

# TRACKING CONTROL OF 1D SCALAR CONSERVATION LAWS IN THE PRESENCE OF SHOCKS

RODRIGO LECAROS<sup>1,3</sup> AND ENRIQUE ZUAZUA<sup>1,2</sup>

ABSTRACT. We analyze a model tracking problem for a 1D scalar conservation law. It consists in optimizing the initial datum so to minimize a weighted distance to a given target during a given finite time horizon.

Even if the optimal control problem under consideration is of classical nature, the presence of shocks is an impediment for classical methods, based on linearization, to be directly applied.

We adapt the so-called alternating descent method that exploits the generalized linearization that takes into account both the sensitivity of the shock location and of the smooth components of solutions. This method was previously applied successfully on the inverse design problem and that of identifying the nonlinearity in the equation.

The efficiency of the method in comparison with more classical discrete methods is illustrated by several numerical experiments.

## CONTENTS

1. Introduction	2
2. Existence of Minimizers	3
3. The Discrete Minimization Problem	4
4. Sensitivity analysis: the continuous approach	6
4.1. Sensitivity without shocks	6
4.2. Sensitivity of the state in the presence of shocks	8
4.3. Sensitivity of the cost in the presence of shocks	9
5. The alternating descent method	12
6. Numerical approximation of the descent direction	13
6.1. The discrete approach	14
6.2. The alternating descent method	15
7. Numerical experiments	16
7.1. Experiment 1	17
7.2. Experiment 2	17
8. Conclusions and perspectives	20
References	21

---

Partially supported by Grants MTM2008-03541 and MTM2011-29306 of MICINN Spain, Project PI2010-04 of the Basque Government, ERC Advanced Grant FP7-246775 NUMERIWAVES and ESF Research Networking Programme OPTPDE.

This work was finished while the second author was visiting the Laboratoire Jacques Louis Lions with the support of the Paris City Hall “Research in Paris” program.

## 1. INTRODUCTION

There is an extensive literature on the control, optimization and inverse design of partial differential equations. But, when dealing with nonlinear models, most often, the analysis requires to linearize the system under consideration. This is why most of the existing results do not apply to hyperbolic conservation laws since the shock discontinuities of solutions are an impediment to linearize the system under consideration in a classical manner.

This paper is devoted to analyze this issue for 1D scalar conservation laws. To complement previous related works by our group we consider the tracking problem in which the goal is to optimally choose the initial datum so that the solution is as close as possible to a prescribed trajectory. Our previous works are devoted to the problem of inverse design in which the goal is to determine the initial datum so that the solution achieves a given target at the final time (see, for instance, [8], [9]) or the nonlinearity entering in the conservation law [10]).

To be more precise, given a finite time  $T > 0$ , a target function  $u^d \in L^2(\mathbb{R} \times (0, T))$ , and a positive weight function  $\rho \in L^\infty(\mathbb{R} \times (0, T))$  with compact support in  $\mathbb{R} \times (0, T)$ , we consider the functional cost to be minimized  $J$ , over a suitable class of initial data  $\mathcal{U}_{ad}$ , defined by

$$J(u^0) = \frac{1}{2} \int_0^T \int_{\mathbb{R}} \rho(x, t) |u(x, t) - u^d(x, t)|^2 dx dt, \quad (1.1)$$

where  $u : \mathbb{R}_x \times \mathbb{R}_t \rightarrow \mathbb{R}$  is the unique entropy solution of the scalar conservation law

$$\partial_t u + \partial_x(f(u)) = 0, \text{ in } \mathbb{R} \times (0, T); \quad u(x, 0) = u^0(x), \quad x \in \mathbb{R}. \quad (1.2)$$

Thus, the problem under consideration reads: To find  $u^{0, min} \in \mathcal{U}_{ad}$  such that

$$J(u^{0, min}) = \min_{u^0 \in \mathcal{U}_{ad}} J(u^0). \quad (1.3)$$

Here the flux  $f : \mathbb{R} \rightarrow \mathbb{R}$  is assumed to be smooth:  $f \in C^1(\mathbb{R}, \mathbb{R})$ . The initial datum  $u^0$  will be assumed to belong to a suitable admissible class  $\mathcal{U}_{ad}$  to ensure the existence of a minimizer. As a preliminary fact we remind that for  $u^0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}) \cap BV(\mathbb{R})$ , there exists a unique entropy solution in the sense of Kruřkov (see [15]) in the class  $C^0([0, T]; L^1(\mathbb{R})) \cap L^\infty(\mathbb{R} \times [0, T]) \cap L^\infty([0, T]; BV(\mathbb{R}))$ .

As we will see, the existence of minimizers can be established under some natural assumptions on the class of admissible data  $\mathcal{U}_{ad}$ , using the well-posedness and compactness properties of solutions of the conservation law (1.2). The uniqueness of the minimizers is false, in general, due, in particular, to the possible presence of discontinuities in the solutions of (1.2).

One of our main goals in this paper is to compare how the discrete approach and the alternating descent methods perform in this new prototypical optimization problem. This alternating descent method was introduced and developed in the previous works by our group to deal with and exploit the sensitivity of shocks. In other words, the solutions of the conservation law are seen as a multi-physics object integrated by both the solution itself but also by the geometric location of the shock. Perturbing the initial datum of the conservation law produces perturbations on both components. Accordingly it is natural to

analyze and employ both sensitivities to develop descent algorithms leading to an efficient computation of the minimizer. This is the basis of the alternating descent method ([8]).

As we shall see, in agreement with the results achieved in previously considered examples, the alternating descent method performs better than the classical discrete one which consists simply in applying a classical gradient descent algorithm to a discrete version of the functional and the conservation law.

This paper is limited to the one dimensional case but the alternating descent method can also be extended to the multi-dimensional frame, although this requires a much more careful geometric treatment of shocks since, for instance in  $2 - d$ , these are curves evolving in time whose location has to be carefully determined and their motion and perturbation handled carefully to avoid numerical instabilities (see [16]).

There are other possible methods that could be employed to deal with these problems. In the present  $1 - d$  setting, for instance, whenever the flux is convex, one could use the techniques developed in [1] based on the Lax-Oleinik representation formula of solutions. One could also employ the nudging method developed by J. Blum and coworkers ([2], [3]). It would be interesting to compare in a systematic manner the results obtained in this article with those that could be achieved by these other methods.

The rest of this paper is organized as follows. In Section 2 we formulate the problem under consideration more precisely and prove the existence of minimizers.

In Section 3 we introduce the discrete approximation of the continuous optimal control problem. We prove the existence of minimizers for the discrete problem and their  $\Gamma$ -convergence towards the continuous ones as the mesh-size parameters tend to zero.

As we shall see, purely discrete approaches based on the minimization of the resulting discrete functionals by descent algorithms lead to very slow iterative processes. We thus need to introduce an alternated descent algorithm that takes into account the possible presence of shock discontinuities in solutions. For doing this the first step is to develop a careful sensitivity analysis. This is done in Section 4.

In Section 5 we present the alternating descent method which combines the advantages of both the discrete approach and the sensitivity analysis in the presence of shocks.

In Section 6 we explain how to implement both descent algorithms: The discrete approach consists mainly in applying a descent algorithm to the discrete version  $J_\Delta$  of the functional  $J$ , while the alternating descent method, by the contrary, is a continuous method based on the analysis of the previous section on the sensitivity of shocks.

In Section 7 we present some numerical experiments illustrating the overall efficiency of the alternating descent method. We conclude discussing some possible extensions of the results and methods presented in the paper.

## 2. EXISTENCE OF MINIMIZERS

In this section we prove that, under certain conditions on the set of admissible initial data  $\mathcal{U}_{ad}$ , there exists at least one minimizer of the functional  $J$ , given in (1.1). To do this, we consider the class of admissible initial data  $\mathcal{U}_{ad}$  as:

$$\mathcal{U}_{ad} = \{u^0 \in L^\infty(\mathbb{R}) \cap BV(\mathbb{R}), \text{supp}(u^0) \subset K, \|u^0\|_{L^\infty} + TV(u^0) \leq C\}, \quad (2.1)$$

where  $K \subset \mathbb{R}$  is a given compact set and  $C > 0$  a given constant. Note, however, that the same theoretical results and descent strategies we shall develop here can be applied to a much wider class of admissible sets.

**Theorem 2.1.** *Assume that  $u^d \in L^2(\mathbb{R} \times (0, T))$ , let  $\mathcal{U}_{ad}$  be defined in (2.1) and  $f$  be a  $C^1$  function. Then the minimization problem,*

$$\min_{u^0 \in \mathcal{U}_{ab}} J(u^0), \quad (2.2)$$

*has at least one minimizer  $u^{0,min} \in \mathcal{U}_{ad}$ . Moreover, uniqueness is false in general.*

The proof follows the classical strategy of the Direct Method of the Calculus of Variations. We refer the reader to [8] for a similar proof in the case where the functional  $J$  is replaced by the one in which the distance to a given target at the final time  $t = T$  is minimized. Note that, in particular, for the class  $\mathcal{U}_{ad}$  of admissible initial data considered, solutions enjoy uniform  $BV$  bounds allowing to prove the needed compactness properties to pass to the limit. In the case of convex fluxes, using the well-known one-sided Lipschitz condition, the class of admissible initial data can be further extended.

