

CONTROL OF 2D SCALAR CONSERVATION LAWS IN THE PRESENCE OF SHOCKS

RODRIGO LECAROS AND ENRIQUE ZUAZUA

ABSTRACT. We analyze a model optimal control problem for a 2D scalar conservation law—the so-called inverse design problem—with the goal being to identify the initial datum leading to a given final time configuration. The presence of shocks is an impediment for classical methods, based on linearization, to be directly applied. We develop an 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. A numerical implementation is proposed using splitting and finite differences. The descent method we propose is of alternating nature and combines variations taking account of the shock location and those that take care of the smooth components of the solution. The efficiency of the method is illustrated by numerical experiments.

1. INTRODUCTION

There is an extensive literature on the control and inverse design of partial differential equations. 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 analyzing this issue for 2D scalar conservation laws. To fix ideas, we consider the problem of inverse design aiming to identify the initial datum so that the solution, at the final time, is close to a given final target.

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

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

Received by the editor April 30, 2014 and, in revised form, October 3, 2014 and November 15, 2014.

2010 *Mathematics Subject Classification*. Primary 35L67, 49J20, 90C31, 49M30, 35L65.

The first author was partially supported by Basal-CMM project, PFB 03.

This work was done while the second author was visiting the CIMI (Centre International de Mathématiques et Informatique) of Toulouse (France) and the University of Erlangen-Nürnberg within the Humboldt Research Award program.

This work was supported by the Advanced Grants NUMERIWAVES/FP7-246775 of the European Research Council Executive Agency, FA9550-14-1-0214 of the EOARD-AFOSR, PI2010-04 and the BERC 2014-2017 program of the Basque Government, the MTM2011-29360-C02-00 and SEV-2013-0323 Grants of the MINECO.

where $S : L^1(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2) \cap BV(\mathbb{R}^2) \rightarrow L^1(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2) \cap BV(\mathbb{R}^2)$, is the semigroup

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

which associates to the initial condition $u_0 \in L^1(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2) \cap BV(\mathbb{R}^2)$ the unique entropy solution $u : \mathbb{R}_x^2 \times \mathbb{R}_t \rightarrow \mathbb{R}$ of the scalar conservation law

$$(1.2) \quad \partial_t u + \operatorname{div}_x f(u) = 0, \text{ in } \mathbb{R}^2 \times (0, T); \quad u(x, 0) = u^0(x), \quad x \in \mathbb{R}^2.$$

Here the flux $f : \mathbb{R} \rightarrow \mathbb{R}^2$ is assumed to be a smooth function $f \in C^1(\mathbb{R}, \mathbb{R}^2)$. Thus, the problem under consideration reads: To find $u^{0,\min} \in \mathcal{U}_{ad}$ such that

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

the initial datum u^0 will be assumed to belong to a suitable class. But as a preliminary fact we remind the reader that for $u^0 \in L^1(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2) \cap BV(\mathbb{R}^2)$, there exists an unique entropy solution in the sense of Kruřkov (see [28]) in the class $C^0([0, T]; L^1(\mathbb{R}^2)) \cap L^\infty(\mathbb{R}^2 \times [0, T])$.

The inverse design problem under consideration is one of the classical optimization problems which is often addressed in the context of optimal aerodynamic design (see, for example, [21]).

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).

Although this paper is devoted to this particular choice of J , that can also be handled by other methods, for instance, by backward resolution of the equation out of the final target, as most of our analysis and numerical algorithms can be adapted to many other functionals and control problems, requiring the implementation of descent algorithms, which is the main content of the present paper. In particular our methods could allow us to handle, for instance, the 2D version of the problems addressed in [27] and [14], where the control variable is the nonlinearity of the scalar conservation law. Our analysis can also be extended to higher space dimensions.

One of the classical approaches to effectively compute an accurate approximation of the minimizer is the so-called *continuous approach*, which consists in building a descent algorithm for the continuous optimization problem (1.3) to later discretize it. But this requires the computation of the Gâteaux derivative of J . However, when the constraint (1.2) involves a nonsmooth solution, this classical method does not apply. And this often occurs in the context of conservation laws where solutions may present shocks even for regular initial data $u^0 \in C^1$ (see, for instance, L. C. Evans [18, Subsection 3.4.1], for details). When this occurs, the formal linearization of the equation (1.2)

$$(1.4) \quad \partial_t \delta u + \operatorname{div}_x (f'(u) \delta u) = 0$$

is not rigorously justified.

This problem was addressed in [12] in one space dimension for the inviscid Burgers' equation

$$(1.5) \quad \partial_t u + \partial_x \left(\frac{u^2}{2} \right) = 0, \text{ in } \mathbb{R} \times (0, T); \quad u(x, 0) = u^0(x), \quad x \in \mathbb{R},$$

where the existence of minimizers was proved and an efficient descent algorithm based on a new optimization strategy was introduced, the so-called *alternating descent direction method*, which exploits the generalized gradient and linearization of the system.

Note that when minimizing the functional (1.1), we are looking for an initial datum u^0 so that the solution of the conservation law reaches, or gets as close as possible to, the final target u^d . Thus, the problem can be viewed as a controllability one, the control being the initial datum. The problem can be recast in terms of identifying the set of reachable states at the final time $t = T$ for the semigroup generated by the conservation law. The latter was addressed in [19] for parabolic problems and in [1] for 1D conservation laws. Much less is known in the context of hyperbolic conservation laws.

In 1D, assuming the flux function is strictly convex, a characterization of the set of attainable states is given in [3] (see [2] for systems).

Furthermore, in [1], using the Hopf and the Lax-Olenik formulas, a method to obtain a solution of this optimal control problem was derived by projecting the target u^d into the set of states satisfying the well-known one-sided Lipschitz condition. This leads to iterative algorithms similar to the back-and-forth one developed in [4].

In practical applications and in order to perform numerical computations and simulations, one has to replace the above continuous optimization problem and methods by discrete approximations. Then, it is natural to consider a discretization of system (1.2) and the functional J . If this is done in an appropriate way, the discrete optimization problem has minimizers, that are often taken, for small enough mesh-sizes, as approximations of the continuous minimizers. This is the so-called *discrete or direct approach*: “discretize first and then optimize”. There are however few results in the context of hyperbolic conservation laws proving rigorously the convergence of the discrete optimal controls towards the continuous ones as the mesh-size goes to zero.

One of the main results of this paper ensures the Γ -convergence property of a numerical approximation of the inverse design problem based on a numerical approximation scheme for (1.1) and (1.2).

When optimal solutions have shock discontinuities, the purely discrete approach based on minimizing a discrete version of J obtained by means of a discretization of the conservation law (1.2) produces highly oscillating minimizing sequences, and the effective descent rate is very weak. As a remedy we introduce the 2D version of the *alternating descent method*. It uses a 2D extension of generalized tangent vectors (introduced in [10] and [9] for 1D problems, for instance), which enables us to linearize (1.2) around discontinuous solutions. The alternating method distinguishes between and alternates the descent directions that move the shock and those that perturb the profile of the solution away from it, producing overall a very efficient and fast descent algorithm.

Thus, one of the main contributions of this paper consists in developing a careful sensitivity analysis of the continuous (1.2) and functional (1.1) equations, paying special attention to the possible presence of shock discontinuities in solutions. This extends to the 2D case the 1D analysis based on generalized tangent vectors [9, 10].

Once this is done, we then develop the computational version of the alternating descent strategy, which requires important further developments to cope with the

2D case, integrating tools, in particular, from image processing. We finally prove its efficiency by several numerical experiments.

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

In Section 3 we introduce the discrete approximation of the continuous inverse design problem and prove the existence of minimizers and their Γ -convergence in the sense that discrete inverse designs converge to 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 alternating 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 two descent algorithms. The discrete approach consists mainly in applying a descent algorithm to the discrete version J_Δ of the functional J . And the alternating descent method, in contrast, is a continuous method based on the analysis of the previous section.

Section 7 is devoted to presenting details of numerical implementation. In Section 8 we present some numerical experiments illustrating the overall efficiency of the method. We conclude with a discussion 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:

$$(2.1) \quad \mathcal{U}_{ad} = \{f \in L^\infty(\mathbb{R}^2) \cap BV(\mathbb{R}^2), \text{supp}(f) \subset K, \|f\|_{L^\infty} + TV(f) \leq C\},$$

where $K \subset \mathbb{R}^2$ is a given compact set and $C > 0$ is a given constant. Here $TV(f)$ represents the total variation of f in \mathbb{R}^2 .

Note, however, the same theoretical results and descent strategies that we shall develop here can be applied to a much wider class of admissible sets.

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

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

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

Proof. First, we prove the existence of a minimizer.

Let $\{u_n^0\}_{n \in \mathbb{N}} \subset \mathcal{U}_{ad}$ be a minimizing sequence of J .

In view of the definition of \mathcal{U}_{ad} , the minimizing sequence $\{u_n^0\}_{n \in \mathbb{N}}$ is necessarily bounded in BV . Thanks to the compactness of the embedding from BV_{loc} into L^1_{loc} , we deduce that up to the extraction of subsequences u_n^0 weakly* converges to u_*^0 in $BV(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$ but, simultaneously, $u_n^0 \rightarrow u_*^0$ strongly in $L^1_{loc}(\mathbb{R}^2)$. Actually, this strong convergence holds in $L^p_{loc}(\mathbb{R}^2)$ for all $1 \leq p < +\infty$.

We also have $u_*^0 \in \mathcal{U}_{ad}$. Let $u_n(x, t)$, and let $u_*(x, t)$ be the entropy solutions of (1.2) with initial data u_n^0 and u_*^0 , respectively. We claim that

$$(2.3) \quad u_n(\cdot, T) \rightarrow u_*(\cdot, T), \text{ strongly in } L^2(\mathbb{R}^2).$$

Assuming for the moment that the claim is true, we deduce that

$$\inf_{u^0 \in \mathcal{U}_{ad}} J(u^0) = \lim_{n \rightarrow \infty} J(u_n^0) = J(u_*^0),$$

and we conclude that u_*^0 is a minimizer of J .

Let us now prove (2.3). Taking the structure of \mathcal{U}_{ad} into account and using the maximum principle and the finite velocity of propagation that entropy solutions satisfy, it is easy to see that the support of all solutions for all $t \in [0, T]$ is uniformly included in the same compact set of \mathbb{R}^2 . The needed compactness property is then a consequence of the compactness of the embedding $BV_{loc} \subset L^1_{loc}$. Therefore, it is sufficient to prove convergence in L^2_{loc} . This is obtained from the L^1 -contraction property of the equation (we refer Kruřkov, Theorem 4.1 in [22]):

$$(2.4) \quad \|u_n(\cdot, t) - u_*(\cdot, t)\|_{L^1} \leq \|u_n^0 - u_*^0\|_{L^1}, \quad \forall t \in [0, T],$$

which guarantees a uniform convergence of u_n to u_* in L^1 . Using the uniform bound in L^∞ , we obtain (2.3).

This completes the proof of the existence of minimizers.

The uniqueness of the minimizer is in general false for this type of optimization problem. In fact, there are examples of target functions u^d for which there exist two different minimizers u_1^0 and u_2^0 such that the corresponding solutions u_j , $j = 1, 2$, satisfy $u_j(T) = u^d$, $j = 1, 2$, in such a way that the minimal value of J vanishes. We refer to [12] in the proof of Theorem 2.1 for an example in 1D which is easy to extend to 2D. Indeed, we can consider

$$(2.5) \quad u_1^0(x) = \begin{cases} 1 & x \in (-T, 0) \\ 0 & \text{otherwise,} \end{cases} \quad u_2^0 = \begin{cases} 1 & x \in (-T, -T/2) \\ \frac{1}{2} - \frac{x}{T} & x \in (-T/2, T/2) \\ 0 & \text{otherwise,} \end{cases}$$

$$(2.6) \quad u^d = \begin{cases} \frac{x}{T} + 1 & x \in (-T, 0) \\ 1 & x \in (0, T/2) \\ 0 & \text{otherwise,} \end{cases}$$

and we consider a 2D extension

$$(2.7) \quad \tilde{u}_i^0(x, y) = \begin{cases} u_i^0(x) & y \in [-1, 1] \\ 0 & \text{otherwise,} \end{cases} \quad \tilde{u}^d(x, y) = \begin{cases} u^d(x) & y \in [-1, 1] \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, using the flux $f(z) = (z^2/2, 0)$, we obtain that \tilde{u}_1^0 and \tilde{u}_2^0 are both different minimizers (see Figure 1).

Here and in what follows we shall often use the notation u^T for the value of the solution u of the conservation law under consideration at the final time $t = T$: $u^T(x) = u(x, T)$. □

Remark 2.2. The above proof is quite general and can be adapted to other optimization problems with different functionals and admissible sets, and in particular to the functional

$$(2.8) \quad J(u^0) = \frac{1}{2} \int_0^T \int_{\mathbb{R}^2} |u(x, t) - u^d(x, t)|^2 dx dt,$$

with $u^d \in L^2(\mathbb{R}^2 \times (0, T))$.

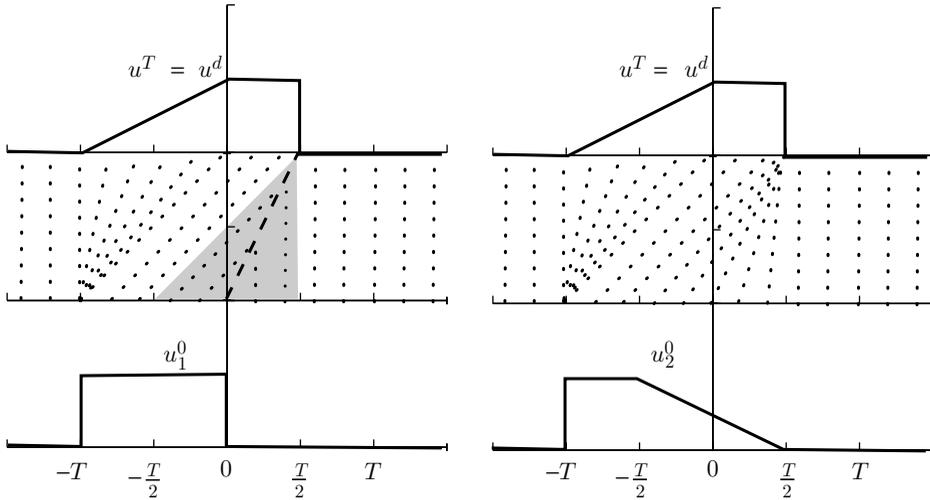


FIGURE 1. Plots of $u_1^0, u_2^0, u^d = u^T$ and the corresponding characteristic lines.

Remark 2.3. The nonuniqueness issue which is displayed cannot be fixed by further penalizing any of L^1, L^∞ or BV norms as shown by the explicit example on the bottom line of Figure 1.

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 $\mathbb{R}^2 \times [0, T]$ given by $(x_i, y_j, t^n) = (i\Delta x, j\Delta y, n\Delta t)$ ($i = -\infty, \dots, \infty; j = -\infty, \dots, \infty; n = 0, \dots, N + 1$, so that $(N + 1)\Delta t = T$), and let $u_{i,j}^n$ be a numerical approximation of $u(x_i, y_j, t^n)$ obtained as a solution of a suitable discretization of equation (1.2).

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

$$(3.1) \quad J_\Delta(u_\Delta^0) = \frac{\Delta x \Delta y}{2} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (u_{i,j}^{N+1} - u_{i,j}^d)^2,$$

where $u_\Delta^0 = \{u_{i,j}^0\}$ is the discrete initial datum and $u_\Delta^d = \{u_{i,j}^d\} = \Pi_\Delta u^d$ is the discretization of the target u^d at x_i, y_j , respectively, and Π_Δ is a discretization operator. A common choice consists in taking

$$(3.2) \quad \Pi_\Delta u^d = u_{i,j}^d = \frac{1}{\Delta x \Delta y} \int_{x_{i-1/2}}^{x_{i+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} u^d(x, y) dy dx,$$

where $x_{i\pm 1/2} = x_i \pm \Delta x/2$ and $y_{j\pm 1/2} = y_j \pm \Delta y/2$.

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 $\varphi_\Delta = \{\varphi_{i,j}\}_{i,j \in \mathbb{Z}}$ for which the

associated piecewise constant interpolation function, that we still denote by φ_Δ , defined by

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

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, j such that $x_i, y_j \in K$ and for which the discrete L^∞ -norm and TV are bounded above by the same constant C .

Let us consider $S_\Delta : l^1(\mathbb{Z}^2) \rightarrow l^1(\mathbb{Z}^2)$, an explicit numerical scheme for (1.2), where

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

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 $S : L^1(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2) \cap BV(\mathbb{R}^2) \rightarrow L^1(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2) \cap BV(\mathbb{R}^2)$, is the semigroup (solution operator of (1.2))

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

which associates to the initial condition $u_0 \in L^1(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2) \cap BV(\mathbb{R}^2)$ the entropy solution u of (1.2).

