Numerical Method for the Computation of Tangent Vectors to Hyperbolic Systems of Conservation Laws

Michael Herty and Benedetto Piccoli

Institut für Geometrie und Praktische Mathematik
RWTH Aachen

Templergraben 55, D–52056 Aachen (Germany)

RWTH Aachen University, IGPM, Templergraben 55, 52056 Aachen, Germany, herty@igpm.rwth-aachen.de

Department of Mathematics, Rutgers University, Camden, NJ, 08102, USA, piccoli@camden.rutgers.edu

NUMERICAL METHOD FOR THE COMPUTATION OF TANGENT VECTORS TO HYPERBOLIC SYSTEMS OF CONSERVATION LAWS

MICHAEL HERTY AND BENEDETTO PICCOLI

ABSTRACT. We are interested in the development of a numerical method for solving optimal control problems governed by hyperbolic systems of conservation laws. The main difficulty of computing the derivative in the case of shock waves is resolved in the presented scheme. Our approach is based a combination of a relaxation approach in combination with a numerical scheme to resolve the evolution of the tangent vectors. Numerical results for optimal control problems are presented.

AMS 35L65, 49K20, 49K40

Keywords. Conservation laws, optimization, tangent vectors.

1. INTRODUCTION

We are concerned with a numerical approach to optimization problems governed by (systems of) hyperbolic partial differential equations in a single spatial dimension. As a prototype, we consider a tracking type problem for a terminal state y_d prescribed at some given time $t = T$ and the control acts as initial condition u_0. A mathematical formulation of this optimal control problem is reduced to minimizing a functional, and, for instance, it can be stated as follows:

$$(1.1) \qquad \min_{u_0} J(y(\cdot, \cdot, T); y_d(\cdot)),$$

where J is the given cost functional and $y \in \mathbb{R}^n$ is the unique entropy solution of the nonlinear conservation law

$$(1.2) \qquad \begin{aligned} y_t + f(y)_x &= 0, & x \in \mathbb{R}, \quad t > 0, \\ y(0, x) &= u_0(x), & x \in \mathbb{R}. \end{aligned}$$

There has been tremendous progress in both analytical and numerical studies of problems of type (1.1), (1.2), see, e.g., [1, 2, 3, 7, 26, 11, 8, 12, 18, 23, 28, 27, 19, 17, 21]. Its solution relies on the property of the evolution operator $\mathcal{S}_t : u_0(\cdot) \to y(\cdot, t) = \mathcal{S}_t u_0(\cdot)$ for (1.2). It is known that the semi-group $\mathcal{S}_t$ generated by a nonlinear hyperbolic conservation law is generically nondifferentiable in L^1 even in the scalar one-dimensional (1-D) case (see, e.g., [11, Example 1]). A calculus for the first-order variations of $\mathcal{S}_t u_0$ with respect to u_0 has been established in [11, Theorems 2.2 and 2.3] for general 1-D systems of conservation laws with a piecewise Lipschitz continuous u_0 that contains finitely many discontinuities. Therein, the concept of generalized first order tangent vectors has

been introduced to characterize the evolution of variations with respect to u_0, see [11, equations (2.16)–(2.18)]. It has been further extended in [10] to establish continuous dependence of $\mathcal{S}_t$ on the initial data u_0. This result has been extended to BV initial data in [7, 3] and lead to the introduction of a differential structure for $u_0 \to \mathcal{S}_t u_0$, called shift-differentiability, see e.g. [3, Definition 5.1]. Further extensions have also been discussed for example [16]. Related to that equations for the generalized cotangent vectors have been introduced for 1-D systems in [9, Proposition 4]. These equations (also called adjoint equations) consists of a nonconservative transport equation [9, equation (4.2)] and an ordinary differential equation [9, equations (4.3)–(4.5)] for the tangent vector and shift in the positions of possible shocks in $y(x, t)$, respectively. Necessary conditions for a general optimal control problem have been established in [9, Theorem 1]. However, this result was obtained using strong assumptions on u_0 (see [9, Remark 4] and [3, Example 5.5]), which in the 1-D scalar case can be relaxed as shown for example in [28, 12]. We note that the nonconservative transport part of the adjoint equation has been intensively studied also independently from the optimal control context. In the scalar case we refer to [28, 5, 6, 25] for a notion of solutions and properties of solutions to those equations. Analytical results for optimal control problems in the case of a scalar hyperbolic conservation law with a convex flux have also been developed using a different approach in [28]. The relation to the weak formulation has been discussed in [2] in the case of Burger's equation.

Numerical methods for the optimal control problems have been discussed in [1, 19]. In [18], the adjoint equation has been discretized using a Lax-Friedrichs-type scheme, obtained by including conditions along shocks and modifying the Lax-Friedrichs numerical viscosity. Convergence of the modified Lax-Friedrichs scheme has been rigorously proved in the case of a smooth convex flux function. Convergence results have also been obtained in [28] for the class of schemes satisfying the one-sided Lipschitz condition (OSLC) and in [1] for implicit-explicit finite-volume methods. Other examples of finite volume methods and Lagrangian methods are given in [13, 22].

In [12], analytical and numerical results for the optimal control problem (1.1) coupled with the 1-D inviscid Burgers equation have been presented in the particular case of a least-square cost functional J. Therein, existence of a minimizer u_0 was proven, however, uniqueness could not be obtained for discontinuous functions y. This result was also extended to the discretized optimization problem provided that the numerical schemes satisfy either the OSLC or discrete Oleinik's entropy condition. Furthermore, convergence of numerical schemes was investigated in the case of convex flux functions and with a–priori known shock positions, and numerical resolution of the adjoint equations in both the smooth and non–smooth cases was studied. In [20] perturbations of initial data are studied using an additional spatial dimension. Numerical results as well as a formalism to derive the linearized equations have been presented therein. In the scalar case of a production model coupled to ordinary differential equations has been studied in [15]. Therein, convergence of the wave–front tracking approximation to the tangent vector equation is proven.

We contribute to the discussion by introducing a novel scheme which allows to include the arising discontinuities in an optimization framework and *without* an a–priori assumption on the location of the discontinuities. This is possible and computationally

efficient under three basic assumptions: first, we only compute derivatives with respect to piecewise constant controls u_0; second, the system (1.2) is not solved directly, but an $\epsilon-$relaxation approximation (2.4) is solved instead and last, we compute the exact derivative for the $\epsilon-$approximation of the system (2.4) using tangent vectors. The number of discontinuities in u_0 may herein be as large as $\frac{1}{\Delta x}$, where Δx is the spatial width of the numerical grid. The overall algorithm requires the solution of two additional hyperbolic partial differential equations. The motivation and theoretical investigations are presented in section 2 and numerical results in section 3.

2. Motivation and theoretical results

In order to derive a numerical scheme we consider the relaxation approximation [24] to (1.2). For simplicity we consider only the Jin–Xin relaxation in the case $n = 1$ and on the full real line $x \in \mathbb{R}$. Then, the hyperbolic relaxation for

$$(2.3) \qquad y_t^{(1)} + f(y^{(1)})_x = 0, \ y^{(1)}(0, x) = u_0(x),$$

is given by

$$(2.4) \qquad y_t + \begin{pmatrix} 0 & 1 \\ a^2 & 0 \end{pmatrix} y_x = \begin{pmatrix} 0 \\ \frac{1}{\epsilon}(f(y^{(1)}) - y^2) \end{pmatrix}$$

and by the initial data

$$(2.5) \qquad y^{(1)}(0, x) = u_0(x), \ y^{(2)}(0, x) = f(u_0(x)).$$

We assume that the value a^2 fulfills the sub characteristic condition

$$(2.6) \qquad a^2 - f'(y^{(1)})^2 \geq 0 \ \forall y^{(1)}.$$

For positive ϵ it is known that a solution y yields an approximation to (1.2) in the following sense. For ϵ sufficiently small we have up to second order in ϵ

$$y_t^{(1)} + f(y^{(1)})_x = \epsilon \left(\left(a^2 - f'(y^{(1)})^2 \right) y_x^{(1)} \right)_x.$$

Formally we obtain the original conservation law for $y^{(1)}$ in the limit $\epsilon = 0$. We refer to [4] for a detailed analysis. For $\epsilon > 0$ the method is known as a relaxing scheme. The main advantage of (2.4) over (1.2) is the linear transport which greatly simplifies the computation of associated tangent vectors (in particular the numerical resolution of the finite dimensional component is trivial for (2.4)).

