On Convergence of Numerical Methods for Optimization Problems Governed by Scalar Hyperbolic Conservation Laws

Michael Herty, Alexander Kurganov and Dmitry Kurochkin

Abstract We consider optimization problems governed by scalar hyperbolic conservation laws in one space dimension and study numerical schemes for the solution to arising linear adjoint equations. We analyze convergence properties of adjoint and gradient approximations on an unbounded domain $x \in \mathbb{R}$ with a strictly convex flux. This paper provides the theoretical foundation of the scheme introduced in [14]. We also demonstrate that using a higher-order temporal discretization helps to substantially improve both the efficiency and accuracy of the overall numerical method.

1 Introduction

We are concerned with numerical methods for optimization problems governed by scalar hyperbolic conservation laws in one space dimension, which also could be further generalized to nonlinear hyperbolic systems. These types of problems arise in a variety of applications where inverse problems for the corresponding initial value problems (IVP) are to be solved. We focus on numerical methods related to those presented in [14]. In this paper, we discuss convergence properties of adjoint and gradient approximations in one-dimensional (1-D) scalar problems on an unbounded domain $x \in \mathbb{R}$ with a strictly convex flux. The discussion will be based

Michael Herty
RWTH Aachen University, Department of Mathematics, Templergraben 55, D-52056 Aachen, Germany e-mail: nherty@igpm.rwth-aachen.de

Alexander Kurganov
Department of Mathematics, Southern University of Science and Technology of China, Shenzhen, 518055, China and Mathematics Department, Tulane University, New Orleans, LA 70118, USA; e-mail: kurganov@math.tulane.edu

Dmitry KurochkinMathematics Department, Tulane University, New Orleans, LA 70118, USA; e-mail: dkurochk@math.tulane.edu

on a general existence results for first-order schemes presented in [20] utilized to establish existence for a variety of other methods, see, e.g., [1, 7, 10, 21].

The optimization problem is formulated as follows: Find an optimal initial condition $u_0(x)$ (control) such that the objective functional J is minimized:

$$\min_{u_0} J(u(\cdot,T);u_d(\cdot)), \tag{1}$$

Here, the objective functional J is

$$J(u(\cdot,T);u_d(\cdot)) := \frac{1}{2}\int\limits_{-\infty}^{\infty} (u(x,T)-u_d(x))^2\,dx \tag{2}$$

and $u(x,t)$ is the unique entropy solution of the following IVP for the 1-D scalar hyperbolic conservation law:

$$\begin{aligned} &u_t + f(u)_x = 0, \qquad && x\in\mathbb{R},\ t\in(0,T],\\ &u(x,0) = u_0(x), \qquad && x\in\mathbb{R}. \end{aligned} \tag{3}$$

Here, $u\colon \mathbb{R}\times[0,T]\to\mathbb{R}$, $u_0(x)$ is an arbitrary bounded measurable function on $\mathbb{R}$, the corresponding nonlinear flux is denoted by $f(u)$, and the terminal state $u_d(x)$ is prescribed at time $t=T$.

In case of sufficiently smooth solutions the formal adjoint equation is given by

$$p_t + f'(u(x,t))p_x = 0, \quad x\in\mathbb{R},\ t\in[0,T), \tag{4}$$

subject to the following terminal state

$$p(x,T) = p_T(x), \quad p_T(x) := u(x,T)-u_d(x), \quad x\in\mathbb{R}. \tag{5}$$

The coupled systems (3) and (4), (5) together with

$$p(x,0) = 0 \ \text{ a.e. } \ x\in\mathbb{R} \tag{6}$$

represent the first-order optimality system for smooth solutions of the problem (1)–(3), in which (3) should be solved forward in time from $t=0$ to $t=T$, while the adjoint equation (4) should be solved backward in time from $t=T$ to $t=0$.

There has been extensive literature on PDE-constrained problems of type (1)–(3) both analytically and numerically. The semi-group generated by the conservation law is not differentiable in L^1 and therefore the usual notion of derivatives has to be extended to tangent vectors consisting of an L^1-part and real part for the variation in shock position, see [4]. The studied equations (3)–(6) only capture the L^1-variation leaving the variations in the shock positions aside. A first-order optimality system that includes shock variations is presented in [5]. The most recent review of existing literature can be found in [10, 8, 7, 14, 11, 6, 20, 15].

Even though only the L^1-part of the optimality system is captured by (3)–(6), it has been shown in [14] that a suitable numerical implementation allows to solve the optimization problem (1)–(3). In this paper, we formalize the observed behavior by proving convergence of a numerical scheme based on solving (3) and (4)–(6). We only prove convergence outside the regions influenced by shocks. The key discussion will be on the numerical discretization of the nonconservative transport equation (4). Theoretical discussion on transport equations can be found, e.g., in [2, 3, 18]. In order to obtain a well-posed adjoint problem, we follow [20] and assume a one-sided Lipschitz condition (OSLC) to be satisfied. The OSLC condition for $v \in L^\infty(\mathbb{R} \times (0,T))$, where $v(x,t) := f'(u(x,t))$, reads

$$v_x(\cdot,t) \leq \alpha(t), \quad \alpha \in L^1(0,T). \tag{7}$$

The adjoint equation is then well-posed for Lipschitz terminal data p_T in the sense that there exists a unique reversible solution of (4), (5). The OSLC condition for equation (4) holds, for example, if the flux in (3) is strictly convex, that is, if

$$f'' \geq c > 0 \text{ for some } c > 0. \tag{8}$$

In this paper, we consider the scheme, which is second-order in time and first-order in space and use the results from [20] to establish its convergence. The convergence proof for a second-order in time scheme is the novel contribution of this work, which provides a theoretical base for the numerical results presented in [14].

2 Numerical Method

In this section, we introduce the iterative optimization algorithm for the problem (1)–(3) based on the formal optimality system. The algorithm is a simplified version of Algorithm 3.1 from [14] and may be seen as a block Gauß-Seidel iteration. From now on the optimal solution u_0 of the problem (1)–(3) will be called the *recovered initial data*, while the corresponding solution of the system (3) will be referred to as the *recovered solution*.

2.1 Iterative Algorithm

Assuming two tolerance parameters, ε_J and $\varepsilon_{\Delta J}$ (the second parameter is needed since the optimal value of the objective functional may be strictly positive), are chosen a priori, we implement the iterative algorithm to generate a sequence $\{u_0^{(m)}(x)\}, m = 0,1,2,\ldots$ of recovered initial data as follows.