We also observe that, as indicated in [8], the uniqueness of the minimizer is in general false for this type of optimization problems.

### 3. THE DISCRETE MINIMIZATION PROBLEM

The purpose of this section is to show that discrete minimizers obtained by a numerical approximation of (1.1) and (1.2), converge to a minimizer of the continuous problem as the mesh-size tends to zero. This justifies the usual engineering practice of replacing the continuous functional and model by discrete ones to compute an approximation of the continuous minimizer.

Let us introduce a mesh in  $\mathbb{R} \times [0, T]$  given by  $(x_i, t^n) = (i\Delta x, n\Delta t)$  ( $i = -\infty, \dots, \infty$ ;  $n = 0, \dots, N + 1$ , so that  $(N + 1)\Delta t = T$ ), and let  $u_{i,j}^n$  be a numerical approximation of  $u(x_i, t^n)$  obtained as solution of a suitable discretization of the equation (1.2).

Let us consider the following approximation of the functional  $J$  in (1.1):

$$J_\Delta(u_\Delta^0) = \frac{\Delta x \Delta t}{2} \sum_{n=0}^N \sum_{i=-\infty}^{\infty} \rho_i^n (u_i^n - (u^d)_i^n)^2, \quad (3.1)$$

where  $u_\Delta^0 = \{u_i^0\}$  is the discrete initial datum and  $u_\Delta^d = \{(u^d)_i^n\} = \Pi_\Delta u^d$  is the discretization of the target  $u^d$  at  $(x_i, t^n)$ ,  $\Pi_\Delta$  being a discretization operator. A common choice consists in taking

$$\Pi_\Delta u^d = (u^d)_i^n = \frac{1}{\Delta x \Delta t} \int_{x_{i-1/2}}^{x_{i+1/2}} \int_{t^{n-1/2}}^{t^{n+1/2}} u^d(x, t) dx dt, \quad (3.2)$$

where  $x_{i\pm 1/2} = x_i \pm \Delta x/2$  and  $t^{n\pm 1/2} = t^n \pm \Delta t/2$ .

Moreover, we introduce an approximation of the class of admissible initial data  $\mathcal{U}_{ad}$  denoted by  $\mathcal{U}_{ad,\Delta}$  and constituted by sequences  $\varphi_\Delta = \{\varphi_i\}_{i \in \mathbb{Z}}$  for which the associated

piecewise constant interpolation function, that we still denote by  $\varphi_\Delta$ , defined by

$$\varphi_\Delta(x) = \varphi_i, \quad x \in (x_{i-1/2}, x_{i+1/2}),$$

satisfies  $\varphi_\Delta \in \mathcal{U}_{ad}$ . Obviously,  $\mathcal{U}_{ad,\Delta}$  coincides with the class of discrete vectors with support on those indices  $i$  such that  $x_i \in K$  and for which the discrete  $L^\infty$  and  $TV$ -norms are bounded above by the same constant  $C$ .

Let us consider  $S_\Delta : l^1(\mathbb{Z}) \rightarrow l^1(\mathbb{Z})$ , an explicit numerical scheme for (1.2), where

$$u_\Delta^n = S_\Delta^n u_\Delta^0, \quad (3.3)$$

is the approximation of the entropy solution  $u(\cdot, t) = S(t)u^0$  of (1.2), i.e.  $u_\Delta^n \simeq S(t)u^0$ , with  $t = n\Delta t$ . Here  $S : L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}) \cap BV(\mathbb{R}) \rightarrow L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}) \cap BV(\mathbb{R})$ , is the semigroup (solution operator)

$$S : u_0 \rightarrow u = Su_0,$$

which associates to the initial condition  $u_0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}) \cap BV(\mathbb{R})$  the entropy solution  $u$  of (1.2).

For each  $\Delta = \Delta t$  (with  $\lambda = \Delta t/\Delta x$  fixed, typically given by the corresponding CFL-condition for explicit schemes), it is easy to see that the discrete analogue of Theorem 2.1 holds. In fact this is automatic in the present setting since  $\mathcal{U}_{ad,\Delta}$  only involves a finite number of mesh-points. But passing to the limit as  $\Delta \rightarrow 0$  requires a more careful treatment. In fact, for that to be done, one needs to assume that the scheme under consideration, (3.3), is a contracting map in  $l^1(\mathbb{Z})$ .

Thus, we consider the following discrete minimization problem: Find  $u_\Delta^{0,min}$  such that

$$J_\Delta(u_\Delta^{0,min}) = \min_{u_\Delta^0 \in \mathcal{U}_{ad,\Delta}} J_\Delta(u_\Delta^0). \quad (3.4)$$

The following holds

**Theorem 3.1.** *Assume that  $u_\Delta^n$  is obtained by a numerical scheme (3.3), which satisfies the following:*

- For a given  $u^0 \in \mathcal{U}_{ad}$ ,  $u_\Delta^n = S_\Delta^n \Pi_\Delta u^0$  converges to  $u(x, t)$ , the entropy solution of (1.2). More precisely, if  $n\Delta t = t$  and  $T_0 > 0$  is any given number,

$$\max_{0 \leq t \leq T_0} \|u_\Delta^n - u(\cdot, t)\|_{L^1(\mathbb{R})} \rightarrow 0, \quad \text{as } \Delta \rightarrow 0. \quad (3.5)$$

- The map  $S_\Delta$  is  $L^\infty$ -stable, i.e.

$$\|S_\Delta u_\Delta^0\|_{L^\infty} \leq \|u_\Delta^0\|_{L^\infty}. \quad (3.6)$$

- The map  $S_\Delta$  is a contracting map in  $l^1(\mathbb{Z})$  i.e.

$$\|S_\Delta u_\Delta^0 - S_\Delta v_\Delta^0\|_{L^1} \leq \|u_\Delta^0 - v_\Delta^0\|_{L^1}. \quad (3.7)$$

Then:

- For all  $\Delta$ , the discrete minimization problem (3.4) has at least one solution  $u_\Delta^{0,min} \in \mathcal{U}_{ad,\Delta}$ .
- Any accumulation point of  $u_\Delta^{0,min}$  with respect to the weak-\* topology in  $L^\infty$ , as  $\Delta \rightarrow 0$ , is a minimizer of the continuous problem (2.2).

The proof of this result can be developed as in [8]. We also refer to [16] for an extension to the multi-dimensional case.

This convergence result applies for 3-point conservative numerical approximation schemes, where  $S_\Delta$  is given by

$$(S_\Delta v_\Delta)_i = v_i - \frac{\Delta t}{\Delta x} (g(v_{i+1}, v_i) - g(v_i, v_{i-1})), \quad (3.8)$$

and  $g$  is the numerical flux. This scheme is consistent with the corresponding equation (1.2) when  $g(u, u) = f(u)$ .

When the discrete semigroup  $S_\Delta(u, v, w) = v - \lambda(g(u, v) - g(v, w))$ , with  $\lambda = \Delta t / \Delta x$  is monotonic increasing with respect to each argument, the scheme is also monotone. This ensures the convergence to the weak entropy solutions of the continuous conservation law, as the discretization parameters tend to zero, under a suitable CFL condition (see Ref. [12], Chap. 3, Th. 4.2).

All this analysis and results apply to the classical Godunov, Lax-Friedrichs and Engquist-Osher schemes, the corresponding numerical flux being:

$$g^G(u, v) = \begin{cases} \min_{w \in [u, v]} f(w), & \text{if } u \leq v, \\ \max_{w \in [u, v]} f(w), & \text{if } u \geq v, \end{cases} \quad (3.9)$$

$$g^{LF}(u, v) = \frac{(f(u) + f(v))}{2} - \frac{(v - u)}{2\lambda^x}, \quad (3.10)$$

$$g^{EO}(u, v) = \frac{f(u) + f(v) - \int_u^v |f'(\tau)| d\tau}{2}. \quad (3.11)$$

See Chapter 3 in [12] for more details.

These 1D methods satisfy the conditions of Theorem 3.1.

#### 4. SENSITIVITY ANALYSIS: THE CONTINUOUS APPROACH

We divide this section in three subsections. Specifically, in the first one we consider the case where the solution  $u$  of (1.2) has no shocks. In the second and third subsections we analyze the sensitivity of the solution and the functional in the presence of a single shock.

**4.1. Sensitivity without shocks.** In this subsection we give an expression for the sensitivity of the functional  $J$  with respect to the initial datum based on a classical adjoint calculus for smooth solutions. First we present a formal calculus and then we show how to justify it when dealing with a classical smooth solution for (1.2).