In the following we will therefore discuss the optimization problem (1.1) with respect to (2.4). Still, the numerical computation of tangent vectors in general poses severe challenges addressed below. Therefore, we further simplify by only considering piecewise constant controls u_0. In the following $TV(\cdot)$ denotes the total variation.

Given some $C > 0$, we consider problem 1.1 subject to (2.4) and (2.5) for controls $u_0 \in \mathcal{U}$.

Definition 2.1. *We indicate by $\mathcal{U} := \{u : \mathbb{R} \to \mathbb{R} : u \text{ measurable }, TV(u) \leq C, u \text{ piecewise constant }\}$ the set of admissible controls. For every $u \in \mathcal{U}$ we indicate by $x_k = x_k(u), k = 1, \ldots, N(u)$ the points of discontinuity of u.*

For $y_d \in L^1(\mathbb{R})$, some $T > 0$ and a bounded interval $I \subset \mathbb{R}$ we consider as a prototype example an unregularized cost functional of tracking type

$$(2.7) \qquad J(y(\cdot,\cdot), y_d(\cdot)) = \int \chi_I(x) \left(y(T,x) - y_d(x) \right)^2 dx.$$
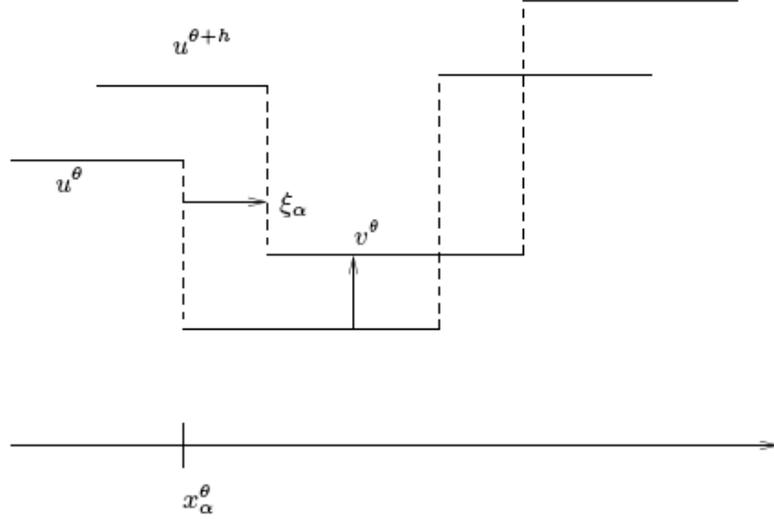
FIGURE 2.1.   Construction of a tangent vector (ξ_α, v_θ) to u^θ having a discontinuity at x_α^θ

We now introduce the notion of tangent vectors, see [11] and [10]. In (2.4) we assume $\epsilon > 0$ fixed and $f \in C^4(\mathbb{R})$ and a^2 fulfills the subcharacteristic condition. For a function $u \in \mathcal{U}$ a generalized tangent vector consists of two components (v, ξ) where $v \in L^1(\mathbb{R})$ describes the L^1 infinitesimal displacement and $\xi \in \mathbb{R}^{N(u)}$ describes the infinitesimal displacement of $N(u)$ discontinuities. A distance on the space of tangent vectors $T_u := L^1(\mathbb{R}; \mathbb{R}^n) \times \mathbb{R}^{N(u)}$ is given by

$$(2.8) \qquad \|(v, \xi)\| := \|v\|_{L^1} + \sum_{i=1}^{N(u)} |\Delta_i u| \, |\xi_i|$$

where $\Delta_i u = u(x_i+) - u(x_i-)$ denotes the jump in u. The distance depends u through the number of points of discontinuity. Tangent vectors may be used to describe variations of u, see Figure 2.1. For $\delta > 0$ an infinitesimal displacement u_δ of u is given by

$$(2.9) \qquad u_\delta = u + \delta v + \sum_{i=1}^{N(u)} \chi_{[x_i + \delta\xi_i, \, x_{i+1} + \delta\xi_{i+1}]}(x) u(x_i+)$$

u_δ is obtained from u by shifting the function values by δv and the ith discontinuities by $\delta\xi_i$. For δ sufficiently small u_δ has the same number of discontinuities as u. Note that if $\xi \neq 0$ then the function $\delta \to u_\delta$ is *not* differentiable in L^1 as the ratio $\frac{u_{\delta+h} - u_\delta}{h}$ does not

converge to any limit in L^1 for $h \to 0$. However, the previous limit remains meaningful if interpreted as a weak limit in a space of measures with a singular point mass located at x_i and having magnitude $|\Delta_i u|\xi_i$. Therefore, in [11] a class of variations $\delta \to u_\delta$ is described up to first order by (generalized) tangent vectors (v, ξ).

We introduce some notation and definitions (see [11, 10]) already in view of the special system (2.4). Let $u \in L^1(\mathbb{R}; \mathbb{R}^n)$ be a piecewise Lipschitz continous function with N jumps. Consider Σ_u the family of all continuous paths $\gamma : [0, \delta_0] \to L^1_{loc}$ with $\gamma(0) = u$ with δ_0 possibly depending on γ.

Definition 2.2. [11, Definition 1,3] *The space of generalized tangent vectors to a piecewise Lipschitz function u with jumps at the points $x_1 < x_2 \cdots < x_N$ is $T_u := L^1(\mathbb{R}; \mathbb{R}^n) \times \mathbb{R}^N$. A continuous path $\gamma \in \Sigma_u$ generates a tangent vector $(v, \xi) \in T_u$ if*

$$\lim_{\delta \to 0} \frac{1}{\delta} \|\gamma(\delta) - \bar{\gamma}(\delta)\|_{L^1} = 0$$

for

$$\bar{\gamma} := u + \delta v - \sum_{i:\xi_i>0} \Delta_i u \ \chi_{[x_i, x_i+\delta\xi_i]} + \sum_{i:\xi_i<0} \Delta_i u \ \chi_{[x_i+\delta\xi_i, x_i]}.$$

Let u be a piecewise Lipschitz function with simple discontinuities [11, Definition 2]. Then, a path $\gamma \in \Sigma_u$ is a regular variation for u if additionally all function $\gamma(\delta) = u_\delta$ are piecewise Lipschitz with simple discontinuities and the jumps x_i^δ depend continuously on δ.