For each $\Delta = \Delta t$ (with $\lambda^x = \Delta t / \Delta x$ and $\lambda^y = \Delta t / \Delta y$ 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 contraction in $l^1(\mathbb{Z}^2)$ satisfying a discrete version of (2.4).

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

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

The following holds.

Theorem 3.1. *Assume that u_Δ^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,

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

- The map S_Δ is L^∞ -stable, i.e.

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

- The map S_Δ is a contraction in $l^1(\mathbb{Z}^2)$, i.e.

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

Then

- For all Δ , 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^∞ , as $\Delta \rightarrow 0$, is a minimizer of the continuous problem (2.2).

Proof of Theorem 3.1. The existence of discrete minimizers in the first statement of the theorem is obvious in this case since we are dealing with a finite-dimensional problem. Actually, at this point only the continuity property (3.7) of S_Δ is necessary.

The second statement is less trivial. And the other properties are required to guarantee the compactness of the numerical solutions as $\Delta \rightarrow 0$.

Let us assume that the following property holds: If $u_\Delta^0 \in \mathcal{U}_{ad,\Delta}$ satisfies $u_\Delta^0 \rightharpoonup u_*^0$ in $L^\infty(\mathbb{R}^2)$ with respect to the weak-* topology, then

$$(3.8) \quad S_\Delta^n u_\Delta^0 \rightarrow S(t)u_*^0, \quad \text{strongly in } L^2,$$

for all $t = n\Delta t$. This is a discrete version of (2.3).

We now follow a standard Γ -convergence argument. Using (3.8), we obtain the continuity property

$$(3.9) \quad J_\Delta(u_\Delta^0) \rightarrow J(u^0).$$

Now, let $\hat{u}^0 \in \mathcal{U}_{ad}$ be an accumulation point of $u_\Delta^{0,\min}$ with respect to the weak-* topology of L^∞ . To simplify the notation, we still denote by $u_\Delta^{0,\min}$ the subsequence for which $u_\Delta^{0,\min} \rightharpoonup \hat{u}^0$ is weakly-* in $L^\infty(\mathbb{R}^2)$, as $\Delta \rightarrow 0$. Let $v^0 \in \mathcal{U}_{ad}$ be any other function. We are going to prove that

$$(3.10) \quad J(\hat{u}^0) \leq J(v^0).$$

To do this, we construct a sequence $v_\Delta^0 \in \mathcal{U}_{ad,\Delta}$ such that $v_\Delta^0 \rightarrow v^0$, in $L^1(\mathbb{R}^2)$, as $\Delta \rightarrow 0$ (we can consider in particular the approximation $\Pi_\Delta v^0$ in (3.2)). Taking the continuity property (3.9) into account, we have

$$J(v^0) = \lim_{\Delta \rightarrow 0} J_\Delta(v_\Delta^0) \geq \lim_{\Delta \rightarrow 0} J_\Delta(u_\Delta^{0,\min}) = J(\hat{u}^0),$$

which proves (3.10).

Finally we need to prove (3.8). The class of initial data $\mathcal{U}_{ad,\Delta}$ we are considering guarantees uniform local BV bounds on the discrete solutions. This implies local compactness in L^1 , and using that the supports are in the same compact set, we obtain that $u_\Delta^0 \rightarrow u_*^0$ in $L^1(\mathbb{R}^2)$. Now, setting $t = n\Delta t$, we have

$$(3.11) \quad \|S_\Delta^n u_\Delta^0 - S(t)u_*^0\|_{L^1} \leq \|S_\Delta^n u_\Delta^0 - S_\Delta^n \Pi_\Delta u_*^0\|_{L^1} + \|S_\Delta^n \Pi_\Delta u_*^0 - S(t)u_*^0\|_{L^1}.$$

Using the contraction property (3.7) in (3.11) and the convergence of the scheme under consideration (3.5), we obtain

$$(3.12) \quad \max_{0 \leq t \leq T_0} \|S_\Delta^n u_\Delta^0 - S(t)u_*^0\|_{L^1} \rightarrow 0, \quad \text{as } \Delta \rightarrow 0.$$

We note that $\|S(t)u_*^0\|_{L^\infty} \leq \|u_*^0\|_{L^\infty}$ and using the L^∞ -stability property (3.6), we have

$$(3.13) \quad \max_{0 \leq t \leq T_0} \|S_\Delta^n u_\Delta^0 - S(t)u_*^0\|_{L^2}^2 \leq C \max_{0 \leq t \leq T_0} \|S_\Delta^n u_\Delta^0 - S(t)u_*^0\|_{L^1},$$

where $C = \max_{0 \leq t \leq T_0} \|S_\Delta^n u_\Delta^0 - S(t)u_*^0\|_{L^\infty(\mathbb{R})}$. Therefore, using (3.12) and (3.13), we obtain (3.8). \square

We need now to introduce a numerical approximation scheme for (1.2). This can be done directly by introducing a 2D finite difference scheme. But here we shall rather employ a dimensional splitting method.

Let us then introduce one-dimensional nonlinear difference numerical approximation schemes, H_Δ^x and H_Δ^y , where $v_\Delta^n = (H_\Delta^x)^n v_\Delta^0$ and $w_\Delta^n = (H_\Delta^y)^n w_\Delta^0$ are the numerical solutions at time $t = n\Delta t$ for the one dimensional problems

$$(3.14) \quad \begin{aligned} v_t + (f_1(v))_x &= 0, & w_t + (f_2(w))_y &= 0, \\ v(x, y, 0) &= v^0(x, y), & w(x, y, 0) &= w^0(x, y), \end{aligned}$$

where f_1 and f_2 are the two components of the flux $f(z) = (f_1(z), f_2(z))$. For example, we can consider 3-point conservative numerical approximation schemes, where H_Δ^x and H_Δ^y are given by

$$(3.15) \quad (H_\Delta^x v_\Delta)_{i,j} = v_{i,j} - \frac{\Delta t}{\Delta x} (g_1(v_{i+1,j}, v_{i,j}) - g_1(v_{i,j}, v_{i-1,j})),$$

$$(3.16) \quad (H_\Delta^y w_\Delta)_{i,j} = w_{i,j} - \frac{\Delta t}{\Delta y} (g_2(v_{i+1,j}, v_{i,j}) - g_2(v_{i,j}, v_{i-1,j}))$$

and g_1, g_2 are the numerical fluxes. Those schemes are consistent with the corresponding equation in (3.14) when $g_1(u, u) = f_1(u)$ and $g_2(u, u) = f_2(u)$.

When the functionals $H^x(u, v, w) = v - \lambda^x (g_1(u, v) - g_1(v, w))$ and $H^y(u, v, w) = v - \lambda^y (g_2(u, v) - g_2(v, w))$ with $\lambda^x = \Delta t / \Delta x$ and $\lambda^y = \Delta t / \Delta y$ are monotone increasing with respect to each argument, the schemes are also monotone. Those are particularly useful schemes in one space dimension, since the discrete solutions obtained with them converge to weak entropy solutions of the continuous conservation law, as the discretization parameters tend to zero, under a suitable CFL condition (see Ref. [22], Chap. 3, Th. 4.2).

Thus, we can consider a numerical approximation scheme for (1.2) combining splitting and finite differences, leading to the numerical approximation

$$(3.17) \quad u_\Delta^n = (H_\Delta^y H_\Delta^x)^n u_\Delta^0.$$

The convergence result for this scheme was established in Theorem 2 of [15]. More precisely, for a given $u^0 \in L^1(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$, when $\Delta \rightarrow 0$ with λ^x, λ^y fixed (satisfying a suitable CFL condition), u_Δ^n converges to $u(x, t)$, the entropy solution of (1.2): If $n\Delta t = t$ and $T_0 > 0$ is fixed,

$$(3.18) \quad \max_{0 \leq t \leq T_0} \|(H_\Delta^y H_\Delta^x)^n \Pi_\Delta u^0 - S(t)u^0\|_{L^1} \rightarrow 0, \quad \text{as } \Delta t \rightarrow 0,$$

provided the schemes H_Δ^x and H_Δ^y used component-wise are monotone, of conservation form, consistent with the 1D operators in (3.14), and have continuous numerical fluxes. Thus, the scheme (3.17) satisfies (3.5).

In addition to the convergence property in [15] it is important to underline that the numerical scheme obtained by splitting as above also fulfills the contraction property (3.7) and the L^∞ -stability property (3.6). Note that according to the Crandall-Tartar Lemma (see Lemma 5.2 in [22]) this is so since each of the 1D schemes H_Δ^x and H_Δ^y employed map $l^1(\mathbb{Z}^2)$ into $l^1(\mathbb{Z}^2)$, being monotonic and conservative.

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

$$(3.19) \quad g_1^G(u, v) = \begin{cases} \min_{w \in [u, v]} f_1(w), & \text{if } u \leq v, \\ \max_{w \in [u, v]} f_1(w), & \text{if } u \geq v, \end{cases}$$

$$(3.20) \quad g_1^{LF}(u, v) = \frac{f_1(u) + f_1(v)}{2} - \frac{(v - u)}{2\lambda^x},$$

$$(3.21) \quad g_1^{EO}(u, v) = \frac{f_1(u) + f_1(v) - \int_u^v |f'(\tau)| d\tau}{2}.$$

See Chapter 3 in [22] for more details.

These 1D methods, combined with dimensional splitting, satisfy the conditions of Theorem 3.1.

Remark 3.2. One may also consider the Strang-type splitting

$$u_\Delta^n = \left(H_{\frac{\Delta}{2}}^y H_\Delta^x H_{\frac{\Delta}{2}}^y \right)^n u_\Delta^0,$$

which is second order accurate in time for sufficiently smooth solutions. We have employed (3.17) since it is simpler to implement and because, for conservation laws, the second order accuracy of the Strang-type splittings can be lost for oblique shocks. We refer to [15] for an example of the lost accuracy.

Remark 3.3. We could also use other schemes to approximate the solution of (1.2) in two space dimensions, such as the genuinely 2D numerical scheme introduced in [8, 24]. These schemes satisfy the conditions of the Theorem 3.1.

4. SENSITIVITY ANALYSIS: THE CONTINUOUS APPROACH

We divide this section into four 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 located on a regular surface without boundary.

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 $C_0^1(\mathbb{R}^2)$ be the set of C^1 functions with compact support, and let $u^0 \in C_0^1(\mathbb{R}^2)$ 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 $\delta u^0 \in C_0^1(\mathbb{R}^2)$ 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 $\mathbb{R}^2 \times [0, T]$. 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 $u^{\varepsilon, 0}(x) = u^0(x) + \varepsilon \delta u^0(x)$, is also a classical solution in $(x, t) \in \mathbb{R}^2 \times (0, T)$, and $u^\varepsilon \in C^1(\mathbb{R}^2 \times [0, T])$ can be written as

$$(4.1) \quad u^\varepsilon = u + \varepsilon \delta u + o(\varepsilon), \text{ with respect to the } C^1 \text{ topology,}$$

where δu is the solution of the linearized equation,

$$(4.2) \quad \partial_t \delta u + \operatorname{div}_x (f'(u) \delta u) = 0, \text{ in } \mathbb{R}^2 \times (0, T); \quad \delta u(x, 0) = \delta u^0(x), \quad x \in \mathbb{R}^2.$$

Now, we introduce the adjoint system

$$(4.3) \quad -\partial_t p - f'(u) \cdot \nabla_x p = 0, \text{ in } \mathbb{R}^2 \times (0, T); \quad p(x, T) = p^T(x), \quad x \in \mathbb{R}^2,$$

where $p^T(x) = u(x, T) - u^d(x)$. Thus, if we denote by δJ the Gâteaux derivative of J at u^0 in the direction δu^0 , we have

$$(4.4) \quad \delta J(u^0)[\delta u^0] = \int_{\mathbb{R}^2} p(x, 0) \delta u^0 dx.$$

Therefore, once we have computed the adjoint state, we can obtain a descent direction for the continuous functional J . We just take:

$$(4.5) \quad \delta u^0 = -p(x, 0).$$

Under the assumptions above on u^0 , u , δu and p can be obtained from their data $u^0(x)$, $\delta u^0(x)$ and $p^T(x)$ 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

$$(4.6) \quad x'(t) = f'(u(x(t), t)) = f'(u^0(x^0)), \quad t \in (0, T); \quad x(0) = x^0 \in \mathbb{R}^2.$$

They are straight lines whose slopes depend on the initial data. 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 $\mathbb{R}^2 \times [0, T + \tau]$. This allows us to define u in $\mathbb{R}^2 \times [0, T + \tau]$ in a unique way from the initial data.

For $\varepsilon > 0$ sufficiently small, the solution $u^\varepsilon(x, t)$ corresponding to the initial datum $u^{\varepsilon, 0}(x)$ has similar characteristics to those of u . This allows us to guarantee that two different characteristic lines do not intersect for $0 \leq t \leq T$ if $\varepsilon > 0$ is small enough. Note that u^ε may possibly be discontinuous for $t \in (T, T + \tau]$ if u^0 generates a discontinuity at $t = T + \tau$; this is irrelevant for the analysis in $[0, T]$ we are carrying out. Therefore $u^\varepsilon(x, t)$ is also a classical solution in $(x, t) \in \mathbb{R}^2 \times [0, T]$, and it is easy to see that the solution u^ε can be written as (4.1) where δu satisfies (4.2). Thus, the solution δu along a characteristic line can be obtained from δu^0 , i.e.

$$\delta u(x(t), t) = \delta u^0(x^0) \exp\left(-\int_0^t \operatorname{div}_x(f'(u))(x(s), s) ds\right).$$

Finally, the adjoint system (4.3) is also solved by characteristics. This yields the steepest descent direction in (4.5) for the continuous functional

$$p(x, 0) = u^0(x) - u^d(x + T f'(u^0(x))).$$

Remark 4.1. Note that for classical solutions, the Gâteaux derivative of J at u^0 is given by (4.4), and this provides an obvious descent direction for J at u^0 , given by (4.5). 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 [5–7, 10, 23, 33]), we present an extension to the 2D case.

We focus on the particular case of solutions having a single surface of shock without boundary. For more details about the structure of discontinuities in two space dimensions, see [35].

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 surface, of class C^2 ,

$$\Sigma = \{(\varphi(s, t), t) \mid (s, t) \in \mathbb{R} \times [0, T]\} = \bigcup_{t \in [0, T]} \Sigma^t \subset \mathbb{R}^3,$$

where $\Sigma^t = \{(\varphi(s, t), t), s \in \mathbb{R}\} \subset \mathbb{R}^3$, is the shock curve at time t and $\varphi(\cdot, t)$ is a regular parametrization. The solution $u(x, t)$ is Lipschitz continuous outside Σ .

We call \hat{s} the classical tangent vector to Σ^t , i.e.

$$(4.7) \quad \hat{s} = \frac{(\partial_s \varphi, 0)}{h_s}, \text{ on } \Sigma^t,$$

where $h_s = \|\partial_s \varphi\|$. We note that for all $t \in [0, T]$, Σ^t is a flat curve, i.e. $\hat{s} \cdot \hat{k} = 0$, with $\hat{k} = (0, 0, 1)$. And we define the normal unit vector at the curve Σ^t as

$$\hat{\nu} = \hat{s} \times \hat{k}, \text{ on } \Sigma.$$

On the other hand, $u(x, t)$ satisfies the *Rankine-Hugoniot* condition on Σ

$$(4.8) \quad [u]_{\Sigma^t} n_{\Sigma}^t + [f(u)]_{\Sigma^t} \cdot n_{\Sigma}^x = 0, \text{ on } \Sigma,$$

where $n_{\Sigma} = (n_{\Sigma}^x, n_{\Sigma}^t) \in \mathbb{R}^2 \times \mathbb{R}$ is the normal vector on the surface Σ which satisfies

$$(4.9) \quad \hat{\nu} \cdot n_{\Sigma} > 0, \text{ on } \Sigma.$$

This sets the direction of the normal vector n_{Σ} .

Here and in the sequel we use the following notation for the jump across a surface:

$$(4.10) \quad [u]_{\Sigma^t}(y) = \lim_{\varepsilon \rightarrow 0} u((y, t) + \varepsilon \hat{\nu}(y, t)) - u((y, t) - \varepsilon \hat{\nu}(y, t)), \quad \forall y \in \Sigma^t.$$

In this form, the surface Σ divides the space $\mathbb{R}^2 \times (0, T)$ into two parts, to the left and right of Σ , respectively, Q_- and Q_+ (see Figure 2), where

$$Q_{\pm} = \{(x, t) \in Q_{\pm}^t \times (0, T)\},$$

with $\partial Q_{\pm}^t = \Sigma^t$, the outward normal to Q_{\pm}^t being $\mp \hat{\nu}$ for all $t \in [0, T]$.