Algorithm 2.1

Step 1. Choose an initial guess $u_0^{(0)}(x)$ for the initial data $u_0(x)$. Set $m := 0$.

Step 2. Numerically solve (3) with the initial state $u_0(x) = u_0^{(m)}(x)$ forward in time from $t = 0$ to $t = T$ by the semi-discrete version of the Engquist-Osher scheme described in §2.2. We denote the obtained solution by $u^{(m)}(x,t)$.

Step 3. Compute the objective functional

$$J(u^{(m)}(\cdot,T);u_d(\cdot)) := \frac{1}{2}\int\limits_{-\infty}^{\infty} \left(u^{(m)}(x,T) - u_d(x)\right)^2 dx.$$

Step 4. If either

$$J(u^{(m)}(\cdot,T);u_d(\cdot)) \le \varepsilon_J,$$

or

$$m > 0 \text{ and } \left|J(u^{(m)}(\cdot,T);u_d(\cdot)) - J(u^{(m-1)}(\cdot,T);u_d(\cdot))\right| \le \varepsilon_{\Delta J},$$

stop the iteration process. The obtained $u_0^{(m)}(x)$ will be the approximation to the optimal control.

Step 5. Numerically solve the adjoint system (4), (5) subject to the terminal condition $p(x,T) = u_0^{(m)}(x) - u_d(x)$ backward in time from $t = T$ to $t = 0$ using a semi-discrete upwind scheme described below. The solution is denoted by $p^{(m)}(x,t)$.

Step 6. Update the control $u_0^{(m)}(x)$ using either a gradient descent or quasi-Newton method [17].

Step 7. Set $m := m + 1$. Go to Step 2.

2.2 Numerical Schemes

In Step 2 of the Algorithm described in §2.1, the conservation law (3) is being solved using the semi-discrete version of the Engquist-Osher scheme, which is described in this section. We consider the IVP (3) and solve it numerically on a uniform spatial grid with $x_\alpha := \alpha\Delta x$. We denote by $\lambda := \Delta t/\Delta x$ and introduce the computed cell averages over the cells $[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$:

$$u_j(t) := \frac{1}{\Delta x}\int\limits_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u(x,t)\,dx, \;\; u_j^n := u_j(t^n),$$

where $t^n := n\Delta t$. The cell averages are then evolved in time using the following semi-discrete scheme:

$$\frac{d\,u_j(t)}{dt} = -\frac{\mathcal{F}_{j+\frac{1}{2}}(t) - \mathcal{F}_{j-\frac{1}{2}}(t)}{\Delta x}, \tag{9}$$

where $\mathcal{F}_{j+\frac{1}{2}}$ denotes the Engquist-Osher numerical flux: $\mathcal{F}_{j+\frac{1}{2}}(t) = \int\limits_0^{u_j(t)} f'(\xi)^+ \, d\xi + \int\limits_0^{u_{j+1}(t)} f'(\xi)^- \, d\xi$. The semi-discretization (9) is a system of ODEs, which should be integrated using a (nonlinearly) stable and sufficiently accurate ODE solver.

For example, the system (9) can be solved using the second-order strong stability preserving (SSP) Runge-Kutta method [12, 13] also known as the Heun method [16]:

$$u_j^{n+1} = u_j^n - \lambda \left(H_{j+\frac{1}{2}}^{n+\frac{1}{2}} - H_{j-\frac{1}{2}}^{n+\frac{1}{2}} \right), \tag{10}$$

where

$$H_{j+\frac{1}{2}}^{n+\frac{1}{2}} := \frac{1}{2} \left(\mathcal{F}_{j+\frac{1}{2}}^n + \widehat{\mathcal{F}}_{j+\frac{1}{2}}^{n+1} \right), \tag{11a}$$

$$\mathcal{F}_{j+\frac{1}{2}}^n = \int\limits_0^{u_j^n} f'(\xi)^+ \, d\xi + \int\limits_0^{u_{j+1}^n} f'(\xi)^- \, d\xi, \tag{11b}$$

$$\widehat{\mathcal{F}}_{j+\frac{1}{2}}^{n+1} = \int\limits_0^{\widehat{u}_j^{n+1}} f'(\xi)^+ \, d\xi + \int\limits_0^{\widehat{u}_{j+1}^{n+1}} f'(\xi)^- \, d\xi. \tag{11c}$$

Here, we denote by $(\cdot)^+ := \max\{\cdot, 0\}$ and $(\cdot)^- := \min\{\cdot, 0\}$, and the intermediate value $\widehat{u}_j^{n+1}$ is defined by

$$\widehat{u}_j^{n+1} := u_j^n - \lambda \left(\mathcal{F}_{j+\frac{1}{2}}^n - \mathcal{F}_{j-\frac{1}{2}}^n \right), \tag{12}$$

and is, in fact, a solution obtained after a forward Euler step.

In the following, we describe the semi-discrete upwind scheme used in Step 5 of Algorithm 2.1 for solving the adjoint equation (4). Since $u(x,t)$ has been computed in Step 2, the adjoint problem (4) is, in fact, the following linear equation with variable coefficients:

$$p_t + v(x,t) p_x = 0, \quad x \in \mathbb{R}, \ t \in [0, T), \tag{13}$$

subject to the terminal conditions (5), where

$$v(x,t) := f'(u(x,t)). \tag{14}$$

According to [20], Algorithm 2.1 will converge provided the numerical method for the adjoint problem (4), (5) is induced by the numerical method for the conservation law (3). Introducing the notation $p_j^n := p(x_j, t^n)$, the corresponding discretization of the adjoint problem can be written as follows:

$$p_j^n = p_j^{n+1} + \lambda \sum_{k=-1}^{2} v^n_{j-k+\frac{1}{2},k}\left(p^{n+1}_{j-k+1} - p^{n+1}_{j-k}\right), \tag{15}$$

where

$$v^n_{j+\frac{1}{2},k} = \frac{\partial}{\partial u^n_{j+k}} H^{n+\frac{1}{2}}_{j+\frac{1}{2}}\left(u^n_{j-1},\dots,u^n_{j+2}\right), \tag{16}$$

and the numerical flux $H^{n+\frac{1}{2}}_{j+\frac{1}{2}}$ is defined in (11), (12). The coefficients $v^n_{j+\frac{1}{2},k}$ can be obtained explicitly by substituting (11b), (11c) and (12) into (11a) and then computing the partial derivatives in (16), which result in