A regular variation γ for u generates a tangent vector (v, ξ) by

(2.10)
$$\xi_i = \lim_{\delta \to 0} \frac{x_i^\delta - x_i}{\delta}, \ \lim_{\delta \to 0} \int_a^b \|\frac{u^\delta(x_i^\delta + y) - u(x_i + y)}{\delta} - v(x_i + y) - \xi_i u_x(x_i + y)\| dy = 0$$

whenever $[x_i + a, x_i + b]$ does not contain any other point of discontinuity of u except x_i. Further, the length of a regular path γ can be computed by (2.8). We now consider the initial data u_0 and a regular variation generating the tangent vector $(v, \xi) \in T_u$. Under suitable regularity assumptions [11, Theorem 2.2] regular variations are locally preserved by the system (2.4) and linearized equations for the time evolution of the tangent vector $(v(t, \cdot), \xi(\cdot))$ can be derived. We have:

Lemma 2.1. [11, Theorem 2.2] *Let $y(\cdot, \cdot)$ be a piecewise Lipschitz continuous solution to (2.4) and initial data (2.5) $y(0, \cdot) = \bar{y}$ piecewise Lipschitz with N simple discontinuities. Let $(\bar{v}, \bar{\xi}) \in T_{\bar{y}}$ be a tangent vector to $\bar{y}$ generated by the regular variation $\delta \to \bar{y}_\delta$. Let $y_\delta(t, x)$ be the solution to (2.4) and initial data (2.5) $y_\delta(0, x) = \bar{y}_\delta(\cdot)$. Then, there exists a time $t_0 > 0$ such that the path $\delta \to y_\delta(t, \cdot)$ is a regular variation of $y(t, \cdot)$ generating the tangent vector $(v(t), \xi(t)) \in T_{y(t, \cdot)}$ and (v, ξ) is the unique broad solution to*

(2.11)
$$v(0, \cdot) = \bar{v}(\cdot), v_t + \begin{pmatrix} 0 & 1 \\ a^2 & 0 \end{pmatrix} v_x = \frac{1}{\epsilon} \begin{pmatrix} 0 \\ f'(y^{(1)})v^{(1)} - v^{(2)} \end{pmatrix},$$

where $v = (v^{(1)}, v^{(2)})$ and outside of the discontinuities of y. For $i = 1, \ldots, N$ we have

(2.12)
$$\xi_i(t) = \bar{\xi}_i \ and \ l_j \cdot (\Delta_i v + \Delta_i y_x \xi_i) = 0 \ j \neq k.$$

along each line of discontinuity $x_i(t)$ where u has a discontinuity in the kth characteristic family. Here, $\Delta_i v = v(x_i(t)+, t) - v(x_i(t)-, t)$ and l_j denotes the jth left eigenvector of the matrix $\begin{pmatrix} 0 & 1 \\ a^2 & 0 \end{pmatrix}$.

Some remarks are in order. According to [11, Definition 2] $\bar{u} \in \mathcal{U}$ is piecewise Lipschitz with simple discontinuities. The equations (2.11) and (2.12) are particularly simple due to the linear transport in the hyperbolic relaxation. In particular, the equation for ξ_i is solved without any effort. Further, if we diagonalize (2.4) first and then apply the tangent vector calculus, the second equation (2.12) simplifies and we come back to this point later in the numerical scheme. It is important to note that in the previous result it is assumed that all variations u_δ posses the same number of discontinuities which in the case of problem (1.1) is unknown a priori.

Tangent vectors and their property (2.10) can be used to compute the derivative of the cost functional (2.7).

Lemma 2.2. *Assume the assertions of Lemma 2.1 hold true and let J be given by (2.7) and assume I is sufficiently large. Then, the variation of J with respect to initial data $y(0, x) = (u_0(x), f(u_0(x))$ is given by*

$$\nabla_{u_0} J(y, y_d) = 2 \int \chi_I(x) \left(y^{(1)}(T, x) - y_d(x) \right) v^{(1)}(T, x) dx +$$

$$(2.13) \qquad \sum_{i=1}^{N(u_0)} \left(\left(y^{(1)}(T, x_i+) - y_d(x_i+) \right) + \left(y^{(1)}(T, x_i-) - y_d(x_i-) \right) \right) \Delta_i y^{(1)}(T, \cdot) \xi_i(T).$$

The proof is similar to [15, Proposition 1] and omitted.

Lemma 2.2 and equation (2.9) already suggests a numerical method for solving (1.1) with cost (2.7). Given some control u_0 and a stepsize $\rho > 0$ we obtain a new control $\tilde{u}_0$ corresponding to smaller value of the cost functional J by the following variation

$$(2.14) \qquad \tilde{u}_0(x) = u_0(x) - \left(\rho v(0, x) - \sum_{i=1}^{N(u_0)} \chi_{[x_i+\rho\xi_i(0), x_{i+1}+\rho\xi_{i+1}(0)]}(x) u_0(x_i+) \right),$$

where $v(0, x)$ is the solution at time $t = 0$ to (2.11) for *terminal* data

$$(2.15) \qquad v^{(1)}(T, x) = \left(y^{(1)}(T, x) - y_d(x) \right), \quad v^{(2)}(T, x) = 0,$$

and $\xi_i(0)$ the solution to (2.12) with *terminal* data

$$(2.16) \qquad \xi_i(T) = \left(\left(y^{(1)}(T, x_i+) - y_d(x_i+) \right) + \left(y^{(1)}(T, x_i-) - y_d(x_i-) \right) \right) \Delta_i y^{(1)}(T).$$

Obviously, the given choice for $(v(T, x), \xi(T))$ yields $J(\tilde{u}_0) < J(u_0)$. Note that this requires to solve (2.11) backwards in time and to fulfill (2.12). The previous computations motivate a numerical scheme for approximately solving (1.1), (2.7) and (2.4).

Before stating the full discrete algorithm we reformulate and comment on some parts of the method. The system (2.4) is diagonalisable with eigenvalues $\lambda_{1,2} = \pm a$ and characteristic variables

$$(2.17) \qquad \eta^{(1)} = y^{(2)} + a y^{(1)} \text{ and } \eta^{(1)} = y^{(2)} - a y^{(1)}.$$

Also, in view of condition (2.12) it is numerically advantageous to consider the minimization problem for J in characteristic variables $\eta = (\eta^{(1)}, \eta^{(2)})$. Furthermore, in view of (2.14) equation (2.11) will be solved backwards in time for given terminal data $v(T, x)$. We obtain for $\tilde{v}(t, x) = v(T - t, x)$ the system

$$(2.18) \qquad \tilde{v}(0, \cdot) = v(T, \cdot), \tilde{v}_t - \begin{pmatrix} 0 & 1 \\ a^2 & 0 \end{pmatrix} \tilde{v}_x = \frac{1}{\epsilon} \begin{pmatrix} 0 \\ \tilde{v}^{(2)} - f'(y^{(1)}(T - t, \cdot))\tilde{v}^{(1)}. \end{pmatrix}$$

and eigenvalues $\lambda_{1,2} = \mp a$ and corresponding characteristic variables

$$\varphi^{(1)} = \tilde{v}^{(2)} + a\tilde{v}^{(1)} \text{ and } \varphi^{(2)} = \tilde{v}^{(2)} - a\tilde{v}^{(1)}.$$

Let $I = [0, 1]$, $T > 0$, y_d and $\epsilon > 0$ be given. The cost functional J is given by (2.7). We discuss the numerical discretization of problem (2.19) using a first-order finite volume scheme with periodic boundary conditions. Note that we leave (for the moment) *both* components of the initial data y_0 subject to optimization. For ϵ sufficiently small a solution to (2.4) will due to the relaxation term satisfy $\|f(y_0^{(1)}) - y_0^{(2)}\| << 1$.

$$(2.19) \qquad \min_{y_0 = (y_0^{(1)}, y_0^{(2)})} J \text{ sbj to } (2.4), \ y(0, x) = y_0(x), \ y(t, 1) = y(t, 0), \ x \in I, \ t \geq 0.$$

