Chapter 6
Filtered gradient algorithms for inverse design problems of one-dimensional Burgers equation

Laurent Gosse and Enrique Zuazua

Abstract Inverse design for hyperbolic conservation laws is exemplified through the 1D Burgers equation which is motivated by aircraft's sonic-boom minimization issues. In particular, we prove that, as soon as the target function (usually a N-wave) isn't continuous, there is a whole convex set of possible initial data, the backward entropy solution being possibly its centroid. Further, an iterative strategy based on a gradient algorithm involving "reversible solutions" solving the linear adjoint problem is set up. In order to be able to recover initial profiles different from the backward entropy solution, a filtering step of the backward adjoint solutoin is inserted, mostly relying on scale-limited (wavelet) subspaces. Numerical illustrations, along with profiles similar to F-functions, are presented.

6.1 An elementary modeling of sonic boom

The study of sound waves propagation in the atmosphere appears to trace back to Whitham's famous paper [47], where the so-called "F-function", representing an approximate, explicit pressure profile around a supersonic bullet, was derived. With some adjustments, that F-function was taken as the standard macroscopic pressure shape located in a close neighborhood of a plane jet: see for instance, [3, 19, 37, 38]. However, in order to propagate down to the ground the imprint a given, even approximate, pressure profile, one was still needing to solve inviscid Euler equations of gas dynamics. In Cleveland's Ph.D. thesis [13, eq. (2.69), page 36] (see also [40]), a so-called "augmented Burgers equation" was derived, reading,

Laurent Gosse
IAC, CNR, via dei Taurini, 19, 00185 Roma (Italia), e-mail: l.gosse@ba.iac.cnr.it

Enrique Zuazua
Universidade Autonoma di Madrid (Spain), e-mail: enrique.zuazua@uam.es

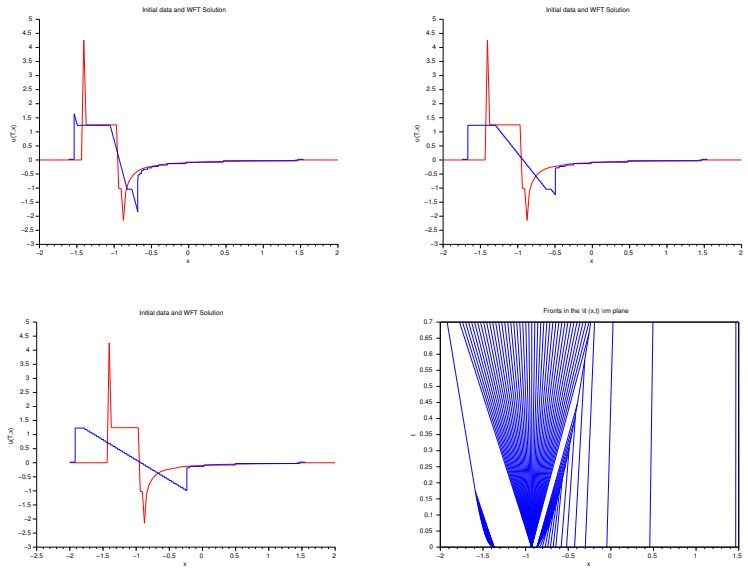

Fig. 6.1 Decay of a F-function onto a N-wave by solving a Burgers equation like (6.1).

$$\partial_x P - P\partial_t P = \frac{1}{\Gamma}\frac{\partial^2 P}{\partial t^2} - \frac{1}{2A}\frac{\partial A}{\partial x}\cdot P, \tag{6.1}$$

the terms on the right-hand side accounting for absorption and ray-tube spreading, respectively. Hence, in the simplest case, it reduces to a Burgers equation with opposite velocity and the roles of space and time being exchanged. With empirical data of a standard F-function at hand, one can visualize the evolution induced by such an homogeneous, inviscid, Burgers equation, see Fig. 6.1; the simulation is performed by means of the wavefront-tracking algorithm [28], which doesn't contain any numerical viscosity, being not based on finite-differences. The dominant process is the spreading of the big rarefaction wave (induced by the entropy-violating downward jump initially located around $x \simeq -1$) separated by two shocks. Both the upward (in $x \simeq -1.5$) and downward ($x \simeq -0.8$) spikes, initially present in the F-function, are quickly dissipated in order to let the so–called N-wave develop. Such a N-wave stands for the self-similar large-time behavior (also called "intermediate asymptotics") of one-dimensional, genuinely nonlinear, scalar conservation laws.

This situation raises very delicate issues when it comes to trying to optimize the pressure profile around a jet plane (by slightly correcting its aerodynamic design) based on the constraint of minimizing its sound imprint on the ground. Such an optimal control problem is usually referred to as to both "sonic boom minimization", and the "inverse design problem". Despite the simplicity of the Burgers model (6.1), it appeared that the hopes of aiming at a rigorous inverse design procedure are discouragingly thin, as a consequence of irreversibility expressed by entropy dis-

sipation beyond shock onset. This is illustrated, for instance, on Fig. 6.2, where very different initial data end up yielding the same shock at time $t \simeq 10$, even after quite complicated intermediate dynamics inside which a lot of entropy dissipates. Aware of that drawback, engineers decided to reduce some inverse design problems in aerodynamics to the more tractable optimization of a finite number of parameters inside the F-function, as advocated in [19] and followers, like e.g. [38]. Hereafter, we aim at taking a step back, by still planning to **perform a gradient-based optimization of a Burgers equation's initial with respect to a given target at time $t = T$, but modulating its iterations by means of a specific filtering, driving the iterates toward minimizers belonging to one or another linear subspace of L^2.** Accordingly different shapes of initial profiles can be reconstructed: see for instance Fig. 6.7, where two different wavelet filters are applied during to totality of the optimization process in order to drive iterates onto rather different initial data, each one belonging to a different "scale-limited subspace" of L^2. Adequate filters can be drawn out of signal processing books, like [27, 33]: in particular, Prolate Spheroids [42, 46] and Father Wavelet functions [16, 35] furnish good candidates toward an efficient filtering of a gradient optimization algorithm. Strategies can be combined with one another in order to produce "hybridized algorithms": see for instance Fig. 6.9, where different wavelet filters are applied to the iterates. Moreover, PSWF and wavelets can even be merged into each other as in [45] (see also [25, §3.4]). Of course, applying filters in order to restore stability in the context of a control problem isn't new: see e.g. [23, 36]. Theoretical approaches to optimal design problems for hyperbolic equations in the presence of shocks are presented in [9, 10].

6.2 Some analytical aspects of 1D inverse design

6.2.1 Irreversibility versus reverse Poincaré inequality

As a preliminary step, consider the 1D linear wave equation,

$$\partial_{tt} y - \partial_{xx} y = 0, \qquad t > 0, x \in \mathbb{R},$$

with enough decay at infinity so as to ensure conservation of energy:

$$E(t = T) = \int_{\mathbb{R}} |\partial_t y(T,x)|^2 + |\partial_x y(T,x)|^2 \, dx = E(t = 0),$$

The equation, being well-posed in the backward sense, the inverse design problem has a unique solution, belonging to the same space as the prescribed target. Instead,

$$\partial_t y - \partial_{xx} y = 0, \qquad t > 0, x \in \mathbb{R}.$$

brings out a different type of energy identity since, in that case,

$$\frac{d}{dt}\int_{\mathbb{R}}\frac{y^2}{2}dx=-\int_{\mathbb{R}}\mid\partial_x y\mid^2 dx=-\Lambda(t)\parallel y(t)\parallel^2_{L^2(\mathbb{R})}. \tag{6.2}$$

The so-called "reverse Poincaré inequality" for the heat equation reads,

$$\forall t\geq 0,\qquad \Lambda(t)=\frac{\|\partial_x y(t)\|^2_{L^2}}{\|y(t)\|^2_{L^2}}\leq\Lambda(0)=\Lambda_0,$$

also called "Cacciopoli's inequality", [34, 44], and so that,

$$\frac{d}{dt}\int_{\mathbb{R}}\frac{y^2}{2}dx+\Lambda_0\parallel y(t)\parallel^2_{L^2(\mathbb{R})}\geq 0,$$

hence the ratio of initial data to final (target) data grows exponentially with time:

$$\boxed{\forall t>0,\qquad ||y(0)||^2\leq\exp(2\Lambda_0 t)||y(t)||^2,} \tag{6.3}$$

expressing strong ill-conditioning of the inverse design problem. This estimate is sharp, being the energy version of the Fourier series representation for an equation posed in a bounded domain. Most linear systems enjoy the property of backward uniqueness (except transport equations in bounded domains). For Burgers equation (more generally for genuinely nonlinear scalar laws) irreversibility manifests itself through gradient steepening and shock formation: as a (decreasing, $u^+\leq u^-$) jump splits $\mathbb{R}$ into open sets $Q^{\pm}$, the pair (u,φ)= (flow solution, shock location) solves:

$$\begin{cases}\partial_t u+u\partial_x u=0, & \text{in } Q^-\cup Q^+,\\ \dot{\varphi}(t)=\dfrac{u^+(\varphi(t),t)+u^-(\varphi(t),t)}{2}, & t\in(0,T),\\ \varphi(0)=\varphi^0\in\mathbb{R},\ u(x,0)=u^0(x), & \text{in } \{x<\varphi^0\}\cup\{x>\varphi^0\}.\end{cases}$$

A less singular problem is the viscous Burgers equation, see (6.1), for which the Hopf-Cole transform [11, 14, 29] applies; consider $u^{\varepsilon}(t,x)$ a classical solution of,

$$\partial_t u^{\varepsilon}-\varepsilon\partial_{xx}u^{\varepsilon}+\partial_x|u^{\varepsilon}|^2=0,\qquad \mid u^{\varepsilon}(x,t)\mid+\mid\partial_x u^{\varepsilon}(x,t)\mid\to 0 \text{ as } \mid x\mid\to\infty. \tag{6.4}$$

Since an antiderivative of u^{ε},

$$v(x,t)=\int_{-\infty}^{x}u^{\varepsilon}(s,t)ds\quad\text{ solves }\quad\partial_t v_t-\varepsilon\partial_{xx}v+\mid\partial_x v\mid^2=0.$$

Then by rescaling the time variable, one gets that

$$z(x,t)=v(x,t/\varepsilon)/\varepsilon\quad\text{ solves }\quad\partial_t z-\partial_{xx}z+\mid\partial_x z\mid^2=0.$$

At last, by defining

$$\eta(x,t)=\exp(-z(t,x))=\exp(-v(x,t/\varepsilon)/\varepsilon),$$

one (maybe surprisingly!) recovers the linear heat equation,

$$\boxed{\partial_t \eta - \partial_{xx}\eta = 0. \quad \text{from which} \quad u^\varepsilon(x,t) = -\varepsilon \frac{\partial_x \eta(x,\varepsilon t)}{\eta(x,\varepsilon t)}.} \tag{6.5}$$

This classical argument suggests that, even in the absence of shocks as $u^\varepsilon \in C^\infty$, the inverse design problem for (6.4) is exponentially ill-conditioned, too, as a consequence of (6.3). In the singular limit $\varepsilon \to 0$, the situation deteriorates even more, because as soon as a shock appears, all the information about any initial profile which gave rise to it is irremediably lost: this loss of information is quantified through the dissipation of entropy. In a numerical perspective, two main issues emerge:

- Build efficient numerical solvers to find one (generally smooth) inverse design;
- Further, try to recover as many other possible inverse designs to study their set.