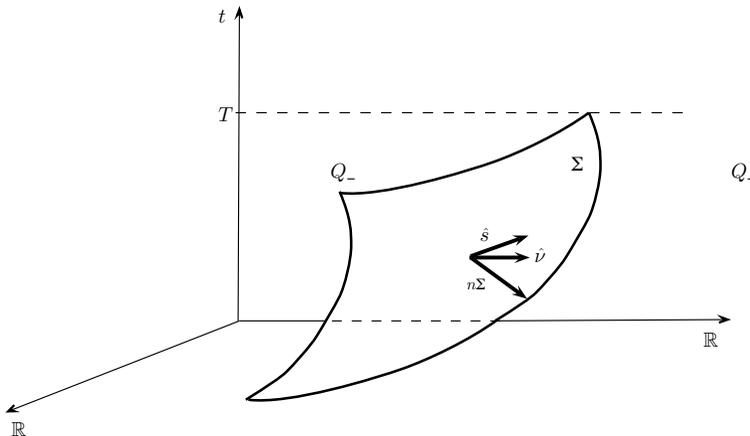


FIGURE 2. Scheme of the sets Q_- , Q_+ and the shock surface Σ , with the vectors $n_{\Sigma}, \hat{\nu}, \hat{s}$ and n_{Σ}^x in the space $\mathbb{R}^2 \times [0, T]$.

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, Σ) combining the solution of (1.2) and the shock location Σ . This is relevant in the analysis of sensitivity of solutions below and when applying descent algorithms.

Then the pair (u, Σ) satisfies the system

$$(4.11) \quad \begin{aligned} \partial_t u + \operatorname{div}_x(f(u)) &= 0, & \text{in } Q_- \cup Q_+, \\ [u]n_\Sigma^t + [f(u)]n_\Sigma^x &= 0, & \text{on } \Sigma, \\ u(x, 0) &= u^0(x), & x \in \mathbb{R}^2 \setminus \Sigma^0, \end{aligned}$$

where Σ^0 is the shock curve at $t = 0$. In other words, the conservation law that is fulfilled to both sides of the shock is to be complemented with the Rankine-Hugoniot condition.

Let us now analyze the sensitivity of (u, Σ) with respect to perturbations of the initial datum, in particular, with respect to variations δu^0 of the initial profile u^0 and $\delta \Sigma^0$ of the shock location Σ^0 . To be more precise, we introduce a notion of generalized tangent vectors in 2D.

Definition 4.3. Let $v : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a piecewise Lipschitz continuous function in $L^1(\mathbb{R}^2)$ with a single discontinuity on the curve $\sigma \subset \mathbb{R}^2$. We define \mathcal{P}_v as the family of all continuous paths $\gamma : [0, \varepsilon_0] \rightarrow L^1(\mathbb{R}^2)$ with

- (1) $\gamma(0) = v$ and $\varepsilon_0 > 0$ possibly depending on γ .
- (2) For any $\varepsilon \in [0, \varepsilon_0]$ the functions $v^\varepsilon = \gamma(\varepsilon)$ are piecewise Lipschitz with a single discontinuity on the curve σ^ε dividing the space \mathbb{R}^2 into two parts: Ω_-^ε and Ω_+^ε , the subdomains of \mathbb{R}^2 to the left and to the right of σ^ε , respectively,

$$\Omega_-^\varepsilon \cup \sigma^\varepsilon \cup \Omega_+^\varepsilon = \mathbb{R}^2.$$

- (3) The curve σ^ε depends continuously on ε and there exists a constant L independent of $\varepsilon \in [0, \varepsilon_0]$ such that

$$|v^\varepsilon(x) - v^\varepsilon(x')| \leq L|x - x'|,$$

for all x, x' such that, $\sigma^\varepsilon \cap [x, x'] = \emptyset$, when $[x, x'] = \{\tau x' + (1 - \tau)x, \tau \in [0, 1]\}$.

Furthermore, we define the set T_v of generalized tangent vectors of v as the space of $(\delta v, \delta \varphi) \in L^1(\mathbb{R}^2) \times W^{1,1}(\sigma)$, for which the path $\gamma_{(\delta v, \delta \varphi)}$ given by

$$(4.12) \quad \gamma_{(\delta v, \delta \varphi)}(\varepsilon)(x) = v(x) + \varepsilon \delta v(x) + \psi_{(\sigma, \delta \varphi)}^\varepsilon(x), \quad x \in \mathbb{R}^2,$$

where, in local coordinates, i.e. $x = y + \lambda \delta \varphi(y) n_\sigma(y)$, for all $(y, \lambda) \in \sigma \times [0, \varepsilon]$, we have

$$\psi_{(\sigma, \delta \varphi)}^\varepsilon(x) = \begin{cases} \operatorname{sgn}(\delta \varphi(y)) [v]_\sigma(y), & (y, \lambda) \in \sigma \times [0, \varepsilon] \\ 0, & \text{otherwise,} \end{cases}$$

satisfies $\gamma_{(\delta v, \delta \varphi)} \in \mathcal{P}_v$.

Finally, we define the equivalence relation \sim on \mathcal{P}_v by

$$\gamma \sim \gamma' \text{ if and only if } \lim_{\varepsilon \rightarrow 0} \frac{\|\gamma(\varepsilon) - \gamma'(\varepsilon)\|_{L^1}}{\varepsilon} = 0,$$

and we say that a path $\gamma \in \mathcal{P}_v$ generates the generalized tangent vector $(\delta v, \delta \varphi) \in T_v$ if it is equivalent to $\gamma_{(\delta v, \delta \varphi)}$ as in (4.12).

Remark 4.4. The path $\gamma_{(\delta v, \delta \varphi)} \in \mathcal{P}_v$ in (4.12) represents, at first order, the variation of a function v by adding a perturbation function $\varepsilon \delta v$ and by shifting the discontinuity by $\varepsilon \delta \varphi$. We note that $\text{supp}(\psi_{(\sigma, \delta \varphi)}^\varepsilon) \subset [\sigma, \sigma^\varepsilon]$, meaning that the support is localized in the region between the original curve σ and the new curve $\sigma^\varepsilon = \sigma + \varepsilon \delta \varphi n_\sigma$.

Let u^0 be the initial datum in (4.11) that we assume to be Lipschitz continuous to both sides of a single discontinuity located on the curve Σ^0 , and consider a generalized tangent vector $(\delta u^0, \delta \varphi^0) \in C_c^1(\mathbb{R}^2) \times C_c^{1,1}(\Sigma^0)$. Let $u^{0,\varepsilon} \in \mathcal{P}_{u^0}$ be a path which generates $(\delta u^0, \delta \varphi^0)$. For ε sufficiently small the solution $u^\varepsilon(\cdot, t)$ of (4.11) is Lipschitz continuous with a single discontinuity on the curve $\Sigma^{t,\varepsilon}$, for all $t \in [0, T]$. Therefore, $u^\varepsilon(\cdot, t)$ generates a generalized tangent vector $(\delta u(\cdot, t), \delta \varphi(\cdot, t)) \in L^1(\mathbb{R}^2) \times W^{1,1}(\Sigma^t)$. Thus, we note that the new position of the shock after perturbations of the form $(\delta u^0, \delta \varphi^0)$ is described by a function $\delta \varphi : \Sigma \rightarrow \mathbb{R}$ which denotes the variation of Σ as

$$(4.13) \quad \Sigma^\varepsilon = \Sigma + \varepsilon \delta \Sigma = \{(y, t) + \varepsilon \delta \varphi \hat{\nu}, \quad (y, t) \in \Sigma\}.$$

Then $\delta \varphi$ represents the displacement in the normal direction to Σ^t .

Now, let us introduce some definitions and notations. The divergence on the surfaces Σ of the vector field G , i.e. $\text{div}_\Sigma(G)$, is defined by

$$(4.14) \quad \int_\Sigma \text{div}_\Sigma(G) \psi \, d\Sigma = - \int_\Sigma G \cdot \nabla_\Sigma \psi \, d\Sigma, \quad \forall \psi \in C_c^1(\mathbb{R}^2 \times (0, T)),$$

where ∇_Σ is tangential gradient (covariant derivative) defined as

$$(4.15) \quad \nabla_\Sigma \psi = \nabla_{x,t} \psi - (\nabla_{x,t} \psi \cdot n_\Sigma) n_\Sigma, \quad \text{on } \Sigma.$$

The following result describes the nature of the linearized system governing the dynamics of the generalized tangent vector $(\delta u, \delta \varphi)$.

Theorem 4.5. *The equations for the new variables δu and $\delta \varphi$ are*

$$(4.16) \quad \partial_t \delta u + \text{div}_x(f'(u) \delta u) = 0, \quad \text{in } Q_- \cup Q_+,$$

$$(4.17) \quad \text{div}_\Sigma(\delta \varphi \|n_\Sigma^\# \| ([f(u)]_{\Sigma^t}, [u]_{\Sigma^t})) = ([f'(u) \delta u]_{\Sigma^t}, [\delta u]_{\Sigma^t}) \cdot n_\Sigma, \quad \text{on } \Sigma,$$

$$(4.18) \quad \delta u(x, 0) = \delta u^0(x), \quad x \in \mathbb{R}^2 \setminus \Sigma^0,$$

$$(4.19) \quad \delta \varphi(x, 0) = \delta \varphi^0(x), \quad x \in \Sigma^0.$$

Proof of of Theorem 4.5. First let us define the function $G : \mathbb{R} \rightarrow \mathbb{R}^3$, as $G(z) = (f(z), z)$. The system (4.11), including the Rankine-Hugoniot condition on the shock, can be rewritten in the form

$$(4.20) \quad \int_{\mathbb{R}^2 \times (0, T)} G(u) \cdot \nabla_{x,t} \psi \, dx dt = 0, \quad \forall \psi \in C_c^1(\mathbb{R}^2 \times (0, T)).$$

Using Reynolds' transport theorem, we obtain that the linearization of (4.20) is

$$(4.21) \quad \int_{Q_- \cup Q_+} \delta u G'(u) \cdot \nabla_{x,t} \psi \, dx dt - \int_\Sigma \delta \varphi (\nu \cdot n_\Sigma) [G(u)]_{\Sigma^t} \cdot \nabla_{x,t} \psi \, d\Sigma = 0, \quad \forall \psi \in C_c^1(\mathbb{R}^2 \times (0, T)).$$

Considering ψ with support outside Σ in (4.21), we obtain

$$(4.22) \quad \text{div}_{x,t}(\delta u G'(u)) = \partial_t \delta u + \text{div}_x(f'(u) \delta u) = 0, \quad \text{in } Q_- \cup Q_+.$$

On other hand, the Rankine-Hugoniot condition (4.8) can be rewritten as

$$(4.23) \quad [G(u)]_{\Sigma^t} \cdot n_{\Sigma} = 0, \quad \text{on } \Sigma,$$

and

$$(4.24) \quad [G(u)]_{\Sigma^t} \cdot \nabla_{x,t} \psi = [G(u)]_{\Sigma^t} \cdot \nabla_{\Sigma} \psi, \quad \forall \psi \in C_c^1(\mathbb{R}^2 \times (0, T)).$$

Using (4.22), (4.24) and (4.14) in (4.21), we have

$$(4.25) \quad 0 = - \int_{\Sigma} \psi [\delta u G'(u)]_{\Sigma^t} \cdot n_{\Sigma} d\Sigma + \int_{\Sigma} \operatorname{div}_{\Sigma} (\delta \varphi (\nu \cdot n_{\Sigma}) [G(u)]_{\Sigma^t}) \psi d\Sigma, \quad \forall \psi \in C_c^1(\mathbb{R}^2 \times (0, T)).$$

Therefore, we obtain

$$(4.26) \quad \operatorname{div}_{\Sigma} (\delta \varphi (\nu \cdot n_{\Sigma}) [G(u)]_{\Sigma^t}) = [\delta u G'(u)]_{\Sigma^t} \cdot n_{\Sigma}, \quad \text{on } \Sigma.$$

Thus, using the choice of the normal vector (4.9), we have $\|n_{\Sigma}^x\| = \hat{\nu} \cdot n_{\Sigma}$. Then, replacing this in (4.26), we obtain (4.17).

This concludes the proof. □

Remark 4.6. The linearized system (4.16)-(4.19) is obtained, at least formally, by a perturbation argument. We introduce a perturbation of the data

$$(u_{\varepsilon}^0, \Sigma_{\varepsilon}^0) = (u^0, \Sigma^0) + \varepsilon(\delta u^0, \delta \varphi^0)$$

and compute the equations for the first order perturbation of the solution. This is in fact the usual approach in the study of the linearized stability of shocks. We refer to [23] for a detailed description of this method in the scalar case and [29] for more general systems of conservation laws in higher dimensions. This analysis leads to the formal expansion,

$$(u_{\varepsilon}, \Sigma_{\varepsilon}) = (u, \Sigma) + \varepsilon(\delta u, \delta \varphi) + o(\varepsilon).$$

In 1D, this expansion can be justified for general scalar 1D conservation laws of the form

$$\partial_t u + \partial_x (f(u)) = 0,$$

when the flux $f \in C^1$ is convex in the context of duality solutions using the one-sided Lipschitz condition (OSLC) (see [7]). This differentiability has been proved for more general situations, for instance in [10] for systems of conservation laws, or [33] for scalar equations in 1D but under more restrictive structural conditions on the solutions. In this paper we follow the latter approach, and our linearization is not justified in the general frame of Kruřkov’s entropy solutions.

4.3. The linearized transport equation on the shock surface Σ . The equations (4.17) and (4.19) on the function $\delta \varphi$ form a transport equation on the surface Σ . Here we explain how this equation can be solved.

Let us consider Σ , as in Hypothesis 4.2, and its parametrization $(\varphi(s, t), t)$. Then there exists $\varepsilon_0 > 0$, such that $H : \mathbb{R} \times (0, T) \times (-\varepsilon_0, \varepsilon_0) \rightarrow \mathbb{R}^3$ given by

$$(4.27) \quad H(s, t, \lambda) = (\varphi(s, t), t) + \lambda n_{\Sigma}(s, t),$$

is invertible, i.e. H defines a local change of coordinates around of Σ , and the set $\Theta = H(\mathbb{R} \times (0, T) \times (-\varepsilon_0, \varepsilon_0)) \subset \mathbb{R}^2 \times \mathbb{R}$ is a neighborhood of Σ . Thus, we can write the gradient of any function $\psi \in C^1(\Theta)$ in this local coordinates, and we obtain

$$(4.28) \quad \nabla_{x,t}\psi = \frac{1}{\det(DH)} \left(\partial_s \tilde{\psi} \partial_t H \times \partial_\lambda H + \partial_t \tilde{\psi} \partial_\lambda H \times \partial_s H + \partial_\lambda \tilde{\psi} \partial_s H \times \partial_t H \right), \text{ in } \Theta,$$

where $\tilde{\psi}(s, t, \lambda) = \psi(H(s, t, \lambda))$ and DH is the differential of H , i.e. $DH = (\partial_s H, \partial_t H, \partial_\lambda H)$. Then, if we consider (4.28) on Σ , i.e. $\lambda = 0$, the definition of tangent vector (4.7) and the normal vector n_Σ in terms of the parametrization $H(s, t, 0) = (\varphi(s, t), t)$ of Σ , i.e.

$$n_\Sigma = \frac{\partial_s H \times \partial_t H}{\|\partial_s H \times \partial_t H\|} = \frac{\hat{s} \times \partial_t H}{\|\hat{s} \times \partial_t H\|} = \|n_\Sigma^x\| \hat{\nu} + n_\Sigma^t \hat{k}, \text{ on } \Sigma,$$

we obtain

$$\nabla_{x,t}\psi = \frac{1}{h_s \|\hat{s} \times \partial_t H\|} \left(\partial_s \tilde{\psi} \partial_t H \times n_\Sigma + h_s \partial_t \tilde{\psi} n_\Sigma \times \hat{s} \right) + \partial_\lambda \tilde{\psi} n_\Sigma, \text{ on } \Sigma.$$

With the notation $\|n_\Sigma^x\| = \frac{1}{\|\hat{s} \times \partial_t H\|}$ on Σ , we have

$$(4.29) \quad \nabla_\Sigma \psi = \frac{\|n_\Sigma^x\|}{h_s} \left(\partial_s \tilde{\psi} \partial_t H \times n_\Sigma + h_s \partial_t \tilde{\psi} n_\Sigma \times \hat{s} \right), \text{ on } \Sigma.$$