Fix a^2 such that the subcharacteristic condition (2.6) is fulfilled Introduce an equidistant spatial grid $\{x_i\}_{i=0}^{N_x}$ on I with $\Delta x = x_{i+1} - x_i$. We choose Δt such that the CFL condition holds, i.e., $\Delta t |a| = \Delta x$, and denote by $t^n = \Delta t \, n$ for $n = 0, \ldots, N_t$. We write $x_{i+\frac{1}{2}} = x_i + \frac{\Delta x}{2}$ and for simplicity assume $x_{N_x} = 1$ and $t^{N_t} = T$. Also for notational convenience we denote by $x_{-\frac{1}{2}} = x_{N_x - \frac{1}{2}}$ and $x_{N_x + \frac{1}{2}} = x_{\frac{1}{2}}$. Let $\mathcal{T}^{-1} \in \mathbb{R}^{2 \times 2}$ be the transformation to characteristic variables (2.17) , i.e.,

$$\mathcal{T}^{-1} \begin{pmatrix} 0 & 1 \\ a^2 & 0 \end{pmatrix} \mathcal{T} = \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix}, \ \eta = \mathcal{T}^{-1} y.$$

The cell average on $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ at time t^n for any function $u(t, x)$ is denoted by $u_i^n = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u(t^n, x)dx$. As in [24] an operator splitting is used to discretize (2.4) and similarly for (2.11).

Then, a first–order Upwind discretization of problem (2.19) with an exact integration of the source term for the solution of (2.4) is given by (2.20) for $i = 0, \ldots, N_x$ and $n = 1, \ldots, N_t$.

$$(2.20a) \qquad y_i^0 = (y_0)_i$$

$$(2.20b) \qquad \eta_i^{(1)} = (\mathcal{T}^{-1} y_{i-1}^{n-1})^{(1)}, \ \eta_0^{(1)} = (\mathcal{T}^{-1} y_{N_x}^{n-1})^{(1)}, \ \eta_i^{(2)} = (\mathcal{T}^{-1} y_i^{n-1})^{(2)},$$

$$(2.20c) \qquad \tilde{y}_i^{(1)} = (\mathcal{T} \eta_i)^{(1)}, \ \tilde{y}_i^{(2)} = \exp(-\frac{\Delta t}{\epsilon})(\mathcal{T} \eta_i)^{(2)} + (1 - \exp(-\frac{\Delta t}{\epsilon}))f(\tilde{y}_i^{(1)}),$$

$$(2.20d) \qquad \eta_i^{(2)} = (\mathcal{T}^{-1} \tilde{y}_{i+1})^{(2)}, \ \eta_{N_x}^{(2)} = (\mathcal{T}^{-1} \tilde{y}_0)^{(2)}, \ \eta_i^{(1)} = (\mathcal{T}^{-1} \tilde{y}_i)^{(1)},$$

$$(2.20e) \qquad y_i^n = \mathcal{T} \eta_i.$$

The discretization (2.20) uses a different splitting compared to [24] which leads to more complicated update formulas above but will be advantageous later on. In the current splitting we do first transport in the first characteristic variable, then apply the source

term and finally transport the second characteristic variable. The transformation to characteristic variables η and the CFL condition allows to resolve the transport exactly.

As explained above instead of equation (2.11) we discretize equation (2.18) to solve for the variations $\tilde{v}$. Similarly to (2.20), we transform (2.18) to characteristic variables φ and resolve the linear transport exactly. If the discretized initial data is denoted by $\tilde{v}_i^0$, then we obtain for $i = 0, \ldots, N_x$ and $n = 1, \ldots, N_t$

$$(2.21a) \qquad \varphi_i^{(2)} = (\mathcal{T}^{-1}\tilde{v}_{i-1}^{n-1})^{(2)}, \ \varphi_0^{(2)} = (\mathcal{T}^{-1}\tilde{v}_{N_x}^{n-1})^{(2)}, \ \varphi_i^{(1)} = (\mathcal{T}^{-1}\tilde{v}_i^{n-1})^{(1)}$$

$$(2.21b) \qquad \bar{v}_i^{(1)} = (\mathcal{T}\varphi_i)^{(1)}, \ \bar{v}_i^{(2)} = \exp(\frac{\Delta t}{\epsilon})(\mathcal{T}\varphi_i)^{(2)} + (1 - \exp(\frac{\Delta t}{\epsilon})f'(y_i^{(1),N_t-n})\bar{v}_i^{(1)}$$

$$(2.21c) \qquad \varphi_i^{(1)} = (\mathcal{T}^{-1}\bar{v}_{i+1})^{(1)}, \ \varphi_{N_x}^{(1)} = (\mathcal{T}^{-1}\bar{v}_0)^{(1)}, \ \varphi_i^{(2)} = (\mathcal{T}^{-1}\bar{v}_i)^{(2)},$$

$$(2.21d) \qquad \tilde{v}_i^n = \mathcal{T}\varphi_i.$$

Next we turn to the discretization of equation (2.12). Within a first–order finite volume scheme a piecewise constant approximation is used to recover the solution, i.e.,

$$(2.22) \qquad\qquad y(t,x) \approx \sum_{i=0}^{N_x} \chi_{[t^n,t^{n+1}] \times [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]}(t,x) y_i^n,$$

and similarly for the initial data. Therefore, numerically solving problem (2.19) naturally leads to consider piecewise constant controls $y_0 \in \mathcal{U}$ having discontinuities at possibly each cell boundary $x_{i+\frac{1}{2}}$. Therefore, a shift in the position of the discontinuity ξ_i may occur on each boundary $x_{i+\frac{1}{2}}$. As long as the spatial resolution is not modified the number of discontinuities is however fixed being a crucial assumption in Theorem 2.1. Since the conservative y and characteristic variables η are equivalent upon the linear transformation $\mathcal{T}$ we optimize in (2.19) for $\eta_0 = \mathcal{T}^{-1}y_0$ instead of y_0. Furthermore, we consider as admissible controls piecewise constant $\eta_0^{(i)} \in \mathcal{U}$ for $i = 1, 2$ having each only $\frac{N_x}{2}$ points of discontinuity which are of the following type: the first component $\eta_0^{(1)}$ may have a discontinuity only at $x_{i+\frac{1}{2}}$ for some odd value i and the second component $\eta_0^{(2)}$ may only have a discontinuity at $x_{i+\frac{1}{2}}$ for some even value i. We therefore choose $\eta_0^{(i)}$ according to (2.22) but such that

$$(2.23) \qquad\qquad (\eta_0^{(1)})_{2i} = (\eta_0^{(1)})_{2i+1}, \ (\eta_0^{(2)})_{2i-1} = (\eta_0^{(2)})_{2i}, \ i = 0, \ldots, \frac{N_x}{2}.$$

For fixed N_x the set of all admissible controls $\mathcal{U}_{ad} \subset \mathcal{U}$ consists of all piecewise constant functions $\eta_0(x)$ given by

$$\eta_0^{(j)}(x) = \sum_{i=0}^{N_x} \chi_{[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]}(x) \eta_{0,i}^{(j)}, \ j = 1, 2,$$

which additionally fulfill (2.23). Note that for N_x sufficiently large the condition (2.23) still allows to approximate any piecewise constant function.