Let  $C_0^1(\mathbb{R})$  be the set of  $C^1$  functions with compact support and let  $u^0 \in C_0^1(\mathbb{R})$  be a given initial datum for which there exists a classical solution  $u(x, t)$  of (1.2) that can be extended to a classical solution in  $t \in [0, T + \tau]$  for some  $\tau > 0$ . Note that this imposes some restrictions on  $u^0$  other than being smooth.

Let  $\delta u^0 \in C_0^1(\mathbb{R})$  be any possible variation of the initial datum  $u_0$ . Due to the finite speed of propagation, this perturbation will only affect the solution in a bounded set of  $\mathbb{R} \times [0, T]$ . This simplifies the argument below that applies in a much more general setting provided solutions are smooth enough.

Then, for  $\varepsilon > 0$  sufficiently small, the solution  $u^\varepsilon(x, t)$  corresponding to the initial datum

$$u^{\varepsilon,0}(x) = u^0(x) + \varepsilon \delta u^0(x), \quad (4.1)$$

is also a classical solution in  $(x, t) \in \mathbb{R} \times (0, T)$  and  $u^\varepsilon \in C^1(\mathbb{R} \times [0, T])$  can be written as

$$u^\varepsilon = u + \varepsilon \delta u + o(\varepsilon), \text{ with respect to the } C^1 \text{ topology,} \quad (4.2)$$

where  $\delta u$  is the solution of the linearized equation,

$$\partial_t \delta u + \partial_x (f'(u) \delta u) = 0, \text{ in } \mathbb{R} \times (0, T); \quad \delta u(x, 0) = \delta u^0(x), \quad x \in \mathbb{R}. \quad (4.3)$$

Let  $\delta J$  be the Gateaux derivative of  $J$  at  $u^0$  in the direction  $\delta u^0$ . We have,

$$\delta J(u^0)[\delta u^0] = \int_0^T \int_{\mathbb{R}} \rho(x, t) (u(x, t) - u^d(x, t)) \delta u(x, t) dx dt \quad (4.4)$$

where  $\delta u$  solves the linearized system above (4.3). Now, we introduce the adjoint system,

$$-\partial_t p - f'(u) \partial_x p = \rho(u - u^d), \text{ in } \mathbb{R} \times (0, T); \quad p(x, T) = 0, \quad x \in \mathbb{R}. \quad (4.5)$$

Multiplying the equations satisfied by  $\delta u$  by  $p$ , integrating by parts, and taking into account that  $p$  satisfies (4.5), we easily obtain

$$\int_0^T \int_{\mathbb{R}} \rho(x, t) (u(x, t) - u^d(x, t)) \delta u(x, t) dx dt = \int_{\mathbb{R}} p(x, 0) \delta u^0(x) dx. \quad (4.6)$$

Thus,  $\delta J$  in (4.4) can be written as,

$$\delta J(u^0)[\delta u^0] = \int_{\mathbb{R}} p(x, 0) \delta u^0(x) dx. \quad (4.7)$$

This expression provides an easy way to compute a descent direction for the continuous functional  $J$ , once we have computed the adjoint state. We just take:

$$\delta u^0(x) = -p(x, 0). \quad (4.8)$$

Under the assumptions above on  $u^0, u, \delta u$  and  $p$  can be obtained from their data  $u^0(x), \delta u^0(x)$  and  $\rho(u - u^d)$  by using the characteristic curves associated to (1.2). For the sake of completeness we briefly explain this below.

The characteristic curves associated to (1.2) are defined by

$$x'(t) = f'(u(x(t), t)), \quad t \in (0, T), \quad x(0) = x^0. \quad (4.9)$$

They are straight lines whose slopes depend on the initial data:

$$x(t) = x^0 + t f'(u^0(x^0)), \quad t \in (0, T).$$

As we are dealing with classical solutions,  $u$  is constant along such curves and, by assumption, two different characteristic curves do not meet each other in  $\mathbb{R} \times [0, T + \tau]$ . This allows to define  $u$  in  $\mathbb{R} \times [0, T + \tau]$  in a unique way from the initial data.

For  $\varepsilon > 0$  sufficiently small, the solution  $u^\varepsilon(x, t)$  corresponding to the initial datum (4.1) has similar characteristics to those of  $u$ . This allows guaranteeing that two different characteristic lines do not intersect for  $0 \leq t \leq T$  if  $\varepsilon > 0$  is small enough. Note that  $u^\varepsilon$  may possibly be discontinuous for  $t \in (T, T + \tau]$  if  $u^0$  generates a discontinuity at  $t = T + \tau$

but this is irrelevant for the analysis in  $[0, T]$  we are carrying out. Therefore  $u^\varepsilon(x, t)$  is also a classical solution in  $(x, t) \in \mathbb{R} \times [0, T]$  and it is easy to see that the solution  $u^\varepsilon$  can be written as (4.2) where  $\delta u$  satisfies (4.3).

System (4.3) can be solved again by the method of characteristics. In fact, as  $u$  is a regular function, the first equation in (4.3) can be written as

$$\partial_t \delta u + f'(u) \partial_x \delta u = -\partial_x(f'(u)) \delta u, \quad (4.10)$$

i.e.

$$\frac{d}{dt} \delta u(x(t), t) = -\partial_x(f'(u)) \delta u, \quad (4.11)$$

where  $x(t)$  are the characteristic curves defined by (4.9). Thus, the solution  $\delta u$  along a characteristic line can be obtained from  $\delta u^0$  by solving this differential equation, i.e.

$$\delta u(x(t), t) = \delta u^0(x^0) \exp\left(-\int_0^t \partial_x(f'(u))(x(s), s) ds\right).$$

Finally, the adjoint system (4.5) is also solved by characteristics, i.e.

$$-\frac{d}{dt} p(x(t), t) = \rho(x(t), t)(u(x(t), t) - u^d(x(t), t)).$$

This yields the steepest descent direction in (4.8) for the continuous functional:

$$p(x^0, 0) = u^0(x^0) \int_0^T \rho(x(s), s) ds - \int_0^T \rho(x(s), s) u^d(x(s), s) ds.$$

**Remark 4.1.** *Note that for classical solutions the Gateaux derivative of  $J$  at  $u^0$  is given by (4.7) and this provides an obvious descent direction for  $J$  at  $u^0$ , given by (4.8). However this fact is not very useful in practice since, even when we initialize the iterative descent algorithm with a smooth  $u^0$ , we cannot guarantee that the solution remains classical along the iterative process.*

**4.2. Sensitivity of the state in the presence of shocks.** Inspired in several results on the sensitivity of solutions of conservation laws in the presence of shocks in one-dimension (see [7, 4, 5, 6, 17, 13]), we focus on the particular case of solutions having a single shock. But the analysis can be extended to consider more general one-dimensional systems of conservation laws with a finite number of noninteracting shocks. We introduce the following hypothesis:

**Hypothesis 4.2.** *Assume that  $u(x, t)$  is a weak entropy solution of (1.2) with a discontinuity along a regular curve  $\Sigma = \{(\varphi(t), t), t \in (0, T)\}$ , which is Lipschitz continuous outside  $\Sigma$ . In particular, it satisfies the Rankine-Hugoniot condition on  $\Sigma$*

$$\varphi'(t)[u]_{\Sigma^t} = [f(u)]_{\Sigma^t}.$$

Here we have used the notation:  $[v]_{\Sigma^t} = \lim_{\varepsilon \searrow 0} v(\varphi(t) + \varepsilon, t) - v(\varphi(t) - \varepsilon, t)$ , for the jump at  $\Sigma^t = (\varphi(t), t)$  of any piecewise continuous function  $v$  with a discontinuity at  $\Sigma^t$ .

Note that  $\Sigma$  divides  $\mathbb{R} \times (0, T)$  into two parts:  $Q_-$  and  $Q_+$ , the sub-domains of  $\mathbb{R} \times (0, T)$  to the left and to the right of  $\Sigma$  respectively.

As we will see, in the presence of shocks, to deal correctly with optimal control and design problems, the state of the system needs to be viewed as constituted by the pair  $(u, \varphi)$  combining the solution of (1.2) and the shock location  $\varphi$ . This is relevant in the analysis of sensitivity of functions below and when applying descent algorithms.

We adopt the functional framework based on the generalized tangent vectors (see [7] and Definition 4.1 in [8]).