As an illustration, Fig. 6.2 displays several, quite different, initial profiles u_0 decaying, after having produced more or less intricate interaction patterns, onto a single entropy shock connecting 1 to 0 at time $t = 10$. Theoretical results are available by Adimurthi et al. [1], using the explicit Lax-Oleinik representation formula: see also related works by Ancona and Cannarsa [4] for Hamilton-Jacobi equations.

6.2.2 *Convexity of the set of inverse designs*

Consider the simplest occurrence of (6.1), where, for ease of reading, space and time variables are recast into their usual setup: in smooth areas, $u(t,x)$ solves,

$$\partial_t u + u\partial_x u = 0, \qquad u(t=0,x) = u_0(x).$$

The method of characteristics stems on curves passing through each point (t,x),

$$\dot{X}(t) = u(t,X(t)), \quad \dot{u}(t,X(t)) = 0, \Rightarrow X(t) = X(t=0) + t \cdot u_0(X(t=0)).$$

However, as $u_0(X(t=0)) = u(t,X(t))$, the solution of Burgers equation reads,

$$u(t,x) = \frac{x-y}{t}, \qquad y \text{ the foot of the characteristic passing at } (t,x). \tag{6.6}$$

Actually, beyond shock onset, (6.6) still holds almost everywhere, [17, p.145],

$$u(t,x) = \frac{x-\xi_u}{t}, \quad \inf_y \left(\frac{|x-y|^2}{2t} + \int^y u_0(s)ds \right) = \frac{|x-\xi_u|^2}{2t} + \int^{\xi_u} u_0(s)ds. \tag{6.7}$$

The functional which is minimized is the Lax function: according to [15, Chap. 11], given a point (t,x), the minimum on $\mathbb{R}$ is unique (and the backward characteristic is a shock-free straight line) when $u(t,x)$ is continuous. Oppositely, if $u(t,x)$ displays an entropy shock, there are two distinct minima $\xi_u^\pm$, with $\xi_u^- < \xi_u^+$, another

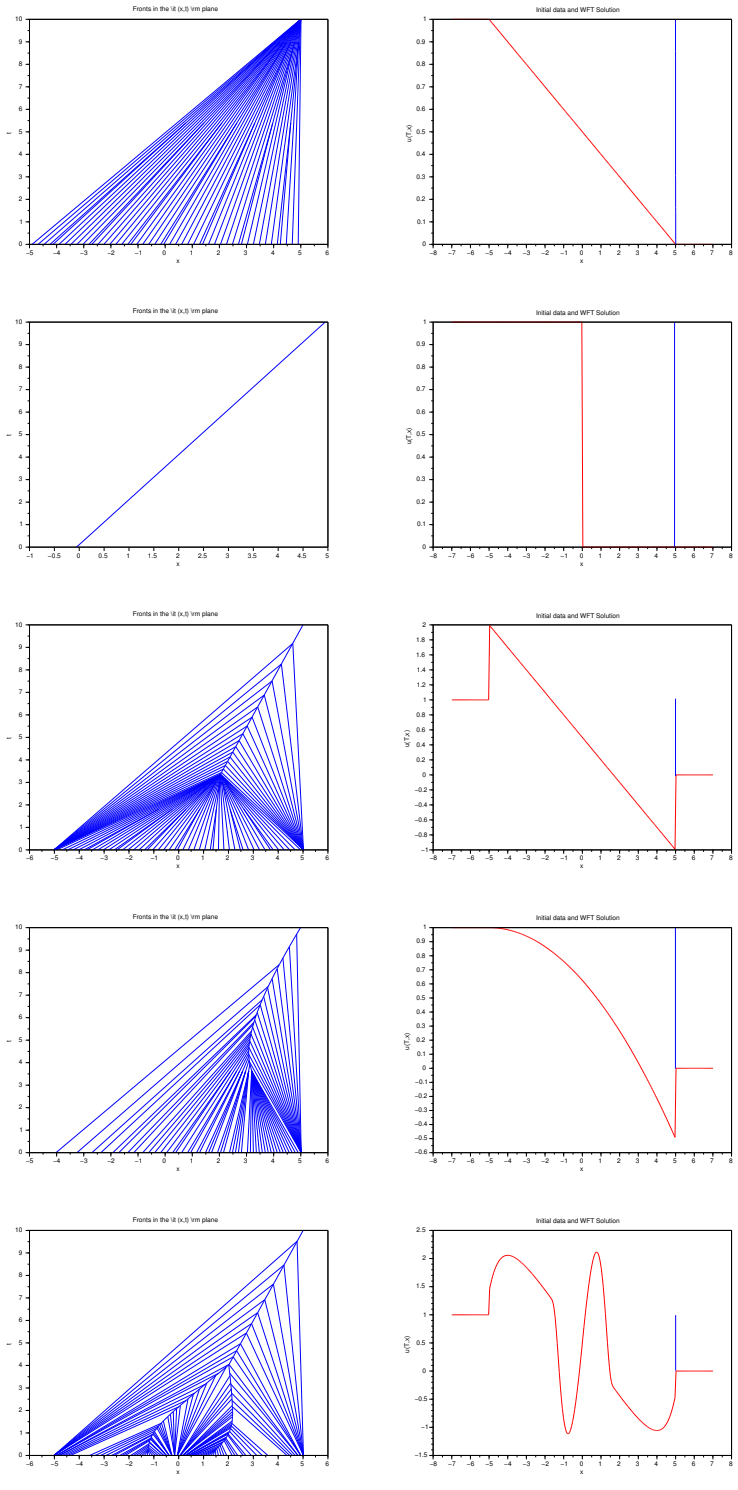

Fig. 6.2 Initial profiles yielding a single shock in $x = 5$ at $t = 10$.

manifestation of Oleinik's condition: as $u(t,x^{\pm}) = u_0(\xi^{\pm})$, by differentiating,

$$\frac{u_0(\xi^+) - u_0(\xi^-)}{\xi^+ - \xi^-} = \frac{u(t,x^+) - u(t,x^-)}{\xi^+ - \xi^-} = \frac{1}{t}.$$

An illustration is provided by Fig. 6.3, for the special case of Burgers equation with

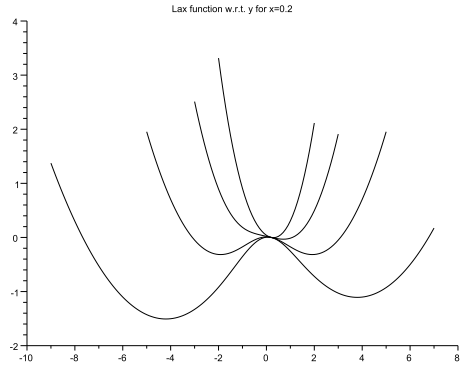

Fig. 6.3 Lax function with $u_0(x) = 0.1 - \tanh(x)$ and $x = 0.2$, $t = 0.5, 1, 2, 4$.

initial data $u_0(x) = 0.1 - \tanh(x)$, so that the Lax function reads,

$$G(y;t,x) = \frac{|x-y|^2}{2t} - \log\cosh y + 0.1\, y.$$

The figure displays the different minima for $x \equiv 0.2$ with $t = 0.5, 1, 2, 4$. At time $t \simeq 2$, the shock appears and G is endowed with two distinct minimizers.

The existence of inverse designs, given a specific target at time $T > 0$ supported inside an interval of $\mathbb{R}$, was studied by the authors of [1, Theorem 1.2]:

Theorem 6.1. *Let ρ be a locally bounded, non-decreasing, possibly discontinuous function, and, given a time $T > 0$, consider the BV_{loc}, entropy admissible, target,*

$$\forall x \in (X_1, X_2), \qquad \bar{u}(x) = \frac{x - \rho(x)}{T}, \qquad \rho' \geq 0.$$

Then under the following condition,

$$\rho(X_1) \geq A_1 > Y_1, \qquad \rho(X_2) \leq A_2 < Y_2,$$

there exists an inverse design $u_0 \in L^{\infty}(Y_1, Y_2)$ yielding at time T an entropy solution of Burgers equation, which restriction to the interval (X_1, X_2) is exactly $\bar{u}$.

Being the difference of two locally bounded functions, the target $\bar{u}$ is in $BV_{loc}(\mathbb{R})$; by construction it is entropy admissible because it cannot contain increasing shocks.

Theorem 6.2. *Let u_0, v_0 be two $L^\infty(\mathbb{R})$ inverse designs; assume that, at a given time $T>0$, their respective, piecewise-continuous, targets $\bar{u}(\cdot)$ and $\bar{v}(\cdot)$ coincide, i.e. $u(T,x)=v(T,x)$. Then, for any $\alpha\in(0,1)$, any convex combination $w_0^\alpha=\alpha u_0+(1-\alpha)v_0$ yields the same target at the same time, too,*

$$\forall x\in\mathbb{R},\qquad w^\alpha(T,x)=u(T,x)=v(T,x).$$

Given any reachable target, the union of its inverse designs is a convex set of $L^\infty(\mathbb{R})$.

