned} B^n_{j,k} &= \delta_{0,k} + \lambda(f^{h,n}_{y_{k+1}, j-k-\frac{1}{2}} - f^{h,n}_{y_k, j-k+\frac{1}{2}}), \quad -K < k < K, \\ B^n_{j,-K} &= \lambda f^{h,n}_{y_{1-K}, j+K-\frac{1}{2}}, \quad B^n_{j,K} = -\lambda f^{h,n}_{y_K, j-K+\frac{1}{2}}, \end{aligned} \tag{53}$$

and by (3.2), (20)

$$\begin{aligned} C^n_{j,k} &= \delta_{0,k} + \lambda(f^{h,n}_{y_{k+1}, j-k+\frac{1}{2}} - f^{h,n}_{y_k, j-k+\frac{1}{2}}), \quad -K < k < K, \\ C^n_{j,-K} &= B^n_{j+1,-K}, \quad C^n_{j,K} = B^n_{j,K}. \end{aligned} \tag{54}$$

We need the following properties of the numerical flux function.

Assumption 2 *$f^h \in C^{1,1}_{loc}(\mathbb{R}^{2K})$ and is consistent with f, i.e., (13) holds.*

With constants $h_0, M_y > 0$ and the entropy solution $y = y(u)$ of (1) for all $h = \Delta t/\lambda \leq h_0$ there holds

$$\|y_h\|_\infty \leq M_y, \quad y_h(t,\cdot) \to y(t,\cdot) \quad \text{in } L^1_{loc}(\mathbb{R}) \text{ for all } t \in [0,T] \text{ as } h \to 0, \tag{55}$$

$$f^h_{y_k} \text{ are on } [-M_y, M_y]^{2K} \text{ nondecreasing in each argument.} \tag{56}$$

With a function $\gamma \in L^1(0,T)$ and some $h_0 > 0$ for all $h = \Delta t/\lambda \leq h_0$ the discrete OSLC holds

$$\Delta^+ y^n_j \leq \frac{h}{\Delta t} \int_{t_n}^{t_{n+1}} \gamma(t)\, dt \quad \forall\, j \in \mathbb{Z}, \; n = 0, \ldots, N_T - 1. \tag{57}$$

We show now that Assumption 2 implies Assumption 1.

Lemma 6 (i) *If* (13) *holds for* f^h *and* y_h *satisfies* (55)*, then the coefficients* $a^n_{j+1/2,k} = f^n_{y_k,j+1/2}$ *satisfy* (30) *and* (31) *and we can choose*

$$M_a = \sup_{\substack{y\in[-M_y,M_y]^{2K}\\ -K<k\le K}} |f^h_{y_k}(y)|. \tag{58}$$

(ii) *If* (56) *holds for* f^h *and* y_h *satisfies* (55)*,* (57) *then the coefficients* $a^n_{j+1/2,k} = f^n_{y_k,j+1/2}$ *satisfy the discrete OSLC* (32)*.*

Proof (i): By (55) we have $\|y_h\|_\infty, \|y\|_\infty \le M_y$ for $h \le h_0$. Moreover, $f^h_{y_k}$ are continuous on $[-M_y, M_y]^{2K}$ by (13). Thus, M_a in (58) is bounded and is obviously an upper bound for $|a^n_{j+1/2,k}| = |f^n_{y_k,j+1/2}|$ if $h \le h_0$. This yields (30). It remains to show (31). Since (13) ensures $f^h(y,\dots,y) = f(y)$, we have

$$\sum_{k=1-K}^{K} f^h_{y_k}(y,\dots,y) = f'(y).$$

Therefore, we obtain for $(t,x) \in \Omega_T$

$$\begin{aligned}\left|\sum_{k=1-K}^{K}\Big(f'(y(t,x)) - (f^h_{y_k})_h(t,x)\Big)\right| = \Bigg|\sum_{k=1-K}^{K}\Big(&f^h_{y_k}(y(t,x),\dots,y(t,x))\\ &-f^h_{y_k}(y_h(t,x-(K-1)h),\dots,y_h(t,x+Kh))\Big)\Bigg|.\end{aligned} \tag{59}$$

(55) yields for all $I = [-R,R]$, $R > 0$, and all $h = \Delta t/\lambda \le h_0$

$$\begin{aligned}\|y_h(\cdot,\cdot+kh) - y\|_{1,(0,T)\times I} &\le \|y_h - y\|_{1,(0,T)\times(-R-1,R+1)}\\ &\quad + \|y(\cdot,\cdot+kh) - y\|_{1,(0,T)\times I} \to 0 \quad \text{as } h\to 0.\end{aligned}$$

Since $f^h_{y_k}$ are by (13) Lipschitz continuous on $[-M_y,M_y]^{2K}$, we see that the right hand side of (59) tends to zero in $L^1_{loc}(\Omega_T^{cl})$, which shows (31).

(ii): Let y_h satisfy (55) and (57). By (13), $f^h_{y_k}$ has a Lipschitz constant L_k on $[-M_y,M_y]^{2K}$. We use the notation $\alpha\vee\beta := \max(\alpha,\beta)$. Then the monotonicity of $f^h_{y_k}$ in all arguments on $[-M_y,M_y]^{2K}$, ensured by (56), yields, by (57), where we assume without restriction that $\gamma \ge 0$

$$\begin{aligned}\Delta^+ f^{h,n}_{y_k,j-k+\frac12} &\le f^h_{y_k}(y^n_{j-k-K+2}\vee y^n_{j-k-K+1},\dots) - f^{h,n}_{y_k,j-k+\frac12}\\ &\le L_k\sum_{l=1-K}^{K}\max(\Delta^+ y^n_{j-k+l},0) \le \frac{h}{\Delta t}\int_{t_n}^{t_{n+1}} L_k\gamma(t)\,dt.\end{aligned}$$

Therefore, (32) holds with $\alpha = (L_{1-K} + \ldots + L_K)\gamma$, and (57) yields $\alpha \in L^1(0,T)$. □

The previous lemma shows that the convergence results of Theorem 6 can be applied if the coefficients $B^n_{j,k}$ in (53) and $C^n_{j,k}$ in (54) satisfy (52). This condition is examined in the following lemma.