Let  $u^0$  be the initial datum, that we assume to be Lipschitz continuous to both sides of a single discontinuity located at  $x = \varphi^0$ , and consider a generalized tangent vector  $(\delta u^0, \delta \varphi^0) \in L^1(\mathbb{R}) \times \mathbb{R}$  for all  $0 \leq T$ . Let  $u^{0,\varepsilon}$  be a path which generates  $(\delta u^0, \delta \varphi^0)$ . For  $\varepsilon$  sufficiently small, the solution  $u^\varepsilon(\cdot, t)$  of (1.2) is Lipschitz continuous with a single discontinuity at  $x = \varphi^\varepsilon(t)$ , for all  $t \in [0, T]$ . Therefore,  $u^\varepsilon(\cdot, t)$  generates a generalized tangent vector  $(\delta u(\cdot, t), \delta \varphi(t)) \in L^1(\mathbb{R}) \times \mathbb{R}$ . Moreover, in [8] it is proved that it satisfies the following linearized system:

$$\partial_t \delta u + \partial_x (f'(u) \delta u) = 0, \quad \text{in } Q_- \cup Q_+ \quad (4.12)$$

$$\frac{d}{dt} ([u]_{\Sigma^t} \delta \varphi) = [f'(u) \delta u]_{\Sigma^t} - [\delta u]_{\Sigma^t} \frac{d}{dt} \varphi, \quad t \in (0, T) \quad (4.13)$$

$$\delta u(x, 0) = \delta u^0(x), \quad \{x < \varphi^0\} \cup \{x > \varphi^0\} \quad (4.14)$$

$$\delta \varphi(0) = \delta \varphi^0. \quad (4.15)$$

**4.3. Sensitivity of the cost in the presence of shocks.** In this section we study the sensitivity of the functional  $J$  with respect to variations associated with the generalized tangent vectors defined in the previous section. We first define an appropriate generalization of the Gateaux derivative of  $J$ .

**Definition 4.3.** *Let  $J : L^1(\mathbb{R}) \rightarrow \mathbb{R}$  be a functional and  $u^0 \in L^1(\mathbb{R})$  be Lipschitz continuous with a discontinuity in  $\Sigma^0$ , an initial datum for which the solution of (1.2) satisfies hypothesis (4.2).  $J$  is Gateaux differentiable at  $u^0$  in a generalized sense if for any generalized tangent vector  $(\delta u^0, \delta \varphi^0)$  and any family  $u^{0,\varepsilon}$  associated to  $(\delta u^0, \delta \varphi^0)$  the following limit exists,*

$$\delta J = \lim_{\varepsilon \rightarrow 0} \frac{J(u^{0,\varepsilon}) - J(u^0)}{\varepsilon},$$

*and it depends only on  $(u^0, \varphi^0)$  and  $(\delta u^0, \delta \varphi^0)$ , i.e. it does not depend on the particular family  $u^{0,\varepsilon}$  which generates  $(\delta u^0, \delta \varphi^0)$ . The limit is the generalized Gateaux derivative of  $J$  in the direction  $(\delta u^0, \delta \varphi^0)$ .*

The following result easily provides a characterization of the generalized Gateaux derivative of  $J$  in terms of the solution of the associated adjoint system (4.17), (4.18), (4.19), (4.20), (4.21) and (4.22).

**Proposition 4.4.** *The Gateaux derivative of  $J$  can be written as follows*

$$\delta J(u^0)[\delta u^0, \delta \varphi^0] = \int_{\mathbb{R}} p(x, 0) \delta u^0(x) dx - q(0) [u]_{\Sigma^0} \delta \varphi^0, \quad (4.16)$$

where the adjoint state pair  $(p, q)$  satisfies the system

$$-\partial_t p - f'(u)\partial_x p = \rho(u - u^d), \quad \text{in } Q_- \cup Q_+ \quad (4.17)$$

$$[p]_{\Sigma^t} = 0, \quad t \in (0, T) \quad (4.18)$$

$$q(t) = p(\varphi(t), t), \quad t \in (0, T) \quad (4.19)$$

$$-\frac{d}{dt}q = \frac{(1 + (\dot{\varphi})^2)^{1/2}[\rho(u - u^d)^2]_{\Sigma^t}}{2[u]_{\Sigma^t}}, \quad t \in (0, T) \quad (4.20)$$

$$p(x, T) = 0, \quad \{x < \varphi(T)\} \cup \{x > \varphi(T)\} \quad (4.21)$$

$$q(T) = 0. \quad (4.22)$$

Let us briefly comment the result of Proposition 4.4 before giving its proof.

System (4.17)-(4.22) has a unique solution. In fact, to solve the backward system (4.17)-(4.22) we first define the solution  $q$  on the shock  $\Sigma$  from the conditions for  $q$  (4.20) and (4.22). This determines the value of  $p$  along the shock. We then propagate this information, together with (4.17) and (4.21), to both sides of  $\Sigma$ , by characteristics (see Figure 1 where we illustrate this construction).

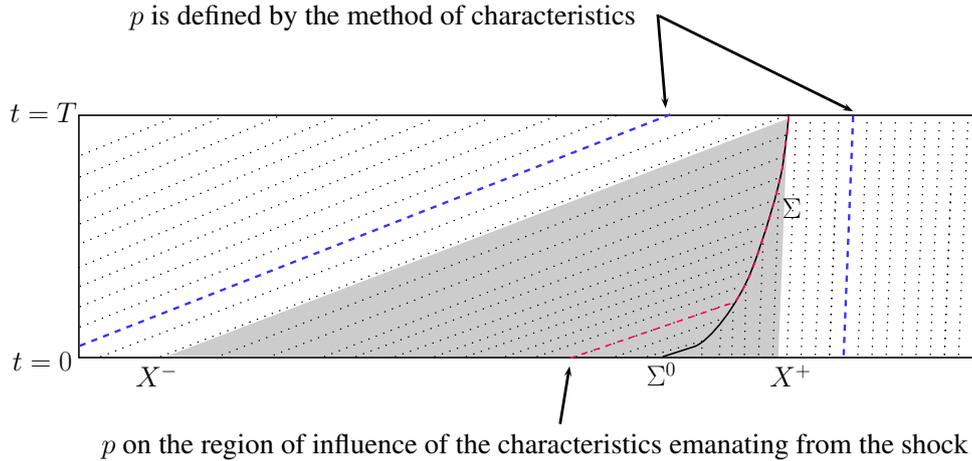


FIGURE 1. Characteristic lines entering on a shock and how they may be used to build the solution of the adjoint system both away from the shock and on its region of influence.

Formula (4.16) provides an obvious way to compute a first descent direction of  $J$  at  $u^0$ . We just take

$$(\delta u^0, \delta \varphi^0) = (-p(\cdot, 0), q(0)[u]_{\Sigma^0}). \quad (4.23)$$

Here, the value of  $\delta \varphi^0$  must be interpreted as the optimal infinitesimal displacement of the discontinuity of  $u^0$ .

In [8], when considering the inverse design problem, it was observed that the solution  $p$  of the corresponding adjoint system at  $t = 0$  was discontinuous, with two discontinuities,

one in each side of the original location of the discontinuity at  $\Sigma^0$ . This was a reason not to use this descent direction and for introducing the alternating descent method. In the present setting, however, the adjoint state  $p$  obtained is typically continuous. This is due to the fact that  $p$  at both side of the discontinuity is defined by the method of characteristics and that, on the region of influence of the characteristics emanating from the shock, the continuity is preserved by the fact that, on one hand,  $q = q(t)$  itself is continuous as the primitive of an integrable function and that the data for  $p$  and  $q$  at  $t = T$  are continuous too. Despite of this, as we shall see, the implementation of the alternating descent direction method is worth since it significantly improves the results obtained by the purely discrete approach.

We finish this section with the proof of Proposition 4.4.

*Proof.* (of Proposition 4.4) A straightforward computation shows that  $J$  is Gateaux differentiable in the generalized sense of Definition 4.3 and that the generalized Gateaux derivative of  $J$  in the direction of the generalized tangent vector  $(\delta u^0, \delta \varphi^0)$  is given by

$$\begin{aligned} \delta J(u^0)[\delta u^0, \delta \varphi^0] &= \int_0^T \int_{\{x > \varphi(t)\} \cup \{x < \varphi(t)\}} \rho(x, t)(u(x, t) - u^d(x, t)) \delta u(x, t) dx dt \\ &\quad - \int_{\Sigma} \left[ \rho \frac{(u - u^d)^2}{2} \right]_{\Sigma^t} \delta \varphi(t) d\Sigma(t), \end{aligned} \quad (4.24)$$

where the pair  $(\delta u, \delta \varphi)$  solves the linearized problem (4.12)-(4.15) with initial data  $(\delta u^0, \delta \varphi^0)$ .