If the target is a non-negative N-wave, the results in [32] ensure that a Dirac atom can be considered an inverse design (in a weaker sense of bounded measures), too.

Proof. It splits into three main steps:

- First, there is no restriction in considering regularized initial data $u_0^\varepsilon, v_0^\varepsilon\in C^\infty$ because standard mollifying kernels can be applied to any pair $u_0, v_0\in L^\infty$. Denoting $u^\varepsilon(t,\cdot)=S_t u_0^\varepsilon$ and similar for $v^\varepsilon(t,\cdot)$, Kružkov contraction estimate yields

$$\forall t>0,\qquad \|u(t,\cdot)-u^\varepsilon(t,\cdot)\|_{L^1}\leq\|u_0-u_0^\varepsilon\|_{L^1}\to 0,\quad \varepsilon\to 0.$$

- Yet, consider the representation (6.7) for both $u(t,\cdot)$ and $v(t,\cdot)$: accordingly, ξ_u and ξ_v are mappings $(t,x)\mapsto\xi_u(t,x)$ or $\xi_v(t,x)$ (yielding the foot of the corresponding characteristic where the target is smooth). At any point $t=T$, $x\in\mathbb{R}$,

$$u(t,x)=v(t,x)\Rightarrow\xi_u(t,x)=\xi_v(t,x)=\text{ some value }\xi\in\mathbb{R}.$$

  If T,x is a shock location, each minimum is reached at two distinct locations,

$$\xi_u^\pm(t=T,x)=\xi_v^\pm(t=T,x)=\xi^\pm\in\mathbb{R}^2.$$

- Now, applying again (6.7) to $w_0^\alpha=\alpha u_0+(1-\alpha)v_0$ should bring $\xi_w=\xi$ with,

$$\begin{aligned}\inf_y\left(\frac{|x-y|^2}{2t}+\int^y(\alpha u_0(s)+(1-\alpha)v_0(s))ds\right)\\ =\frac{|x-\xi_w|^2}{2t}+\int^{\xi_w}(\alpha u_0(s)+(1-\alpha)v_0(s))ds\\ =\alpha\left(\frac{|x-\xi_w|^2}{2t}+\int^{\xi_w}u_0(s)\right)+(1-\alpha)\left(\frac{|x-\xi_w|^2}{2t}+\int^{\xi_w}v_0(s))ds\right).\end{aligned}$$

  Consequently, since $\inf(A+B)\geq\inf(A)+\inf(B)$, where

$$A=\alpha\left(\frac{|x-y|^2}{2t}+\int^y u_0(s)ds\right),\qquad B=(1-\alpha)\left(\frac{|x-y|^2}{2t}+\int^y v_0(s)ds\right).$$

and both A, B are minimized at the same point $\xi_u = \xi_v = \xi$, then equality holds:

$$\inf_y \left(\frac{|x-y|^2}{2t} + \int^y (\alpha u_0(s) + (1-\alpha) v_0(s)) ds \right) \geq$$
$$\alpha \left(\frac{|x-\xi|^2}{2t} + \int^\xi u_0(s) \right) + (1-\alpha) \left(\frac{|x-\xi|^2}{2t} + \int^\xi v_0(s)) ds \right) =$$
$$\frac{|x-\xi|^2}{2t} + \int^\xi (\alpha u_0(s) + (1-\alpha) v_0(s)) ds.$$

That inequality shows that the value of $G_{w^\alpha}(t,x,\xi)$ is smaller than its infimum on y; hence $y = \xi$ is automatically a critical point for $w^\alpha(t,x)$. The same holds for $\xi^\pm$.

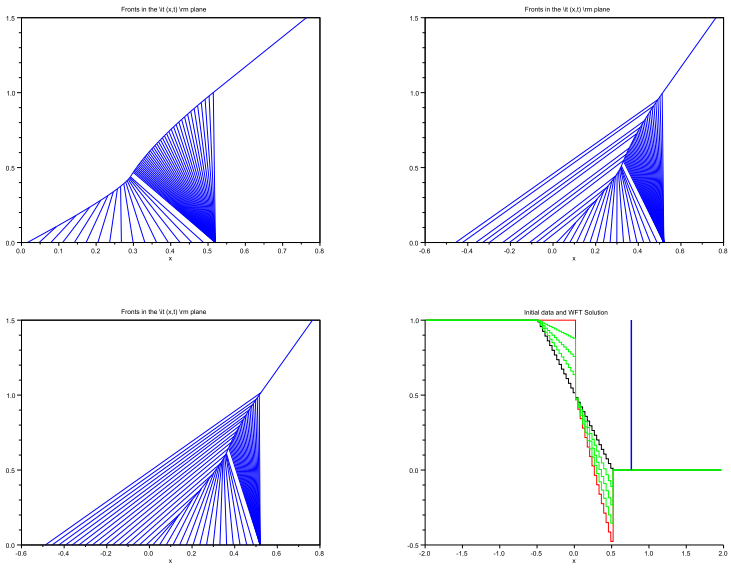

Fig. 6.4 Time evolution of convex combinations of continuous and discontinuous "ramps".

Besides Fig. 6.2, which already illustrates the variety of inverse designs associated to the simple target $\chi_{x<5}$, at time $T = 10$, Fig. 6.4 shows several convex combinations of two elementary inverse designs, (χ_A being the indicator of the set A)

$$u_0(x) = \chi_{x<0} + (\frac{1}{2} - 2x)\chi_{0<x<\frac{1}{2}}, \qquad v_0(x) = \min(1, \frac{1}{2} - x)\chi_{x<\frac{1}{2}},$$

which all reach the same target at time, namely a single shock at time $t \simeq 1.2$.

6.2.3 Backward entropy solution as the centroid

An elementary observation is that, given a smooth target $\bar{u} \in C^\infty(\mathbb{R})$ and a time $T > 0$ small enough, the inverse design problem is solved by "tracing characteristics backwards". More precisely, such a solution is $u(t, X(t)) = \bar{u}(X(T))$ with,

$$\dot{X}(t) = u(t, X(t)), \text{ so that } X(t) = X(T) + (t-T)u(T, X(T)), \quad 0 \leq t < T.$$

Such a process is actually equivalent to solving Burgers equation backward in time,

$$\partial_t u - \partial_x(u^2/2) = 0, \qquad u(T,x) = \bar{u}(x). \tag{6.8}$$

This situation corresponds to the convex set of Theorem 6.2 being reduced to a

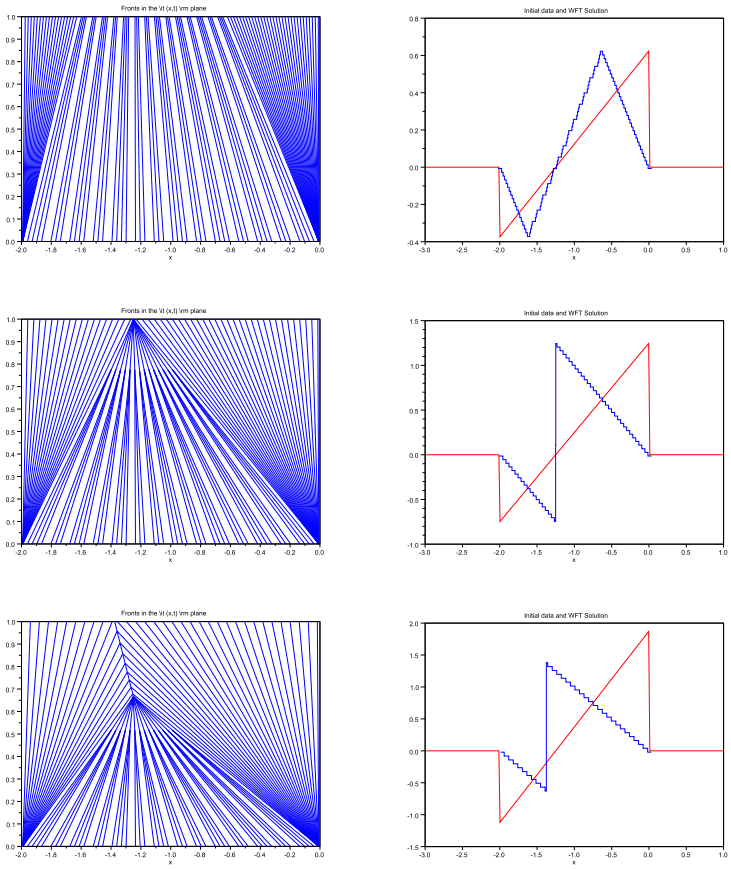

Fig. 6.5 Dissipation in backward entropy solutions of (6.8) with $\bar{u}(x)$ three bigger N-waves

single point. Yet, when the target $\bar{u}$ is an admissible entropy (hence one-sided Lipschitz) BV function, more elaborate techniques (generalized characteristics, Lax formula, ...) allow to derive an inverse design by means of the entropy solution of (6.8): see first row in Fig. 6.2. Clearly, as the flux function changes its sign, any entropy admissible (decreasing) jump in $\bar{u}$ becomes entropy-violating for (6.8) and is instantaneously converted in a (smooth) backward expansion wave. In contrast, the other solutions shown on Fig. 6.2 all display increasing (entropy-violating) jumps at time $t = 0$. This suggests that this particular inverse design from (6.8) may be the centroid of the convex set $K \in L^\infty(\mathbb{R})$ containing all the possible inverse designs of a discontinuous target: by definition, the centroid is the barycenter,

$$C = \frac{\int_K y\,dy}{\int_K dy}, \qquad C(x) \text{ an inverse design of } \bar{u}(x).$$

Let η stand for any convex entropy (for instance $\eta(u) = u^2$): by Jensen's inequality,

$$\int_{\mathbb{R}} \eta(C(x))dx = \int_{\mathbb{R}} \eta\left(\frac{\int_K y\,dy}{\int_K dy}\right) dx \leq \int_K \frac{\int_{\mathbb{R}} \eta(y(x))\,dx}{\int_K dy}\,dy,$$

so that the centroid $C(x)$ is the inverse design of the discontinuous, entropy admissible, target $\bar{u}$ minimizing a convex entropy η with respect to all the others in K.