Thus, for any $F \in C^1(\Theta)$, using (4.29) and

$$(4.30) \quad d\Sigma = \frac{h_s}{\|n_\Sigma^x\|} ds dt,$$

in the definition (4.14), we obtain

$$(4.31) \quad \operatorname{div}_\Sigma(F) = \frac{\|n_\Sigma^x\|}{h_s} \left(\partial_s(F \cdot (\partial_t H \times n_\Sigma)) + \partial_t(h_s F \cdot (n_\Sigma \times \hat{s})) \right), \text{ on } \Sigma.$$

Let us denote $n^\perp = n_\Sigma \times \hat{s}$, which corresponds to the ninety degree rotation of n_Σ around \hat{s} . If $n_\Sigma = (n_\Sigma^x, n_\Sigma^t) \in \mathbb{R}^2 \times \mathbb{R}$, we obtain

$$n^\perp = \left(-\frac{n_\Sigma^t}{\|n_\Sigma^x\|} n_\Sigma^x, \|n_\Sigma^x\| \right).$$

On the other hand, we can consider F , such that $F_n = F \cdot n_\Sigma = 0$, because by definition of tangent divergence (4.14), we have

$$(4.32) \quad \int_\Sigma \operatorname{div}_\Sigma(F) \psi d\Sigma = \int_\Sigma \operatorname{div}_\Sigma(\tilde{F}) \psi d\Sigma, \quad \forall \psi \in C^1(\mathbb{R}^2 \times (0, T)),$$

where $\tilde{F} = F - (F \cdot n_\Sigma) n_\Sigma$ on Σ . Therefore, assuming

$$(4.33) \quad F \cdot n_\Sigma = 0, \text{ on } \Sigma,$$

we obtain

$$(4.34) \quad F \cdot n^\perp = \frac{1}{\|n_\Sigma^x\|} F \cdot \hat{k}, \text{ on } \Sigma,$$

and an explicit computation gives

$$(4.35) \quad \partial_t H \times n_\Sigma = ((\partial_t \varphi, 0) \cdot \hat{s}) n^\perp + \frac{1}{\|n_\Sigma^x\|} \hat{s}, \text{ on } \Sigma.$$

Then, replacing (4.34) and (4.35) in (4.31), we obtain

$$(4.36) \quad \operatorname{div}_\Sigma(F) = \frac{\|n_\Sigma^x\|}{h_s} \left(\partial_s \left(\frac{((\partial_t \varphi, 0) \cdot \hat{s})}{\|n_\Sigma^x\|} F \cdot \hat{k} + \frac{1}{\|n_\Sigma^x\|} F \cdot \hat{s} \right) + \partial_t \left(\frac{h_s}{\|n_\Sigma^x\|} F \cdot \hat{k} \right) \right), \text{ on } \Sigma.$$

Now if we consider $F = \delta\varphi \|n_\Sigma^x\| ([f(u)]_{\Sigma^t}, [u]_{\Sigma^t})$ on Σ , in (4.36), we have

$$(4.37) \quad \operatorname{div}_\Sigma(\delta\varphi \|n_\Sigma^x\| ([f(u)]_{\Sigma^t}, [u]_{\Sigma^t})) = \frac{\|n_\Sigma^x\|}{h_s} (\partial_s (B \delta\varphi) + \partial_t (h_s [u]_{\Sigma^t} \delta\varphi)), \text{ on } \Sigma,$$

where the coefficient B is given by $B = ((\partial_t \varphi, 0) \cdot \hat{s})[u]_{\Sigma^t} + ([f(u)]_{\Sigma^t}, 0) \cdot \hat{s}$.

We note that the coefficient $h_s[u]_{\Sigma^t}$ does not vanish. Then the equation (4.17) with initial condition in $t = 0$ is well-posed.

Thus using (4.37), we have that under the hypothesis (4.2), the coefficients in the equation (4.17) are Lipschitz continuous (i.e. $\|n_\Sigma^x\|[G(u)]_{\Sigma^t}$ is Lipschitz on Σ). The linearized system (4.16)-(4.19) has a unique solution which can be computed in two steps. The method of characteristics determines δu in $Q^- \cup Q^+$, i.e. outside Σ , from the initial data δu^0 by the method of characteristics (note that system (4.16)-(4.19) has the same characteristics as (4.11)). This yields the value of u and δu to both sides of the shock Σ , and allows us to determine the coefficients and the right side of the equation that $\delta\varphi$ satisfies. Then, we solve this transport equation on Σ to obtain $\delta\varphi$.

4.4. Sensitivity of J 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 Gâteaux derivative of J .

Definition 4.7. Let $J : L^1(\mathbb{R}^2) \rightarrow \mathbb{R}$ be a functional, and let $u^0 \in L^1(\mathbb{R}^2)$ be Lipschitz continuous with a discontinuity along a regular curve Σ^0 , of class C^2 , an initial datum for which the solution of (1.2) satisfies hypothesis (4.2) and (4.13). J is Gâteaux derivative differentiable at u^0 in a generalized sense if for any generalized tangent vector $(\delta u^0, \delta\varphi^0)$ and any family $u^{0,\varepsilon} \in \mathcal{P}_{u^0}$ associated to $(\delta u^0, \delta\varphi^0)$ the limit

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

exists, and it depends only on (u^0, Σ^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 Gâteaux derivative of J in the direction $(\delta u^0, \delta\varphi^0)$.

The following result easily provides a characterization of the generalized Gâteaux derivative of J in terms of the solution of the associated adjoint system (4.39), (4.40), (4.41), (4.42), (4.43), and (4.44).

Proposition 4.8. *The Gâteaux derivative of J can be written as*

$$(4.38) \quad \delta J(u^0)[\delta u^0, \delta \varphi^0] = \int_{\mathbb{R}^2} p(x, 0) \delta u^0 dx - \int_{\Sigma^0} q(x, 0) [u]_{\Sigma^0} \delta \varphi^0 d\sigma,$$

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

$$(4.39) \quad \partial_t p + f'(u) \cdot \nabla p = 0, \quad \text{in } Q_- \cup Q_+,$$

$$(4.40) \quad [p]_{\Sigma^t} = 0, \quad \text{on } \Sigma,$$

$$(4.41) \quad q(x, t) = p(x, t), \quad (x, t) \in \Sigma,$$

$$(4.42) \quad ([f(u)]_{\Sigma^t}, [u]_{\Sigma^t}) \cdot \nabla_{\Sigma} q = 0, \quad \text{on } \Sigma,$$

$$(4.43) \quad p(x, T) = u(x, T) - u^d(x), \quad x \in \mathbb{R}^2 \setminus \Sigma^T,$$

$$(4.44) \quad q(x, T) = \frac{[(u(\cdot, T) - u^d)^2 / 2]_{\Sigma^T}}{[u]_{\Sigma^T}}, \quad x \in \Sigma^T.$$

Let us briefly comment on the result of Proposition 4.8 before giving its proof. Formula (4.38) provides an obvious way to compute a first descent direction of J at u^0 . We just take

$$(4.45) \quad (\delta u^0, \delta \varphi^0) = (-p(\cdot, 0), q(\cdot, 0)[u]_{\Sigma^0}).$$

Here, the value of $\delta \varphi^0$ must be interpreted as the optimal infinitesimal displacement of the discontinuity of u^0 . However, it is important to emphasize that this $(\delta u^0, \delta \varphi^0)$ is not a generalized tangent vector in T_{u^0} since $p(x, 0)$ is not continuous away from Σ^0 . In fact in 1D, $p(x, t)$ takes the constant value $q(T)$ in the whole triangular region occupied by the characteristics of (1.2) which meet the shock Σ . Thus, p has, in general, two discontinuities at the boundary of this region and so will be for $p(x, 0)$ (see [12]) and the scheme in Figure 3.

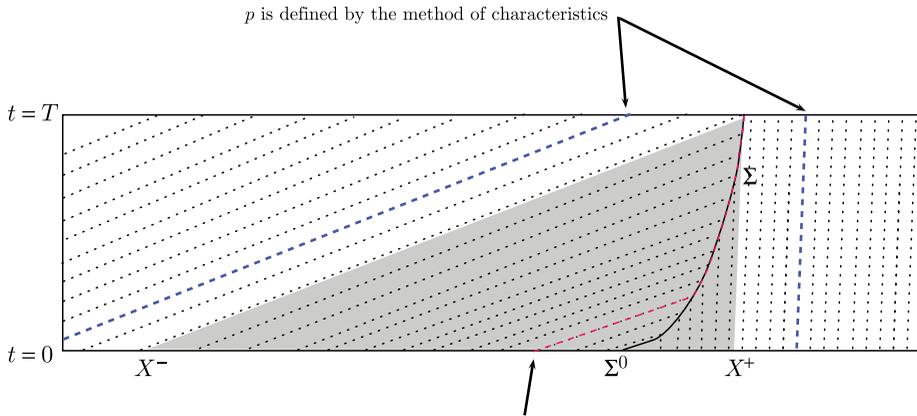


FIGURE 3. Characteristic lines entering on a shock, and how they may be used to build the solution of the adjoint system both away from the shock and on its region of influence.

This is an important drawback in developing a descent algorithm for J . Indeed, according to Definition 4.3, if $(\delta u^0, \delta \varphi^0)$ is a descent direction belonging to T_{u^0} , the new datum $u^{0,new}$ should be obtained from u^0 following a path associated to this descent direction (see (4.12)).

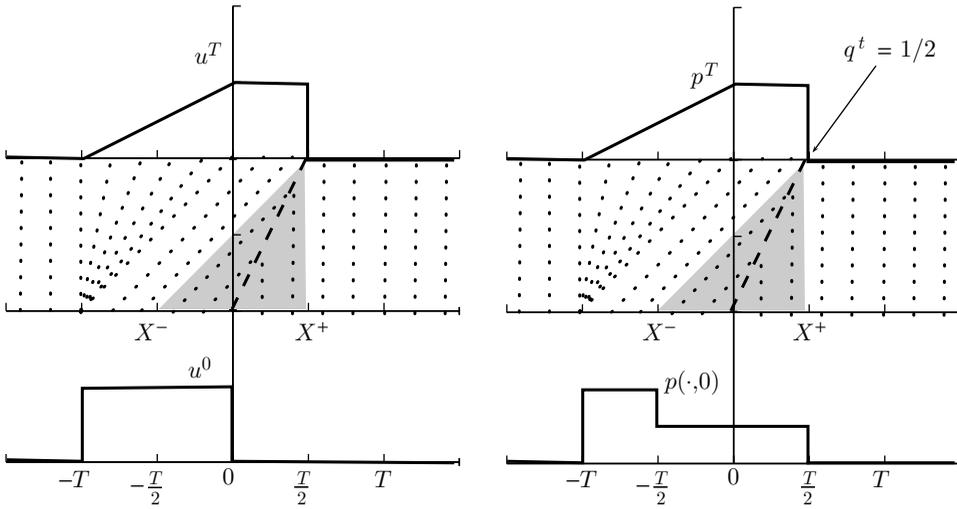


FIGURE 4. Schematic representation of the increase of the number of shocks in Burgers' equation (1.5) when using the adjoint system. Here $u^0 \in \mathcal{U}_{ad}$, $u^d \equiv 0$, and $q^T = -[u^T]_{\Sigma^T}$.

Note that, if we take (4.45) as descent direction $(\delta u^0, \delta \varphi^0)$, which is not a generalized tangent vector as explained above, the new datum $u^{0,new}$ will have three discontinuities; one comes from the displacement of the discontinuity of u^0 at Σ^0 and two more are produced by the discontinuities of $p(x, 0)$. Thus, in an iterative process, the descent algorithm will create more and more discontinuities increasing artificially the complexity of solutions. This motivates the alternating descent method which we propose here that, based on this notion of generalized gradient, develops a descent algorithm that keeps the complexity of solutions bounded. This will be done in the following section.

We finish this section with the proof of Proposition 4.8.

Proof of Proposition 4.8. A straightforward computation shows that J is Gâteaux differentiable in the generalized sense of Definition 4.7, and by using Reynolds' transport theorem, we obtain that the generalized Gâteaux derivative of J in the direction of the generalized tangent vector $(\delta u^0, \delta \varphi^0)$ is given by

$$(4.46) \quad \delta J(u^0)[\delta u^0, \delta \varphi^0] = \int_{\mathbb{R}^2 \setminus \Sigma^T} (u(x, T) - u^d(x)) \delta u(x, T) dx - \int_{\Sigma^T} \left[\frac{(u(\cdot, T) - u^d(\cdot))^2}{2} \right] (x) \delta \varphi(x, T) d\Sigma^T(x),$$

where the pair $(\delta u, \delta \varphi)$ solves the linearized problem (4.16)-(4.19) with initial data $(\delta u^0, \delta \varphi^0)$.

Let us now introduce the adjoint system (4.39)-(4.44). Multiplying the equations of δu by p and integrating, we get

$$\begin{aligned}
 (4.47) \quad 0 &= \int_{\mathbb{R}^2 \times (0, T)} (\partial_t \delta u + \operatorname{div}_x (f'(u) \delta u)) p \, dx dt \\
 &= - \int_{\mathbb{R}^2 \times (0, T)} \delta u (\partial_t p + f'(u) \nabla_x p) \, dx dt + \int_{\mathbb{R}^2 \setminus \Sigma^T} \delta u(x, T) p(x, T) \, dx \\
 &\quad - \int_{\mathbb{R}^2 \setminus \Sigma^0} \delta u^0(x) p(x, 0) \, dx - \int_{\Sigma} ([\delta u p]_{\Sigma^t} n_{\Sigma}^t + [f'(u) \delta u p]_{\Sigma^t} \cdot n_{\Sigma}^x) \, d\Sigma.
 \end{aligned}$$

The first term in the right-hand side vanishes since p satisfies the adjoint equation (4.39).

Therefore, replacing (4.43), we obtain

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

The last two terms in the right-hand side of (4.48) 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 Σ^t , i.e.

$$\bar{g}(x, t) = \frac{1}{2} \lim_{\varepsilon \rightarrow 0} (g((x, t) + \varepsilon \hat{\nu}) + g((x, t) - \varepsilon \hat{\nu})), \quad \forall (x, t) \in \Sigma.$$

Thus we have

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

Using (4.17) and (4.40), we have

$$\begin{aligned}
 (4.49) \quad \int_{\Sigma} ([\delta u p]_{\Sigma^t} n_{\Sigma}^t + [f'(u) \delta u p]_{\Sigma^t} \cdot n_{\Sigma}^x) \, d\Sigma \\
 = \int_{\Sigma} \bar{p} \operatorname{div}_{\Sigma} (\delta \varphi \|n_{\Sigma}^x\| ([f(u)]_{\Sigma^t}, [u]_{\Sigma^t})) \, d\Sigma.
 \end{aligned}$$

On the other hand, we have the identity

$$(4.50) \quad \operatorname{div}_{\Sigma} (g G) = g \operatorname{div}_{\Sigma} (G) + \nabla_{\Sigma} g \cdot G,$$

and from (4.30) and (4.37) we have the property

$$(4.51) \quad \int_{\Sigma} \operatorname{div}_{\Sigma} (\|n_{\Sigma}^x\| G) \, d\Sigma = \int_{\Sigma^T} G \cdot \hat{k} \, d\Sigma^T - \int_{\Sigma^0} G \cdot \hat{k} \, d\Sigma^0,$$

for any $G \in (W^{1,1}(\Theta))^3$, with $\Theta \subset \mathbb{R}^2 \times \mathbb{R}$ an open set such that $\Sigma \subset \Theta$.

Thus, using (4.50) and (4.51) in (4.49), we obtain

$$\begin{aligned}
 \int_{\Sigma} ([\delta u p]_{\Sigma^t} n_{\Sigma}^t + [f'(u)\delta u p]_{\Sigma^t} n_{\Sigma}^x) d\Sigma &= \int_{\Sigma} \operatorname{div}_{\Sigma} (\bar{p} \delta\varphi \|n_{\Sigma}^x\| ([f(u)]_{\Sigma^t}, [u]_{\Sigma^t})) d\Sigma \\
 &\quad - \int_{\Sigma} \delta\varphi \|n_{\Sigma}^x\| ([f(u)]_{\Sigma^t}, [u]_{\Sigma^t}) \cdot \nabla_{\Sigma} \bar{p} d\Sigma \\
 &= \int_{\Sigma^T} \bar{p}(x, T) \delta\varphi(x, T) [u]_{\Sigma^T} d\Sigma^T \\
 &\quad - \int_{\Sigma^0} \bar{p}(x, 0) \delta\varphi(x, 0) [u]_{\Sigma^0} d\Sigma^0 \\
 (4.52) \qquad \qquad \qquad &\quad - \int_{\Sigma} \delta\varphi \|n_{\Sigma}^x\| ([f(u)]_{\Sigma^t}, [u]_{\Sigma^t}) \cdot \nabla_{\Sigma} \bar{p} d\Sigma.
 \end{aligned}$$

Therefore, replacing (4.41), (4.42), (4.44), and (4.52) in (4.48), we obtain

$$\delta J(u^0)[\delta u^0, \delta\varphi^0] = \int_{\mathbb{R}^2 \setminus \Sigma^0} \delta u^0(x) p(x, 0) dx - \int_{\Sigma^0} q(x, 0) \delta\varphi^0(x) [u]_{\Sigma^0} d\Sigma^0.$$

This concludes the proof. □

5. THE ALTERNATING DESCENT METHOD

As explained above, when dealing with discontinuities, one of the main drawbacks of the continuous approach is the increase in the complexity of solutions one gets along the iterative process, due to the use of optimal descent directions (see Figure 5 for a scheme of the solution of the adjoint system, in this scheme we represent the behavior of the adjoint system in a toy 1D case; then in the new step we increase the number of discontinuities in $u^{0,\varepsilon}$, see Figure 6). A possible remedy is to use the generalized tangent vectors as descent directions for J .