$$v^n_{j+\frac{1}{2},-1} = \frac{\lambda}{2} f'(\widehat{u}^{n+1}_j)^+ f'(u^n_{j-1})^+, \tag{17a}$$

$$v^n_{j+\frac{1}{2},0} = \frac{1}{2} f'(u^n_j)^+ + \frac{1}{2} f'(\widehat{u}^{n+1}_j)^+ \left(1 - \lambda |f'(u^n_j)|\right) + \frac{\lambda}{2} f'(\widehat{u}^{n+1}_{j+1})^- f'(u^n_j)^+, \tag{17b}$$

$$v^n_{j+\frac{1}{2},1} = \frac{1}{2} f'(u^n_{j+1})^- + \frac{1}{2} f'(\widehat{u}^{n+1}_{j+1})^- \left(1 - \lambda |f'(u^n_{j+1})|\right) - \frac{\lambda}{2} f'(\widehat{u}^{n+1}_j)^+ f'(u^n_{j+1})^-, \tag{17c}$$

$$v^n_{j+\frac{1}{2},2} = -\frac{\lambda}{2} f'(\widehat{u}^{n+1}_{j+1})^- f'(u^n_{j+2})^-. \tag{17d}$$

Notice that when we solve the adjoint problem, both the second-order, $\{u^n_j\}$, and first-order, $\{\widehat{u}^n_j\}$, solutions of the forward problem are available for all n since they have been computed in Step 2 of Algorithm 2.1.

Other numerical flux functions $\mathcal{F}_{j+\frac{1}{2}}$ are possible. We require that their derivative obtained by equation (16) yields a discretization of (13). Suitable conditions are stated in the following section.

Remark 1. The scheme (15), (17) can be derived in an alternative way. Consider the following semi-discrete upwind scheme for the adjoint equation (13):

$$\frac{dp_j(t)}{dt} = -\left[f'(u_j(t))^+ \frac{p_{j+1}(t) - p_j(t)}{\Delta x} + f'(u_j(t))^- \frac{p_j(t) - p_{j-1}(t)}{\Delta x}\right], \tag{18}$$

where $p_j(t) := p(x_j,t)$. Using the vector notations, the ODE system (18) can be written as

$$\frac{d\boldsymbol{p}(t)}{dt} = \boldsymbol{g}(\boldsymbol{u}(t),\boldsymbol{p}(t)), \tag{19}$$

where $\boldsymbol{p}(t) = \{p_j(t)\}^T$ and $\boldsymbol{g}(\cdot,\cdot) = \{g_j(\cdot,\cdot)\}^T$ is the right-hand side (RHS) of (18).

We now apply the second-order Heun method to the system (19) and obtain the following fully discretize scheme for (13):

$$\boldsymbol{p}^n = \boldsymbol{p}^{n+1} - \frac{\Delta t}{2}\left[\boldsymbol{g}(\boldsymbol{u}^{n+1},\boldsymbol{p}^{n+1}) + \boldsymbol{g}(\boldsymbol{u}^n,\tilde{\boldsymbol{p}}^n)\right], \tag{20}$$

where $\boldsymbol{p}^n := \boldsymbol{p}(t^n)$ and the intermediate value $\tilde{\boldsymbol{p}}^n$ is a solution obtained after one step of forward Euler method (applied backward in time) and defined as

$$\tilde{\boldsymbol{p}}^n := \boldsymbol{p}^{n+1} - \Delta t \boldsymbol{g}(\boldsymbol{u}^{n+1}, \boldsymbol{p}^{n+1}). \tag{21}$$

One can see that the adjoint scheme, introduced in (15), (17), can be written (see Appendix 5) in the form very similar to (20), (21):

$$\boldsymbol{p}^n = \boldsymbol{p}^{n+1} - \frac{\Delta t}{2}\Big[\boldsymbol{g}(\widehat{\boldsymbol{u}}^{n+1}, \boldsymbol{p}^{n+1}) + \boldsymbol{g}(\boldsymbol{u}^n, \widehat{\boldsymbol{p}}^n)\Big], \tag{22}$$

$$\widehat{\boldsymbol{p}}^n = \boldsymbol{p}^{n+1} - \Delta t \boldsymbol{g}(\widehat{\boldsymbol{u}}^{n+1}, \boldsymbol{p}^{n+1}). \tag{23}$$

Therefore, the induced scheme (22), (23) can bee seen as a modification of the scheme (20), (21). Since $\widehat{u}_j^{n+1}$ is the first-order approximation of $u(x_j, t^{n+1})$, while u_j^{n+1} is the second-order one, the approximation used in (20), (21) should be a little more accurate. However, the formal order of accuracy of both schemes is the same (they are first-order in space and second-order in time) and the convergence proof presented in the next section is valid for the induced scheme (22), (23) only.

Remark 2. We have tested both the second order-scheme (20), (21) and its modification (22), (23) in numerical experiments with a variety of initial guesses $u_0^{(0)}(x)$ and terminal states $u_d(x)$, both smooth and discontinuous ones. Because the difference in the numerical results obtained by the schemes (20), (21) and (22), (23) is negligible, for the sake of brevity we omit the results obtained by the scheme (22), (23).

Remark 3. Note that the first-order version of both the scheme (20), (21) and (22), (23) is simply given by

$$\boldsymbol{p}^n = \boldsymbol{p}^{n+1} - \Delta t \boldsymbol{g}(\widehat{\boldsymbol{u}}^{n+1}, \boldsymbol{p}^{n+1}). \tag{24}$$

3 Convergence Analysis

In this section, we discuss convergence properties of the numerical method, introduced in §2. We assume that (8) holds and without loss of generality we set

$$f(0) = 0.$$

We assume that the scheme (10)–(12) yields entropy solutions of (3). (In fact, the semi-discrete version of the Engquist-Osher scheme (9) is entropy stable as it has been proved in [19].) To be more precise, let

$$u^\Delta(x,t) := \sum_j u_j(t)\chi_j(x), \tag{25}$$

where $\chi_j(x)$ is a characteristic function of the interval $[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, be a numerical solution computed at time t. We assume that there are positive constants M and Δ_0 and an entropy solution $u(x,t)$ such that for all $\Delta t = \lambda \Delta x \le \Delta_0$

$$\|u^\Delta\|_\infty \le M,\ u^\Delta(\cdot,t) \to u(\cdot,t) \text{ in } L^1(\mathbb{R})\ t > 0, \Delta t \to 0. \tag{26}$$