Let us now introduce the adjoint system (4.17)-(4.22). Multiplying the equations of  $\delta u$  by  $p$ , and integrating we get

$$\begin{aligned} 0 &= \int_0^T \int_{\{x > \varphi(t)\} \cup \{x < \varphi(t)\}} (\partial_t \delta u + \partial_x (f'(u) \delta u)) p dx dt = - \int_0^T \int_{\{x > \varphi(t)\} \cup \{x < \varphi(t)\}} \delta u (\partial_t p + f'(u) \partial_x p) dx dt \\ &\quad + \int_{\{x > \varphi(T)\} \cup \{x < \varphi(T)\}} \delta u(x, T) p(x, T) dx - \int_{\{x > \varphi^0\} \cup \{x < \varphi^0\}} \delta u^0(x) p(x, 0) dx \\ &\quad - \int_{\Sigma} ([\delta u p]_{\Sigma^t} n_{\Sigma}^t + [f'(u) \delta u p]_{\Sigma^t} \cdot n_{\Sigma}^x) d\Sigma(t), \end{aligned} \quad (4.25)$$

where  $(n^x, n^t)$  are the Cartesian components of the normal vector to the curve  $\Sigma$ .

Therefore, since  $p$  satisfies the adjoint equation (4.17), (4.21), from (4.24) we obtain

$$\begin{aligned} \delta J(u^0)[\delta u^0, \delta \varphi^0] &= \int_{\{x > \varphi^0\} \cup \{x < \varphi^0\}} \delta u^0(x) p(x, 0) dx \\ &\quad + \int_{\Sigma} ([\delta u p]_{\Sigma^t} n_{\Sigma}^t + [f'(u) \delta u p]_{\Sigma^t} \cdot n_{\Sigma}^x) d\Sigma(t) - \int_{\Sigma} \left[ \rho \frac{(u - u^d)^2}{2} \right]_{\Sigma^t} \delta \varphi(t) d\Sigma(t). \end{aligned} \quad (4.26)$$

The last two terms in the right hand side of (4.26) will determine the conditions that  $p$  must satisfy on the shock.

Observe that for any functions  $f, g$  we have

$$[fg]_{\Sigma^t} = \bar{f}[g]_{\Sigma^t} + \bar{g}[f]_{\Sigma^t},$$

where  $\bar{g}$  represents the average of  $g$  to both sides of the shock  $\Sigma^t$ , i.e.

$$\bar{g}(t) = \frac{1}{2} \lim_{\varepsilon \searrow 0} (g(\varphi(t) + \varepsilon, t) + g(\varphi(t) - \varepsilon, t)), \quad \forall t \in (0, T).$$

Thus we have

$$\begin{aligned} \int_{\Sigma} ([\delta u p]_{\Sigma^t} n_{\Sigma}^t + [f'(u) \delta u p]_{\Sigma^t} \cdot n_{\Sigma}^x) d\Sigma &= \int_{\Sigma} [p]_{\Sigma^t} \left( \overline{\delta u n_{\Sigma}^t} + \overline{f'(u) \delta u} \cdot n_{\Sigma}^x \right) d\Sigma \\ &+ \int_{\Sigma} \bar{p} ([\delta u]_{\Sigma^t} n_{\Sigma}^t + [f'(u) \delta u]_{\Sigma^t} \cdot n_{\Sigma}^x) d\Sigma. \end{aligned} \quad (4.27)$$

Now, we note that the Cartesian components of the normal vector to  $\Sigma$  are given by

$$n^t = \frac{-\varphi'(t)}{\sqrt{1 + (\varphi'(t))^2}}, \quad n^x = \frac{1}{\sqrt{1 + (\varphi'(t))^2}},$$

and  $d\Sigma(t) = \sqrt{1 + (\varphi'(t))^2} dt$ . Using (4.13), (4.18), we have

$$\begin{aligned} \int_{\Sigma} ([\delta u p]_{\Sigma^t} n_{\Sigma}^t + [f'(u) \delta u p]_{\Sigma^t} n_{\Sigma}^x) d\Sigma &= \int_0^T \bar{p} (-[\delta u]_{\Sigma^t} \varphi'(t) + [f'(u) \delta u]_{\Sigma^t}) dt \\ &= \int_0^T \bar{p} \frac{d}{dt} ([u]_{\Sigma^t} \delta \varphi) dt. \end{aligned} \quad (4.28)$$

Therefore, replacing (4.19), (4.20), (4.22) and (4.28) in (4.26), we obtain (4.16). This concludes the proof.  $\square$

## 5. THE ALTERNATING DESCENT METHOD

Here we explain how the *alternating descent method* introduced in [8] can be adapted to the present optimal control problem (1.3).

First let us introduce some notation. We consider two points

$$X^- = \varphi(T) - T f'(u^-(\varphi(T), T)), \quad X^+ = \varphi(T) - T f'(u^+(x, T)), \quad (5.1)$$

$(p, q)$  the solutions of (4.17)-(4.22) and

$$p^{0,-} = \lim_{x \nearrow X^-} p(x, 0), \quad p^{0,+} = \lim_{x \searrow X^+} p(x, 0).$$

Now, we introduce the two classes of descent directions we shall use in our descent algorithm.

**First directions:** With this set of directions, we mainly modify the profile of  $u^0$ . We set the first directions  $d_1 = (\delta u^0, \delta \varphi^0)$ , given by:

$$\delta u^0 = \begin{cases} -p(x, 0), & x < X^-, \\ -p^{0,-}, & x \in (X^-, \varphi^0), \\ -p^{0,+}, & x \in (\varphi^0, X^+), \\ -p(x, 0), & x > X^+. \end{cases}$$

$$\delta \varphi^0 = \begin{cases} 0, & \text{if } \int_{X^-}^{X^+} p(x, 0) \delta u^0(x) dx \leq 0, \\ \frac{\int_{X^-}^{X^+} p(x, 0) \delta u^0(x) dx}{[u^0]_{\Sigma^0} q(0)}, & \text{if } \int_{X^-}^{X^+} p(x, 0) \delta u^0(x) dx > 0, \text{ and } q(0) \neq 0. \end{cases} \quad (5.2)$$

We note that, if  $q(0) = 0$  and  $\int_{X^-}^{X^+} p(x, 0) \delta u^0(x) dx > 0$ , these directions are not defined.

**Second directions:** They are aimed to move the shock without changing the profile of the solution to both sides. Then  $d_2 = (\delta u^0, \delta \varphi^0)$ , with:

$$\delta u^0 \equiv 0, \quad \delta \varphi^0(x) = [u^0]_{\Sigma^0} q(0). \quad (5.3)$$

We observe that  $d_1$  given by (5.2) satisfies

$$\begin{aligned} \delta J(u^0)[d_1] &= - \int_{x \notin [X^-, X^+]} |p(x, 0)|^2 dx \\ &\quad - p^{0,-} \int_{X^-}^{\varphi^0} p(x, 0) dx - p^{0,+} \int_{\varphi^0}^{X^+} p(x, 0) dx - [u^0]_{\Sigma^0} q(0) \delta \varphi^0 \leq 0. \end{aligned} \quad (5.4)$$

Note that in those cases where  $d_1$  is not defined, as indicated above, this descent direction is simply not employed in the descent algorithm.

And for  $d_2$  given by (5.3) we have

$$\delta J(u^0)[d_2] = -([u^0]_{\Sigma^0} q(0))^2 \leq 0. \quad (5.5)$$

Thus, the two classes of descent directions under consideration have three important properties:

- (1) They are both descent directions.
- (2) They allow to split the design of the profile and the shock location.
- (3) They are true generalized gradients and therefore keep the structure of the data without increasing its complexity.

## 6. NUMERICAL APPROXIMATION OF THE DESCENT DIRECTION

We have computed the gradient of the continuous functional  $J$  in several cases ( $u$  smooth or having shock discontinuities) but, in practice, one has to work with discrete versions of

the functional  $J$ . In this section we discuss two possible ways of searching discrete descent directions based either on the discrete or the continuous approaches and in the later on the alternating descent method.

The discrete approach consists mainly in applying directly a descent algorithm to the discrete version  $J_\Delta$  of the functional  $J$  by using its gradient. The alternating descent method, by the contrary, is a method inspired on the continuous analysis of the previous section in which the two main classes of descent directions that are, first, identified and later discretized.

Let us first discuss the discrete approach.

**6.1. The discrete approach.** Let us consider the approximation of the functional  $J$  by  $J_\Delta$  defined (1.1) and (3.1) respectively. We shall use the Engquist-Osher scheme which is a 3-point conservative numerical approximation scheme for (1.2). More explicitly we consider:

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x} (g(u_{i+1}^n, u_i^n) - g(u_i^n, u_{i-1}^n)), \quad i \in \mathbb{Z}, \quad n = 0, \dots, N, \quad (6.1)$$

where  $g$  is the numerical flux defined in (3.11).

The gradient of the discrete functional  $J_\Delta$  requires computing one derivative of  $J_\Delta$  with respect to each node of the mesh. This can be done in a cheaper way using the discrete adjoint state. We illustrate it for the Engquist-Osher numerical scheme. However, as the discrete functionals  $J_\Delta$  are not necessarily convex the gradient methods could possibly provide sequences that do not converge to a global minimizer of  $J_\Delta$ . But this drawback and difficulty appears in most applications of descent methods in optimal design and control problems.

As we will see, in the present context, the approximations obtained by discrete gradient methods are satisfactory, although convergence is slow due to unnecessary oscillations that the descent method introduces.