First let us introduce some notation. We consider two curves

$$(5.1) \qquad X^- = \{x - T f'(u^-(x, T)) \mid \forall x \in \Sigma^T\}, \quad X^+ = \{x - T f'(u^+(x, T)) \mid \forall x \in \Sigma^T\},$$

and the set $[X^-, \Sigma^0]$ constituted by the points in \mathbb{R}^2 between the curves X^- and Σ^0 . We can also define similarly the set $[\Sigma^0, X^+]$.

The set $[X^-, X^+]$ represents the basis of the black triangle in Figures 3 and 4, where shocks collapse at time $t = T$.

5.1. The algorithm in 2D. Following the previous ideas, we introduce descent directions that do not increase the complexity of the solution in each iteration of the optimization process.

We now introduce the three classes of perturbations we shall use in our descent algorithm.

First class of perturbations: With this set of perturbations we want to change the profile of u^0 only at one side of the shock. We set the first perturbation $d_1 = (\delta u^0, \delta\varphi^0)$, given by

$$(5.2) \qquad \delta\varphi^0 \equiv 0, \quad \delta u^0(x) = \begin{cases} -p(x, 0), & x \in Q_-^0 \setminus [X^-, \Sigma^0], \\ r(x), & x \in [X^-, \Sigma^0], \\ 0, & \text{otherwise,} \end{cases}$$

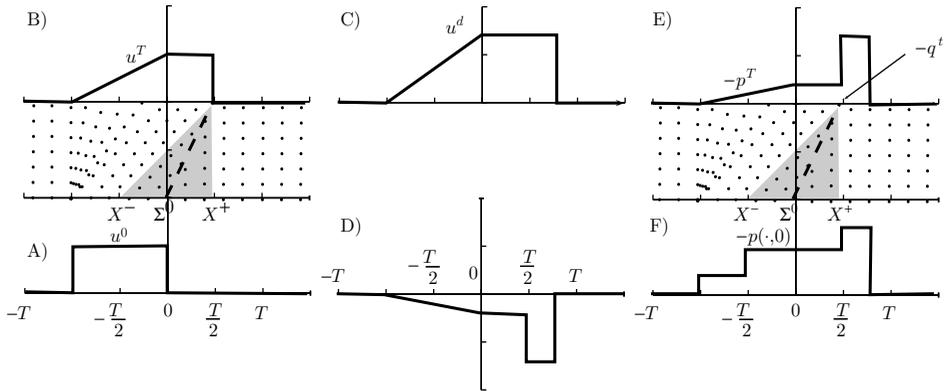


FIGURE 5. Scheme of the solution of the adjoint system in a toy case. **A)** The initial datum u^0 . **B)** The solution of Burgers' equation at time T . **C)** The target function u^d . **D)** The value at $t = T$ of the adjoint state $p^T(x) = u(x, T) - u^d(x)$. **E)** The initial conditions p^T and q^T . **F)** The solution of the adjoint system at time $t = 0$.

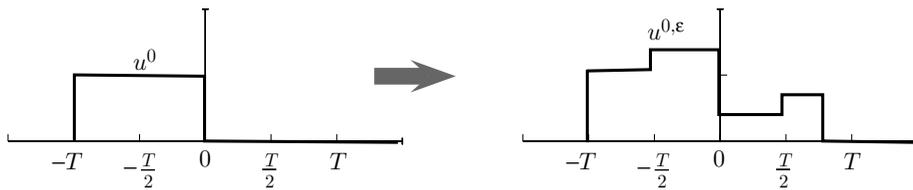


FIGURE 6. Scheme of the next step in the descent process.

where p is the solution of the adjoint system (4.39)-(4.43) and $r(x)$ is any extension of $-p(x, 0)$ to the set $[X^-, \Sigma^0]$ such that δu^0 is only discontinuous on Σ^0 ; see Figure 7

Second class of perturbations: In this case, we only want to move the shock without changing the profile of the function. We set the second perturbation $d_2 = (\delta u^0, \delta \varphi^0)$, given by

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

where q is the solution of the adjoint system (4.42) and (4.44).

Third class of perturbations: With this set of perturbations we want to change the profile of u^0 only at one side of the shock, without modifying the position of the shock at time $t = 0$. The third perturbation $d_3 = (\delta u^0, \delta \varphi^0)$ is chosen to be

$$(5.4) \quad \delta \varphi^0 \equiv 0, \quad \delta u^0(x) = \begin{cases} -p(x, 0) & x \in Q_+^0 \setminus [\Sigma^0, X^+], \\ r(x) & x \in [\Sigma^0, X^+], \\ 0 & \text{otherwise,} \end{cases}$$

where p is the solution of the adjoint system (4.39)-(4.43) and $r(x)$ is any extension of $-p(x, 0)$ to the set $[\Sigma^0, X^+]$ such that δu^0 is only discontinuous on Σ^0 .

In this way the first and the third perturbations will conserve the shock structure and only change the profile of the solution to both sides. Therefore, d_1 and d_3 given by (5.2) and (5.4) satisfy

$$(5.5) \quad \delta J(u^0)[d_1] = - \int_{Q_-^0 \setminus [X^-, \Sigma^0]} |p(x, 0)|^2 dx + \int_{[X^-, \Sigma^0]} r(x) p(x, 0) dx,$$

$$(5.6) \quad \delta J(u^0)[d_3] = - \int_{Q_+^0 \setminus [\Sigma^0, X^+]} |p(x, 0)|^2 dx + \int_{[\Sigma^0, X^+]} r(x) p(x, 0) dx.$$

Then, in general, these classes of perturbations do not necessarily produce the decrease of the functional J , but they preserve the structure of the function u^0 and, simultaneously, give the possibility of moving from a local minimizer to a global one, exploring other profiles outside of the shock.

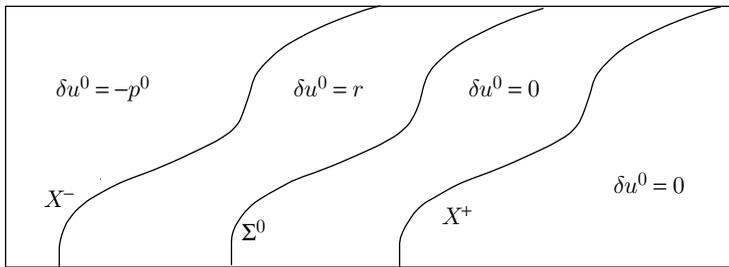


FIGURE 7. Scheme of the curves Σ^0 , X^- , X^+ , and the extension of δu^0 in the first perturbation.

However, the second perturbation d_2 given by (5.3) satisfies

$$(5.7) \quad \delta J(u^0)[d_2] = - \int_{\Sigma^0} |q(x, 0)|^2 d\Sigma^0 \leq 0.$$

Thus d_2 produces the decrease of the functional J and preserves the structure of the function u^0 .

Remark 5.1. When considering the first and third perturbations, we need to choose r . This corresponds to the *inpainting* (image interpolation) problem in image restoration. Following the approaches by Rudin, Osher and Fatemi [31] and the object-edge model by Mumford and Shah [30], we may take r as the solution of a general variational inpainting model, minimizing a variational problem of the form

$$(5.8) \quad \min_{\psi \in W^{1,2}(\mathbb{R}^2)} E(\psi) = \frac{\alpha}{2} \int_{\Omega \setminus [X^-, X^+]} |\psi + p^0|^2 dx + \frac{\beta}{2} \int_{\Omega} |\nabla \psi|^2 dx,$$

$\Omega = Q_-^0$ being fixed for the first perturbation and $\Omega = Q_+^0$ for the third perturbation. Here α and β are positive constants. In (5.8) the first term is commonly called the *data fidelity term*, which represents the data assimilation of the model,

and the last term gives the regularity of the solution, producing a good interpolation. Whether there are other, better suited, methods to extend δu^0 near the shock is an interesting open problem.

We observe that the choice of r is not relevant at first order outside the shock. Indeed, u^0 depends on r only in $[X^-, X^+]$ and in the linear system (4.16), (4.17), (4.18), (4.19), this region at time T collapses on the shock curve Σ^T . Therefore, the choice of r only affects the motion of the curve Σ^T at time T . Then when we fix r , we loose the optimal descent direction for the functional J , but we do not change essentially the profile of the solution $u(\cdot, T)$ at both sides of the shock. We mainly affect the position of the shock Σ^T , because this region $[X^-, X^+]$ collapses at time T .

5.2. Descent strategy. Here we propose a new method built on the results in Proposition 4.8 and the discussion thereafter, which is an extension to 2D of the method introduced in [12]. We shall refer to this new method as the alternating descent method in 2D.

For a given initialization of u^0 , in each step of the descent iteration process, we proceed in the following three substeps:

- (1) **First direction:** We change the profile in Q_-^0 .
 - Compute (5.2) and find the optimal step size ε such that $p(x, 0)$, the solution of the adjoint system (4.39)-(4.43), and its L^2 -norm diminish in $Q_-^0 \setminus [X^-, \Sigma^0]$. In this way we obtain the best profile in Q_-^0 for u^0 .
 - Compute (5.3) and find the optimal step size ε for which this datum must be modified as in (5.3). In this way we obtain the best location of the discontinuity for this u^0 .
- (2) **Second direction:** We only move the shock. Compute (5.3) and find the optimal step size ε for which this datum must be modified in perturbations given by (5.3).
- (3) **Third direction:** We change the profile in Q_+^0 .
 - Compute (5.4) and find the optimal step size ε such that $p(x, 0)$, the solution of the adjoint system (4.39)-(4.43) and its L^2 -norm diminish in $Q_+^0 \setminus [\Sigma^0, X^+]$. In this way we obtain the best profile in Q_+^0 for the new u^0 .
 - Compute (5.3) and find the optimal step size ε for which this datum must be modified in the perturbation given by (5.3).

In the above procedure, finding the step size ε involves a 1D optimization problem that we can solve with a classical method (bisection, Armijo’s rule, etc.) In the first and third directions we choose the step size ε such that the L^2 -norm of $p(\cdot, 0)$ decreases, trying to move to a critical point of the functional J .

If we consider u^0 given by (5.2) or (5.4), we can solve (4.16) and (4.18) and compute the right-hand side of the equation (4.17). Therefore, using (4.37), we have that, under the hypothesis (4.2), the coefficients in the equation (4.17) are Lipschitz continuous. Then we observe that there exists $\delta\varphi^0$ such that

$$\begin{aligned}
 (5.9) \quad \operatorname{div}_\Sigma (\delta\varphi \|n_\Sigma^x\| ([f(u)]_{\Sigma^t}, [u]_{\Sigma^t})) &= ([f'(u)\delta u]_{\Sigma^t}, [\delta u]_{\Sigma^t}) \cdot n_\Sigma, \quad \text{on } \Sigma, \\
 \delta\varphi(x, 0) &= \delta\varphi^0(x), \quad x \in \Sigma^0,
 \end{aligned}$$

satisfies the terminal condition

$$(5.10) \quad \delta\varphi(x, T) = 0, \quad x \in \Sigma^T.$$

Therefore, the first and third directions produce the decrease of the functional J at first order.

Thus, one has to iterate this procedure to assure a simultaneous better placement of the shock and a better fitting of the value of the solution away from it.

The main difference of our 2D algorithm with the 1D one proposed in [12] relies on the definition of the first and third classes of directions. In [12] $\delta\varphi^0$ does not necessarily vanish, but is rather taken for the shock at time T not to move. The analogue in 2D would be to find $\delta\varphi^0$ in such a way that the solution of (5.9) satisfies $\delta\varphi^T \equiv 0$ (5.10). This can be viewed as a null control problem, the control being $\delta\varphi^0$. And for this choice of $\delta\varphi^0$, $\delta J(u^0)[\delta u^0, \delta\varphi^0] \leq 0$. But the computation of the solution of this control problem (5.9) is expensive, since it requires us to obtain the coefficients of (5.9), i.e. u and δu at both sides of Σ and then to develop a numerical solver for the control of (5.9). For that reason we only consider the perturbations introduced in (5.2) and (5.4). These directions conserve the structure of the shock and change the profile of the solution. See Figures 8, 9, 10 and 11 for a numerical example where the first perturbation is computed.

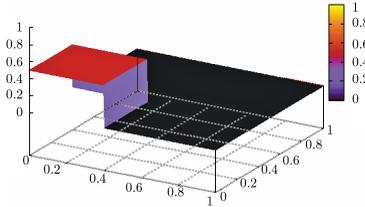


FIGURE 8. Initial condition u^0 .

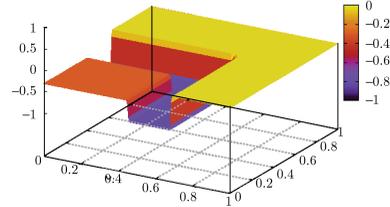


FIGURE 9. Datum of the adjoint system at time $t = T$, p^T .

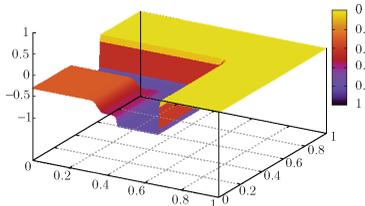


FIGURE 10. Solution of the adjoint system p^0 .

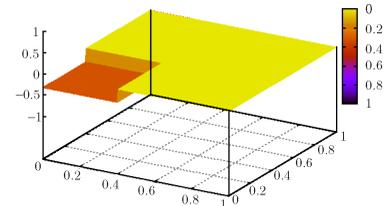


FIGURE 11. First direction δu^0 .

In the next section we explain how to implement a descent algorithm following these ideas.

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 look for descent directions for a discrete version of the functional J . In this section we discuss two possibilities for searching them based either on the discrete or the continuous approaches.

The discrete approach consists mainly in applying a descent algorithm to the discrete version J_Δ of the functional J . The alternating descent method, in contrast, is a continuous method based on the analysis of the previous section in which the two main classes of directions are identified.

Let us first discuss the discrete approach.

6.1. The discrete approach. Let us consider the approximation of the functional J by J_Δ defined (1.1) and (3.1), respectively. Consider a 3-point conservative numerical approximation scheme for (1.2), using a dimensional splitting in both variables, i.e. we use (3.17) and (3.15) with in each component H_Δ^x and H_Δ^y . We shall use a 1D Engquist-Osher scheme, with the following numerical flux:

$$\begin{aligned}
 g_1(a, b) &= \frac{1}{2} \left(f_1(a) + f_1(b) - \int_b^a |f_1'(z)| dz \right), \\
 g_2(a, b) &= \frac{1}{2} \left(f_2(a) + f_2(b) - \int_b^a |f_2'(z)| dz \right).
 \end{aligned}
 \tag{6.1}$$

More explicitly, replacing (3.15) in (3.17), we obtain

$$u_{i,j}^{n+\frac{1}{2}} = u_{i,j}^n - \frac{\Delta t}{\Delta x} \left(g_1(u_{i+1,j}^n, u_{i,j}^n) - g_1(u_{i,j}^n, u_{i-1,j}^n) \right),$$

$i \in \mathbb{Z}, n = 0, \dots, N,$

$$u_{i,j}^{n+1} = u_{i,j}^{n+\frac{1}{2}} - \frac{\Delta t}{\Delta y} \left(g_2(u_{i,j+1}^{n+\frac{1}{2}}, u_{i,j}^{n+\frac{1}{2}}) - g_2(u_{i,j}^{n+\frac{1}{2}}, u_{i,j-1}^{n+\frac{1}{2}}) \right),$$

$j \in \mathbb{Z}, n = 0, \dots, N,$

where g_1, g_2 are the numerical fluxes defined in (6.1).

The gradient of the discrete functional J_Δ requires computing one derivative of J_Δ with respect to each node of the mesh. This can be done in a cheaper way using the adjoint state. We illustrate it for the Engquist-Osher numerical scheme. However, as the discrete functionals J_Δ are not necessarily convex, the gradient methods could possibly provide sequences that do not converge to a global minimizer of J_Δ . 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 gradient methods are satisfactory, although convergence is slow due to unnecessary oscillations that the descent method introduces. The gradient of J_Δ , rigorously speaking, requires the linearization of the numerical scheme (6.2) and (6.3) used to approximate the equation (1.2).