LEMMA 7 *Let* (13) *hold.*

(i) *The condition*

$$B^n_{j,k} \geq 0 \quad \textit{for all } y^n_j \in [-M_y, M_y]$$

with $B^n_{j,k}$ *in* (53) *is satisfied if and only if the finite difference scheme* (12) *is monotone on* $[-M_y, M_y]$ *in the sense of* (14).

(ii) *If* $\|y_h\|_\infty \leq M_y$ *and the coefficients* $B^n_{j,k}$ *in* (53) *satisfy* $B^n_{j,k} \geq \beta > 0$, $-K < k < K$ *then the coefficients* $C^n_{j,k}$ *in* (54) *satisfy automatically* $C^n_{j,k} \geq 0$ *under the Courant-Friedrichs-Lewy (CFL) condition*

$$\lambda = \frac{\Delta t}{h} \leq \frac{\beta}{2M_a}, \quad M_a \textit{ as in } (58).$$

PROOF (i): It is easy to check that the partial derivative of the scheme (12) is given by $B^n_{j+k,k}$. Hence, (i) is obvious.

(ii): Now assume that $B^n_{j,k} \geq \beta > 0$, $-K < k < K$. Then we have, by (3.2), (3.2), $C^n_{j,K} = B^n_{j,K} \geq 0$ and $C^n_{j,-K} = B^n_{j+1,-K} \geq 0$, and by (39)

$$C^n_{j,k} = B^n_{j,k} + \lambda \Delta^- a^n_{j-k+\frac{1}{2},k+1} \geq \beta - 2\lambda M_a, \quad -K < k < K$$

with M_a from (58). Therefore, $C^n_{j,k} \geq 0$ is ensured under the CFL condition $\lambda \leq 2M_a/\beta$. □

REMARK 4 *In the case of a monotone three-point scheme, i.e.,* $K = 1$*, one has only to check* $C^n_{j,0} \geq 0$.

3.3. Convergence of the adjoint scheme for discontinuous end data

Consider the situation of Theorem 1, where $y(\bar{t}, \cdot; u)$ is piecewise C^1 and has finitely many shocks at $x_1 < x_2 < \ldots < x_K$. Then, the end data of the adjoint equation (6), (7) for the adjoint-based derivative representation (5) of the objective functional (3) are given by

$$p^{\bar{t}}(x) = \begin{cases} \gamma(x)\psi_y(y(\bar{t},x), y_d(x)) & \text{if } x \notin \{x_1, \ldots, x_K\} \\ \gamma(x)\frac{[\psi(y(\bar{t},x), y_d(x))]}{[y(\bar{t},x)]} & \text{if } x \in \{x_1, \ldots, x_K\} \end{cases}, \quad x \in \mathbb{R}. \tag{60}$$

and are thus discontinuous and contained in $B_{Lip}(\mathbb{R})$, see (11). Even more, they have particular values at the points $x \in \{x_1, \ldots, x_K\}$, which are propagated in the whole shock funnel, see Remark 1.

It has been demonstrated in Giles and Ulbrich (2010a,b) that the convergence of the discrete adjoint, given by (24), (25), is not ensured within the shock funnels if the numerical scheme has numerical viscosity $O(h)$. However, convergence was proven in Giles and Ulbrich (2010a,b) for the end data (25) and a modified Lax-Friedrichs scheme with numerical viscosity $O(h^\beta)$ for appropriate $0 < \beta < 1$.

We present now a new approach that ensures convergence of the discrete adjoint by using (24) with slightly modified end data. To this end, we use the fact that the correct end data of the adjoint equation are given by (60) and that we can use $y_h(\bar{t}, \cdot; u_h)$ and the pointwise convergence result of Theorem 5 to compute convergent approximations of the left and right limits $y(\bar{t}, x\pm; u)$ at $x \in \{x_1, \dots, x_K\}$.

To approximate the shock locations $x_1, \dots, x_K$ we determine the K regions, where $\Delta^+ y_j^{N_{\bar{t}}} = -O(\sqrt{h})$ and choose x_k^h as the middle point x_{j_k} of the k-th region. Then, we approximate $p^{\bar{t}}(x_k)$ in (60) by

$$p^{\bar{t}}_{x_k^h} = \gamma(x_k^h) \frac{[\psi(y_h(\bar{t}, x_k^h + h^{1/3}), y_d(x_k^h)) - \psi(y_h(\bar{t}, x_k^h - h^{1/3}), y_d(x_k^h))]}{[y_h(\bar{t}, x_k^h + h^{1/3}) - y_h(\bar{t}, x_k^h - h^{1/3})]}.$$

Now, let $r > 0$ with $r < \min_{1 \le k < K} |x_{k+1}^h - x_k^h|/8$ and define the weighting function

$$\omega^r(x) = \begin{cases} 1 & \text{if } |x| \le r, \\ \max\left\{\frac{2r - |x|}{r}, 0\right\} & \text{if } |x| > r. \end{cases}$$

Next, we approximate (60) by

$$p_j^{N_{\bar{t}}, r} = \begin{cases} \gamma_j \psi_y(y_j^{N_{\bar{t}}}, y_{d,j}) & \text{if } |x_j - x_k^h| > 2r, 1 \le k \le K, \\ \omega^r(x_j - x_k^h) p^{\bar{t}}_{x_k^h} + & \text{otherwise.} \\ (1 - \omega^r(x_j - x_k^h)) \gamma_j \psi_y(y_j^{N_{\bar{t}}}, y_{d,j}) & \end{cases} \tag{61}$$

We have the following result.