Remark 6.1. The relevance of the backward entropy solution of (6.8) is limited to **"small targets" or "small times"**, an assumption tacitly made in Theorem 6.1:

$$\frac{d\rho}{dx} \geq 0 \quad \Rightarrow \quad \frac{d\bar{u}}{dx} \leq \frac{1}{T} \text{ where the target } \bar{u}(x) = \frac{x - \rho(x)}{T}. \tag{6.9}$$

Indeed, if $\bar{u}$ contains entropy jumps, the backward evolution cannot induce wave cancellations (see first and second rows of Fig. 6.5), so the backward entropy solution realizes an inverse design. Oppositely, the third row of Fig. 6.5 shows that, when such a smallness assumption is broken, that solution isn't an inverse design anymore because too much entropy, like the L^2 norm, was dissipated in the discontinuity.

Corollary 6.1. *Under the assumptions of Theorem 6.1, let $u(t,x)$ be the "backward entropy solution" solving (6.8) with final data $u(T,\cdot) = \bar{u}$, then*

- *for all $t \in (0,T)$, $u(t,\cdot)$ is Lipschitz continuous;*
- *entropy is preserved during the backward evolution;*
- *$u(0,\cdot)$ is an inverse design, so that $u_0 = u(0,\cdot)$.*

However, $\bar{u}$ can display decreasing shocks, and u_0, increasing ones.

Proof. First, as $\bar{u}$ is entropy-admissible, it may contain only decreasing discontinuities which are all instantaneously regularized as rarefaction waves in the backward evolution. Then, the smallness assumption (6.9) implies that the entropy solution of (6.8) cannot blow up in the open interval $(0,T)$ because its blowup time satisfies,

$$\bar{T} = \frac{1}{\sup_x \bar{u}(x)} \geq T.$$

A numerical discussion of inverse designs for "big targets" was given in [2, §6.4].

6.3 A PDE-constraint optimization process

6.3.1 Reduced objective functional and control-to-state map

Inverse design consists in reconstructing an initial profile u_0 at time $t = 0$ able to match (or to get as close as possible to) some given target function $\bar{u}$ at time $t = T \in \mathbb{R}^+$. It can be formalized as a PDE-constrained optimization process for an "objective functional" $\mathscr{J}_T$ and a control $u_0 \in U$, namely

$$\inf_U \mathscr{J}_T(u) = \frac{1}{2}\int |u(T,x) - \bar{u}(x)|^2 dx, \qquad T > 0,$$

subject to the (presumably infinite set of) constraints,

$$\partial_t u + \partial_x(u^2/2) = 0, \qquad u(t=0,\cdot) = u_0, \qquad \partial_t \eta(u) + \partial_x q(u) \leq 0.$$

Here, the entropy η stands for any convex function and q, the entropy flux is such that $q'(u) = u \cdot \eta'(u)$. Since Burgers' flux is strictly convex, only one convex entropy is actually needed because just one entropy inequality yields all the other ones. So the optimization process contains only one equality and one inequality, say for $\eta(u) = u^2$. The PDE rewrites $e(u, u_0) = 0$ and is referred to as an "equation of state". That equation of state $e(\cdot,\cdot)$ rewrites as a mapping such that,

$$e : L^1 \cap BV(\mathbb{R}) \times L^1 \cap L^\infty(\mathbb{R}) \to \mathscr{M}(\mathbb{R}),$$

where, for any $\phi(t,x) \in C^1(\mathbb{R} \times \mathbb{R}^+)$ with compact support,

$$\int\int_{\mathbb{R}} u\partial_t\phi(t,x) + \frac{u^2}{2}\partial_x\phi(t,x)dx.dt + \int_{\mathbb{R}} u_0(x)\phi(t=0,x)dx = 0.$$

For any $t \geq 0$, let $u(t,\cdot) = S_t(u_0)$ stand for the entropy solution (in the sense of Kružkov) of the inviscid 1D scalar Burgers equation,

$$\partial_t u + \partial_x(u^2/2) = 0, \qquad \partial_t(u^2/2) + \partial_x(u^3/3) \leq 0, \qquad x \in \mathbb{R},\ t > 0.$$

Being a semi-group of L^1 contractions, it implies a BV estimate,

$$\forall h > 0, \qquad \|u(t,\cdot + h) - u(t,\cdot)\|_{L^1} \leq \|u_0(\cdot + h) - u_0(\cdot)\|_{L^1} \leq h \cdot TV(u_0),$$

where $TV(u_0)$ stands for the total variation of u_0. So, $BV \subset L^\infty$ is a stable space for this convex scalar law: $u_0 \in L^1 \cap BV(\mathbb{R}) \subset L^2$ (by classical interpolation inequalities) produces an entropy solution lying in $L^2(\mathbb{R})$, too. Accordingly, the "control-to-state map" S_T can be defined along with a "reduced objective functional",

$$\tilde{\mathscr{J}}_T(u_0) = \frac{1}{2}\int |S_T(u_0)(x) - \bar{u}(x)|^2 dx = \mathscr{J}_T(\text{“}u(u_0)\text{”}), \tag{6.10}$$

inside which the PDE dynamics are already encoded (in the form of $S_T(u_0)$). In particular, for a given $T > 0$, one may consider the minimization problem,

$$\text{given } \bar{u}, \qquad \inf_{u_0 \in U} \tilde{\mathscr{J}}_T(u_0), \qquad U = L^1 \cap BV(\mathbb{R}) \subset L^2(\mathbb{R}),$$

which can be expected to admit many solutions (see Theorem 6.2) when the target function $\bar{u}$ contains entropy shocks.

6.3.2 Regularized objective functionals and duality solutions

Duality solutions is a peculiar notion of weak solutions to one-dimensional continuity equations enjoying both uniqueness and weak stability properties; they were proposed in [8], then applied to differentiability issues for convex scalar conservation laws in [9]. Finally their numerical analysis was tackled in [26]. Following both [12] and [9], one can consider the "sensitivity problem" for Burgers equation: let $u_0 \in L^1 \cap L^\infty(\mathbb{R})$ be initial data, and $u(T,\cdot) = S_T(u_0)$ its entropy solution at time $t = T > 0$. A small deviation δu_0 and a parameter $0 < \eta \ll 1$, yield new data,

$$u_0^\eta = u_0 + \eta \cdot \delta u_0, \qquad u^\eta(T,\cdot) = S_T(u_0^\eta) \doteq u(T,\cdot) + \eta \cdot \delta u,$$

Clearly, by defining $\delta u(t,\cdot)$ as the perturbation at time $0 \le t \le T$,

$$\delta u(t,\cdot) = \frac{S_t(u_0^\eta) - S_t(u_0)}{\eta}, \quad \partial_t(\delta u) + \partial_x(a(t,x)\cdot \delta u) = 0, \text{ with } a = \frac{|u^\eta|^2 - |u|^2}{2(u^\eta - u)},$$

one finds that it satisfies a linear continuity equation, with a possibly discontinuous coefficient, though, so that uniqueness is a delicate issue.

Definition 6.1. • Let $m(t=0,\cdot) = m_0$ a bounded measure on $\mathbb{R}$ and a bounded velocity field $a(t,x) \in L^\infty$, a weak solution $m(t,\cdot)$ is a duality solution to

$$\partial_t m + \partial_x(a\cdot m) = 0, \qquad \partial_x a(t,\cdot) \le \alpha(t), \text{ with } \alpha \in L^1(]0,T[),$$

if it satisfies, for any $0 < \tau \le T$ and any $p(t,x)$ reversible solution in $(0,\tau)\times\mathbb{R}$ with compact support in x,

$$\frac{d}{dt}\int_{\mathbb{R}} p(t,x)m(t,x)dx = 0, \qquad t \in [0,\tau]. \tag{6.11}$$

- under the same assumptions, given a Lipschitz final datum p^T, a reversible solution to the dual transport equation is a Lipschitz solution of,

$$\partial_t p + a \cdot \partial_x p = 0, \qquad p(T,\cdot) = p(\cdot),$$

which is equivalently characterized by $TV(p(t,\cdot)) = TV(p^T)$ or $p = p_1 - p_2$, a difference of two non-decreasing Lipschitz solutions.

Thanks to these notions, Bouchut and James stated in [9] the following:

Theorem 6.3. *Let u_0 be entropy admissible, i.e. $\partial_x u_0(\cdot) \leq C \in \mathbb{R}$, then as $\eta \to 0$, the perturbation $\delta u(t,\cdot) \to m(t,\cdot)$, which is the duality solution to,*

$$\partial_t m + \partial_x (S_t(u_0) \cdot m) = 0, \qquad m(t=0,\cdot) = \delta u_0.$$

Accordingly, $m(t,\cdot)$ is the directional derivative of $S_t(u_0)$ in the direction δu_0.

Especially, that duality solution satisfies, for any (Lipschitz) reversible solution p,

$$\int_{\mathbb{R}} p(T,x) m(T,x) dx = \int_{\mathbb{R}} p(t=0,x) m(t=0,x) dx = \int_{\mathbb{R}} p(0,x) \delta u_0(x) dx. \quad (6.12)$$

Yet, following [12], let an objective functional and its Gâteaux derivative along δu_0,

$$\mathscr{J}_T(u^\eta) = \frac{1}{2} \int_{\mathbb{R}} |S_T(u_0^\eta) - \bar{u}(x)|^2 dx, \qquad \delta \mathscr{J}_T = \int_{\mathbb{R}} \delta u(T,\cdot) \cdot (S_T(u_0^\eta) - \bar{u}) dx.$$

An essential property for inverse design feasibility is (6.12) holding in the case,

$$p(T,\cdot) = S_T(u_0^\eta) - \bar{u}, \text{ so that } \delta \mathscr{J}_T = \int_{\mathbb{R}} p(0,x) \delta u_0(x) dx = \langle p(0,\cdot), \delta u_0 \rangle .$$

Theorem 6.3 states that, being δu in general a (measure) duality solution, such a property can be expected to hold mostly for p being a (Lipschitz) reversible solution, see (6.11). As $S_T(u_0^\eta) - \bar{u}$ is likely to be discontinuous, setting up "regularized objective functionals", involving some convolution kernel, secures the Lipschitz regularity of each $p(T,\cdot)$ and the crucial property (6.12), too: for instance, let

$$0 \leq \mu^\varepsilon(x) = \frac{1}{\varepsilon\sqrt{\pi}} \exp(-|x/\varepsilon|^2) \in C^\infty(\mathbb{R}),$$

in order to define "ε-regularized objective functions",

$$\mathscr{J}_T^\varepsilon(u) = \frac{1}{2} \int \left| [\mu^\varepsilon * u(T,\cdot)] - \bar{u}(x) \right|^2 dx = \frac{1}{2} \int \left| [\mu^\varepsilon * S_T(u_0)] - \bar{u}(x) \right|^2 dx. \quad (6.13)$$

Clearly, for $\varepsilon \ll 1$, this type of perturbation of the L^2 norm remains negligible, especially when it comes to numerics. More precisely, for most applications, a grid size $\Delta x > 0$ is fixed, so that the correction (6.13) becomes transparent if $\varepsilon \ll 1$.