Then the linearization corresponds to

$$\begin{aligned}
 \delta u_{i,j}^{n+\frac{1}{2}} &= \delta u_{i,j}^n - \frac{\Delta t}{\Delta x} (\partial_a g_1(u_{i+1,j}^n, u_{i,j}^n) \delta u_{i+1,j}^n - \partial_b g_1(u_{i,j}^n, u_{i-1,j}^n) \delta u_{i-1,j}^n) \\
 &\quad - \frac{\Delta t}{\Delta x} (\partial_b g_1(u_{i+1,j}^n, u_{i,j}^n) - \partial_a g_1(u_{i,j}^n, u_{i-1,j}^n)) \delta u_{i,j}^n, \\
 (6.4) \quad &i, j \in \mathbb{Z}, \quad n = 0, \dots, N,
 \end{aligned}$$

$$\begin{aligned}
 \delta u_{i,j}^{n+1} &= \delta u_{i,j}^{n+\frac{1}{2}} - \frac{\Delta t}{\Delta y} (\partial_a g_2(u_{i,j+1}^{n+\frac{1}{2}}, u_{i,j}^{n+\frac{1}{2}}) \delta u_{i,j+1}^{n+\frac{1}{2}} - \partial_b g_2(u_{i,j}^{n+\frac{1}{2}}, u_{i,j-1}^{n+\frac{1}{2}}) \delta u_{i,j-1}^{n+\frac{1}{2}}) \\
 &\quad - \frac{\Delta t}{\Delta y} (\partial_b g_2(u_{i,j+1}^{n+\frac{1}{2}}, u_{i,j}^{n+\frac{1}{2}}) - \partial_a g_2(u_{i,j}^{n+\frac{1}{2}}, u_{i,j-1}^{n+\frac{1}{2}})) \delta u_{i,j}^{n+\frac{1}{2}}, \\
 (6.5) \quad &i, j \in \mathbb{Z}, \quad n = 0, \dots, N.
 \end{aligned}$$

In view of this, the discrete adjoint system of (6.4) and (6.5) can also be written for the differentiable flux functions (6.1):

$$(6.6) \quad p_{i,j}^{N+1} = p_{i,j}^T, \quad i, j \in \mathbb{Z},$$

$$\begin{aligned}
 (6.7) \quad p_{i,j}^{n+\frac{1}{2}} &= p_{i,j}^{n+1} + \frac{\Delta t}{\Delta y} \partial_b g_2(u_{i,j+1}^{n+\frac{1}{2}}, u_{i,j}^{n+\frac{1}{2}}) (p_{i,j+1}^{n+1} - p_{i,j}^{n+1}) \\
 &\quad + \frac{\Delta t}{\Delta y} \partial_a g_2(u_{i,j}^{n+\frac{1}{2}}, u_{i,j-1}^{n+\frac{1}{2}}) (p_{i,j-1}^{n+1} - p_{i,j}^{n+1}), \\
 &i, j \in \mathbb{Z}, \quad n = 0, \dots, N,
 \end{aligned}$$

$$\begin{aligned}
 (6.8) \quad p_{i,j}^n &= p_{i,j}^{n+\frac{1}{2}} + \frac{\Delta t}{\Delta x} \partial_b g_1(u_{i+1,j}^n, u_{i,j}^n) (p_{i+1,j}^{n+\frac{1}{2}} - p_{i,j}^{n+\frac{1}{2}}) \\
 &\quad + \frac{\Delta t}{\Delta x} \partial_a g_1(u_{i,j}^n, u_{i-1,j}^n) (p_{i,j}^{n+\frac{1}{2}} - p_{i-1,j}^{n+\frac{1}{2}}), \\
 &i, j \in \mathbb{Z}, \quad n = 0, \dots, N.
 \end{aligned}$$

In fact, when multiplying the equations in (6.5) by $p_{i,j}^{n+1}$ and adding in $i, j \in \mathbb{Z}$ and $n = 0, \dots, N$, the following identity is easily obtained:

$$(6.9) \quad \Delta x \Delta y \sum_{i,j \in \mathbb{Z}} p_{i,j}^T \delta u_{i,j}^{N+1} = \Delta x \Delta y \sum_{i,j \in \mathbb{Z}} p_{i,j}^0 \delta u_{i,j}^0.$$

This is the discrete version of formula (4.4) which allows us to simplify the derivative of the discrete cost functional.

Thus, for any variation δu_{Δ}^0 , the Gâteaux derivative of the cost functional defined in (3.1) is given by

$$(6.10) \quad \delta J_{\Delta} = \Delta x \Delta y \sum_{i,j \in \mathbb{Z}} (u_{i,j}^{N+1} - u_{i,j}^d) \delta u_{i,j}^{N+1},$$

where $\delta u_{i,j}^n$ solves the linearized system (6.4) and (6.5). If we consider $p_{i,j}^n$ the solution of (6.6), (6.8), (6.9) with final datum

$$p_{i,j}^T = u_{i,j}^{N+1} - u_{i,j}^d, \quad \forall i, j \in \mathbb{Z},$$

then δJ_{Δ} in (6.10) can be written as

$$\delta J_{\Delta} = \Delta x \Delta y \sum_{i,j \in \mathbb{Z}} p_{i,j}^0 \delta u_{i,j}^0,$$

and this allows us to obtain easily the steepest descent direction for J_Δ by considering

$$(6.11) \quad \delta u_\Delta^0 = -p_\Delta^0.$$

Remark 6.1. 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 (4.39), (4.40), (4.41), (4.42), (4.43), and (4.44), constitutes an interesting problem for future research. We refer to [25] and [34] for preliminary work on this direction in 1D.

6.2. The alternating descent method in 2D. Now we explain how we implement the method proposed in Subsection 5.2. The main idea is to approximate a minimizer of J alternating with three directions: First we perturb the initial datum u^0 at side Q_-^0 and find the motion of the curve which produces the decrease of J . Second we move the shock curve without altering the profile of u^0 at both sides of Σ . Finally the third direction perturbs the profile in Q_+^0 and we find the motion of the curve which produces the decrease of J .

More precisely, for a given initialization u^0 and target function u^d , we implement the following Algorithm 1, iterating it until we reach the stopping criterion of the minimization process.

Algorithm 1 Description of the alternating descent method in 2D

- 1: From u_Δ^0 , we compute u_Δ , solving the numerical system (6.2), (6.3).
 - 2: From u_Δ and u_Δ^d , we compute p_Δ^0 , solving the numerical system (6.6), (6.8), (6.9).
 - 3: From u_Δ^T , we compute the curve Σ_Δ^T , and then X_Δ^- and X_Δ^+ .
 - 4: From u_Δ^0 , we compute the curve Σ_Δ^0 .
 - 5: From p_Δ^0 , X_Δ^- and Σ_Δ^0 , we compute $\delta(u_-^0)_\Delta$ (which corresponds to δu_Δ^0 in the first perturbations (5.2)).
 - 6: Find the optimal step size ε for which the datum u_Δ^0 must be modified in the direction given by $\delta(u_-^0)_\Delta$.
 - 7: Update $u_\Delta^0 = u_\Delta^{0,\varepsilon}$.
 - 8: Repeat the steps (1), (2), and (4).
 - 9: From p_Δ^0 and Σ_Δ^0 , evaluating p_Δ^0 on Σ_Δ^0 , we compute q_Δ^0 first and, then, the direction (5.3).
 - 10: Find the optimal step size ε for which this datum u_Δ^0 must be modified in the direction given by (5.3) $\Sigma_\Delta^{0,\varepsilon} = \Sigma_\Delta^0 + \varepsilon \delta \varphi_\Delta^0$ (we move the curve Σ_Δ^0 in $\varepsilon \delta \varphi_\Delta^0$).
 - 11: Repeat (7), (1), (2), (3), and (4).
 - 12: From p_Δ^0 , X_Δ^+ , and Σ_Δ^0 , we compute $\delta(u_+^0)_\Delta$ (which corresponds to δu_Δ^0 in the third class of perturbations (5.4)).
 - 13: Find the optimal step size ε for which the datum u_Δ^0 must be modified in the direction given by $\delta(u_+^0)_\Delta$.
 - 14: Repeat (7).
 - 15: Stopping criterion: The algorithm is iterated, starting from the new initial datum, until the functional J_Δ takes a value smaller than a given tolerance.
-

The main advantage of this method 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 by modifying the old one, in the direction of this generalized tangent vector, will have again a single discontinuity. We have presented here the method in the particular case in which both the target u^d and the initial datum u^0 that initializes the process have one single shock discontinuity. But these ideas can be applied in a much more general context in which the number of shocks does not necessarily coincide. In particular, in 1D this method is able both to generate shocks and to destroy them, if any of these facts contribute 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 in the present setting would correspond to controlling the location of the shock) are combined with classical shape deformations (that would correspond to simply shaping the solution away from the shock in the present setting); see [20].

7. NUMERICAL IMPLEMENTATION

In this section we explain the main computational ingredients entering in the implementation of the alternating descent method in 2D, and its solutions from a numerical point of view.

The first difficulty arises when determining the boundary conditions on the numerical scheme (especially if the shock interacts with the boundary): The behavior of boundary conditions is well understood for 1D scalar problems, but their treatment is less clear for systems and in 2D (Dubois and LeFloch, 1988 [17]). For that reason we employ a ghost cell strategy in our numerical schemes. Using the maximum principle and the finite velocity of propagation that entropy solutions satisfy, we obtain an efficient numerical implementation. The ghost cell strategy consists in considering a larger computational domain, containing the support of the solution $\{u_\Delta^n\}_{n=0,\dots,N}$ of (6.2) and (6.3) for all $n = 0, \dots, N + 1$. Its size can be estimated from the initial condition u_Δ^0 , thanks to the finite velocity of propagation. The functional J_Δ is then localized in a smaller domain such that the boundary effects in the larger one do not affect its value.

On the other hand, we did not develop a specific method to approximate the continuous adjoint system. We rather use the discrete adjoint as an approximation, although its validity is not analytically established.

Now, we comment on the main difficulties encountered in the implementation of Algorithm 1 and the way we overcome them:

Edge detector: One of the key steps in the implementation of the alternating descent method is the identification of the shock location. This requires a numerical method to recover the curves Σ^T and Σ^0 . We describe a subroutine to find the discontinuities of a vector $u_\Delta = \{u_{i,j}\}_{i,j=1,\dots,N}$ based on a shift condition. Roughly, we introduce a parameter α , and we look for the indexes $i, j \in \{1, \dots, N - 2\}$, where

$$(7.1) \quad D_\Delta(u_\Delta, i, j) > \alpha, \text{ and } \left(D_\Delta(u_\Delta, i + 1, j) \leq \alpha, \text{ or } D_\Delta(u_\Delta, i, j + 1) \leq \alpha \right),$$

where $D_\Delta(u_\Delta, i, j) = |u_{i+1, j} - u_{i, j}| + |u_{i, j+1} - u_{i, j}|$. The condition

$$D_\Delta(u_\Delta, i, j) > \alpha$$

means that the discrete gradient is greater than α in the node i, j . In this way, we obtain a region of nodes fulfilling this condition. The other condition in (7.1) is to identify a curve and not a region. To simplify the presentation, we consider the case in which only one discontinuity curve is relevant in the numerical experiment, and we identify it on the discrete collection of the indexes by the above criterium. See Algorithm 2 for more details. For a more sophisticated algorithm see [11].

Algorithm 2 Edge detector

$$\{u_{i, j}\}_{i, j=1, \dots, N} \rightarrow (\sigma_\Delta, u_-, u_+)$$

```

1: input  $\Delta x, \Delta y, \{u_{i, j}\}_{i, j=1, \dots, N}, \alpha$  ▷  $\alpha$  is the jump sensibility parameter
2:  $\sigma_\Delta := \emptyset, u_- := \emptyset, u_+ = \emptyset$  ▷ Initialize the curve
3: for  $i = 0$  to  $N - 2$  do
4:   for  $j = 0$  to  $N - 2$  do
5:     if  $D_\Delta(\{u_{i, j}\}_{i, j=1, \dots, N}, i, j) > \alpha$  then
6:       if  $D_\Delta(\{u_{i, j}\}_{i, j=1, \dots, N}, i+1, j) \leq \alpha$  and  $D_\Delta(\{u_{i, j}\}_{i, j=1, \dots, N}, i, j+1) \leq$ 
        $\alpha$  then
7:          $\sigma_\Delta \leftarrow \{i, j\}$  ▷ Storage of the index of the shock curve
8:          $u_- \leftarrow u_{i, j}$  ▷  $u$  at the left of the curve
9:          $u_+ \leftarrow (u_{i+1, j} + u_{i, j+1})/2$  ▷  $u$  at the right of the curve
10:        else if  $D_\Delta(\{u_{i, j}\}_{i, j=1, \dots, N}, i+1, j) \leq \alpha$  then
11:           $\sigma_\Delta \leftarrow \{i, j\}$  ▷ Storage of the index of the shock curve
12:           $u_- \leftarrow u_{i, j}$  ▷  $u$  at the left of the curve
13:           $u_+ \leftarrow u_{i+1, j}$  ▷  $u$  at the right of the curve
14:          else if  $D_\Delta(\{u_{i, j}\}_{i, j=1, \dots, N}, i, j+1) \leq \alpha$  then
15:             $\sigma_\Delta \leftarrow \{i, j\}$  ▷ Storage of the index of the shock curve
16:             $u_- \leftarrow u_{i, j}$  ▷  $u$  at the left of the curve
17:             $u_+ \leftarrow u_{i, j+1}$  ▷  $u$  at the right of the curve
18:          end if
19:        end if
20:      end for
21:    end for

```

Segmentation problem: In lines 5 and 12 of Algorithm 1 it is necessary to restrict the function p^0 to the regions $[X^-, \Sigma^0]$ and $[\Sigma^0, X^+]$, respectively. This requires an algorithm to identify whether a point is to the right or the left side of a given curve. This is a segmentation problem in image processing that we solve using the winding number algorithm. $w(a)$, the winding number of a closed curve (polygon in the discrete case) C about a point a in the plane, is defined as a contour integral in the complex plane

$$W(a) = \frac{1}{2\pi i} \oint_C \frac{1}{z - a} dz$$

that measures not only whether C encloses a , but also how many times and in which orientation C “winds around” a . In particular

$$w(a) = \begin{cases} 0 & \text{if } a \text{ is not inside } C, \\ n > 0 & \text{if } C \text{ winds around } a \text{ } n \text{ times counterclockwise,} \\ n < 0 & \text{if } C \text{ winds around } a \text{ } -n \text{ times clockwise.} \end{cases}$$

Note that the winding number is not defined when the point a is on the curve C . The winding number algorithm consists in computing, for a given point a , the winding number $w(a)$ with respect to the curve C (oriented counterclockwise). If the winding number $w(a)$ is positive, the point lies inside the curve. There exists various versions of this algorithm employing different discretization of the above integral. For more details see [16], [26].

Inpainting problem: In steps 5 and 12 in Algorithm 1 we need to extend the function $p(\cdot, x)$ in the corresponding regions. As in Remark 5.1 we see there exists a rich literature on this problem. The common strategy is to consider the minimization of a discrete version of a variational problem (5.8) and to apply a descent algorithm, for example the steepest descent algorithm.

For our examples in the next section, $-p^0$ is a constant step function. And a strategy to extend $-p^0$ on $[X^-, \Sigma^0]$ and over $[\Sigma^0, X^+]$, such that the extension only has a shock on Σ^0 , consists in extending the constant value inside the corresponding regions.

Motion and deformation of the shock curve: To produce the motion of the curve, we use the Fast Marching Method introduced by J. A. Sethian; see [32]. After moving the shock curve, we need to change the value of the profile u^0 at both sides of the new curve, following the normal lines, using a discrete version of (4.12).

8. NUMERICAL EXPERIMENTS

In this section we present two numerical experiments illustrating the results obtained in an optimization model problem with each one of the numerical methods described in the previous section. We have chosen as the computational domain the square $(0, 1) \times (0, 1)$.