When computing the tangent vector to η_0 we now have the L^1-variations φ_0 and the variation in the position of the discontinuities ξ_i. We denote by ξ_i, $i = 0, \ldots, N_x$ the variation of the discontinuity at position $x_{i+\frac{1}{2}}$. Hence, ξ_i for i odd (even) is the

variation of the discontinuity in the first (second) component of η_0. The discretization (2.20) introduces a splitting of the dynamics (2.4) and may be written for $t \in [t^n, t^{n+1}]$ and in characteristic variables as follows

$$\partial_t \eta^{(1)} + a \partial_x \eta^{(1)} = 0, \partial_t \eta^{(2)} = 0,$$
$$\partial_t \eta^{(1)} = S(\eta^{(1)}, \eta^{(2)}), \ \partial_t \eta^{(2)} = S(\eta^{(1)}, \eta^{(2)}),$$
$$\partial_t \eta^{(2)} - a \partial_x \eta^{(2)} = 0, \partial_t \eta^{(1)} = 0.$$

Due to the discretization of the linear transport term the initial discontinuities in η_0 propagate with speed $\pm a$ and therefore during after one time step Δt they are again located at the cell interfaces $x_{i+\frac{1}{2}}$. Further within Δt the splitting does *not* introduce additional discontinuities due the action of the source term $S(\eta^{(1)}, \eta^{(2)}) = \frac{1}{\epsilon} \left(f((\mathcal{T}\eta)^{(1)}) - (\mathcal{T}\eta)^{(2)} \right)$ provided that η_0 fulfills (2.23). Note that the same holds true for φ provided that $\varphi_i^0 = \mathcal{T}^{-1} \tilde{v}_i^0$ fulfills (2.23). The position of the discontinuity in the first component and second component of η at time t^n are given by

$$(2.24) \qquad x_{2i-1}(t^n) = x_{2i-1}(0) + a \ t^n, \ x_{2i}(t^n) = x_{2i}(0) - a \ t^n, \ i = 0, \ldots, N_x,$$

where $x_j(0) = x_{j+\frac{1}{2}}$ and where the discontinuities exiting at $x = 1$ (or $x = 0$) enter again at $x = 0$ ($x = 1$). We turn to the discussion of the second part of equation (2.12). In characteristic variables ℓ_j is the jth unit vector. Let i be odd and η be computed by the previous splitting. Then, we observe $\Delta_i \eta_x^{(2)} = 0$ since $\eta^{(2)}$ is constant across the position of the discontinuity in the first family $x_{2i-1}(t^n)$. Provided we discretize J such that $\varphi_j^0 := \mathcal{T}^{-1} \tilde{v}_j^0, \ j = 0, \ldots, N_x$ in (2.21) also fulfills (2.23), then, we have $\Delta_i \varphi^{(2)} = 0$. The same is true for i even and, therefore, the second part of equation (2.12) is fulfilled trivially within the previous scheme.

Finally, we consider $J(\eta, y_d) = J(\mathcal{T}y, y_d)$ and obtain the gradient of J in terms of the characteristic variables η and its associated tangent vectors (ϕ, ξ) as

(2.25)

$$\nabla_{\eta_0} J(\eta, y_d) = \frac{1}{2} \int \chi_I(x) \left(\frac{\eta^{(1)}(T, x) - \eta^{(2)}(T, x)}{2a} - y_d(x) \right) \phi^{(1)}(T, x) dx +$$

$$(2.26) \qquad - \frac{1}{2} \int \chi_I(x) \left(\frac{\eta^{(1)}(T, x) - \eta^{(2)}(T, x)}{2a} - y_d(x) \right) \phi^{(2)}(T, x) dx +$$

$$(2.27) \qquad \frac{1}{2} \sum_{i=1}^{N(u_0)} \left\{ \left(\frac{\eta^{(1)}(T, x_i+) - \eta^{(2)}(T, x_i+)}{2a} - y_d(x_i+) \right) + \right.$$

$$(2.28) \qquad \left. \left(\frac{\eta^{(1)}(T, x_i-) - \eta^{(2)}(T, x_i-)}{2a} - y_d(x_i-) \right) \right\} \xi_i(T) \left(\Delta_i \eta^{(1)}(T, \cdot) - \Delta_i \eta^{(2)}(T, \cdot) \right).$$

Let $\varphi(t,x) = \phi(T-t,x)$. Then, similarly to (2.14) we may use (2.25) to determine a descent direction for J by the following discretization for $i = 0, \ldots, N_x$,

$$(2.29a) \qquad \varphi_{2i}^{(1),0} = \frac{\eta_{2i}^{(1),N_t} - \eta_{2i}^{(2),N_t}}{2a} - (y_d)_{2i}, \ \varphi_{2i+1}^{(1),0} = \varphi_{2i}^{(1),0},$$

$$(2.29b) \qquad \varphi_{2i-1}^{(2),0} = -\left(\frac{\eta_{2i-1}^{(1),N_t} - \eta_{2i-1}^{(2),N_t}}{2a} - (y_d)_{2i-1} \right), \ \varphi_{2i}^{(2),0} = \varphi_{2i-1}^{(2),0},$$

$$(2.29c) \qquad \xi_{2i-1} = \frac{1}{2a} \left(\overline{\Delta}_{2i-1}(\mathcal{T}\varphi^{N_t})^{(1)} - \overline{\Delta}_{2i-1}y_d \right) \Delta_{2i-1}\varphi^{(1),N_t},$$

$$(2.29d) \qquad \xi_{2i} = -\frac{1}{2a} \left(\overline{\Delta}_{2i}(\mathcal{T}\varphi^{N_t})^{(1)} - \overline{\Delta}_{2i}y_d \right) \Delta_{2i}\varphi^{(2),N_t},$$

where $\Delta_j w(\cdot) = w(x_{j+1}) - w(x_j)$ and $\overline{\Delta}_j w(\cdot) = \frac{1}{2}\left(w(x_{j+1}) + w(x_j) \right)$.

Finally, we discretize (2.14) such that $\mathcal{T}^{-1}\eta_0 \in \mathcal{U}$ and fulfills (2.23). The tangent vector $(\varphi(T,\cdot), \xi(T))$ to $\eta_0(x)$ describes the L^1-variation as well as the variation of the position of discontinuities, respectively. Given a current control η_i^0 we obtain a new control $\tilde{\eta}_i^0$ for $i = 0, \ldots, N_x$, and $j \in \{0, \ldots, N_x\}$ odd, and $k \in \{0, \ldots, N_x\}$ even by equation (2.30).

$$(2.30a) \qquad \Xi_{j-1}^1 = \Xi_j^1 = \min\{(-\overline{\xi}_j)^+, \Delta x\}\eta_{j+1}^{(1),0} + \max\{\overline{\xi}_{j-2}^+ - \Delta x, 0\}\eta_{j-1}^{(1),0} +$$

$$(2.30b) \qquad (\Delta x - \min\{(-\overline{\xi}_j)^+, \Delta x\} - \max\{\overline{\xi}_{j-2}^+ - \Delta x, 0\})\eta_j^{(1),0},$$

$$(2.30c) \qquad \Xi_{k-1}^2 = \Xi_k^2 = \min\{(-\overline{\xi}_k)^+, \Delta x\}\eta_{k+1}^{(2),0} + \max\{\overline{\xi}_{k-2}^+ - \Delta x, 0\}\eta_{k-1}^{(2),0} +$$

$$(2.30d) \qquad (\Delta x - \min\{(-\overline{\xi}_k)^+, \Delta x\} - \max\{\overline{\xi}_{k-2}^+ - \Delta x, 0\})\eta_k^{(2),0},$$

$$(2.30e) \qquad \tilde{\eta}_i^{(1),0} = \eta_i^{(1),0} - \varphi_i^{(1),N_t} - \frac{\Xi_i^1}{\Delta x}$$

$$(2.30f) \qquad \tilde{\eta}_i^{(2),0} = \eta_i^{(2),0} - \varphi_i^{(2),N_t} - \frac{\Xi_i^2}{\Delta x}.$$

Here, we denote by $x^+ = \max\{x, 0\}$ and by $\overline{\xi}_i = \mathcal{P}(\xi_i)$ where $\mathcal{P}$ is the projection on $[-2\Delta x, 2\Delta x]$. Note that equation (2.30) corresponds to a gradient step in the L^1-variation but a scaled gradient step in the variation of the shock position in order to prevent shock variations to interact. In (2.30) a piecewise constant reconstruction of $\tilde{\eta}$ is computed where for example in the case of the first component the discontinuity at $x_{i+\frac{1}{2}}$ is moved by ξ_i and at $x_{i-\frac{3}{2}}$ by ξ_{i-2}. Since φ fulfills (2.23) this holds true for $\tilde{\eta}$.