6.3.3 Critical points of reduced objective functionals

Even in the favorable case where $\mathscr{J}_T$ is strictly convex on $L^2(\mathbb{R})$, a numerical algorithm remains too costly as long as it needs to compute gradients in all the directions corresponding to any Hilbertian basis functions $\varphi_i(x)$. So, an operator $\nabla\tilde{\mathscr{J}}_T$ (a linear form in the dual of $L^2(\mathbb{R})$) expressing the first variation of $\mathscr{J}_T$,

$$\forall \varphi_i, \qquad \delta\mathscr{J}_T(u)[\varphi_i] = \left\langle \nabla\tilde{\mathscr{J}}_T(u_0), \varphi_i \right\rangle = \int_{\mathbb{R}} \nabla\tilde{\mathscr{J}}_T(u_0)\,\varphi_i(x)dx,$$

would be very useful. Formally, using (6.10), it comes,

$$\delta\mathscr{J}_T(u = S_T(u_0))[\varphi_i] = \langle \nabla_u \mathscr{J}_T(u), \nabla S_T(\varphi_i)\rangle\,,$$

by means of the chain rule. By taking the adjoint in the duality bracket,

$$\delta\mathscr{J}_T(S_T(u_0))[\varphi_i] = \langle (\nabla S_T)^*[\nabla_u \mathscr{J}_T(u)], \varphi_i\rangle\,,$$

so that our desired operator reads $\nabla\tilde{\mathscr{J}}_T(u_0) = (\nabla S_T)^*[\nabla_u \mathscr{J}_T(u)]$. In this expression, ∇S_T stands for the linearized evolution operator associated to the Burgers equation, that is, the solution of the transport equation

$$\partial_t v + u(t,x)\partial_x v = 0, \qquad t \in (0,T).$$

Yet, the adjoint $(\nabla S_T)^*[\nabla_u \mathscr{J}_T(u)]$ corresponds to the backward problem,

$$-\partial_t p - u(t,x)\partial_x p = 0, \quad p(T,x) = u(T,x) - \bar{u}(x) = S_T(u_0)(x) - \bar{u}(x).$$

In particular, in the dual space of $L^2(\mathbb{R})$, $(\nabla S_T)^*[\nabla_u \mathscr{J}_T(u)] = p(t=0,\cdot)$,

$$\boxed{\forall \varphi_i, \qquad \left\langle \nabla\tilde{\mathscr{J}}_T(u_0), \varphi_i \right\rangle = \langle p(t=0,\cdot), \varphi_i\rangle\,.} \tag{6.14}$$

Accordingly, $\nabla\tilde{\mathscr{J}}_T^\varepsilon$ is a linear form on $L^2(\mathbb{R})$ such that,

$$\forall \varphi_i, \qquad \delta\mathscr{J}_T^\varepsilon(u)[\varphi_i] = \left\langle \nabla\tilde{\mathscr{J}}_T^\varepsilon(u_0), \varphi_i \right\rangle.$$

As μ^ε is an **even function**, similar computations yield for any φ_i,

$$\begin{aligned}\delta\mathscr{J}_T^\varepsilon(u^\varepsilon = \mu^\varepsilon * S_T(u_0))[\varphi_i] &= \langle \nabla_u \mathscr{J}_T(u^\varepsilon), \mu^\varepsilon * [\nabla S_T(\varphi_i)]\rangle \\ &= \langle \mu^\varepsilon * [\nabla_u \mathscr{J}_T(u^\varepsilon)], \nabla S_T(\varphi_i)\rangle \\ &= \langle (\nabla S_T)^*[\mu^\varepsilon * \nabla_u \mathscr{J}_T(u^\varepsilon)], \varphi_i\rangle\,,\end{aligned}$$

and finally, the expression of the modified gradient operator follows,

$$\nabla\tilde{\mathscr{J}}_T^\varepsilon(u_0) = (\nabla S_T)^*[\mu^\varepsilon * \nabla_u \mathscr{J}_T(u^\varepsilon)], \qquad u^\varepsilon(T,x) = [\mu^\varepsilon * S_T(u_0)](x).$$

The backward equation remains unchanged, except for its final data, so

$$\boxed{\left\langle \nabla \tilde{\mathscr{J}}_T^{\varepsilon}(u_0), \varphi_i \right\rangle = \langle p^{\varepsilon}(0,\cdot), \varphi_i \rangle \text{ with } p^{\varepsilon}(T,\cdot) = \mu^{\varepsilon} * [\mu^{\varepsilon} * S_T(u_0) - \bar{u}]} \quad (6.15)$$

The **velocity field** $u(t,\cdot)$ **in** $\partial_t p^{\varepsilon} + u(t,x)\partial_x p^{\varepsilon} = 0$ **is not regularized**: by Oleinik's form of the entropy condition, it may display downward jumps, so that, to ensure stability, $p^{\varepsilon}(t,x)$ should be interpreted as a "reversible solution", see Definition 6.1.

6.4 Minimizing sequences in filtered gradient algorithms

6.4.1 Large-time behavior of numerical schemes

Let us consider now numerical approximation schemes for the inviscid problem :

$$\begin{cases} u_j^{n+1} = u_j^n - \dfrac{\Delta t}{\Delta x}\big(g(u_j^n, u_{j+1}^n) - g(u_{j-1}^n, u_j^n)\big), & j \in \mathbb{Z}, \quad n > 0. \\ u_j^0 = \frac{1}{\Delta x}\int_{(j-1/2)\Delta x}^{(j+1/2)\Delta x} u_0(x)dx, & j \in \mathbb{Z}. \end{cases}$$

The approximate, piecewise-constant, numerical solution $u_\Delta(t,x)$ is given by

$$u_\Delta(t,x) = u_j^n, \qquad (j-1/2)\Delta x < x < (j+1/2)\Delta x, \qquad n\Delta t \le t < (n+1)\Delta t,$$

Are corresponding large-tine dynamics discrete analogues of continuous ones ?

1. Lax-Friedrichs
$$g^{LF}(u,v) = \frac{u^2+v^2}{4} - \frac{\Delta x}{\Delta t}\left(\frac{v-u}{2}\right),$$
2. Engquist-Osher
$$g^{EO}(u,v) = \frac{u(u+|u|)}{4} + \frac{v(v-|v|)}{4}, \quad (6.16)$$
3. Godunov
$$g^{G}(u,v) = \begin{cases} \min\limits_{w\in[u,v]} \frac{w^2}{2}, & \text{if } u \le v, \\ \max\limits_{w\in[v,u]} \frac{w^2}{2}, & \text{if } v \le u. \end{cases}$$

These methods converge in $L^1_{loc}(\mathbb{R}\times\mathbb{R}^+)$**, i.e. on all compact subsets: do they behave correctly as** $t \to \infty$ **? Such a question was already raised in [41,** §**6.4]**. It has consequences for numerical inverse design problems because a gradient descent method needs an iterative resolution of both forward and adjoint dynamics repeatedly, possibly over long time intervals. Three-point monotone schemes rewrite like

$$\frac{u_j^{n+1}-u_j^n}{\Delta t} + \frac{(u_{j+1}^n)^2 - (u_{j-1}^n)^2}{4\Delta x} = R(u_j^n, u_{j+1}^n) - R(u_{j-1}^n, u_j^n)$$

where the numerical viscosity R is defined in a unique manner as

$$R(u,v)=\frac{Q(u,v)(v-u)}{2}=\frac{1}{2\Delta x}\Big(\frac{u^2}{2}+\frac{v^2}{2}-2g(u,v)\Big).$$

Accordingly, the following results were proved in [39]

Theorem 6.4. *Let $u_0\in L^1\cap L^\infty(\mathbb{R})$ and $\Delta t\|u^n\|_{\infty,\Delta}\le\Delta x$: then, for any $p\in[1,\infty)$,*

- *the numerical solution u_Δ given by the Lax-Friedrichs scheme satisfies*

$$\lim_{t\to\infty} t^{\frac{1}{2}(1-\frac{1}{p})}\|u_\Delta(t)-w(t)\|_{L^p(\mathbb{R})}=0, \tag{6.17}$$

  where the diffusion wave $w=w_{M_\Delta}$, $M_\Delta=\int_{\mathbb{R}}u_\Delta^0$, is the unique solution of

$$\begin{cases}\partial_t w+\partial_x\Big(\frac{w^2}{2}\Big)=\frac{(\Delta x)^2}{2\Delta t}\partial_{xx}w, & x\in\mathbb{R},\quad t>0,\\ w(t=0,x)=M_\Delta\delta(x), & \text{the Dirac mass in } x=0.\end{cases} \tag{6.18}$$

- *the ones given by both Engquist-Osher and Godunov schemes satisfy a similar asymptotic behavior with an $N-$wave, $w=w_{p_\Delta,q_\Delta}$, being unique solution of*

$$\begin{cases}\partial_t w+\partial_x\Big(\frac{w^2}{2}\Big)=0, & x\in\mathbb{R},\quad t>0,\\ w(t=0,x)=M_\Delta\delta(x), & \lim\limits_{t\to0}\int_0^x w(t,z)dz=\begin{cases}0, & x<0,\\ -p_\Delta, & x=0,\\ q_\Delta-p_\Delta, & x>0,\end{cases}\end{cases} \tag{6.19}$$

  with $M_\Delta=\int_{\mathbb{R}}u_\Delta^0$, $p_\Delta=-\min_{x\in\mathbb{R}}\int_{-\infty}^x u_\Delta^0(z)dz$, and $q_\Delta=\max_{x\in\mathbb{R}}\int_x^\infty u_\Delta^0(z)dz$.