8.1. **Experiment 1.** In this experiment we consider the conservation law

$$(8.1) \quad \partial_t u + (u^2/2)_x + (u^4/4)_x = 0, \text{ in } \mathbb{R}^2 \times (0, T),$$

the time horizon $T = 0.2$, and the target function u^d given by

$$(8.2) \quad u^d(x, y) = \begin{cases} 0.7 & x \leq 0.87, \ y \leq 0.76715, \\ 0 & \text{otherwise.} \end{cases}$$

To initialize the iterative descent method, we choose an initial datum u^0 in such a way that the solution at time $t = T$ has a profile similar to u^d , i.e. it is a Lipschitz continuous function with a single discontinuity, located on a curve $\Sigma^0 \subset \mathbb{R}^2$. For example,

$$(8.3) \quad u^0(x, y) = \begin{cases} 0.4 & x \leq 0.2, \ y \leq 0.4, \\ 0 & \text{otherwise.} \end{cases}$$

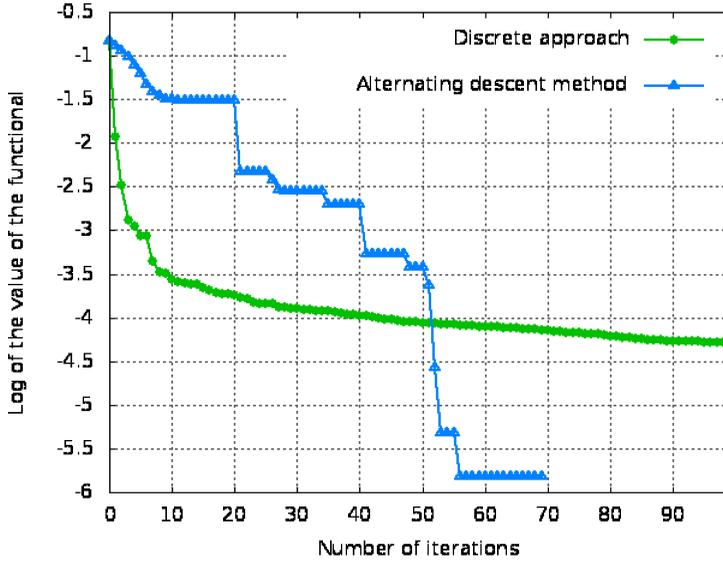


FIGURE 12. Experiment 1. Log of the value of the functional versus the number of iterations in the descent algorithm for the discrete approach scheme and the 2D alternating descent method.

We fix the computational domain $(0, 1) \times (0, 1)$ and the mesh-size $\Delta x = \Delta y = \Delta t = 1/200$.

The CFL-condition for the numerical scheme (6.2) and (6.3) is

$$(8.4) \quad \max \left\{ \frac{\Delta t}{\Delta x} \sup_{x \in \mathbb{R}} |f'_1(u^0(x))|, \frac{\Delta t}{\Delta y} \sup_{x \in \mathbb{R}} |f'_2(u^0(x))| \right\} \leq 1,$$

where $f(z) = (f_1(z), f_2(z))$. Thus, the CFL-condition in our experiment corresponds to

$$(8.5) \quad \max \left\{ \sup_{x \in \mathbb{R}} |u^0(x)|, \sup_{x \in \mathbb{R}} |u^0(x)|^3 \right\} \leq 1.$$

Therefore u^0 , given by (8.3), satisfies (8.5).

In this form, one minimizer is

$$(8.6) \quad u_*^0(x, y) = \begin{cases} 0.7 & x \leq 0.8, \ y \leq 0.75, \\ 0 & \text{otherwise.} \end{cases}$$

In Figure 12 we plot $\log(J)$ with respect to the number of iterations, for both the alternating descent method proposed in this article and the purely discrete approach. We see that the method introduced in this work stabilizes in fewer iterations.

In Figures 16 and 17 we observe the minimizers obtained by the methods above, and the associated solutions in time T in Figures 18 and 19.

The function u^0 obtained by the alternating descent method (Figure 16) is a good approximation of (8.6). The solution given by the discrete approach (Figure 17)

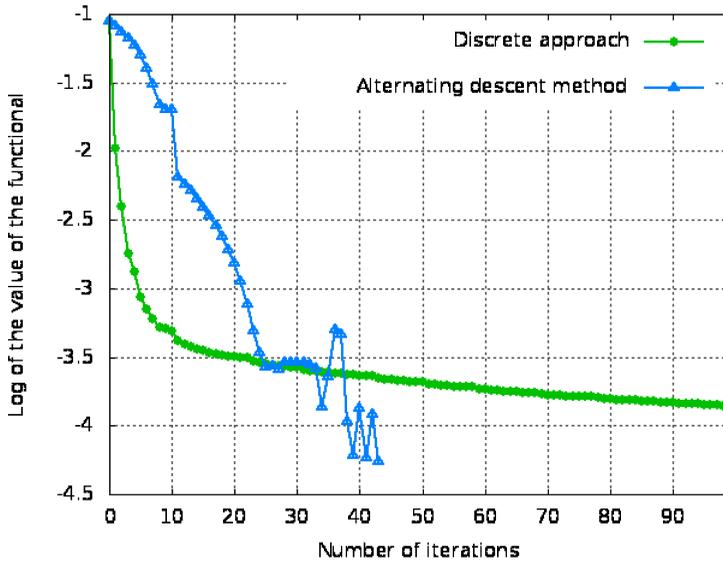


FIGURE 13. Experiment 2. Log of the value of the functional versus the number of iterations in the descent algorithm for the discrete approach scheme and the 2D alternating descent method.

presents high oscillations, and this can produce problems with the CFL-condition (8.5). In Figures 17 and 19 we observe that the solution of the discrete approach preserves a memory of the initialization u^0 in (8.3), i.e. the solution of the discrete approach has oscillations around the shock location of the initial condition u^0 in (8.3).

8.2. Experiment 2. In this experiment we consider the same conservation law (8.1) and the horizon time $T = 0.2$. The target function u^d corresponds to the solution of the numerical scheme (6.2) and (6.3) with the initial condition given by

$$(8.7) \quad u_*^0(x, y) = \begin{cases} 0.7 & (x, y) \notin \mathbb{R}_+ \times \mathbb{R}_+, \ x \leq 0.7, \ y \leq 0.7, \\ 0.7 & (x, y) \in \mathbb{R}_+ \times \mathbb{R}_+, \ x^2 + y^2 \leq (0.7)^2, \\ 0 & \text{otherwise.} \end{cases}$$

To initialize the iterative descent method, we choose an initial datum u^0 given by

$$(8.8) \quad u^0(x, y) = \begin{cases} 0.4 & x \leq 0.3, \ y \leq 0.3, \\ 0 & \text{otherwise,} \end{cases}$$

and the computational domain $(0, 1) \times (0, 1)$, with $\Delta x = \Delta y = \Delta t = 1/400$.

In Figure 13 we plot $\log(J)$ with respect to the number of iterations for both the alternating descent method proposed in this article and the purely discrete approach.

In Figures 22-24, we illustrate the minimizers obtained by the methods above, and the associated solutions in time T .

9. OPEN PROBLEMS

We present some open problems, possible extensions of this work, and the methods we have developed.

A significant number of issues arise even in the 1D case. We mention here some of them:

- Further develop and explore the efficiency of the alternating descent method for interacting shocks.
- Full implementation of the continuous approach developing a specific solver to approximate the adjoint system and to compute numerical approximations of generalized gradients.
- Analysis of the efficiency of a purely discrete method in which the adjoint system is resolved in a coarser grid or, even, by a different more diffusive scheme as Lax-Friedrichs, in opposition to the solver of the flow equation, taken to be, for instance Engquist-Osher, to better resolve the shocks.
- Adaptation of the alternating descent method to viscous equations with small viscosity; see [13] for preliminary work on this direction.
- Adaptation of the alternating descent method to systems of conservation laws.

In the multi-dimensional case the 1D problems above can also be addressed, of course, but some other specific issues related to the multi-dimensional aspects also emerge. We mention here some of them:

- The problem of interacting shocks in the multi-dimensional setting could be of particular interest due to its geometric complexity.
- Take account of the finiteness of the multi-dimension shock and the added terms that its extremes or boundaries may lead to in the linearization of the functionals.
- Develop specific numerical methods for approximation the linearized equations and the adjoint system, paying special attention to the transport equation governing the sensivity of the shocks.

10. FURTHER NUMERICAL EXPERIMENTS

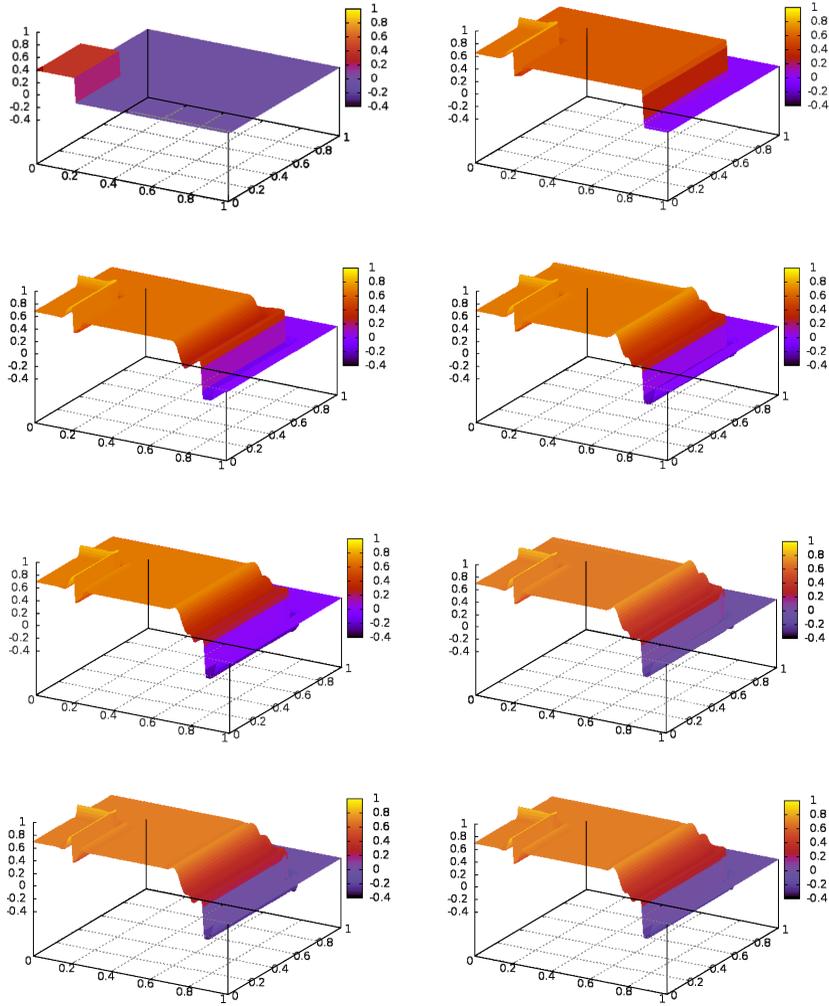


FIGURE 14. Iterations of the discrete approach in Experiment 1.

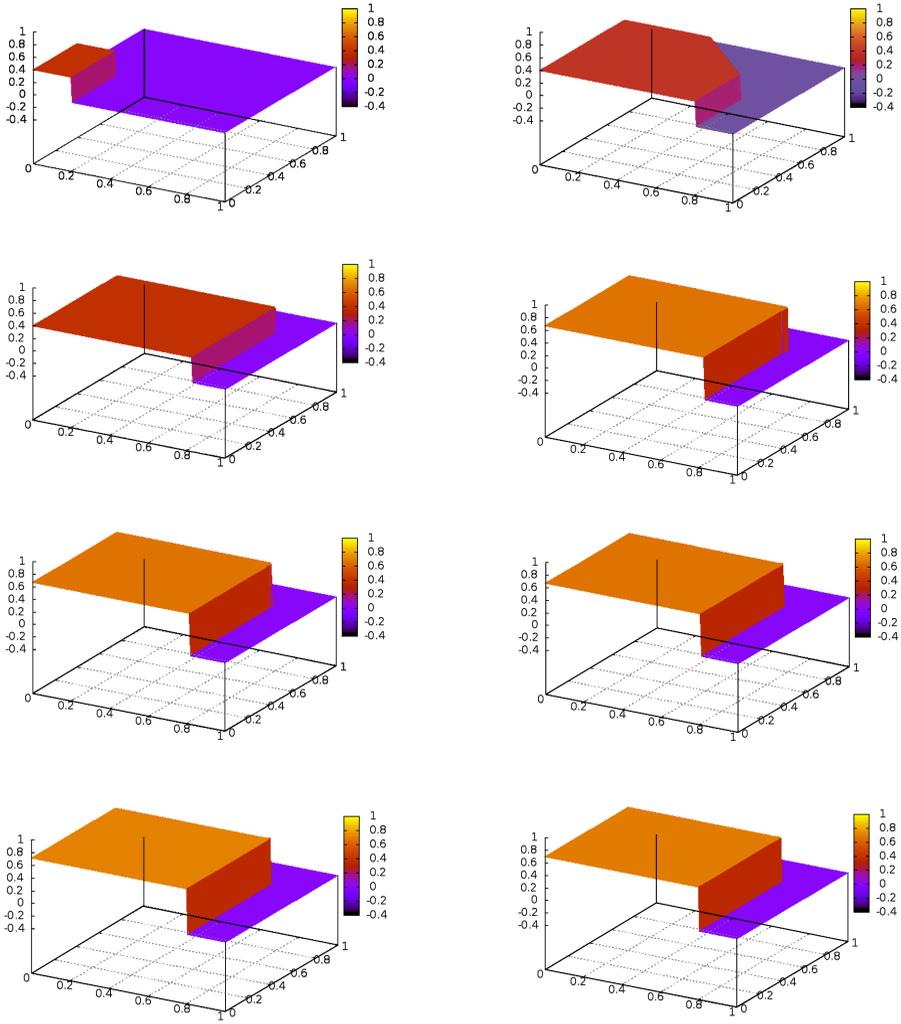


FIGURE 15. Iterations of the alternating descent method in Experiment 1.

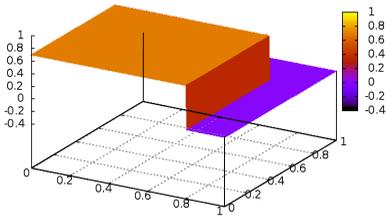


FIGURE 16. u^0 , alternating descent method, iteration $k = 56$.

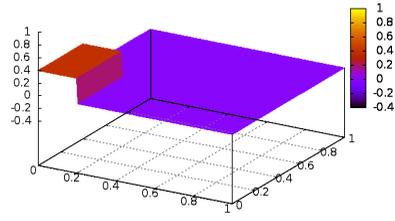


FIGURE 17. u^0 , the discrete approach, iteration $k = 99$.

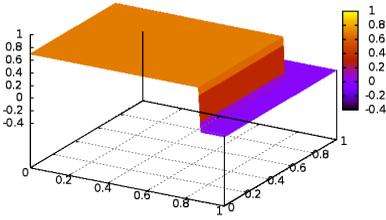


FIGURE 18. u^T , alternating descent method, iteration $k = 56$.

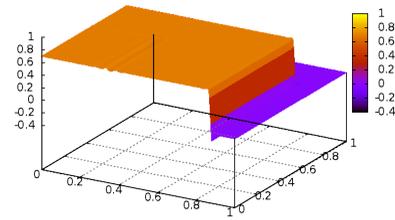


FIGURE 19. u^T , the discrete approach, iteration $k = 99$.