We also assume that for the function α in (7), the discrete OSLC condition,

$$u_{j+1}^n - u_j^n \le \frac{1}{\lambda} \int\limits_{t^n}^{t^{n+1}} \alpha(t)\,dt, \quad \forall j \in \mathbb{Z},\ \forall n, \tag{27}$$

holds. Notice that the condition (27) does not allow for jumps up in the initial data u_0. To allow such jumps, the condition (27) should be relaxed (for a weakened version of (27), see [20, Condition (D3')]). It was proved in [20, Section 6.5.1], that both the conditions (26) and (27) are satisfied for the original Engquist-Osher scheme, which is (12), (11b) with $\widehat{u}_j^{n+1}$ replaced with u_j^{n+1} on the left-hand side of (12). Since the second-order Heun method is in fact a convex combination of two forward Euler steps, the results from [20, Section 6.5.1] are still valid for the scheme (10), (11). We now proceed with the convergence analysis of the scheme (15), (17). We first prove the following monotonicity result needed to establish the convergence proof below.

Lemma 1. *Assume that the L^∞-bound* (26) *holds. Let $H_{j+\frac{1}{2}}^{n+\frac{1}{2}}$ be given by* (11), (12) *and assume that the time step is restricted by the following CFL condition:*

$$\lambda \le \frac{1}{2\max\limits_{|u|\le M} |f'(u)|}. \tag{28}$$

Then, the coefficients $v_{j+\frac{1}{2},k}^n$ defined in (16) *are nondecreasing functions of u_ℓ^n for $\ell \in \{j-1, j, j+1, j+2\}$.*

Proof. Note that by substituting (11a) into (12) and by differentiating (12) one can show that provided the condition (28) is satisfied,

$$\frac{\partial}{\partial u_k^n} \widehat{u}_j^{n+1} \ge 0 \ \text{ for } k \in \{j-1, j, j+1\}.$$

Hence, $\widehat{u}_j^{n+1}$ is nondecreasing with respect to all of its arguments, u_{j-1}^n, u_j^n and u_{j+1}^n. Since f is convex, $f'(\cdot)^\pm$ are nondecreasing functions. Further, (17a) clearly implies that $v_{j+\frac{1}{2},-1}^n$ is nondecreasing with respect to $f'(\widehat{u}_j^{n+1})^+$ and $f'(u_{j-1}^n)^+$. Similarly, from (17d) we obtain that $v_{j+\frac{1}{2},2}^n$ is non-decreasing with respect to $f'(\widehat{u}_{j+1}^{n+1})^-$ and $f'(u_{j+2}^n)^-$. Given the condition (28), one can show that $v_{j+\frac{1}{2},0}^n$ is nondecreasing with respect to $f'(u_j^n)^\pm$, $f'(\widehat{u}_j^{n+1})^+$ and $f'(\widehat{u}_{j+1}^{n+1})^-$ by differentiating (17b), namely:

$$\frac{\partial v^n_{j+\frac{1}{2},0}}{\partial\left(f'(u^n_j)^+\right)} = \frac{1}{4}\left(1-2\lambda f'(\widehat{u}^{n+1}_j)^+\right) + \frac{1}{4}\left(1+2\lambda f'(\widehat{u}^{n+1}_{j+1})^-\right) \geq 0,$$

$$\frac{\partial v^n_{j+\frac{1}{2},0}}{\partial\left(f'(u^n_j)^-\right)} = \frac{\lambda}{2} f'(\widehat{u}^{n+1}_j)^+ \geq 0,$$

$$\frac{\partial v^n_{j+\frac{1}{2},0}}{\partial\left(f'(\widehat{u}^{n+1}_j)^+\right)} = \frac{1}{2}\left(1-\lambda|f'(u^n_j)|\right) \geq 0, \quad \frac{\partial v^n_{j+\frac{1}{2},0}}{\partial\left(f'(\widehat{u}^{n+1}_{j+1})^-\right)} = \frac{\lambda}{2} f'(u^n_j)^+ \geq 0,$$

where the identity $|f'(u^n_j)| = f'(u^n_j)^+ - f'(u^n_j)^-$ is taken into account. Similarly, the differentiation of (17c) shows that $v^n_{j+\frac{1}{2},1}$ is nondecreasing with respect to $f'(u^n_{j+1})^\pm$, $f'(\widehat{u}^{n+1}_j)^+$ and $f'(\widehat{u}^{n+1}_{j+1})^-$, provided (28) is satisfied. Finally, using the chain rule we conclude that $v^n_{j+\frac{1}{2},k}$ are nondecreasing functions of u^n_ℓ for $\ell \in \{j-1, j, j+1, j+2\}$. ■

Next, we rewrite the discrete adjoint scheme (15) as

$$p^n_j = \sum_{k=-2}^{2} B^n_{j,k}\, p^{n+1}_{j-k}, \tag{29}$$

were the corresponding coefficients are

$$B^n_{j,-2} = \lambda v^n_{j+\frac{3}{2},-1}, \quad B^n_{j,-1} = \lambda\left(v^n_{j+\frac{1}{2},0} - v^n_{j+\frac{3}{2},-1}\right), \tag{30a}$$

$$B^n_{j,0} = 1 + \lambda\left(v^n_{j-\frac{1}{2},1} - v^n_{j+\frac{1}{2},0}\right), \quad B^n_{j,1} = \lambda\left(v^n_{j-\frac{3}{2},2} - v^n_{j-\frac{1}{2},1}\right), \tag{30b}$$

$$B^n_{j,2} = -\lambda v^n_{j-\frac{3}{2},2}. \tag{30c}$$

According to [20], the L^∞-stability of the adjoint scheme will follow from the positivity of $B^n_{j,k}$, which will be guaranteed by the following lemma.

Lemma 2. *Let $H^{n+\frac{1}{2}}_{j+\frac{1}{2}}$ be given by* (11), (12) *and the adjoint scheme given by* (29), (30). *Assume that the L^∞-bound* (26) *holds and the time step is restricted by*

$$\lambda \leq \frac{1}{\max\limits_{|u|\leq M} |f'(u)|}. \tag{31}$$

Then, the coefficients $B^n_{j,k}$ are nonnegative: $B^n_{j,k} \geq 0 \quad \forall j \in \mathbb{Z},\ k \in \{-2,-1,0,1,2\}$.