Theorem 7 *Let y satisfy the assumptions of Theorem 1. Consider the scheme* (24) *with end data* (61). *Assume that* (30), (31), *and* (32) *hold. Moreover, let the coefficients $B_{j,k}^n, C_{j,k}^n$ in* (3.2) *and* (3.2) *satisfy*

$$B_{j,k}^n, C_{j,k}^n \ge 0, \quad -K \le k \le K, \quad \text{for all } \ j \in \mathbb{Z}, \ \ 0 \le n \le N_{\bar{t}} - 1.$$

Then, there exists a piecewise constant function $r(h) > 0$ with $r(h) \to 0$ as $h \to 0$ such that with the choice $r = r(h)$ in (61) *the solution of the adjoint scheme* (24), (61) *satisfies*

$$p_h \to p \ \text{ in } C([0, \bar{t}]; L^1_{loc}(\mathbb{R})) \text{ and boundedly everywhere on } \Omega_{\bar{t}}^{cl} \text{ as } h \to 0$$

with the unique reversible solution p of the adjoint equation (6), (7).

REMARK 5 *Note again that Theorem 7 does not require that the state y_h be generated by the scheme, to which the adjoint scheme belongs; y_h has only to ensure* (30), (31), (32) *and that the convergence properties of Theorems 4 and 5 hold. Hence, also an optimize-then-disretize approach is covered.*

PROOF (of Theorem 7) Similarly to (61) we define $p^{\bar{t},r} \in C^{0,1}(\mathbb{R})$ by

$$p^{\bar{t},r}(x) = \begin{cases} \gamma(x)\psi_y(y(\bar{t},x), y_d(x)) & \text{if } |x - x_k| > 2r, 1 \le k \le K, \\ \omega^r(x - x_k)\frac{[\psi(y(\bar{t},x_k),y_d(x_k))]}{[y(\bar{t},x_k)]} & \text{otherwise.} \\ +(1 - \omega^r(x - x_k))\gamma(x)\psi_y(y(\bar{t},x), y_d(x)) & \end{cases}$$

We consider first the case of fixed $r > 0$. Theorem 4, point 5. yields

$$\|y^h(\bar{t},\cdot;u_h) - y(\bar{t},\cdot;u)\|_1 = O(h^{1/2}),$$

and this implies $|x_k - x_k^h| = O(h^{1/2})$. Therefore, for h small enough, $y(\bar{t},\cdot)$ is C^1 outside of $[x_k^h - h^{1/3}, x_k^h + h^{1/3}]$ and thus Theorem 5 yields that $p_h^{\bar{t},r}$ corresponding to (61) converges uniformly to $p^{\bar{t},r}$. If p_h^r and p^r denote the corresponding solution of (24), (61) and the reversible solution of (8) with data $p^{\bar{t},r}$, respectively, then Theorem 6 yields

$$p_h^r \to p^r \quad \text{in } B([0,\bar{t}] \times [-R,R]) \text{ for all } R > 0 \text{ as } h = \Delta t/\lambda \to 0. \tag{62}$$

Moreover, $p^{\bar{t},r}$ converges for $r \searrow 0$ to $p^{\bar{t}} \in B_{Lip}(\mathbb{R})$ in the sense of Theorem 3. Hence, Theorem 3 yields

$$p^r \to p \ \text{ in } C([0,\bar{t}]; L^1_{loc}(\mathbb{R})) \text{ and boundedly everywhere on } \Omega^{cl}_{\bar{t}}, \tag{63}$$

where p is the reversible solution of the adjoint equation (6), (8).

To conclude the proof, we define a piecewise constant function $r(h) > 0$ with $r(h) \to 0$ for $h \to 0$ as follows. Let $0 < h_0 < 1$ be the initial grid size and choose $r_0 > h_0$, for example, as $r_0 = h_0^{1/3}$. First of all, we note that p^r and p_h^r are independent of $0 < r \le r_0$ outside of $[0,\bar{t}] \times [-R,R]$ for $R > 0$ big enough by the finite propagation speed of (8) and of the scheme (24).

Let $(\nu_i)_{i\in\mathbb{N}}$ be a monotone decreasing sequence with $\nu_i \to 0$, for example $\nu_i = \frac{1}{i}$. We construct inductively a sequence $h_0 > h_i \searrow 0$ such that with $r_i = h_{i-1}^{1/3}$ for all $i \in \mathbb{N}$ there holds

$$\|p_h^{r_i} - p^{r_i}\|_{B([0,\bar{t}]\times[-R,R])} \le \nu_i \quad \forall\, 0 < h \le h_i. \tag{64}$$

This is possible by (62). Now, we set

$$r(h) = r_i \text{ for } h \in (h_{i+1}, h_i].$$

Then, (63) and (64) yield

$$p^r(h) \to p \ \text{ in } C([0,\bar{t}]; L^1_{loc}(\mathbb{R})) \text{ and boundedly everywhere on } [0,\bar{t}] \times [-R,R].$$

Since, $p^r = p$ as well as p_h^r are independent of r outside of $[0, \bar{t}] \times [-R, R]$, the convergence there follows from (62). □

Theorem 7 does not give an explicit formula for the choice of $r(h)$. It should be enough if the intervals $[x_k^h - r(h), x_k^h + r(h)]$ cover the numerical shock profile and hence, by Theorem 5, $r(h) = O(h^{1/3})$ should be a safe upper bound for the choice of $r(h)$. While an exact proof of this fact is beyond the scope of this paper, we will give numerical evidence in Section 5 and will provide in Section 4 a proof for the adjoint Engquist-Osher scheme and a simplified structure of the coefficients $a^n_{j+\frac{1}{2},k}$ in (20), see Theorem 8.