These quantitative results confirm, in an accurate manner, the observation formerly made in [41, §6.4]: "*Neither of these two results [Error Estimates in Theorems 6.4.1 and 6.4.2] is uniform with respect to the time and indeed they could not be*".

6.4.2 Reversible solutions of the adjoint problem

A classical tool to build a minimizing sequence $(u_0^n)_{n\in\mathbb{N}}$ is the fixed-step gradient algorithm: let $0<\eta\ll1$ be small enough as an iteration step, (see e.g. [6, 7])

$$u_0^{n+1}=u_0^n-\eta\nabla\tilde{\mathscr{J}}_T(u_0^n)\quad\Rightarrow\quad\tilde{\mathscr{J}}_T(u_0^{n+1})\le\tilde{\mathscr{J}}_T(u_0^n).$$

What saves the day is that, again on $(0,T)$, and under the condition that the function $p_0:=p(t=0,\cdot)$ can be well defined, the minimizing sequence is just,

$$u_0^{n+1}=u_0^n-\eta\cdot p_0^n,\qquad\text{in } L^2(\mathbb{R})^*,$$

where $p_0^n=p^n(t=0,\cdot)=\nabla\tilde{\mathscr{J}}(u_0^n)$ as seen in (6.14). It stops when $p_0^n\simeq0$. For definiteness, let $g(u,v)$ stand for the Engquist-Osher numerical flux (6.16) for the

(forward) Burgers equation, along with n,k,j being the minimizing sequence, time and space indexes, respectively: an inverse design comes from iterating in $n \in \mathbb{N}$,

$$\left.\begin{array}{l} \forall j, \qquad u_j^{n=0,k=0} = 0; \\ \\ \left\{ \begin{array}{l} u_j^{n,k+1} = u_j^{n,k} - \frac{\Delta t}{\Delta x}\left[g(u_j^{n,k}, u_{j+1}^{n,k}) - g(u_{j-1}^{n,k}, u_j^{n,k}) \right], \\ \text{for } k = 0,1,...,\frac{T}{\Delta t} - 1; \end{array} \right. \\ \\ \left\{ \begin{array}{l} \forall j, \qquad p_j^{n,k=T/\Delta t} = u_j^{n,k=T/\Delta t} - \bar{u}_j, \\ \text{or its regularized value given in (6.15);} \end{array} \right. \\ \\ \left\{ \begin{array}{l} p_j^{n,k} = p_j^{n,k+1} + \frac{\Delta t}{\Delta x}\left[\frac{\partial g}{\partial u}(u_j^{n,k}, u_{j+1}^{n,k})(p_{j+1}^{n,k} - p_j^{n,k}) \right. \\ \qquad\qquad \left. + \frac{\partial g}{\partial v}(u_{j-1}^{n,k}, u_j^{n,k})(p_j^{n,k} - p_{j-1}^{n,k}) \right], \\ \text{for } k = \frac{T}{\Delta t} - 1, ..., 1, 0; \end{array} \right. \\ \\ \left\{ \begin{array}{l} \forall j, \qquad u_j^{n+1,0} = u_j^{n,0} - \eta\, p_j^{n,0}, \\ \text{or a filtered version, see (6.22).} \end{array} \right. \end{array}\right\} \tag{6.20}$$

An issue in (6.20) is that the backward solution p may not be unique when u admits discontinuities. However, as long as the final data p^T remains Lipschitz continuous, both uniqueness and stability hold for "reversible solutions" because $u(t,\cdot)$ satisfies a one-sided Lipschitz condition (OSLC). Moreover, usual finite-difference (dual) schemes converge toward reversible solutions as the grid size $\Delta x \to 0$.

Theorem 6.5. *(see [26, Theorem 3.7]) Let p^T be a Lipschitz continuous function with Lipschitz constant $Lip(p^T)$. Let the backward scheme be consistent, satisfy a uniform OSLC and that monotonicity-preservation is ensured by a CFL restriction. Then the sequence $p^{\Delta x}$ of piecewise-constant functions converges as $\Delta x \to 0$ strongly in $L^1_{loc}(\mathbb{R}\times\mathbb{R}^+)$ and almost everywhere toward the reversible solution.*

Obviously, as $\mu^\varepsilon \in C^\infty(\mathbb{R})$, $\mu^\varepsilon * [u^\varepsilon(T,\cdot) - \bar{u}] \in C^\infty$ is also a Lipschitz function. So $p^\varepsilon(t,\cdot)$ is rigorously a reversible solution of the adjoint equation because the velocity field $u(t,\cdot)$ satisfies Oleinik's OSLC. Accordingly, most of the monotonicity-preserving backward finite-difference schemes approximating $p^\varepsilon(t,\cdot)$ converge toward this unique and stable reversible solution under the CFL restriction,

$$\Delta t \max_{j,k,n} |u_j^{n,k}| \leq \frac{\Delta x}{2}.$$

Remark 6.2. One may wonder why (6.20) wasn't initialized with the entropy solution of the "backward Burgers equation" (6.8) with $\bar{u}$ as its final data. Indeed, consider again the situation displayed on Fig. 6.2: that backward solution is the "ramp" displayed on the first row. Starting with these data, (6.20) produces the entropy shock, so that the velocity field acting on $p(T,\cdot) \equiv 0$ (for ε small enough if the

regularization (6.13) is applied) is discontinuous. By definition of reversible solutions, $p(t,\cdot)$ will remain null ($TV(p(t,\cdot)) = TV(p(T,\cdot))$) in the cone of dependence of the shock, so $p(t=0,\cdot) \equiv 0$, too, and the algorithm will always keep the "ramp" as a steady inverse design. Hence, initializing with (6.8) may not be a good strategy when seeking different, more oscillating inverse designs (belonging to K).

6.4.3 *Projection and component-wise gradient algorithm*

The previously introduced method allows to iteratively recover an initial profile which image by the evolution semi-group S_T matches the target $\bar{u}$, just by solving a simple transport equation for p^ε (for which supposedly many schemes may be available). Indeed, it's very easy to numerically check that trying to devise a "non-entropic backward evolution" for (6.8) starting from $\bar{u}$ is strongly ill-posed since entropy dissipation in shocks cancels many details (oscillations), so the backward evolution algorithm would have to guess all those missing data. Accordingly, a majority of (stable) finite-difference schemes converge to the (unique) reversible solution: hence only one initial profile (probably the centroid of the convex set of inverse designs) is generally recovered. Unfortunately different, generally more sharply-varying, profiles are sought in the realm of concrete applications: it may be necessary to circumvent the stability of the reversible solution p^ε by projecting it onto a closed linear subspace V of the Hilbert space $L^2(\mathbb{R})$. Let P_V stand for the orthogonal projection onto V and, still, $0 < \eta \ll 1$ so that,

$$\tilde{\mathscr{J}}_T^\varepsilon(\tilde{u}_0^{n+1}) = \tilde{\mathscr{J}}_T^\varepsilon(\tilde{u}_0^n) - \eta \left\langle \nabla \tilde{\mathscr{J}}_T^\varepsilon(\tilde{u}_0^n), P_V \nabla \tilde{\mathscr{J}}_T^\varepsilon(\tilde{u}_0^n) \right\rangle + o(\eta) \leq \tilde{\mathscr{J}}_T^\varepsilon(\tilde{u}_0^n),$$

because

$$\left\langle \nabla \tilde{\mathscr{J}}_T^\varepsilon(\tilde{u}_0^n), P_V \nabla \tilde{\mathscr{J}}_T^\varepsilon(\tilde{u}_0^n) \right\rangle = \| P_V \nabla \tilde{\mathscr{J}}_T^\varepsilon(\tilde{u}_0^n) \|^2 \geq 0. \tag{6.21}$$

Clearly, the definition (6.14) is quite convenient, because it suffices to apply it to φ_i, $i \in \mathbb{N}$, the Hilbertian base of the linear subspace V. Imposing that the control belongs to the subspace V, i.e., $u_0 \in L^1 \cap L^\infty \subset V \subset L^2(\mathbb{R})$,

$$u_0(x) = \sum_i \alpha_i \varphi_i(x), \qquad (\alpha_i) \in \ell^2(\mathbb{Z}),$$

at each iteration n, by taking all the duality brackets with $p_0^n = \sum_i \beta_i \varphi_i$,

$$\boxed{\left\langle u_0^{n+1}, \varphi_i \right\rangle = \left\langle u_0^n, \varphi_i \right\rangle - \eta \left\langle \nabla \tilde{\mathscr{J}}_T^\varepsilon(u_0^n), \varphi_i \right\rangle \quad \Leftrightarrow \quad \alpha_i^{n+1} = \alpha_i^n - \eta \cdot \beta_i^n.} \tag{6.22}$$

Instead of (6.22), one may apply filtering to p^T when V contains Lipschitz functions.

Remark 6.3. Let K be the convex (see Theorem 6.2) associated to a discontinuous target $\bar{u}$; the inverse design in V, obtained through the filtering (6.22), is not (in general) the projection of the centroid $C(x)$ of K. Since P_V is an orthogonal projection,

$$\int_{\mathbb{R}} |P_V(C)(x)|^2\,dx = \|P_V(C)\|^2 = \|C\|^2 - \|(Id-P_V)(C)\|^2 \leq \int_{\mathbb{R}} |C(x)|^2\,dx.$$

So, being $\eta(u)=|u|^2$ a convex entropy for Burgers equation, $P_V(C)$ would dissipate more entropy than C: in general, the minimizer generated by (6.20)–(6.22) is the projection onto V of an inverse design in K, different from its centroid, see Fig. 6.7.

6.4.4 *F-functions or scale-limited subspaces of* L^2

A natural choice for V is one of the nested scaled-limited subspaces of a multi-resolution analysis of $L^2(\mathbb{R})$, spanned by a so–called "scaling function".

Definition 6.2. Nested subspaces V_j form a Multi-Resolution Analysis of $L^2(\mathbb{R})$ if: $\{0\} \subset \cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots \subset L^2(\mathbb{R})$ and moreover,

- for all $f \in L^2(\mathbb{R})$, $\|P_{V_j} f - f\|_{L^2} \to 0$ as $j \to +\infty$ also, $P_{V_j} f \to 0$ as $j \to -\infty$.
- if $f(x) \in V_j$, then $f(x/2) \in V_{j-1}$ and for all $k \in \mathbb{Z}$, $f(x-2^j k) \in V_j$.
- there exists a shift-invariant orthonormal base of V_0 given by the scaling function $\varphi_i(x) = \varphi(x-i)$ for $i \in \mathbb{Z}$.