Proof. First, we obtain that $B^n_{j,\pm 2} \geq 0$ from their definitions (30a), (30c) and from the definition of $v^n_{j+\frac{1}{2},k}$ given in (17).

We now note that $|\widehat{u}^{n+1}_j| \leq M$, since $|u^n_j| \leq M$ and the solution of the first-order Engquist-Osher scheme (12), (11b) satisfies the maximum principle. Then $B^n_{j,0}$ can be estimated using (17) and (31) as follows:

$$
\begin{aligned}
B^n_{j,0} = 1 - \frac{\lambda}{2}\Big(&|f'(u^n_j)| + |f'(\widehat{u}^{n+1}_j)|(1-\lambda|f'(u^n_j)|) \\
&+\lambda\Big(f'(\widehat{u}^{n+1}_{j-1})^+ f'(u^n_j)^- + f'(\widehat{u}^{n+1}_{j+1})^- f'(u^n_j)^+\Big)\Big) \\
\geq 1 - \frac{\lambda}{2}\Big(&|f'(u^n_j)| + |f'(\widehat{u}^{n+1}_j)|\Big) \geq 0.
\end{aligned}
$$

Similarly, from (30a), (30b), (17) and (31) we obtain:

$$
\begin{aligned}
B^n_{j,-1} &= \frac{\lambda}{2}\left(f'(u^n_j)^+\left(1-\lambda|f'(\widehat{u}^{n+1}_{j+1})|\right) + f'(\widehat{u}^{n+1}_j)^+\left(1-\lambda|f'(u^n_j)|\right)\right) \geq 0, \\
B^n_{j,1} &= \frac{\lambda}{2}\left(-f'(u^n_j)^-\left(1-\lambda|f'(\widehat{u}^{n+1}_{j-1})|\right) - f'(\widehat{u}^{n+1}_j)^-\left(1-\lambda|f'(u^n_j)|\right)\right) \geq 0,
\end{aligned}
$$

which completes the proof of the lemma. ■

We further obtain bounds on the discrete difference approximation $p^n_{j+1} - p^n_j$, computed by the adjoint scheme (15), (17). Using the equivalent form (29), (30) of the adjoint scheme, we rewrite the difference as follows:

$$
p^n_{j+1} - p^n_j = \sum_{k=-2}^{2}\left(B^n_{j+1,k}\,p^{n+1}_{j-k+1} - B^n_{j,k}\,p^{n+1}_{j-k}\right) = \sum_{k=-2}^{2} C^n_{j,k}\left(p^{n+1}_{j-k+1} - p^{n+1}_{j-k}\right),
$$

where the coefficients

$$
C^n_{j,k} := B^n_{j,k} + \lambda\left(v^n_{j-k+\frac{1}{2},k+1} - v^n_{j-k-\frac{1}{2},k+1}\right), \text{ for } -2 \leq k \leq 1, \tag{32a}
$$

$$
C^n_{j,2} := B^n_{j,2}, \tag{32b}
$$

are obtained by simply regrouping the summands. The following lemma shows the positivity of the coefficients $C^n_{j,k}$.

Lemma 3. *Assume that the L^∞-bound* (26) *holds and the CFL condition* (28) *is satisfied. Then, the coefficients $C^n_{j,k}$ given by* (32) *are nonnegative:*

$$
C^n_{j,k} \geq 0 \quad \forall j \in \mathbb{Z},\ k \in \{-2,-1,0,1,2\}. \tag{33}
$$

Proof. Lemma 2 implies $C^n_{j,2} = B^n_{j,2} \geq 0$. Then, (32a) together with (30a) and (17a) give

$$
C^n_{j,-2} = B^n_{j,-2} + \lambda\left(v^n_{j+\frac{5}{2},-1} - v^n_{j+\frac{3}{2},-1}\right) = \lambda v^n_{j+\frac{5}{2},-1} \geq 0.
$$

The estimate on $C^n_{j,0}$ follows from (32a), (17) and the CFL condition (28):

$$\begin{aligned}
C^n_{j,0} &= B^n_{j,-2} + \lambda \left(v^n_{j+\frac{1}{2},1} - v^n_{j-\frac{1}{2},1} \right) = 1 + \lambda \left(v^n_{j+\frac{1}{2},1} - v^n_{j+\frac{1}{2},0} \right) \\
&\geq 1 + \frac{\lambda}{2} \left(f'(u^n_{j+1})^- - f'(u^n_j)^+ + f'(\widehat{u}^{n+1}_{j+1})^- (1 - \lambda |f'(u^n_{j+1})|) - f'(\widehat{u}^{n+1}_j)^+ (1 - \lambda |f'(u^n_j)|) \right) \\
&\geq 1 + \frac{\lambda}{2} \left(f'(u^n_{j+1})^- - f'(u^n_j)^+ + f'(\widehat{u}^{n+1}_{j+1})^- - f'(\widehat{u}^{n+1}_j)^+ \right) \geq 0.
\end{aligned}$$

Similarly, from (32), (30), (17) and (28) we obtain

$$\begin{aligned}
C^n_{j,-1} &= B^n_{j,-1} + \lambda \left(v^n_{j+\frac{3}{2},0} - v^n_{j+\frac{1}{2},0} \right) = \lambda \left(v^n_{j+\frac{3}{2},0} - v^n_{j+\frac{3}{2},-1} \right) \\
&= \frac{\lambda}{2} \left(f'(u^n_{j+1})^+ \left(1 + \lambda f'(\widehat{u}^{n+1}_{j+2})^- \right) + f'(\widehat{u}^{n+1}_{j+1})^+ (1 - \lambda |f'(u^n_{j+1})| - \lambda f'(u^n_j)^+) \right) \geq 0, \\
C^n_{j,1} &= B^n_{j,1} + \lambda \left(v^n_{j-\frac{1}{2},2} - v^n_{j-\frac{3}{2},2} \right) = \lambda \left(v^n_{j-\frac{1}{2},2} - v^n_{j-\frac{1}{2},1} \right) \\
&= \frac{\lambda}{2} \left(-f'(\widehat{u}^{n+1}_j)^- \left(1 + \lambda f'(u^n_{j+1})^- - \lambda |f'(u^n_j)| \right) - f'(u^n_j)^- \left(1 - \lambda f'(\widehat{u}^{n+1}_{j-1})^+ \right) \right) \geq 0,
\end{aligned}$$

so that the proof of the lemma is complete. ■

Finally, we apply the convergence proof of [20, Theorem 6.4.4] to the introduced numerical method.