4. Application to sensitivity and adjoint schemes for standard finite difference schemes

In this section, we apply the convergence results of the previous section to several well known difference schemes (12) and the associated adjoint schemes. We assume throughout that the convexity assumption (2) holds.

4.1. The Engquist-Osher scheme

The Engquist-Osher scheme (EO-scheme) has the monotone numerical flux, see, e.g., Engquist and Osher (1981)

$$f^{EO}(y_0, y_1) = f(\bar{y}) + \int_{\bar{y}}^{y_0} f'(y)^+ \, dy + \int_{\bar{y}}^{y_1} f'(y)^- \, dy,$$

where $f'(y)^+ \stackrel{\text{def}}{=} \max(f'(y), 0)$, $f'(y)^- \stackrel{\text{def}}{=} \min(f'(y), 0)$, and $\bar{y} \in \mathbb{R}$ is fixed. Although f^{EO} does not depend on the choice of $\bar{y}$, it will be convenient to choose $\bar{y}$ as the sonic point, i.e., $f'(\bar{y}) = 0$, if it exists. Thus, the Engquist-Osher scheme (12) reads

$$y_j^{n+1} = y_j^n - \lambda \left(\int_{y_{j-1}^n}^{y_j^n} f'(y)^+ \, dy + \int_{y_j^n}^{y_{j+1}^n} f'(y)^- \, dy \right). \tag{65}$$

In order to apply the convergence results of Theorems 6 and 7 for the associated adjoint scheme we make the following observations:

- The Engquist-Osher flux is consistent and $C^{1,1}_{loc}$. Thus, (13) holds. Moreover, we have

  $$f^{EO}_{y_0} = f'(y_0)^+, \quad f^{EO}_{y_1} = f'(y_1)^-,$$

  and these are nondecreasing functions. Therefore, (56) holds.
- The scheme has the form (12) with $K = 1$. The sensitivity and adjoint scheme are given by (19) and (24), the coefficients (20) are

  $$a^{EO,n}_{j+\frac{1}{2},0} = f^{EO,n}_{y_0,j+\frac{1}{2}} = f'(y_j^n)^+, \quad a^{EO,n}_{j+\frac{1}{2},1} = f^{EO,n}_{y_1,j+\frac{1}{2}} = f'(y_{j+1}^n)^-. \tag{66}$$

This yields, for the coefficients $B^n_{j,k}$ in (33), (3.2),

$$\begin{aligned} B^{EO,n}_{j,-1} &= \lambda f'(y^n_j)^+ \geq 0, \\ B^{EO,n}_{j,0} &= 1 + \lambda(f'(y^n_j)^- - f'(y^n_j)^+) = 1 - \lambda|f'(y^n_j)|, \\ B^{EO,n}_{j,1} &= -\lambda f'(y^n_j)^- \geq 0. \end{aligned}$$

Thus, the scheme is monotone on $[-M_y, M_y]$ according to (14), under the CFL condition

$$\lambda \sup_{|y|\leq M_y} |f'(y)| \leq 1.$$

In particular, we have $B^{EO,n}_{j,k} \geq 0$ on $[-M_y, M_y]$.

- Let the CFL condition hold with $M_y = \|u\|_\infty$. Then, the EO-scheme is convergent in the sense of (55) by Theorem 4 and generates iterates in $[-M_y, M_y]$.
- The coefficients $C^n_{j,k}$ in (36) are

$$\begin{aligned} C^{EO,n}_{j,-1} &= B^{EO,n}_{j+1,-1} \geq 0, \quad C^{EO,n}_{j,1} = B^{EO,n}_{j,1} \geq 0, \\ C^{EO,n}_{j,0} &= 1 + \lambda(f'(y^n_{j+1})^- - f'(y^n_j)^+) \geq 1 - \lambda(|f'(y^n_{j+1})| + |f'(y^n_j)|). \end{aligned} \tag{67}$$

Hence, $C^{EO,n}_{j,k} \geq 0$ is ensured under a 1/2-CFL condition, i.e.,

$$\lambda \sup_{|y|\leq M_y} |f'(y)| \leq \frac{1}{2}. \tag{68}$$

- It can be shown that for the initial data $u \in BV(\mathbb{R})$, satisfying an OSLC $u_x \leq M_{u'}$ with $M_{u'} \in [0, \infty]$, the EO-scheme satisfies the discrete OSLC (57) under a 1/2-CFL condition (68), see Brenier and Osher (1988), Ulbrich (2001).

We thus have the following result.

COROLLARY 1 *Under a 1/2-CFL condition* (68)*, the EO-scheme and its adjoint scheme satisfy Assumption 2. Hence, the convergence results of Theorems 4, 6, and 7 hold.*

We show now, for the simplified case of a piecewise constant state y_h with one stationary shock, that for the adjoint EO-scheme Theorem 7 holds with $r(h) = O(h^\beta)$ for any $\beta \in [1/3, 1/2)$.