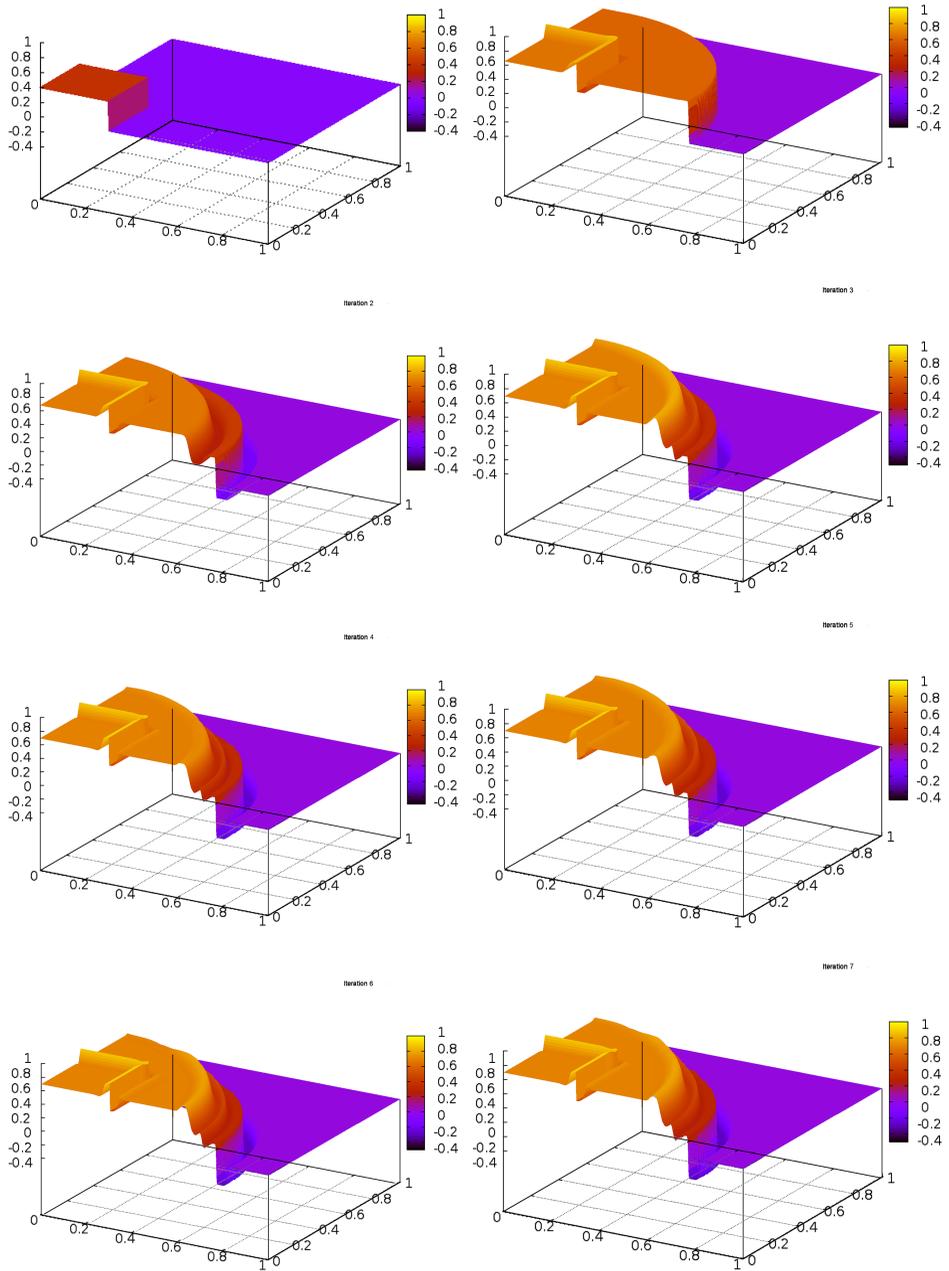


FIGURE 20. Iterations of the discrete approach in Experiment 2.

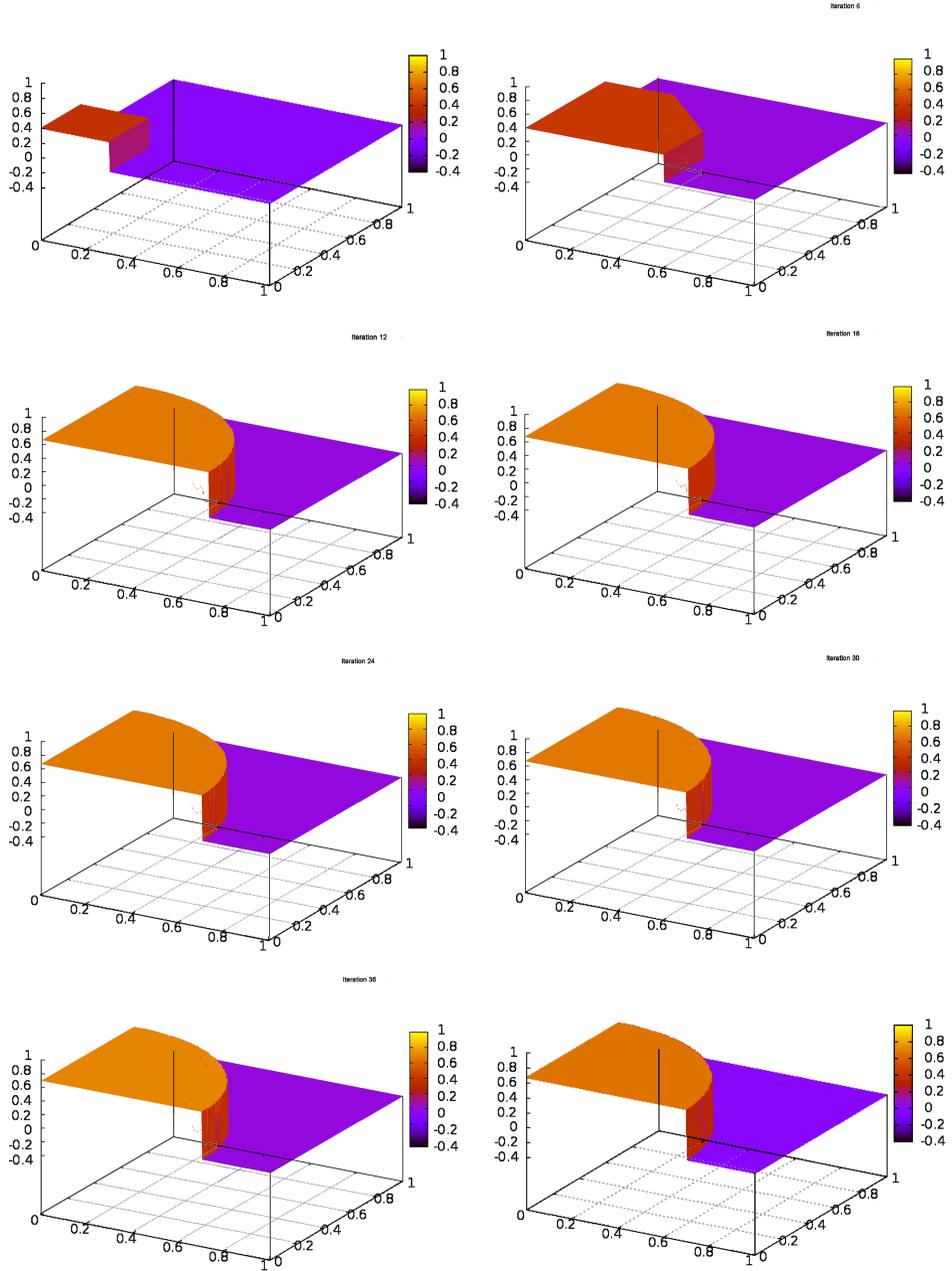


FIGURE 21. Iterations of the alternating descent method in Experiment 2.

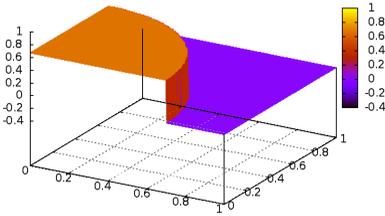


FIGURE 22. u^0 , alternating descent method, iteration $k=43$.

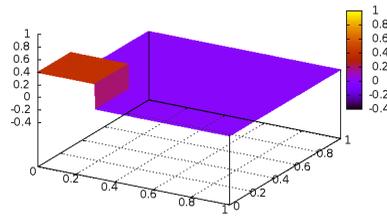


FIGURE 23. u^0 , the discrete approach, iteration $k=99$.

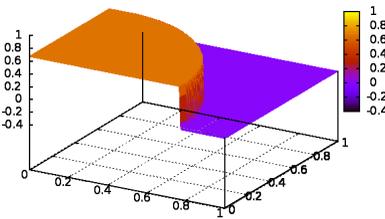


FIGURE 24. u^T , alternating descent method, iteration $k=43$.

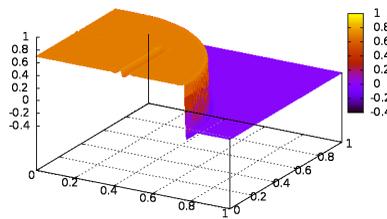


FIGURE 25. u^T , the discrete approach, iteration $k=99$.

ACKNOWLEDGEMENTS

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

REFERENCES

- [1] A. Adimurthi, S. Ghoshal, and V. Gowda, *Exact controllability of scalar conservation law with strict convex flux*, <http://hal.upmc.fr/docs/00/87/35/53/PDF/ExactControl.pdf>, preprint, 2011 [hal-00873553].
- [2] F. Ancona and G. M. Coclite, *On the attainable set for Temple class systems with boundary controls*, *SIAM J. Control Optim.* **43** (2005), no. 6, 2166–2190 (electronic), DOI 10.1137/S0363012902407776. MR2179483 (2006f:93008)
- [3] F. Ancona and A. Marson, *On the attainable set for scalar nonlinear conservation laws with boundary control*, *SIAM J. Control Optim.* **36** (1998), no. 1, 290–312 (electronic), DOI 10.1137/S0363012996304407. MR1616586 (99h:93008)
- [4] D. Auroux and J. Blum, *Back and forth nudging algorithm for data assimilation problems* (English, with English and French summaries), *C. R. Math. Acad. Sci. Paris* **340** (2005), no. 12, 873–878, DOI 10.1016/j.crma.2005.05.006. MR2151776 (2006a:34019)
- [5] C. Bardos and O. Pironneau, *A formalism for the differentiation of conservation laws* (English, with English and French summaries), *C. R. Math. Acad. Sci. Paris* **335** (2002), no. 10, 839–845, DOI 10.1016/S1631-073X(02)02574-8. MR1947710 (2004c:35265)
- [6] F. Bouchut and F. James, *One-dimensional transport equations with discontinuous coefficients*, *Nonlinear Anal.* **32** (1998), no. 7, 891–933, DOI 10.1016/S0362-546X(97)00536-1. MR1618393 (2000a:35243)
- [7] F. Bouchut and F. James, *Differentiability with respect to initial data for a scalar conservation law*, *Hyperbolic problems: theory, numerics, applications*, Vol. I (Zürich, 1998), *Internat. Ser. Numer. Math.*, vol. 129, Birkhäuser, Basel, 1999, pp. 113–118. MR1715739 (2000g:35137)

- [8] T. Boukadida and A. Y. LeRoux, *A new version of the two-dimensional Lax-Friedrichs scheme*, Math. Comp. **63** (1994), no. 208, 541–553, DOI 10.2307/2153282. MR1242059 (95a:65124)
- [9] A. Bressan and A. Marson, *A maximum principle for optimally controlled systems of conservation laws*, Rend. Sem. Mat. Univ. Padova **94** (1995), 79–94. MR1370904 (97c:49025)
- [10] A. Bressan and A. Marson, *A variational calculus for discontinuous solutions of systems of conservation laws*, Comm. Partial Differential Equations **20** (1995), no. 9-10, 1491–1552, DOI 10.1080/03605309508821142. MR1349222 (96g:35120)
- [11] J. Canny, *A computational approach to edge detection*, Pattern Analysis and Machine Intelligence, IEEE Transactions on **PAMI-8** (1986), no. 6, 679–698.
- [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** (2008), no. 3, 369–416, DOI 10.1142/S0218202508002723. MR2397976 (2009c:35392)
- [13] C. Castro, F. Palacios, and E. Zuazua, *Optimal control and vanishing viscosity for the Burgers equation*, Integral methods in science and engineering. Vol. 2, Birkhäuser Boston, Inc., Boston, MA, 2010, pp. 65–90, DOI 10.1007/978-0-8176-4897-8_7. MR2663150
- [14] C. Castro and E. Zuazua, *Flux identification for 1-d scalar conservation laws in the presence of shocks*, Math. Comp. **80** (2011), no. 276, 2025–2070, DOI 10.1090/S0025-5718-2011-02465-8. MR2813348 (2012g:49079)
- [15] M. Crandall and A. Majda, *The method of fractional steps for conservation laws*, Numer. Math. **34** (1980), no. 3, 285–314, DOI 10.1007/BF01396704. MR571291 (81j:65101)
- [16] M. W. D. Davis and M. David, *An algorithm for finding the position of a point relative to a fixed polygonal boundary*, J. Internat. Assoc. Math. Geol. **12** (1980), no. 1, 61–68, DOI 10.1007/BF01039904. MR594012 (81j:51001)
- [17] F. Dubois and P. LeFloch, *Boundary conditions for nonlinear hyperbolic systems of conservation laws*, J. Differential Equations **71** (1988), no. 1, 93–122, DOI 10.1016/0022-0396(88)90040-X. MR922200 (89c:35099)
- [18] L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998. MR1625845 (99e:35001)
- [19] C. Fabre, J.-P. Puel, and E. Zuazua, *On the density of the range of the semigroup for semilinear heat equations*, Control and optimal design of distributed parameter systems (Minneapolis, MN, 1992), IMA Vol. Math. Appl., vol. 70, Springer, New York, 1995, pp. 73–91, DOI 10.1007/978-1-4613-8460-1_4. MR1345629 (97a:35097)
- [20] S. Garreau, P. Guillaume, and M. Masmoudi, *The topological asymptotic for PDE systems: the elasticity case*, SIAM J. Control Optim. **39** (2001), no. 6, 1756–1778 (electronic), DOI 10.1137/S0363012900369538. MR1825864 (2002m:49062)
- [21] M. B. Giles and N. A. Pierce, *Analytic adjoint solutions for the quasi-one-dimensional Euler equations*, J. Fluid Mech. **426** (2001), 327–345, DOI 10.1017/S0022112000002366. MR1819479 (2002a:76037)
- [22] E. Godlewski and P.-A. Raviart, *Hyperbolic Systems of Conservation Laws*, Mathématiques & Applications (Paris) [Mathematics and Applications], vol. 3/4, Ellipses, Paris, 1991. MR1304494 (95i:65146)
- [23] E. Godlewski and P.-A. Raviart, *The linearized stability of solutions of nonlinear hyperbolic systems of conservation laws. A general numerical approach*, Math. Comput. Simulation **50** (1999), no. 1-4, 77–95, DOI 10.1016/S0378-4754(99)00062-2. Modelling '98 (Prague). MR1717658 (2000i:35119)
- [24] L. Gosse, *A two-dimensional version of the Godunov scheme for scalar balance laws*, SIAM J. Numer. Anal. **52** (2014), no. 2, 626–652, DOI 10.1137/130925906. MR3179555
- [25] L. Gosse and F. James, *Numerical approximations of one-dimensional linear conservation equations with discontinuous coefficients*, Math. Comp. **69** (2000), no. 231, 987–1015, DOI 10.1090/S0025-5718-00-01185-6. MR1670896 (2000j:65077)
- [26] J. Hall, *Ptloc-a fortran subroutine for determining the position of a point relative to a closed boundary*, Journal of the International Association for Mathematical Geology **7** (1975), no. 1, 75–79 (English).
- [27] F. James and M. Sepúlveda, *Convergence results for the flux identification in a scalar conservation law*, SIAM J. Control Optim. **37** (1999), no. 3, 869–891 (electronic), DOI 10.1137/S0363012996272722. MR1680830 (2000b:65171)

- [28] S. N. Kružkov, *First order quasilinear equations with several independent variables.* (Russian), *Mat. Sb. (N.S.)* **81 (123)** (1970), 228–255. MR0267257 (42 #2159)
- [29] A. Majda, *The stability of multidimensional shock fronts*, *Mem. Amer. Math. Soc.* **41** (1983), no. 275, iv+95, DOI 10.1090/memo/0275. MR683422 (84e:35100)
- [30] D. Mumford and J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, *Comm. Pure Appl. Math.* **42** (1989), no. 5, 577–685, DOI 10.1002/cpa.3160420503. MR997568 (90g:49033)
- [31] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, *Phys. D* **60** (1992), no. 1-4, 259–268.
- [32] J. A. Sethian, *Level Set Methods and Fast Marching Methods*, *Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*, 2nd ed., *Cambridge Monographs on Applied and Computational Mathematics*, vol. 3, Cambridge University Press, Cambridge, 1999. MR1700751 (2000c:65015)
- [33] S. Ulbrich, *Adjoint-based derivative computations for the optimal control of discontinuous solutions of hyperbolic conservation laws*, *Optimization and control of distributed systems*, *Systems Control Lett.* **48** (2003), no. 3-4, 313–328, DOI 10.1016/S0167-6911(02)00275-X. MR2020647 (2005g:49042)
- [34] 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** (2008), no. 1, 1–22. MR2378582 (2009i:65146)
- [35] Y. Zheng, *Systems of Conservation Laws*, *Two-dimensional Riemann Problems*, *Progress in Nonlinear Differential Equations and Their Applications*, 38, Birkhäuser Boston, Inc., Boston, MA, 2001. MR1839813 (2002e:35155)

BCAM - BASQUE CENTER FOR APPLIED MATHEMATICS, MAZARREDO 14, E-48009, BILBAO, BASQUE COUNTRY, SPAIN — AND — 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

BCAM - BASQUE CENTER FOR APPLIED MATHEMATICS, MAZARREDO 14, E-48009, BILBAO, BASQUE COUNTRY, SPAIN — AND — IKERBASQUE - BASQUE FOUNDATION FOR SCIENCE, MARIA DIAZ DE HARO, 3. 48013 BILBAO, BASQUE COUNTRY, SPAIN

E-mail address: zuazua@bcamath.org

URL: www.bcamath.org/zuazua/