Theorem 1. *Assume* $f \in C^2(\mathbb{R})$ *satisfies* (8)*. Let the terminal state* p_T*, defined in* (5)*, be Lipschitz continuous and* u^Δ *satisfies* (26) *and* (27)*. Assume that the discretization* p^Δ_T *of the terminal state is consistent, that is, there exist constants* $K > 0$ *and* $L > 0$ *such that*

$$\|p^\Delta_T\|_\infty \leq K, \quad \sup_{x \in \mathbb{R}} \left| \frac{p^\Delta_T(x + \Delta x) - p^\Delta_T(x)}{\Delta x} \right| \leq L$$

and

$$p^\Delta_T \to p_T \text{ in } [-R, R] \; \forall R > 0, \Delta x \to 0.$$

Assume the condition (28) *holds. Then, the numerical solution* (15)*,* (17) *converges locally uniformly to the unique reversible solution* $p \in Lip(\mathbb{R} \times (0,T))$ *of the adjoint problem* (13)*,* (14) *as* $\Delta t = \lambda \Delta x \to 0$*.*

Proof. Lemmas 1–3 ensure that all of the assumptions of [20, Theorem 6.4.4] are satisfied and thus the convergence result follows. ■

At the end of this section, we state the convergence result for the discrete gradients justifying the presented algorithm for a smooth version of the objective functional. For a given nonnegative function $\phi_\delta \in Lip_0(\mathbb{R})$ with the support in $[-\frac{\delta}{2}, \frac{\delta}{2}]$ and $\int_{\mathbb{R}} \phi_\delta(x)\, dx = 1$, and for a given $\psi \in C^1_{\text{loc}}(\mathbb{R}^2)$, we define the functional J_δ as

$$J_\delta(u_0) := \int_{-\infty}^{\infty} \psi\big((\phi_\delta * u)(x,T), (\phi_\delta * u_d)(x) \big)\, dx, \tag{34}$$

where, as before, $u(x,t)$ is the entropy solution of the IVP (3), $u_d(x)$ is a terminal state prescribed at time $t=T$, and $*$ denotes a convolution in x. For J_δ to be well-posed we assume $u_d \in L^\infty(\mathbb{R})$. We discretize J_δ by

$$\widetilde{J_\delta}(u_0^\Delta) = \sum_k \psi\big((\phi_\delta * u^\Delta)(x_k,T),(\phi_\delta * u_d^\Delta)(x_k)\big)\Delta x, \tag{35}$$

where u_0^Δ, u_d^Δ and u^Δ denote the corresponding piecewise constant approximations defined in (25).

The gradient of J_δ exists in the sense of Fréchet differentials, see [20, Theorem 5.3.1]. Using Lemmas 1–3 and Theorem 1 the following convergence result immediately obtained from [20, Theorem 6.4.8].

Theorem 2. *Assume that* $f \in C^2(\mathbb{R})$ *satisfies* (8)*. Let* J_δ *be defined by* (34)*, where* $\phi_\delta \in Lip_0(\mathbb{R})$ *is a given nonnegative function with the support in* $[-\frac{\delta}{2},\frac{\delta}{2}]$ *and* $\int_{\mathbb{R}} \phi_\delta(x)\,dx = 1$*, and* $\psi \in C^1_{\rm loc}(\mathbb{R}^2)$*. Assume also that* $u_0 \in L^\infty(\mathbb{R}\times(0,T))$ *such that* $(u_0)_x \le K$*. Let*

$$p^\Delta(x_j,T) = \sum_k \phi_\delta(x_j - x_k)\partial_1\psi\big((\phi_\delta * u^\Delta)(x_k,T),(\phi_\delta * u_d^\Delta)(x_k)\big)\Delta x, \tag{36}$$

where $\partial_1\psi$ *denotes a partial derivative of* ψ *with respect to its first component.*

Let u^Δ *be an approximate solution of* (3) *obtained by* (10)–(12) *and thus satisfies* (26) *and* (27)*. Let* p^Δ *be a piecewise constant approximation of the solution computed by* (15)*,* (17) *subject to the terminal data* (36)*, and assume that the CFL condition* (28) *holds.*

Then, $p^\Delta(\cdot,0)$ *is an approximation to the Fréchet derivative of* J_δ *with respect to* u_0 *in the following sense:*

$$p^\Delta(\cdot,0) \to p(\cdot,0) = \nabla J_\delta(u_0) \text{ in } L^r(\mathbb{R}) \text{ as } \Delta t = \lambda\Delta x \to 0,$$

for all $r \ge 1$*. Herein,* p *is the reversible solution of* (13) *with the terminal data*

$$p_T(x) = \int_{-\infty}^{\infty} \phi_\delta(x-z)\partial_1\psi\big((\phi_\delta * u)(z,T),(\phi_\delta * u_d)(z)\big)\,dz.$$

4 Numerical Results

In this section, we compare the performance of the optimization method described in Section §2.1 using the first-order schemes (12), (11b) and (24) and the second-order schemes (10)–(12) and (20), (21).

We refer the reader to [14] for further numerical results (including a much more complicated case of systems of hyperbolic conservation laws. In particular, in [14] we consider the examples, in which the control u_0 is recovered exactly. Here, on the contrary, we compare the convergence of first- and second-order (in time) schemes.

The convergence analysis yields the convergence of the first-order scheme whereas the numerical results indicate the same qualitative results for its second-order extension. The rate of convergence is improved in the second-order case.

We consider the problem (1), (2) governed by the inviscid Burgers equation

$$u_t + \left(\frac{u^2}{2}\right)_x = 0 \tag{37}$$

with the terminal state

$$u_d(x) = \begin{cases} \sin(6\pi(x-\frac{1}{3})), & \text{if } \frac{1}{3} \le x \le \frac{2}{3}, \\ 0, & \text{otherwise} \end{cases} \tag{38}$$

prescribed at $T = 1/3$. We solve the problem in the interval $[0,1]$ subject to the periodic boundary conditions using the uniform mesh with $\Delta x = 1/400$ and the following initial guess:

$$u_0^{(0)} = \sin(2\pi x). \tag{39}$$

The recovered initial data, $u_0^{(20000)}(x)$, and the corresponding recovered solution $u^{(20000)}(x,T)$, computed using the studied first- and second-order schemes are shown in Figures 1 and 2, respectively.