The previous computation leads to an iterative algorithm for numerically solving (2.19) or equivalently

$$\min_{\eta_0 \in \mathcal{U}} J(\mathcal{T}\eta, y_d) \text{ sbj to } y = \mathcal{T}\eta, (2.4), \eta(0,x) = \eta_0(x), \ \eta(t,0) = \eta(t,1) \text{ and } (2.23).$$

Algorithm

(1) Set terminal time $T > 0$, $a^2 \geq \max_{y^{(1)}}(f'(y^{(1)})^2$ and choose an equidistant spatial discretization with N_x gridpoints. Choose $\Delta t = \frac{\Delta x}{a}$ and $\mathbf{k} = 0$. Let $\eta_{0,i}^{\mathbf{k}} =$

$\left((\eta_{0,i}^{(1)})^{\mathbf{k}},(\eta_{0,i}^{(2)})^{\mathbf{k}}\right)$ for $i = 0,\ldots,N_x$, be an arbitrary initial control such that $\eta_0^{\mathbf{k}}$ fulfills (2.23). Let $(y_d)_i$ be a discretization of the given function $y_d(\cdot)$.

(2) Solve equations (2.20) with $(y_0)_i := \mathcal{T}\eta_{0,i}^{\mathbf{k}}$ to obtain $\eta_i^{N_t} = \mathcal{T}^{-1}y_i^{N_t}$.

(3) Set initial data φ_i^0 and shock variations ξ_i by equation (2.29) where $x_i(t^{N_t})$ are given by equation (2.24).

(4) Solve equations (2.21) for initial data $\tilde{v}_i^0 := \mathcal{T}\varphi_i^0$ to obtain $\varphi_i^{N_t} = \mathcal{T}^{-1}\tilde{v}_i^{N_t}$.

(5) Update the current iterate $\eta_{0,i}^{\mathbf{k}}$ by setting $\eta_{0,i}^{\mathbf{k+1}} := \tilde{\eta}_i^0$ obtained from equation (2.30) with $\eta_i^0 := \eta_{0,i}^{\mathbf{k}}$.

(6) Provided that $J(\mathcal{T}\eta, y_d)$ is sufficiently small we terminate. Otherwise set $k \to k+1$ and continue with step (2).

3. Numerical results

We present numerical results using the previous calculus for two cases. The simplest possible application is the optimal control of a linear system (2.4) *without* source term. Second, we present results on the optimal control for the relaxation system (2.4) for Burgers flux $f(w) = \frac{1}{2}w^2$. All spatial grids are equidistant on $I = [0,1]$ and the temporal discretization is such that the CFL condition [14] is satisfied. We use periodic boundary conditions in all cases. The cost functional J is discretized using the trapezoidal rule. All initial controls are constant with $\eta_0(x) = (\frac{1}{2}, \frac{1}{4})$.
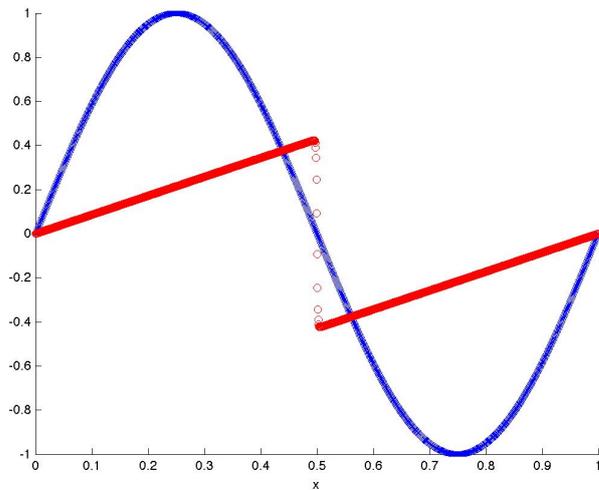
FIGURE 3.2. Solution $y(T = 1, x)^{(1)}$ (red circle) computed by discretization (2.20) with initial data $y_0^{(1)}(x)$ (blue cross) and $y_0^{(2)}(x) = f(y_0^{(1)}(x))$ on a grid with $N_x = 1600$ points in space and $\epsilon = 10^{-6}$.

3.1. **Grid convergence example.** We consider the discretization (2.20) and initial data given by $y_0^{(1)} = \sin(2\pi x)$ and $y_0^{(2)} = f(y_0^{(1)})$. We set $a^2 = 1$ and compute the solution to the proposed scheme (2.20) at time $T = 1$, $\epsilon = 10^{-6}$ and for $N_x = 1600$ points in the domain I. The initial data $y_0^{(1)}$ as well as the solution $y(1, \cdot)^{(1)}$ are depicted in Figure 3.2. Table 3.1 shows the grid convergence behavior of the discretization (2.20) in both components. As expected we observe first-order convergence and remark that the applied modifications do not alter the properties of the original scheme proposed in [24].

N_x	$\|y_*^{(1)}(1, \cdot) - y_{N_x}^{(1)}(1, \cdot)\|_2$	Rate	$\|y_*^{(2)}(1, \cdot) - y_n^{(2)}(1, \cdot)\|_2$	Rate
50	$1.2126e - 02 \ (0.0000e + 00)$		$1.5248e - 03 \ (0.0000e + 00)$	
100	$5.8062e - 03 \ (1.0442e + 00)$		$5.3919e - 04 \ (1.4139e + 00)$	
200	$2.6396e - 03 \ (1.0998e + 00)$		$1.3022e - 04 \ (2.0704e + 00)$	
400	$1.0128e - 03 \ (1.3031e + 00)$		$3.3623e - 05 \ (1.9364e + 00)$	

TABLE 3.1. Grid convergence of the discretization (2.20). y_* is a fine grid solution of (2.20) with $N_x = 1600$ points in space. y_{N_x} is the numerical solution on the grid given in first column of the table.

3.2. **Optimal control of a linear system.** We consider problem (3.31) with periodic boundary conditions on the domain $x \in I$, $T = 0.35$ and $a^2 = \frac{5}{4}$.

(3.31)
$$\min_{u_0} \int_0^1 \left(y^{(1)}(T, x) - y_d^{(1)}(x) \right) dx \text{ sbj to } y_t + \begin{pmatrix} 0 & 1 \\ a^2 & 0 \end{pmatrix} y_x = 0, \ y(0, x) = (u_0(x), f(u_0)).$$

In the example we initialize the control similar to (2.5) and $f(w) = \frac{1}{2}w^2$. The desired state y_d is depicted in Figure 3.3 (along with the obtained optimized state).

The algorithm is started with constant control η_0 and stopped after at most 2000 steps of iteration. The iteration history is depicted in the right part of Figure 3.3. With the proposed method we observe convergence until the grid resolution is reached. The dependence on the spatial grid in Table 3.2. As expected we observe first-order convergence is observed.