THEOREM 8 *Let $u_l > u_r$ with $f(u_l) = f(u_r)$ and $f'(u_l) > 0 > f'(u_r)$. Let $u(x) = u_l$ for $x \leq 0$ and $u(x) = u_r$ for $x > 0$ and let y be the corresponding entropy solution of* (1) *(Riemann problem) given by*

$$y(t,x) = \begin{cases} u_l & x \leq 0, \\ u_r & x > 0. \end{cases}$$

Let a 1/2-CFL condition (68) *hold and let y_h be the grid function corresponding to $y_j^n = y(t_n, x_j)$. Then, under the assumptions of Corollary 1, the convergence result of Theorem 7 holds for the adjoint EQ-scheme for $r(h) = O(h^\beta)$ with any $\beta \in [1/3, 1/2)$.*

Proof The adjoint EQ-scheme reads

$$p_j^n = \lambda f'(y_j^n)^+ p_{j+1}^{n+1} + (1 - \lambda|f'(y_j^n)|)p_j^{n+1} - \lambda f'(y_j^n)^- p_{j-1}^{n+1}. \tag{69}$$

The shock location at $t = \bar{t}$ is $\bar{x} = 0$ and the shock funnel is confined by the characteristics $\xi_l(t) = -f'(u_l)(\bar{t} - t)$ and $\xi_r(t) = -f'(u_r)(\bar{t} - t)$.

We estimate the dependence of $p_{\bar{j}}^{\bar{n}}$ inside the shock funnel on the values $p_j^{N_{\bar{t}}}$ at x_j sufficiently far away from the shock location $\bar{x}$.

Without restriction we consider only $\bar{n} = 0$. Let $(0, x_{\bar{j}})$ be inside the shock funnel and to the left of the shock. Since $f'(y_j^n) = f'(u_l) > 0$ to the left of the shock and $f'(y_j^n) = f'(u_r) < 0$ to the right of the shock, the adjoint EO-scheme (69) to the left of the shock reads with $a_l = f'(u_l) > 0$

$$p_j^n = (1 - \lambda a_l)p_j^{n+1} + \lambda a_l p_{j+1}^{n+1} \tag{70}$$

and with $a_r = f'(u_r) < 0$ on the right hand side of the shock:

$$p_j^n = (1 + \lambda a_r)p_j^{n+1} - \lambda a_r p_{j-1}^{n+1}. \tag{71}$$

Hence, $p_{\bar{j}}^0$ is a convex combination of $p_{\bar{j}+i}^{N_{\bar{t}}}$, $0 \le i \le 1 - \bar{j}$, with weights β_i, and we have

$$\beta_i = \binom{N_{\bar{t}}}{i}(\lambda a_l)^i(1 - \lambda a_l)^{N_{\bar{t}}-i} \quad \text{for } 0 \le i \le -2 - \bar{j}.$$

Hence, the fraction $p_{\bar{j}}^0(l)$ of $p_{\bar{j}}^0$ that depends on $p_j^{N_{\bar{t}}}$ for $j \le l \le -2$ can be estimated by

$$|p_{\bar{j}}^0(l)| \le \sum_{i=0}^{l-\bar{j}} \binom{N_{\bar{t}}}{i}(\lambda a_l)^i(1-\lambda a_l)^{N_{\bar{t}}-i}|p_{\bar{j}+i}^{N_{\bar{t}}}| \le \|p_h^{\bar{t}}\|_\infty \sum_{i=0}^{l-\bar{j}} \binom{N_{\bar{t}}}{i}(\lambda a_l)^i(1-\lambda a_l)^{N_{\bar{t}}-i},$$

where the right hand side vanishes for $l < \bar{j}$ and is an integral over the tail of a binomial distribution X with expected value $E(X) = N_{\bar{t}}\lambda a_l = \frac{\bar{t}f'(y_l)}{h} > 0$ and variance $V(X) = \frac{\bar{t}f'(y_l)}{h}(1 - \lambda f'(y_l))$.

Since $(0, x_{\bar{j}})$ is in the shock funnel, we have $x_{\bar{j}} + \bar{t}f'(y_l) \ge 0$ and thus $\bar{j} + E(X) \ge 0$. Hence, Chebyshev's inequality yields

$$R(l) := \sum_{i=0}^{l-\bar{j}} \binom{N_{\bar{t}}}{i}(\lambda a_l)^i(1-\lambda a_l)^{N_{\bar{t}}-i} \le P(|X - E(X)| \ge -l) \le \frac{V(X)}{l^2} = \frac{O(1)}{hl^2}.$$

If now $-lh = O(h^\beta)$ with $\beta \in [1/3, 1/2)$, then $\frac{1}{l^2 h} = O(h^{1-2\beta}) \to 0$ as $h \to 0$, and the proof is complete, since p_j^0 does not depend on $p_j^{N_{\bar{t}}}$ for $j \geq 2$ as observed above.