Finally, in Figure 3, we show the behavior of the computed objective functional (2) for $m = 1,\dots,20000$ iterations using a logarithmic scale.

The obtained results clearly demonstrate the advantage of a second-order temporal discretization even when the same first-order semi-discrete schemes in space are used.

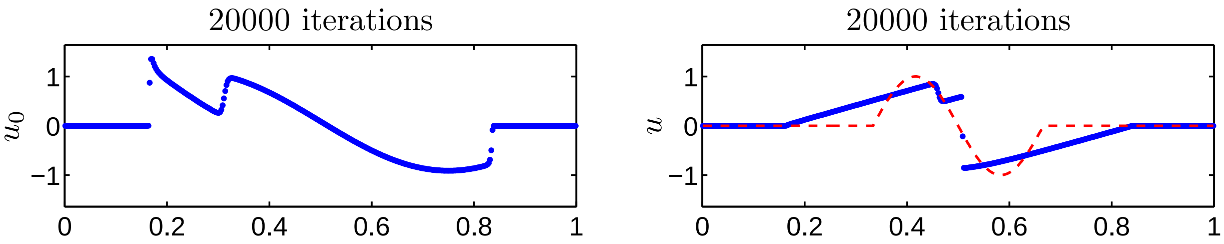

Fig. 1 First-order results. Left: Recovered initial data $u_0^{(20000)}(x)$; Right: Recovered solution $u^{(20000)}(x,T)$ (plotted with points) and the terminal state $u_d(x)$ (dashed line).

Acknowledgements The work of M. Herty was supported by BMBF KinOpt and the cluster of excellence EXC128, HE5386/13,14,15-1. The work of A. Kurganov was supported in part by the

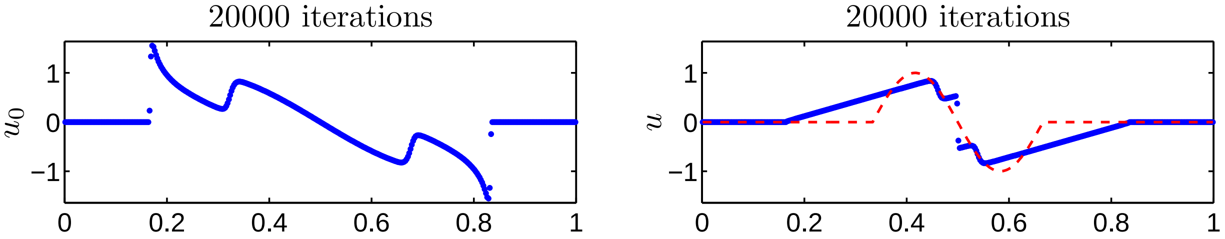

Fig. 2 Second-order results. Left: Recovered initial data $u_0^{(20000)}(x)$; Right: Recovered solution $u^{(20000)}(x,T)$ (plotted with points) and the terminal state $u_d(x)$ (dashed line).

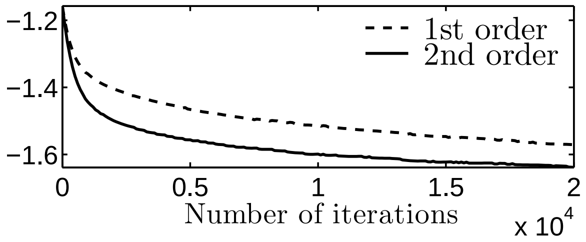

Fig. 3 Dependence of the computed objective functional (measured in a logarithmic scale) on the number of iterations for the first- (dashed line) and second-order (solid line) schemes.

NSF Grants DMS-1115718 and DMS-1521009. The authors also acknowledge the support by the KI-Net research network, NSF RNMS grant DMS-1107444.

5 Appendix

Here, we demonstrate equivalence of the backward scheme (15), (17) and the modified second-order Heun method (22), (23). First, we notice that

$$g_j(\boldsymbol{u},\boldsymbol{p}) = -\left[f'(u_j)^+ \frac{p_{j+1}-p_j}{\Delta x} + f'(u_j)^- \frac{p_j-p_{j-1}}{\Delta x}\right] \tag{40}$$

is linear in its second argument and hence

$$\boldsymbol{g}(\boldsymbol{u}^n,\widehat{\boldsymbol{p}}^n) = \boldsymbol{g}(\boldsymbol{u}^n,\boldsymbol{p}^{n+1}) - \Delta t\,\boldsymbol{g}(\boldsymbol{u}^n,\boldsymbol{g}(\widehat{\boldsymbol{u}}^{n+1},\boldsymbol{p}^{n+1})),$$

where $\widehat{\boldsymbol{p}}^n$ is defined in (23). Then the scheme (22), (23) can be written as follows:

$$\boldsymbol{p}^n=\boldsymbol{p}^{n+1}-\frac{\Delta t}{2}\Big[\boldsymbol{g}(\widehat{\boldsymbol{u}}^{n+1},\boldsymbol{p}^{n+1})+\boldsymbol{g}(\boldsymbol{u}^n,\boldsymbol{p}^{n+1})-\Delta t\,\boldsymbol{g}(\boldsymbol{u}^n,\boldsymbol{g}(\widehat{\boldsymbol{u}}^{n+1},\boldsymbol{p}^{n+1}))\Big]. \tag{41}$$

We then use (40) to rewrite (41) in a componentwise form:

$$\begin{aligned}p_j^n=p_j^{n+1}&+\frac{\lambda}{2}\Big[f'(\widehat{u}_j^{n+1})^+\,(p_{j+1}^{n+1}-p_j^{n+1})+f'(\widehat{u}_j^{n+1})^-\,(p_j^{n+1}-p_{j-1}^{n+1})\Big]\\&+\frac{\lambda}{2}\Big[f'(u_j^n)^+\,(p_{j+1}^{n+1}-p_j^{n+1})+f'(u_j^n)^-\,(p_j^{n+1}-p_{j-1}^{n+1})\Big]\\&+\frac{\lambda^2}{2}f'(u_j^n)^+\Big[f'(\widehat{u}_{j+1}^{n+1})^+\,(p_{j+2}^{n+1}-p_{j+1}^{n+1})+f'(\widehat{u}_{j+1}^{n+1})^-\,(p_{j+1}^{n+1}-p_j^{n+1})\Big]\\&-\frac{\lambda^2}{2}|f'(u_j^n)|\Big[f'(\widehat{u}_j^{n+1})^+\,(p_{j+1}^{n+1}-p_j^{n+1})+f'(\widehat{u}_j^{n+1})^-\,(p_j^{n+1}-p_{j-1}^{n+1})\Big]\\&-\frac{\lambda^2}{2}f'(u_j^n)^-\Big[f'(\widehat{u}_{j-1}^{n+1})^+\,(p_j^{n+1}-p_{j-1}^{n+1})+f'(\widehat{u}_{j-1}^{n+1})^-\,(p_{j-1}^{n+1}-p_{j-2}^{n+1})\Big].\end{aligned} \tag{42}$$