We compare the obtained results also with the case when the shock variations ξ_i are **not** taken into account. Hence, we consider the same example as before and run the same algorithm **but** setting $\Xi_i^j \equiv 0$ for $j = 1, 2$ in equation (2.30). The dependence on the spatial grid in this situation is depicted in Table 3.3. Comparing with Table 3.2 we observe a deterioration in the convergence rate from ≈ 1 when including the shock variations to $\approx \frac{1}{2}$ for the cost functional and $\approx \frac{3}{4}$ for state when neglecting the this contribution.
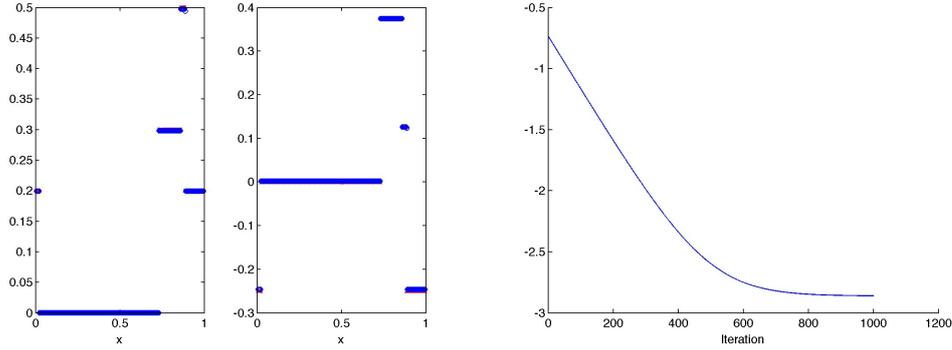
FIGURE 3.3. Desired state (red cross) and optimized state (blue circles) for the linear system with $N_x = 200$ grid points in space for both components $(y^{(1)}, y^{(2)})$ in the left part of the figure. Iteration history in $\log-$scale for the cost. The spatial resolution of the scheme is $\Delta x = 5 \times 10^{-3}$ and $\log_{10}(\Delta x) \approx -2.3$.

N_x	J	Rate	$\|y_0^{(1)} - y_*^{(1)}\|_2$	Rate	$\|y_0^{(2)} - y_*^{(2)}\|_2$	Rate
50	$4.6936e-03$	(0.00)	$5.5364e-03$	(0.00)	$6.9206e-03$	(0.00)
100	$2.5878e-03$	(0.91)	$2.8922e-03$	(0.96)	$3.6152e-03$	(0.96)
200	$1.3510e-03$	(0.96)	$1.4739e-03$	(0.98)	$1.8424e-03$	(0.98)
400	$9.9964e-04$	(0.68)	$8.2176e-04$	(0.90)	$1.2031e-03$	(0.77)

	N_x	$\|y_d^{(1)} - y_*^{(1)}\|_2$	Rate	$\|y_d^{(2)} - y_*^{(2)}\|_2$	Rate	
	50	$5.5364e-03$	(0.00)	$6.9206e-03$	(0.00)	
	100	$2.8922e-03$	(0.96)	$3.6152e-03$	(0.96)	
	200	$1.4739e-03$	(0.98)	$1.8424e-03$	(0.98)	
	400	$7.5945e-04$	(0.97)	$1.2666e-03$	(0.73)	

TABLE 3.2. Convergence history for different spatial grids with N_x equidistant grid points. Reported are the value of the cost functional J after optimization, the L^2-norm difference of both components $(y^{(1)}, y^{(2)})$ for the initial data $y_0 = y(0, x)$ and the desired state y_d. The optimized solution is denoted by $y_*(t, x)$.

N_x	J	Rate	$\|y_0^{(1)} - y_*^{(1)}\|_2$	Rate	$\|y_0^{(2)} - y_*^{(2)}\|_2$	Rate
50	$9.8550e - 04$	(0.00)	$1.9823e - 03$	(0.00)	$3.6662e - 03$	(0.00)
100	$9.8310e - 04$	(0.50)	$1.3402e - 03$	(0.74)	$2.6691e - 03$	(0.69)
200	$9.9839e - 04$	(0.49)	$9.5812e - 04$	(0.70)	$1.8837e - 03$	(0.71)
400	$9.9986e - 04$	(0.50)	$6.7520e - 04$	(0.71)	$1.3360e - 03$	(0.70)

N_x	$\|y_d^{(1)} - y_*^{(1)}\|_2$	Rate	$\|y_d^{(2)} - y_*^{(2)}\|_2$	Rate
50	$1.6294e - 03$	(0.00)	$3.9955e - 03$	(0.00)
100	$1.0328e - 03$	(0.79)	$2.8747e - 03$	(0.69)
200	$7.3280e - 04$	(0.70)	$2.0440e - 03$	(0.70)
400	$4.9368e - 04$	(0.74)	$1.4578e - 03$	(0.70)

TABLE 3.3. Convergence history for different spatial grids with N_x equidistant grid points *neglecting* the effect of the shock variations ξ_i on the optimal control. As in Table 3.2 we report the value of the cost functional J after optimization, the L^2-norm difference of both components $(y^{(1)}, y^{(2)})$ for the initial data $y_0 = y(0, x)$ and the desired state y_d. The optimized solution is denoted by $y_*(t, x)$.

3.3. **Optimal control of the relaxation approximation to Burgers Equation.**
We consider problem (3.31) $T = 0.35$, $\epsilon = 10^{-4}$, $N_x = 400$ and $a = \frac{5}{4}$. The desired state
y_d is the solution at time T to (2.4) and (2.5) for $u_0(x) = \cos(2\pi x)$. The algorithm is
started with constant control η_0 and stopped after the cost functional is below 10^{-4}. We
depict desired and optimized state in the original variables $y(T, \cdot)$ in Figure 3.4. The
obtained optimized control $\eta_0^{\mathbf{k}}$ and the function u_0 are depicted in Figure 3.4 and the
iteration history in Figure 3.5.
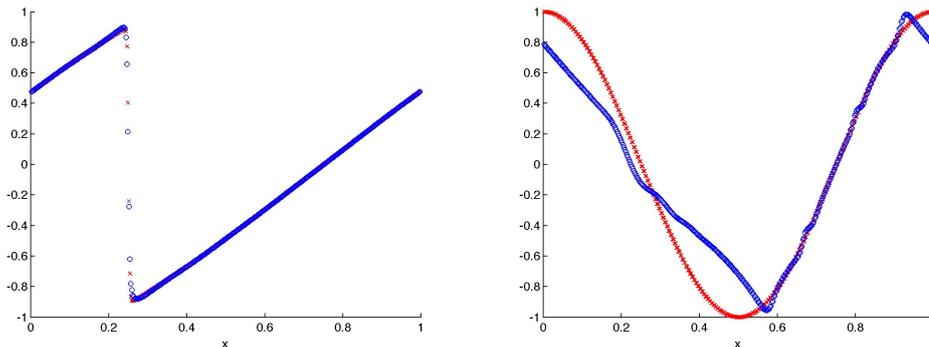
FIGURE 3.4. Desired state (red cross) and optimized state (blue circles)
for the linear system with $N_x = 400$ grid points in space for the variable
$y^{(1)}(T, x)$ in the left part of the figure and for $y^{(1)}(0, x)$ in the right part.
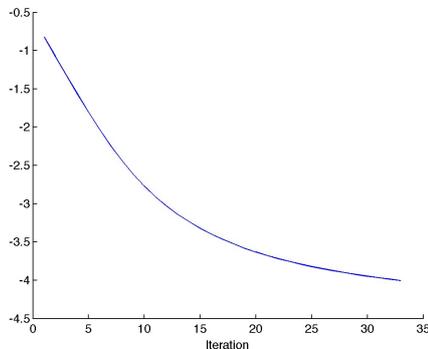
FIGURE 3.5. Iteration history for the example of Figure 3.4 in $\log$ −scale
for the value of the cost functional J. The stopping criteria is $J \leq 10^{-4}$
which is comparable with the spatial resolution of the scheme is $\Delta x = 2.5 \times 10^{-3}$.