The gradient of  $J_\Delta$ , rigorously speaking, requires the linearization of the numerical scheme (6.1) used to approximate the equation (1.2). Then the linearization corresponds to

$$\begin{aligned} \delta u_i^{n+1} &= \delta u_i^n - \frac{\Delta t}{\Delta x} (\partial_a g(u_{i+1}^n, u_i^n) \delta u_{i+1}^n - \partial_b g(u_i^n, u_{i-1}^n) \delta u_{i-1}^n) \\ &\quad - \frac{\Delta t}{\Delta x} (\partial_b g(u_{i+1}^n, u_i^n) - \partial_a g(u_i^n, u_{i-1}^n)) \delta u_i^n, \\ & \quad i \in \mathbb{Z}, \quad n = 0, \dots, N. \end{aligned} \quad (6.2)$$

In view of this, the discrete adjoint system of (6.2) can also be written for the differentiable flux functions:

$$\begin{aligned} p_i^{N+1} &= 0, \quad i \in \mathbb{Z}, \\ p_i^n &= p_i^{n+1} + \frac{\Delta t}{\Delta x} \partial_b g(u_{i+1}^n, u_i^n) (p_{i+1}^{n+1} - p_i^{n+1}) + \frac{\Delta t}{\Delta x} \partial_a g(u_i^n, u_{i-1}^n) (p_i^{n+1} - p_{i-1}^{n+1}) + F_i^n, \\ & \quad i \in \mathbb{Z}, \quad n = 0, \dots, N, \end{aligned} \quad (6.3)$$

where

$$F_i^n = \Delta t \rho_i^n (u_i^n - (u^d)_i^n), \quad i \in \mathbb{Z}, \quad n = 0, \dots, N. \quad (6.4)$$

In fact, when multiplying the equations in (6.2) by  $p_i^{n+1}$  and adding in  $i \in \mathbb{Z}$  and  $n = 0, \dots, N$ , the following identity is easily obtained,

$$\Delta x \sum_{i \in \mathbb{Z}} p_i^0 \delta u_i^0 = \Delta x \sum_{n=0}^N \sum_{i \in \mathbb{Z}} F_i^n \delta u_i^0. \quad (6.5)$$

This is the discrete version of formula (4.6) which allows us to simplify the derivative of the discrete cost functional.

Thus, for any variation  $\delta u_\Delta^0$ , the Gateaux derivative of the cost functional defined in (3.1) is given by

$$\delta J_\Delta = \Delta x \Delta t \sum_{n=0}^N \sum_{i \in \mathbb{Z}} \rho_i^n (u_i^n - (u^d)_i^n) \delta u_i^n, \quad (6.6)$$

where  $\delta u_{i,j}^n$  solves the linearized system (6.2). If we consider  $p_i^n$  the solution of (6.3) with (6.4), we obtain that  $\delta J_\Delta$  in (6.6) can be written as,

$$\delta J_\Delta = \Delta x \sum_{i \in \mathbb{Z}} p_i^0 \delta u_i^0,$$

and this allows to obtain easily the steepest descent direction for  $J_\Delta$  by considering

$$\delta u_\Delta^0 = -p_\Delta^0. \quad (6.7)$$

**Remark 6.1.** *We observe that for the Enquist-Osher's flux (3.11), the system (6.3) is the upwind method for the continuous adjoint system. We do not address here the problem of the convergence of this adjoint scheme towards the solution of the continuous adjoint system. Of course, this is an easy matter when  $u$  is smooth but it is far from being trivial when  $u$  has shock discontinuities. Whether or not this discrete adjoint system, as  $\Delta \rightarrow 0$ , allows reconstructing the complete adjoint system, with the inner Dirichlet condition along the shock (4.17),(4.18),(4.19),(4.20),(4.21) and (4.22), constitutes an interesting problem for future research. We refer to [14] and [18] for preliminary work on this direction in one-dimension.*

**6.2. The alternating descent method.** Now we describe the implementation of the alternating descent method.

The main idea is to approximate a minimizer of  $J$  alternating between two directions of descent: First we perturb the initial datum  $u^0$  using the direction (5.2), which principally changes the profile of  $u^0$ . Second we move the shock curve without altering the profile of  $u^0$  at both sides of  $\Sigma$ , using the direction (5.3).

More precisely, for a given initialization  $u_\Delta^0$  and target function  $u_\Delta^d$ , we compute  $\Sigma_\Delta^0$ , the jump-point of  $u_\Delta^0, u_\Delta$  and  $p_\Delta^0$  as the solutions of (6.2), (6.3) respectively.

As numerical approximation of the adjoint state  $q(0)$  we take the value of  $p_\Delta^0$  over the point  $\Sigma_\Delta^0$ , i.e.

$$q_\Delta^0 = p_\Delta^0(\Sigma_\Delta^0) = p_i^0, \quad \Sigma_\Delta^0 \in (x_{i-1/2}, x_{i+1/2}).$$

The main advantage of this method introduced in [8] is that for an initial datum  $u^0$  with a single discontinuity, the descent directions are generalized tangent vectors, i.e. they introduce Lipschitz continuous variations of  $u^0$  at both sides of the discontinuity and a displacement of the shock position. In this way, the new datum obtained modifying the old one, in the direction of this generalized tangent vector, has again a single discontinuity. The method can be applied in a much more general context in which, for instance, the solution has various shocks since the method is able both to generate shocks and to destroy them, if any of these facts contributes to the decrease of the functional. This method is in some sense close to those employed in shape design in elasticity in which topological derivatives (that would correspond to controlling the location of the shock in our method) are combined with classical shape deformations (that would correspond to simply shaping the solution away from the shock in the present setting) [11].

As mentioned above, in the context of the tracking problem we are considering, the solutions of the adjoint system do not increase the number of shocks. Thus, a direct application of the continuous method, without employing this alternating variant, could make sense. Note that the purely discrete method is a limited version of the continuous one since one simply employs the descent direction indicated by  $p^0$  without paying attention to the value  $q^0$  corresponding to the shock sensitivity. However, the numerical experiments we have done with the full continuous method do not improve the results obtained by the discrete one since the definition of the location of the shocks is lost along the iterative process.

## 7. NUMERICAL EXPERIMENTS

In this section we present some numerical experiments which illustrate the results obtained in an optimization model problem with each one of the numerical methods described in the previous section.

We have chosen as computational domain the interval  $(-4, 4)$  and we have taken as boundary conditions in (6.1), at each time step  $t = t^n$ , the value of the initial data at the boundary. This can be justified if we assume that the initial datum  $u^0$  is constant in a sufficiently large inner neighborhood of the boundary  $x = \pm 4$  (which depends on the size of the  $L^\infty$ -norm of the data under consideration and the time horizon  $T$ ), due to the finite speed of propagation. A similar procedure is employed for the adjoint equation.

We underline once more that the solutions obtained with each method may correspond to global minima or local ones since the gradient algorithm does not distinguish them.

In the experiments we consider the Burgers' equation, i.e.  $f(z) = z^2/2$ ,

$$\partial_t u + \partial_x \left( \frac{u^2}{2} \right) = 0, \quad u(x, 0) = u^0(x). \quad (7.1)$$

The weight function  $\rho$ , under consideration in the experiments is given by

$$\rho(x, t) = \begin{cases} 1 & t \in (T/2, T) \\ 0 & \text{otherwise.} \end{cases}$$

And the time horizon  $T = 1$ .

To compare the efficiency of the different methods we consider a fixed  $\Delta x = 1/20$ ,  $\lambda = \Delta t/\Delta x = 2/3$  (which satisfies the CFL condition). We then analyze the number of iterations that each method needs to attain a prescribed value of the functional.

**7.1. Experiment 1.** We first consider a piecewise constant target profile  $u^d$  given by the solution of (7.1) with the initial condition  $(u^d)^0$  given by

$$(u^d)^0(x) = \begin{cases} 0.7 & x \in [-2, 1] \\ 0 & \text{otherwise.} \end{cases} \quad (7.2)$$

Note that, in this case, (7.2) yields a particular solution of the optimization problem and the minimum value of  $J$  vanishes.

We solve the optimization problem (1.3) with the above described different methods starting from the following initialization for  $u^0$ :

$$u^0(x) = \begin{cases} 0.5 & x \leq 0 \\ -0.1 & x > 0, \end{cases} \quad (7.3)$$

which also has a discontinuity but located on a different point.

In Figure 2 we plot  $\log(J)$  with respect to the number of iterations, for both, the purely discrete method and the alternating descent one. We see that the latter stabilizes in fewer iterations.