A sharper estimate is obtained by using Stirling's formula that yields (see Theorem of Moivre-Laplace)

$$P(X = i) = \frac{1 + o(1)}{\sqrt{2\pi V(X)}} e^{-\frac{(i - E(X))^2}{2V(X)}}.$$

Hence, we obtain, with the substitution $s = \frac{z - E(X)}{\sqrt{V(X)}}$,

$$R(l) \leq C \int_{-\infty}^{l + E(X)} \frac{1}{\sqrt{2\pi V(X)}} e^{-\frac{(z - E(X))^2}{2V(X)}}\, dz = C \int_{-\infty}^{l/\sqrt{V(X)}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt.$$

If now $-lh = O(h^\beta)$ with $\beta \in [1/3, 1/2)$, then $-l = O(h^{\beta - 1})$, while $V(X) = O(h^{-1})$. Hence,

$$R(l) \leq C \int_{-\infty}^{O(-h^{\beta - 1/2})}$$

and the right hand side tends to zero exponentially for $h \to 0$. □

REMARK 6 *If y_h is constant outside of a numerical shock profile, then an analogous proof is possible. One has only to choose x_l outside of the shock profile instead of $l \leq -2$.*

The general case of a nonstationary shock can also be handled, but the proof is quite technical. An extension to general piecewise smooth solutions with shocks should also be possible.

4.2. Modified Lax-Friedrichs scheme

The numerical flux is given by

$$f^{LF}(y_0, y_1) = \frac{1}{2}\left(f(y_0) + f(y_1) - \frac{\gamma}{\lambda}(y_1 - y_0)\right), \quad \gamma \in [\lambda \max |f'(y)|, 1),$$

where the maximum is taken over the whole region in which y_0, y_1 vary. Then, the scheme (12) reads

$$y_j^{n+1} = \frac{1}{2}(\gamma y_{j-1}^n + (2 - 2\gamma) y_j^n + \gamma y_{j+1}^n) - \frac{\lambda}{2}(f(y_{j+1}^n) - f(y_{j-1}^n)). \tag{72}$$

Obviously, the scheme is for any $\gamma \in (0, 1]$ monotone on $[-M_y, M_y]$ under the CFL condition

$$\lambda \sup_{|y| \leq M_y} |f'(y)| \leq \gamma.$$

The original Lax-Friedrichs (L-F) scheme is obtained for $\gamma = 1$ and can be analyzed on the decoupled staggered grids $(n + j) \mod 2 = \text{const}$. In the following, we study the case of $\gamma \in (0, 1)$. In order to apply Theorems 6 and 7 we collect the following properties.

- The Lax-Friedrichs flux is consistent and $C^{1,1}_{loc}$. Thus, (13) holds. Moreover, we have

$$f^{LF}_{y_0} = \frac{1}{2}f'(y_0) + \frac{\gamma}{2\lambda}, \quad f^{LF}_{y_1} = \frac{1}{2}f'(y_1) - \frac{\gamma}{2\lambda},$$

  and these are nondecreasing functions. Therefore, (56) holds.
- The LF-scheme has the form (12) with $K = 1$. The sensitivity and adjoint scheme are given by (19) and (24), and the coefficients (20) are

$$a^{LF,n}_{j+\frac{1}{2},0} = f^{LF,n}_{y_0,j+\frac{1}{2}} = \frac{1}{2}(f'(y^n_j)+\frac{\gamma}{\lambda}), \quad a^{LF,n}_{j+\frac{1}{2},1} = f^{LF,n}_{y_1,j+\frac{1}{2}} = \frac{1}{2}(f'(y^n_{j+1})-\frac{\gamma}{\lambda}).$$

  This yields for the coefficients $B^n_{j,k}$ in (33), (3.2)

$$\begin{aligned} B^{LF,n}_{j,-1} &= \frac{1}{2}(\lambda f'(y^n_j) + \gamma), \\ B^{LF,n}_{j,0} &= 1 - \gamma > 0 \\ B^{LF,n}_{j,1} &= \frac{1}{2}(-\lambda f'(y^n_j) + \gamma). \end{aligned}$$

  Thus, the LF-scheme is monotone on $[-M_y, M_y]$, according to (14), under the CFL condition

$$\lambda \sup_{|y|\le M_y} |f'(y)| \le \gamma.$$

  In particular, we have $B^{LF,n}_{j,0} \ge 0$ on $[-M_y, M_y]$.
- Let the γ-CFL condition hold with $M_y = \|u\|_\infty$. By the monotonicity, the LF-scheme is convergent in the sense of (55) by Theorem 4 and generates iterates in $[-M_y, M_y]$.
- The coefficients $C^n_{j,k}$ in (36) are

$$\begin{aligned} C^{LF,n}_{j,-1} &= B^{LF,n}_{j+1,-1} \ge 0, \quad C^{LF,n}_{j,1} = B^{LF,n}_{j,1} \ge 0, \\ C^{LF,n}_{j,0} &= 1 + \frac{1}{2}(\lambda(f'(y^n_{j+1}) - f'(y^n_j)) - 2\gamma). \end{aligned}$$

  Hence, $C^{LF,n}_{j,k} \ge 0$ is ensured under a $(1 - \gamma)$-CFL condition.

Corollary 2 *Under a $\min(\gamma, 1 - \gamma)$-CFL condition the modified Lax-Friedrichs-scheme and its adjoint scheme satisfy Assumption 2. Hence, the convergence results of Theorems 4, 6, and 7 hold.*

5. Numerical example

We consider the state equation (1) with $f(y) = y^2/2$ and initial data

$$u(x) = \begin{cases} 2 & \text{for } x \le 0, \\ -1 & \text{for } x > 0. \end{cases}$$

The objective function is

$$J(y) = \int_{\mathbb{R}} \gamma(x) \frac{y(1,x)^2}{2}\, dx$$

with $\gamma \in C_c^1(\mathbb{R})$ and $\gamma \equiv 1$ on $[-2,2]$. The entropy solution has a single shock with speed $s = 1/2$ and is given by

$$y(t,x) = \begin{cases} 2 & \text{for } x \le t/2, \\ -1 & \text{for } x > t/2. \end{cases}$$

The reversible solution of the adjoint equation (6) on $[0,T] \times [-2,2]$ is

$$p(t,x) = \begin{cases} 2 & \text{for } -2 \le x < 1/2 - 2(1-t), \\ -1 & \text{for } 1/2 + (1-t) < x \le 2, \\ \frac{1}{2} & \text{for } 1/2 - 2(1-t) \le x \le 1/2 + (1-t). \end{cases}$$

We apply the EO-scheme (65) with $\lambda = 1/4$ to compute y_h and its adjoint scheme (24), (66) to compute p_h. As end data for the adjoint scheme we choose on the one hand (25), which yields the exact discrete adjoint, and on the other hand (61) with $r(h) = h^{9/20}$, which ensures convergence to the correct adjoint state by Theorems 7 and 8.