We observe a good agreement in the recovered desired state y_d. The difference in the
obtained control is due to the fact that the solution to (2.19) is not unique.

In Figure 3.6 and Figure 3.7 we present a similar example but with desired state y_d
obtained as solution to (2.4) and (2.5) for $u_0(x) = \chi_{[0.3,0.45]}(x)$. All other parameters are

as before. Similarly to the previous results we observe that there is no uniqueness in the control u_0 leading to y_d.
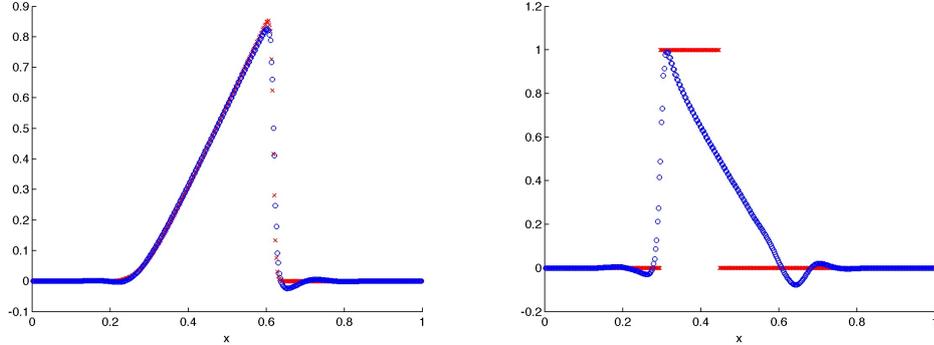
FIGURE 3.6. Desired state (red cross) and optimized state (blue circles) for the linear system with $N_x = 400$ grid points in space for the variable $y^{(1)}(T, x)$ in the left part of the figure and for $y^{(1)}(0, x)$ in the right part.
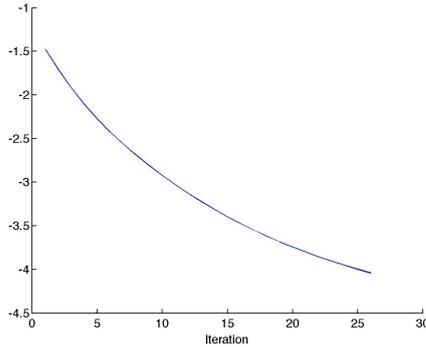
FIGURE 3.7. Iteration history for the example of Figure 3.4 in log −scale for the value of the cost functional J. The stopping criteria is $J \leq 10^{-4}$ which is comparable with the spatial resolution of the scheme is $\Delta x = 2.5 \times 10^{-3}$.

4. SUMMARY

We present a numerical method for solving optimal control problems subject to the relaxation approximation to a scalar hyperbolic conservation law. Equations for the evolution of the corresponding tangent vector of the relaxation system are derived and a numerical discretization has been introduced. The tangent vector has been used to compute the analytical gradient of the reduced cost also in the presence of traveling discontinuities. A numerical discretization of the gradient has been implemented to solve some examples of linear and nonlinear optimal control problems. The computation of

the finite dimensional part of the tangent vector allows to obtain the expected order of convergence.

References

[1] Mapundi K. Banda and Michael Herty. Adjoint IMEX-based schemes for control problems governed by hyperbolic conservation laws. *Comput. Optim. Appl.*, 51(2):909–930, 2012.

[2] Claude Bardos and Olivier Pironneau. A formalism for the differentiation of conservation laws. *C. R. Math. Acad. Sci. Paris*, 335(10):839–845, 2002.

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

[4] Stefano Bianchini. Hyperbolic limit of the Jin-Xin relaxation model. *Comm. Pure Appl. Math.*, 59(5):688–753, 2006.

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

[6] François Bouchut and François James. Differentiability with respect to initial data for a scalar conservation law. In *Hyperbolic problems: theory, numerics, applications, Vol. I (Zürich, 1998)*, volume 129 of *Internat. Ser. Numer. Math.*, pages 113–118. Birkhäuser, Basel, 1999.

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

[8] A. Bressan and A. Marson. A maximum principle for optimally controlled systems of conservation laws. *Rend. Sem. Mat. Univ. Padova*, 94:79–94, 1995.

[9] A. Bressan and W. Shen. Optimality conditions for solutions to hyperbolic balance laws. *Contemporary Mathematics*, 426:129, 2007.

[10] Alberto Bressan, Graziano Crasta, and Benedetto Piccoli. Well-posedness of the Cauchy problem for $n \times n$ systems of conservation laws. *Mem. Amer. Math. Soc.*, 146(694):viii+134, 2000.

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

[12] 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.*, 18:369–416, 2008.

[13] A. Chertok, M. Herty, and A. Kurganov. An eulerian-lagrangian method for optimization problems governed by nonlinear hyperbolic pdes. *Computational Optimization and Applications*, 2014.

[14] R. Courant, K O. Friedrichs, and H. Lewy. Über die partiellen differenzengleichungen der mathematischen physik. *Mathematische Annalen*, 100(1):32–74, 1928.

[15] C. D'Apice, R. Manzo, and B. Piccoli. Numerical Schemes for the Optimal Input Flow of a Supply Chain. *SIAM J. Numer. Anal.*, 51(5):2634–2650, 2013.

[16] Ciro D'Apice, Simone Göttlich, Michael Herty, and Benedetto Piccoli. *Modeling, simulation, and optimization of supply chains*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2010. A continuous approach.

[17] M. Giles and E. Sueli. Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality. *Acta Numerica*, 11:145–236, 2002.

[18] 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.*, 48:882–904, 2010.

[19] Michael Giles. Analysis of the accuracy of shock-capturing in the steady quasi 1d-euler equations. *Int. J. Comput. Fluid Dynam.*, 5:247–258, 1996.

[20] Edwige Godlewski and Pierre-Arnaud Raviart. The linearized stability of solutions of nonlinear hyperbolic systems of conservation laws. A general numerical approach. *Math. Comput. Simulation*, 50(1-4):77–95, 1999. Modelling '98 (Prague).

[21] Martin Gugat, Michaël Herty, Axel Klar, and Gunter Leugering. Conservation law constrained optimization based upon front-tracking. *M2AN Math. Model. Numer. Anal.*, 40(5):939–960 (2007), 2006.

[22] M. Herty, A. Kurganov, and D. Kurochkin. Numerical method for optimal control problems governed by nonlinear hyperbolic systems of pdes. *Communication in Mathematical Sciences*, 2014.

[23] F. James and M. Sepulveda. Convergence results for the flux identification in a scalar conservation law. *SIAM J. Control Optim.*, 37(3):869–891, 1999.

[24] Shi Jin and Zhou Ping Xin. The relaxation schemes for systems of conservation laws in arbitrary space dimensions. *Comm. Pure Appl. Math.*, 48(3):235–276, 1995.

[25] Zheng Liu and Adrian Sandu. On the properties of discrete adjoints of numerical methods for the advection equation. *Internat. J. Numer. Methods Fluids*, 56(7):769–803, 2008.

[26] Nonlinear theory of generalized functions. *Shift differentials of maps in BV spaces*, number 401, Boca Raton, 1999. Chapman & Hall/CRC, Boca Raton, FL.

[27] Niles A. Pierce and Michael Giles. Adjoint and defect error bounding and correction for functional estimates. *J. Comput. Phys.*, 200(2):769–794, 2004.

[28] S. Ulbrich. Adjoint-based derivative computations for the optimal control of discontinuous solutions of hyperbolic conservation laws. *System Control Letters*, 48:313–328, 2003.