In this definition, P_{V_j} stands for the orthogonal projector onto the subspace V_j. Intuitively, it asks for the V_j's to be linear subspaces of functions with increasing spatial resolution: when j decreases, functions in V_j tend to become constants. Oppositely, when j increases, they are allowed to oscillate with high instantaneous frequency. The wavelet spaces W_j are defined as the orthogonal complement of V_j inside V_{j+1}: for all $j \in \mathbb{Z}$, $V_{j+1} = V_j \oplus W_j$. From φ_i, the base of V_0, a base of V_j is deduced by simple dilatation,

$$\varphi_{j,n}(x) = \sqrt{2^j}\varphi_n(2^j x) = \sqrt{2^j}\varphi(2^j x - n). \tag{6.23}$$

Thus, the orthogonal projection of f onto V_j reads:

$$P_{V_j} f = \sum_{n\in\mathbb{Z}} < f, \varphi_{j,n} > \varphi_{j,n}, \qquad < f, \varphi_{j,n} >= \int_{\mathbb{R}} f(x)\varphi_{j,n}(x).dx, \tag{6.24}$$

which is the best approximation of f in V_j in the least-squares sense. At this level, one has all the necessary elements to devise a "scale-limited, relaxed gradient algorithm", in component-wise form, by plugging each $\varphi_{j,n}$ into (6.22) for a given scale $j \in \mathbb{Z}$. Clearly, different choices for φ exist: the Haar function, the multi-fractal Daubechies-4, the smoother Symmlet, etc ...

For specific applications to aerodynamics problems related to sonic-boom imprints, so–called "F-functions" are preferred initial profiles (hence controls). A F-function can be represented by a finite set of parameters,

$$F(x) = \sum_{i=1}^{5} \alpha_i F_i(x), \qquad < F_i, F_j >= \delta_{i,j}.$$

Hence the former algorithm (6.22) applies as soon as the basis functions F_i are orthogonalized, for instance by means of a Gram-Schmidt procedure: Fig. 6.1 shows forward.in-time evolution of F with the control-to-state map S_T.

6.5 Preliminary numerical illustrations

Hereafter, the time at which the target is meant to be reached by the (forward) entropy solution of Burgers equation is normalized at unity: $T = 1$. The smallness assumption raised in Remark 6.1 corresponds to $\|\bar{u}\|_\infty \leq 1$. Our examples are treated with $2^8 = 256$ points, and an adaptive step-size η^n is used in the gradient algorithm:

$$\forall n,k, \qquad \eta^n = 2\left(\frac{\sum_j |p_j^{n,k=0}\, p_j^{n,k=T/\Delta t}|^2}{\sum_j |p_j^{n,k=0}|^2 \sum_j |p_j^{n,k=T/\Delta t}|^2}\right)^{\frac{1}{2}}$$

6.5.1 *Filtered gradient and hybrid methods*

The simplest filtering is obtained through the Haar scaling function, which is just the indicator of $x \in (0,1)$: see Fig. 6.6 for the inverse design of a nonnegative N-wave with increasing resolutions. Differently, Fig. 6.7 displays the inverse design

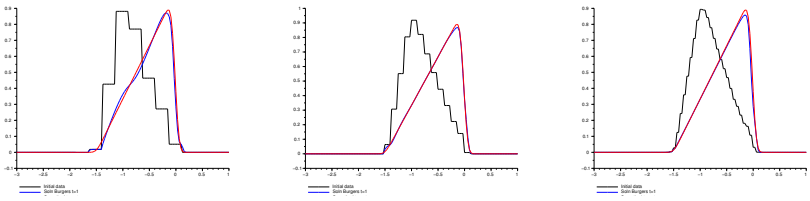

Fig. 6.6 Haar filters of increasing resolution; no filtering corresponds to maximal resolution.

of the same target, but with scaling functions corresponding to Daubechies-4 (left) and Symmlet-6 (right). For completeness, forward evolution by wavefront tracking, highlighting forward wave interactions (and cancellations), is displayed below.

6.5.2 *A target given by a (small) sine wave*

When there exists only one inverse design to a smooth target $\bar{u} \in C^\infty$, wavelet filtering may be used in order to ease the iterations of the gradient algorithm: selecting

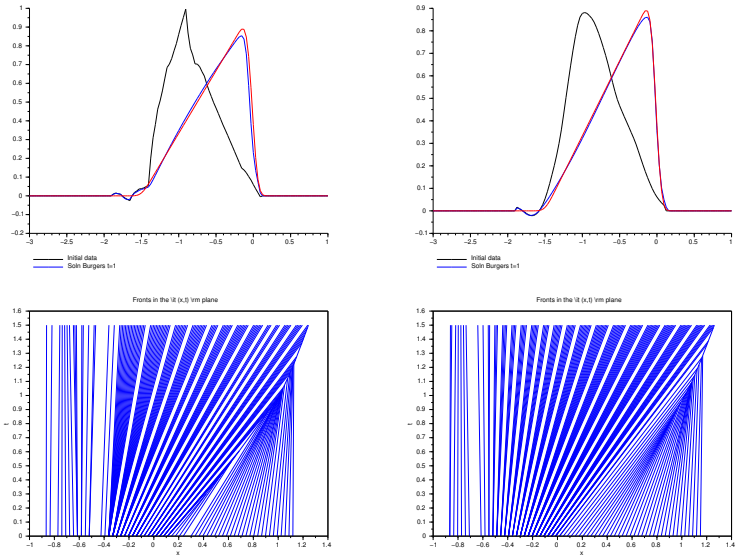

Fig. 6.7 Inversions with Daubechies-4 (left) and Symmlet-6 (right); forward evolution by WFT.

scaling functions $\varphi_{j,n}$ with many vanishing moments, the inverse design at $T=1$ of

$$\bar{u}(x)=\pm 0.35\sin(2\pi\frac{x+1}{3})\chi_{|x+1|<\frac{3}{2}},\qquad T\leq\frac{30}{7\pi}\simeq 1.36,$$

can be recovered, see Fig. 6.8. Despite a limited amount of grid points, the decay of the cost functional $\mathscr{J}_T$ is close to a factor 10^3 with 100 iterations of the adaptive filtered gradient algorithm. Of course, as the inverse design is unique, it matches the smooth solution of the backward Burgers equation (6.8) with final data $\bar{u}$.

6.5.3 Targets given by several N-waves

N-waves realize the large-time behavior (so–called "intermediate asymptotics") of compactly supported initial data: one shouldn't expect uniqueness of inverse designs when prescribing such type of (discontinuous) targets. Hereafter, we consider the final data used in Fig. 6.5, that is,

$$\bar{u}(x)=\beta(\frac{5}{4}+x)\chi_{|x+1|<1},\qquad T\leq\frac{1}{\beta},\qquad \beta>0.$$

On Fig. 6.9, inverse designs corresponding to $\beta=0.45,0.75$ and 0.95 are displayed (top to botton) for $T=1$, along with the decay of the objective functional $\mathscr{J}_T$ and

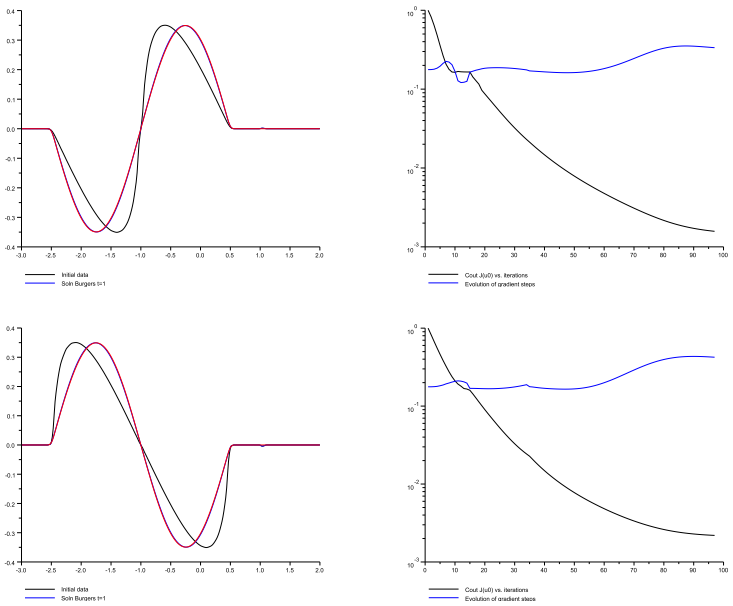

Fig. 6.8 Inverse design of sine waves with an hybrid method involving Coiflet-3 and Symmlet-6.

the values taken by the adaptive step-size. Filtering was made so that the first iterations are filtered with a smooth Coiflet-3 scaling function, but the subsequent ones (details at small scale) are treated with a multi-fractal Daubechies-4. Switching from one another is done between 10 and 20 iterations of the gradient algorithm, as one can see by means of the brutal variation of the adaptive step-size. The general convergence of the algorithm deteriorates with the increasing values of β, so that our inversion time $T = 1$ gets closer to the critical time prescribed by the existence Theorem 6.1. For bigger inversion times, the target ceases to be reachable, and numerical inverse designs display spikes, see also [2, §6.4]. Of course, by using different scaling functions, it is possible to get alternative inverse designs, see Fig. 6.10, along with all convex combinations which can be built from each two of them.