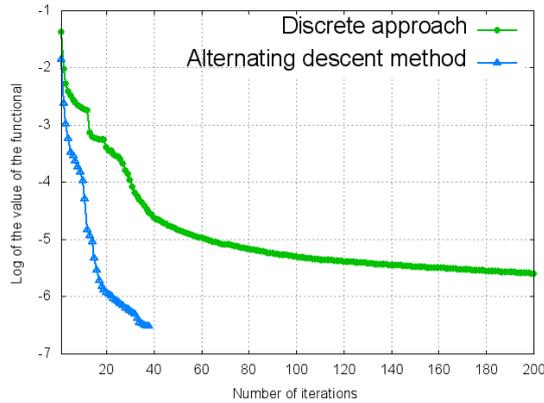


FIGURE 2. Experiment 1.  $\log(J)$  versus the number of iterations in the descent algorithm for the discrete and the alternating descent methods.

In Figures 3 and 4, we present the minimizers obtained by the methods above, and the associated solutions, Figures 5 and 6.

The initial datum  $u^0$  obtained by the alternating descent method (Figures 4) is a good approximation of (7.2). The solution given by the discrete approach (Figures 3) presents added spurious oscillations. Furthermore, the discrete method is much slower and does not achieve the same level of accuracy since the functional  $J_\Delta$  does not decrease so much.

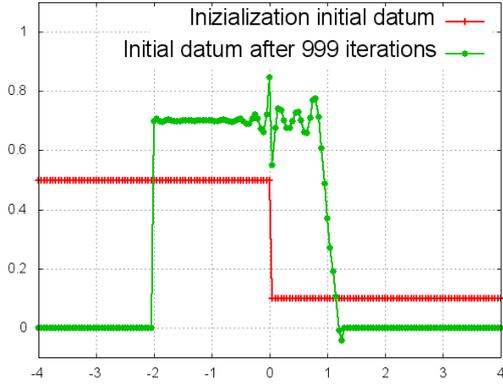


FIGURE 3. Experiment 1:  $u^0$ , discrete method, iteration  $k = 999$

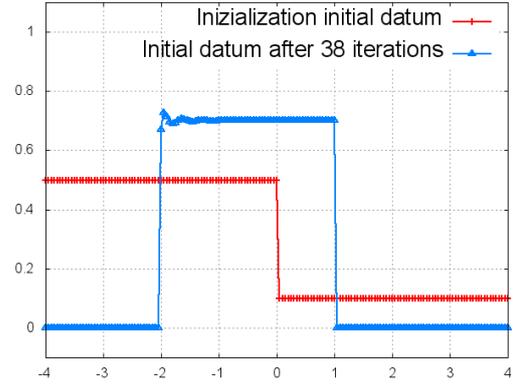


FIGURE 4. Experiment 1:  $u^0$ , alternating descent method, iteration  $k = 38$

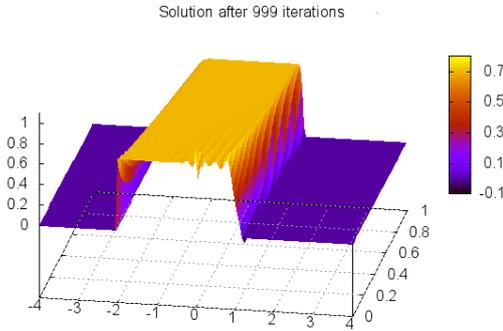


FIGURE 5. Experiment 1: Solution  $u(x,t)$ , discrete method, iteration  $k = 999$

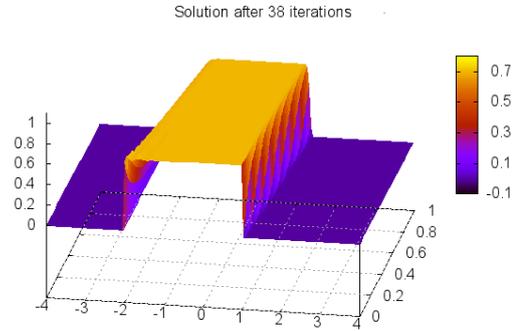


FIGURE 6. Experiment 1: Solution  $u(x,t)$ , alternating descent method, iteration  $k = 38$

**7.2. Experiment 2.** The previous experiment indicates that the alternating descent method performs significantly better. In order to show that this is a systematic fact, which arises independently of the initialization of the method, we consider the target  $u^d$  given by the solution of (7.1) with the initial condition  $(u^d)^0$  given by

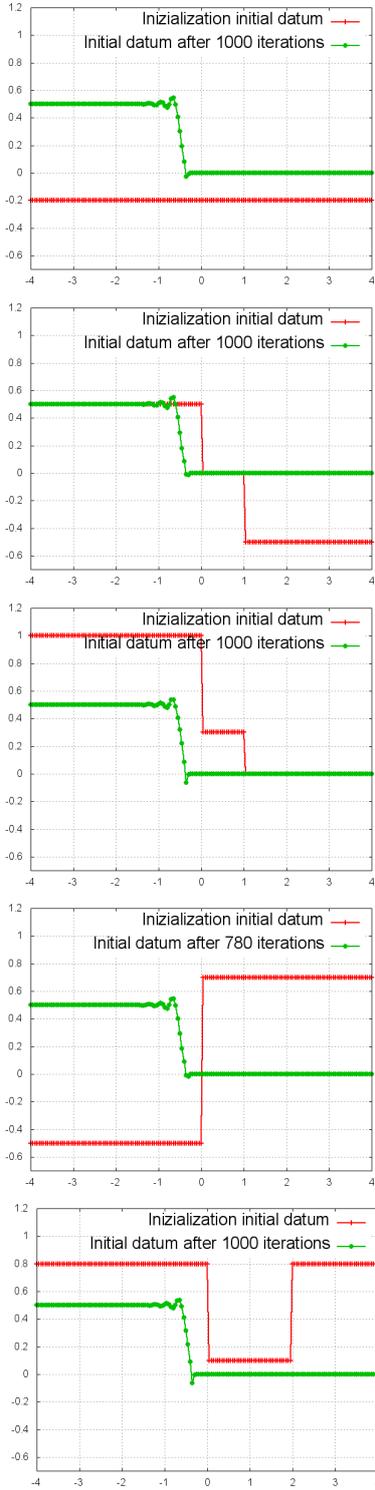
$$(u^d)^0(x) = \begin{cases} 0.5 & x \leq 0.5 \\ 0 & \text{otherwise,} \end{cases} \quad (7.4)$$

but this time we compare the performance of both methods starting from different initializations.

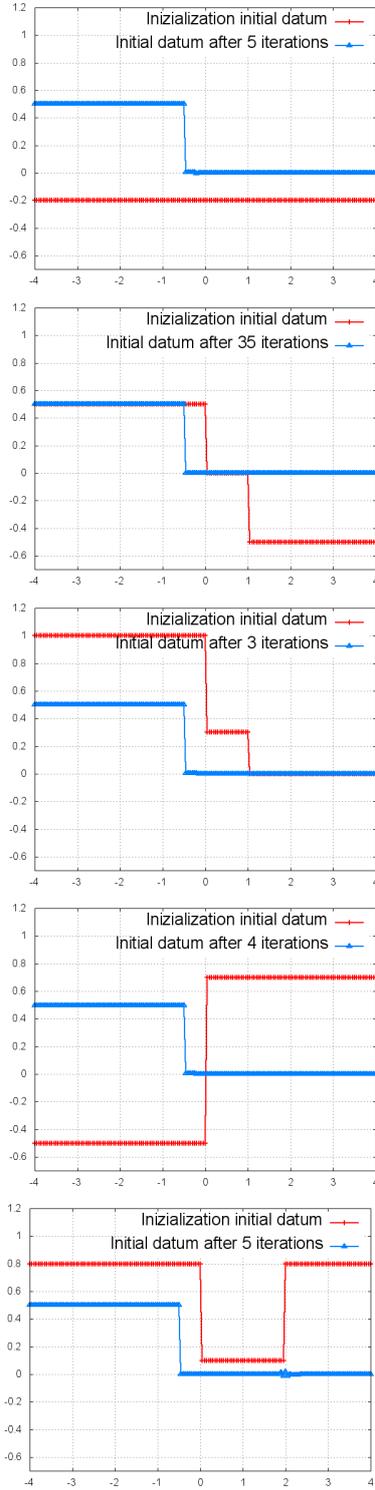
The obtained numerical results are presented in Figure 7.

We see that, regardless the initialization considered, the alternating descent method performs significantly better.