Rearranging the terms in (42), we finally get the adjoint scheme in the following form:

$$\begin{aligned}&p_j^n=p_j^{n+1}+\lambda\left[\frac{\lambda}{2}f'(u_j^n)^+f'(\widehat{u}_{j+1}^{n+1})^+\right](p_{j+2}^{n+1}-p_{j+1}^{n+1})\\&+\lambda\left[\frac{1}{2}f'(\widehat{u}_j^{n+1})^++\frac{1}{2}f'(u_j^n)^++\frac{\lambda}{2}f'(u_j^n)^+f'(\widehat{u}_{j+1}^{n+1})^--\frac{\lambda}{2}|f'(u_j^n)|f'(\widehat{u}_j^{n+1})^+\right](p_{j+1}^{n+1}-p_j^{n+1}),\\&+\lambda\left[\frac{1}{2}f'(\widehat{u}_j^{n+1})^-+\frac{1}{2}f'(u_j^n)^--\frac{\lambda}{2}|f'(u_j^n)|f'(\widehat{u}_j^{n+1})^--\frac{\lambda}{2}f'(u_j^n)^-f'(\widehat{u}_{j-1}^{n+1})^+\right](p_j^{n+1}-p_{j-1}^{n+1})\\&-\lambda\left[\frac{\lambda}{2}f'(u_j^n)^-f'(\widehat{u}_{j-1}^{n+1})^-\right](p_{j-1}^{n+1}-p_{j-2}^{n+1}),\end{aligned} \tag{43}$$

which coincides with (15), (17). Note that the coefficients on the RHS of (43) are the ones given by (17).

References

1. M. K. Banda and M. Herty, *Adjoint IMEX-based schemes for control problems governed by hyperbolic conservation laws*, Comput. Optim. Appl., 51 (2012), pp. 909–930.
2. F. Bouchut and F. James, *Differentiability with respect to initial data for a scalar conservation law*, in Hyperbolic problems: theory, numerics, applications, Vol. I (Zürich, 1998), vol. 129 of Internat. Ser. Numer. Math., Birkhäuser, Basel, 1999, pp. 113–118.
3. ———, *Duality solutions for pressureless gases, monotone scalar conservation laws, and uniqueness*, Comm. Partial Differential Equations, 24 (1999), pp. 2173–2189.
4. A. Bressan and A. Marson, *A variational calculus for discontinuous solutions to conservation laws*, Comm. Partial Differential Equations, 20 (1995), pp. 1491–1552.

5. A. Bressan and W. Shen, *Optimality conditions for solutions to hyperbolic balance laws*, Control methods in PDE-dynamical systems, Contemp. Math., 426 (2007), pp. 129–152.
6. C. Castro, F. Palacios, and E. Zuazua, *An alternating descent method for the optimal control of the inviscid Burgers equation in the presence of shocks*, Math. Models Methods Appl. Sci., 18 (2008), pp. 369–416.
7. A. Chertok, M. Herty, and A. Kurganov, *An eulerian-lagrangian method for optimization problems governed by nonlinear hyperbolic pdes*, Computational Optimization and Applications, 59 (2014), pp. 689–724.
8. C. D'Apice, R. Manzo, and B. Piccoli, *Numerical Schemes for the Optimal Input Flow of a Supply Chain*, SIAM J. Numer. Anal., 51 (2013), pp. 2634–2650.
9. B. Engquist and S. Osher, *One-sided difference approximations for nonlinear conservation laws*, Math. Comp., 36 (1981), pp. 321–351.
10. M. Giles and S. Ulbrich, *Convergence of linearized and adjoint approximations for discontinuous solutions of conservation laws: Part 1:Linearized approximations and linearized output functional*, SIAM J. Numer. Anal., 48 (2010), pp. 882–904.
11. E. Godlewski and P.-A. Raviart, *The linearized stability of solutions of nonlinear hyperbolic systems of conservation laws. A general numerical approach*, Math. Comput. Simulation, 50 (1999), pp. 77–95. Modelling '98 (Prague).
12. S. Gottlieb, D. Ketcheson, and C.-W. Shu, *Strong stability preserving Runge-Kutta and multistep time discretizations*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2011.
13. S. Gottlieb, C.-W. Shu, and E. Tadmor, *Strong stability-preserving high-order time discretization methods*, SIAM Rev., 43 (2001), pp. 89–112.
14. M. Herty, A. Kurganov, and D. Kurochkin, *Numerical method for optimal control problems governed by nonlinear hyperbolic systems of pdes*, Communication in Mathematical Sciences, 51 (2015), pp. 15–48.
15. M. Herty and B. Piccoli, *Numerical method for the computation of tangent vectors to hyperbolic systems of conservation laws*, accepted for publication at Communication in Mathematical Sciences, 2015
16. K. Heun, *Neue Methoden zur approximativen Integration der Differentialgleichungen einer unabhängigen Veränderlichen*, Z. Math. Phys, 45 (1900), pp. 23–38.
17. C. Kelley, *Iterative methods for optimization*, Frontiers in Applied Mathematics. Philadelphia, PA: Society for Industrial and Applied Mathematics. xv, 180 p., 1999.
18. Z. Liu and A. Sandu, *On the properties of discrete adjoints of numerical methods for the advection equation*, Int. J. for Num. Meth. in Fluids, 56 (2008), pp. 769–803.
19. E. Tadmor, *Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems*, Acta Numer., 12 (2003), pp. 451–512.
20. S. Ulbrich, *Optimal control of nonlinear hyperbolic conservation laws with source terms*, Habilitation thesis, Fakultät für Mathematik, Technische Universität München, http://www3.mathematik.tu-darmstadt.de/hp/optimierung/ulbrich-stefan/, 2001.
21. S. Ulbrich, *Adjoint-based derivative computations for the optimal control of discontinuous solutions of hyperbolic conservation laws*, Syst. Control Lett., 48 (2003), pp. 313–328.