Figure 1 shows the discrete state $y_h(1,\cdot)$, the discrete adjoint $p_h(0,\cdot)$ for data (25) and discrete adjoint $p_h(0,\cdot)$ for data (61) when using the Engquist-Osher scheme and its adjoint scheme with $h = 2^{-6}$ and $h = 2^{-10}$, respectively. As already observed in Giles and Ulbrich (2010a,b), the end data (25) corresponding to the exact discrete adjoint do not yield the correct value $1/2$ of p_h in the shock funnel. The sharp shock profile does not allow to propagate the correct value in the shock funnel. In Giles and Ulbrich (2010a,b) it has been shown that a numerical viscosity $O(h^\beta)$ with $\beta < 1$ is necessary to obtain convergence.

The modified end data (61) that have been proposed and analyzed in this paper yield also the correct value $1/2$ up to machine precision without using very dissipative state solvers as required in Giles and Ulbrich (2010a,b). The proposed approach provides an easily applicable remedy to ensure convergence to the correct adjoint also for schemes with sharp shock resolution.

h	$\|(p_h - p)(0,\cdot)\|_{L^1}$ for data (25)	$\|(p_h - p)(0,\cdot)\|_{L^1}$ for data (61)	exp. order of conv.
2^{-6}	1.0749	0.3785	
2^{-7}	1.0194	0.2579	0.5536
2^{-8}	0.9815	0.1856	0.4741
2^{-9}	0.9552	0.1273	0.5443
2^{-10}	0.9369	0.0887	0.5215

Table 1. Left: L^1-error of adjoint (EO) $p_h(0,\cdot)$ for data (25); right: L^1-error of adjoint (EO) $p_h(0,\cdot)$ for data (61) and experimental order of convergence

Table 1 shows for different mesh sizes the L^1-error $\|(p_h - p)(0,\cdot)\|_{L^1(-2,2)}$ for the data (25) (left) and for the data (61) (right). While the error for data (25) remains $O(1)$, it converges for the proposed data (61) to zero with an experimental order of convergence of approximately $h^{1/2}$.

Acknowledgements

This work was supported by *Deutsche Forschungsgemeinschaft* within TRR 154 "Mathematical modelling, simulation and optimization using the example of gas networks", project A02, and within SFB 1194 "Interaction between transport and wetting processes", project B04.

References

Bardos, C. and Pironneau, O. (2005) Data assimilation for conservation laws. *Methods Appl. Anal.* **12**(2), 103-134.

Bouchut, F. and James, F. (1998) One-dimensional transport equations with discontinuous coefficients. *Nonlinear Anal.* **32**(7), 891-933. DOI 10.1016/S0362-546X(97)00536-1.

Brenier, Y. and Osher, S. (1988) The discrete one-sided Lipschitz condition for convex scalar conservation laws. *SIAM J. Numer. Anal.* **25**(1), 8-23.

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

Crandall, M.G. and Majda, A. (1980) Monotone difference approximations for scalar conservation laws. *Math. Comp.* **34**(149), 1-21.

Engquist, B. and Osher, S. (1981) One-sided difference approximations for nonlinear conservation laws. *Math. Comp.* **36**(154), 321-351.

Giles, M. and Ulbrich, S. (2010a) Convergence of linearized and adjoint approximations for discontinuous solutions of conservation laws. Part 1: Linearized approximations and linearized output functionals. *SIAM J. Numer. Anal.* **48**(3), 882-904. DOI10.1137/080727464.

Giles, M. and Ulbrich, S. (2010b) Convergence of linearized and adjoint approximations for discontinuous solutions of conservation laws. Part 2: Adjoint approximations and extensions. *SIAM J. Numer. Anal.* **48**(3), 905-921. DOI10.1137/09078078X.

Gosse, L. and James, F. (2000) Numerical approximations of one-dimensional linear conservation equations with discontinuous coefficients. *Math. Comp.* **69**(231), 987-1015.

Hajian, S., Hintermuller, M. and Ulbrich, S. (2019) Total variation diminishing schemes in optimal control of scalar conservation laws. *IMA J. Numer. Anal.* **39**(1), 105-140.

Homescu, C. and Navon, I.M. (2003) Optimal control of flow with discontinuities. *J. Comput. Phys.* **187**(2), 660-682.

Kruzkov, S.N. (1970) First order quasilinear equations in several independent variables. *Math. USSR Sb.* **10**(2), 217-243.

Kuznecov, N.N. (1976) The accuracy of certain approximate methods for the computation of weak solutions of a first order quasilinear equation. *Z. Vycisl. Mat. i Mat. Fiz.* **16**(6), 1489-1502, 1627.