$u^0$  obtained by the discrete approach



$u^0$  obtained by the alternating descent method



The value of the functional

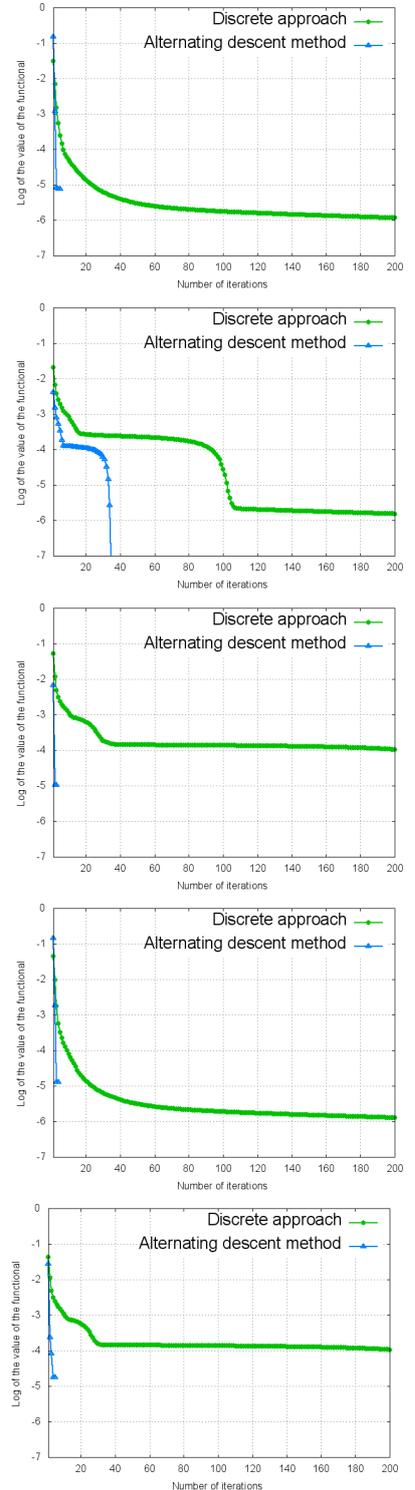


FIGURE 7. Experiment 2. We present a comparison of the results obtained with both methods starting out of five different initialization configurations. In the left column we exhibit the results obtained with the discrete method. In the second one those achieved by the alternating descent method. In the last one we plot the evolution of the functional with both methods.

We observe that in the five experiments the alternating descent method performs better ensuring the descent of the functional in much fewer iterations and yielding smoother, less oscillatory approximation of the minimizer.

Note also that the discrete method, rather than yielding discontinuous approximations of the minimizer as the alternating descent method does, it produces an initial datum with a Lipschitz front. Observe that these are two different configurations that can lead to the same evolution for the Burgers equation after some time, once the front develops the discontinuity. This is in agreement with the fact that the functional to be minimized is only active in the time-interval  $T/2 \leq t \leq T$ .

## 8. CONCLUSIONS AND PERSPECTIVES

In this paper we have adapted and presented the alternating descent method for a tracking problem for a  $1 - d$  scalar conservation law, the goal being to identify an optimal initial datum so that the solution gets as close as possible to a given trajectory. We have exhibited in a number of examples the better performance of the alternating descent method with respect to the classical discrete method. Thus, the paper extends previous works on other optimization problems such as the inverse design or the optimization of the flux function.

The developments in this paper raise a number of interesting problems and questions that will deserve further investigation. We summarize here some of them:

- The alternating descent method and the discrete one do not seem to yield the same minimizer. This should be further investigated in a more systematic manner in other experiments.
- The minimizer that the discrete method yields seems to replace the shock of the initial datum by a Lipschitz function which eventually develops the same dynamics within the time interval  $T/2 \leq t \leq T$  in which the functional we have chosen is active. This is an agreement with the behavior of these methods in the context of the classical problem of inverse design in which one aims to find the initial datum so that the solution at the final time takes a given value. It would be interesting to prove analytically that the two methods may lead to different minimizers in some circumstances.
- The experiments in this paper concern the case where the target  $u^d$  is exactly reachable. It would be interesting to explore the performance of both methods in the case where the target  $u^d$  is not a solution of the underlying dynamics.
- It would be interesting to compare the performance of the methods presented in the paper with the direct continuous method, without introducing the alternating strategy. Our preliminary numerical experiments indicate that the continuous method, without implementing the alternating strategy, does not improve the results that the purely discrete strategy yields.
- In [16] the problem of inverse design has been investigated adapting and extending the alternating descent method to the multi-dimensional case. It would be interesting to extend the analysis of this paper to the multidimensional case too.

- It would be worth to compare the results in this paper with those that could be achieved by the nudging method as in [2], [3].

### Acknowledgements

The authors would like to thank C. Castro, F. Palacios and A. Pozo for stimulating discussions.

### REFERENCES

- [1] A. Adimurthi, S. S. Ghoshal, and V. Gowda. Optimal controllability for scalar conservation laws with convex flux. 12, Jan. 2012. [3](#)
- [2] D. Auroux and J. Blum. Back and forth nudging algorithm for data assimilation problems. *Comptes Rendus Mathématique*, 340(12):873 – 878, 2005. [3](#), [20](#)
- [3] D. Auroux and M. Nodet. The back and forth nudging algorithm for data assimilation problems : theoretical results on transport equations. *ESAIM: Control, Optimisation and Calculus of Variations*, 18(2):318–342, 7 2012. [3](#), [20](#)
- [4] C. Bardos and O. Pironneau. A formalism for the differentiation of conservation laws. *C. R. Math. Acad. Sci. Paris*, 335(10):839–845, 2002. [8](#)
- [5] F. Bouchut and F. James. One-dimensional transport equations with discontinuous coefficients. *Non-linear Anal.*, 32(7):891–933, 1998. [8](#)
- [6] 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)*, volume 129 of *Internat. Ser. Numer. Math.*, pages 113–118. Birkhäuser, Basel, 1999. [8](#)
- [7] A. Bressan and A. Marson. A variational calculus for discontinuous solutions of systems of conservation laws. *Comm. Partial Differential Equations*, 20(9-10):1491–1552, 1995. [8](#), [9](#)
- [8] 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. *Mathematical Models and Methods in Applied Sciences*, 18(03):369–416, 2008. [2](#), [3](#), [4](#), [6](#), [9](#), [10](#), [12](#), [16](#)
- [9] C. Castro, F. Palacios, and E. Zuazua. Optimal control and vanishing viscosity for the burgers equation. In C. Constanda and M. Pérez, editors, *Integral Methods in Science and Engineering, Volume 2*, pages 65–90. Birkhäuser Boston, 2010. [2](#)
- [10] C. Castro and E. Zuazua. Flux identification for 1-d scalar conservation laws in the presence of shocks. *Math. Comp.*, 80(276):2025–2070, 2011. [2](#)
- [11] S. Garreau, P. Guillaume, and M. Masmoudi. The topological asymptotic for pde systems: The elasticity case. *SIAM J. Control Optim.*, 39(6):1756–1778, Dec. 2000. [16](#)
- [12] E. Godlewski and P. Raviart. *Hyperbolic systems of conservation laws, Mathématiques and Applications*. Ellipses, Paris, 1991. [6](#)
- [13] 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(1-4):77–95, 1999. Modelling '98 (Prague). [8](#)
- [14] L. Gosse and F. James. Numerical approximations of one-dimensional linear conservation equations with discontinuous coefficients. *Mathematics of Computation*, 69(231):pp. 987–1015, 2000. [15](#)
- [15] S. N. Kružkov. First order quasilinear equations with several independent variables. *Mat. Sb. (N.S.)*, 81 (123):228–255, 1970. [2](#)
- [16] R. Lecaros and E. Zuazua. Control of 2D scalar conservation laws in the presence of shocks. preprint, 2014. [3](#), [6](#), [20](#)
- [17] S. Ulbrich. Adjoint-based derivative computations for the optimal control of discontinuous solutions of hyperbolic conservation laws. *Systems Control Lett.*, 48(3-4):313–328, 2003. Optimization and control of distributed systems. [8](#)

- [18] X. Wen and S. Jin. Convergence of an immersed interface upwind scheme for linear advection equations with piecewise constant coefficients. I.  $L^1$ -error estimates. *J. Comput. Math.*, 26(1):1–22, 2008. [15](#)

RODRIGO LECAROS<sup>1,3</sup> AND ENRIQUE ZUAZUA<sup>1,2</sup>

<sup>1</sup>BCAM - BASQUE CENTER FOR APPLIED MATHEMATICS,  
MAZARREDO 14, E-48009, BILBAO, BASQUE COUNTRY, SPAIN

<sup>2</sup>IKERBASQUE - BASQUE FOUNDATION FOR SCIENCE,  
ALAMEDA URQUIJO 36-5, PLAZA BIZKAIA, 48011, BILBAO, BASQUE COUNTRY, SPAIN

<sup>3</sup>CMM - CENTRO DE MODELAMIENTO MATEMÁTICO. UNIVERSIDAD DE CHILE (UMI CNRS 2807),  
AVENIDA BLANCO ENCALADA 2120, CASILLA 170-3, CORREO 3, SANTIAGO, CHILE

*E-mail address:* [rlecaros@dim.uchile.cl](mailto:rlecaros@dim.uchile.cl), [zuazua@bcamath.org](mailto:zuazua@bcamath.org)

*URL:* [www.bcamath.org/zuazua/](http://www.bcamath.org/zuazua/)