Malek, J., Necas, J., Rokyta, M. and Ruzicka, M. (1996) *Weak and Measure-Valued Solutions to Evolutionary PDEs. Applied Mathematics and Mathematical Computation*, **13**, Chapman & Hall, London.

Nessyahu, H. and Tadmor, E. (1992) The convergence rate of approximate solutions for nonlinear scalar conservation laws. *SIAM J. Numer. Anal.* **29**(6), 1505-1519.

Oleinik, O.A. (1963) Discontinuous solutions of non-linear differential equations. *Amer. Math. Soc. Transl.* (2) **26**, 95-172.

Pfaff, S. and Ulbrich, S. (2015) Optimal boundary control of nonlinear hyperbolic conservation laws with switched boundary data. *SIAM J. Control and Optimization* **53**(3), 1250-1277. DOI 1137/140995799.

Teng, Z.H. and Zhang, P. (1997) Optimal L1-rate of convergence for the viscosity method and monotone scheme to piecewise constant solutions with shocks. *SIAM J. Numer. Anal.* **34**(3), 959-978.

Ulbrich, S. (2001) Optimal control of nonlinear hyperbolic conservation laws with source terms. Habilitation, Zentrum Mathematik, Technische Universitat Muñchen, Germany.

Ulbrich, S. (2002) A sensitivity and adjoint calculus for discontinuous solutions of hyperbolic conservation laws with source terms. *SIAM J. Control Optim.* **41**(3), 740-797. DOI 10.1137/S0363012900370764. URL http://dx.doi.org/10.1137/S0363012900370764

Ulbrich, S. (2003) Adjoint-based derivative computations for the optimal control of discontinuous solutions of hyperbolic conservation laws. *Systems Control Lett.* **48**(3-4), 313-328. DOI 10.1016/S0167- 6911(02)00275-X.

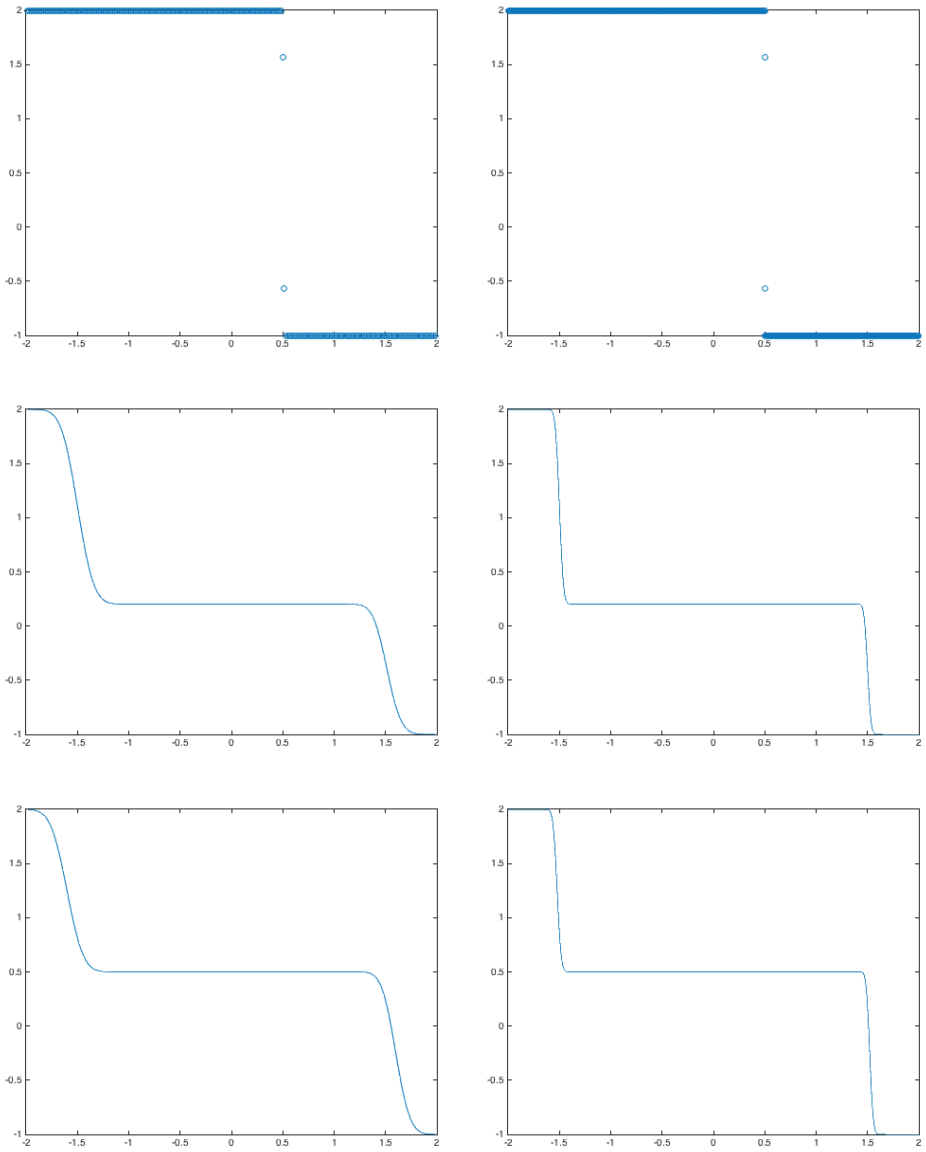

Figure 1. From above: state (EO) $y_h(1,\cdot)$, adjoint (EO) $p_h(0,\cdot)$ for data (25), adjoint (EO) $p_h(0,\cdot)$ for data (61), $h = 2^{-6}$ (left), $h = 2^{-10}$ (right)