6.5.4 F-functions as combinations of "Haar and spikes"

By looking at Fig. 6.1, one may hope to recover interesting inverse designs starting from truncated N-waves at time $T = 1$ like,

$$\bar{u}(x) = \frac{3}{4}\min\left(\frac{4}{5}, \max\left(-\frac{1}{2}, \frac{5}{4}+x\right)\right)\chi_{|x+1|<1}, \qquad T \leq \frac{4}{3}.$$

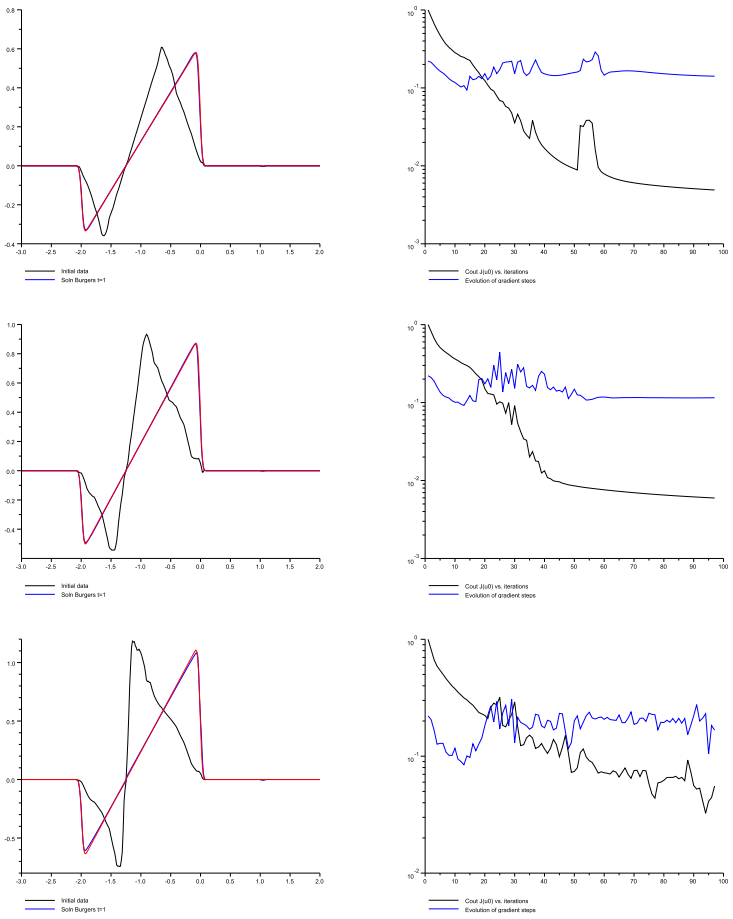

Fig. 6.9 Inverse design of N-waves with an hybrid method involving Coiflet-3 and Daubechies-4.

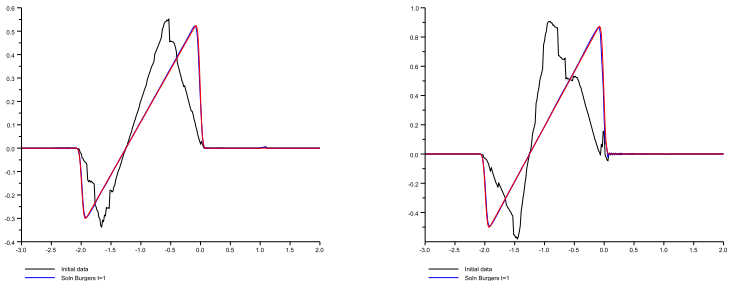

Fig. 6.10 Inverse design for $\beta = 0.45$ with hybrid methods involving Haar and Symmlet wavelets.

Our (mostly empirical, inspired by [43]) computational approach consists in

- setting up a very coarse Haar scaling function during the 10 first iterations,
- then, switching to a slightly finer one to correct smoother regions,
- during the whole process, alternating between a standard wavelet filtering and a "spike filtering", in which cumulative integrals of corrections $\int_{x_0}^{x_1} p(0,x)dx$, are concentrated at a given location $x \in (x_0, x_1)$.

Preliminary results are displayed on Fig. 6.11, to be compared with Fig. 6.1.

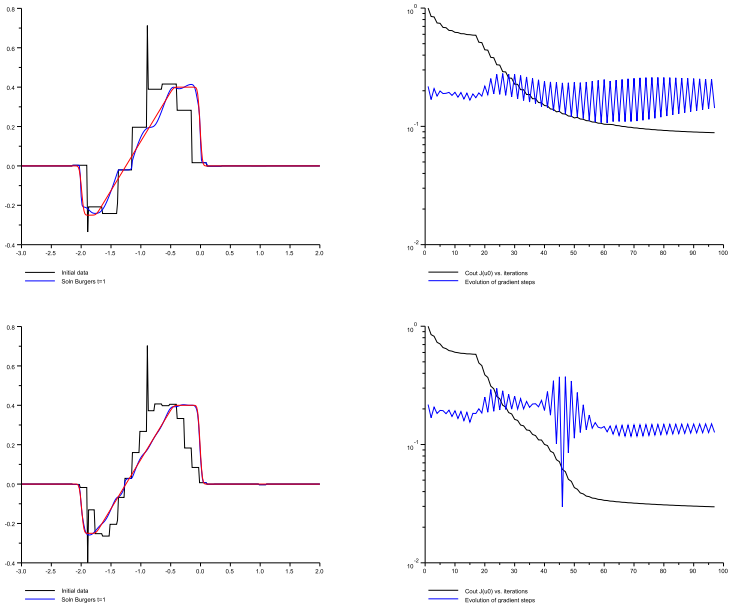

Fig. 6.11 Inverse designs similar to F-functions of a truncated N-wave with "Haar and spikes".

6.6 Conclusions

A few aspects of inverse design for convex scalar conservation laws, exemplified by 1D Burgers equation, were reviewed in the present note. In particular, the huge variety of initial data leading to an entropy-admissible, piecewise discontinuous target is revealed by Theorem 6.2. An iterative gradient algorithm relying on numerically usually produces the "backward entropy solution", mostly because numerically solving linearized backward transport equations brings out corresponding "reversible solutions". In order to circumvent that issue, and to be able to produce different, more steeply-varying numerical inverse designs, a filtering procedure was

applied, by adequate wavelet projectors. The convergence of such a process progressively degrades as targets become close to a non-admissible profile, though.

Acknowledgements Both the authors thank Prof. Andreas Griewank who suggested to study convex combinations of inverse designs during a meeting in september 2015.

References

1. Adimurthi; Ghoshal, Shyam Sundar; Veerappa Gowda, G. D. Exact controllability of scalar conservation laws with strict convex flux. Math. Control Relat. Fields 4 (2014) 401–449.
2. N. Allahverdi, A. Pozo, E. Zuazua, Numerical aspects of large-time optimal control of Burgers equation, ESAIM: M2AN, Forthcoming article, DOI: 10.1051/m2an/2015076.
3. J.J. Alonso, M.R. Colonno, Multidisciplinary Optimization with Applications to Sonic-Boom Minimization, Annual Review of Fluid Mechanics, vol. 44 (2012) 505–506.
4. F. Ancona, P. Cannarsa, K.T. Nguyen, Quantitative compactness estimates for Hamilton-Jacobi equations, Arch. Rational Mech. Anal. (2015),
5. C. Bardos, O. Pironneau, Data assimilation for conservation laws, Meth. Applic. Anal. (2005)
6. Barzilai, J., Borwein, J.M.: Two point step size gradient methods. IMA J. Numer. Anal. 8 (1988) 141–148
7. A.F. Berlinet, Ch. Roland, Geometric interpretation of some Cauchy related methods, Numer. Math. (2011) 119: 437–464.
8. F. Bouchut and F. James, *One-dimensional transport equations with discontinuous coefficients*, Nonlinear Analysis TMA, **32** (1998), 891 – 933.
9. F. Bouchut, F. James, Differentiability with respect to initial data for a scalar conservation law, Proceedings of the 7-th Conference on Hyperbolic Problems, Zürich, 1998 M. Fey & R. Jeltsch, Eds., International Series of Numerical Mathematics, 129, Birkhäuser, Basel, 1999, 113–118
10. A. Bressan, A. Marson, *A variational calculus for discontinuous solutions of systems of conservation laws*, Comm. in Partial Differential Equations **20** (1995) 1491–1552.
11. J.M. Burgers, Application of a model system to illustrate some points of the statistical theory of free turbulence, Proc. Konink. Nederl. Akad. Wetensch. 43 (1940) 2–12.
12. C. Castro, F. Palacios, E. Zuazua, *An alternative descent method for the optimal control of the inviscid Burgers equation in the presence of shocks*, M3AS (2008).
13. R.O. Cleveland, Propagation of sonic-booms through a real stratified atmosphere, Ph.D. dissertation, The University of Texas at Austin (1995).
14. J.D. Cole, On a quasi-linear parabolic equation occurring in aerodynamics, Quart. Appl. Math. 9 (1951) 225–236.
15. C.M. Dafermos, *Hyperbolic Conservation Laws in Continuum Physics,* Grundlehren Math. Wissenschaften Series **325**, Springer-Verlag, 2010.
16. I.C. Daubechies, A.C. Gilbert, *Harmonic analysis, wavelets, and applications*, in **Hyperbolic Equations and Frequency Interactions**, Luis Cafarelli and Weinan E, eds., IAS/Park City Mathematics Series, Vol. 5, 1998.
17. L.C. Evans, Partial Differential Equations, Second edition. Graduate Studies in Mathematics, 19. American Mathematical Society, Providence, RI, 2010..
18. A. Fikl, V. Le Chenadec, T. Sayadi, P.J. Schmid, A Comprehensive Study of Adjoint-Based Optimization of Non-Linear Systems with Application to Burgers' Equation, 46th AIAA Fluid Dynamics Conference – DOI: 10.2514/6.2016-3805.
19. A.R. George, R. Seebass, Sonic boom minimization including both front and rear shocks, AIAA Journal, vol. 9, no. 10 (1971) 2091–2093.
20. M. B. Giles and N. A. Pierce, An introduction to the Adjoint Approach to Design, Turbulence and Combustion, vol. 65 (2001) 393–415

21. M. Giles, S. Ulbricht, Convergence of linearized and adjoint approximations for discontinuous solutions of conservation laws. Part 1: linearized approximations and linearized output functionals, SIAM Journal on Numerical Analysis 48(3) (2010) 882–904
22. M. Giles, S. Ulbricht,Convergence of linearized and adjoint approximations for discontinuous solutions of conservation laws. Part 2: adjoint approximations and extensions, SIAM Journal on Numerical Analysis 48(3) (2010) 905–921.
23. Gilmore, P.; Kelley, C. T., An Implicit Filtering Algorithm for Optimization of Functions with Many Local Minima, SIAM J. Optimization 5 1995.
24. E. Godlewski, P.A. Raviart, *The linearized stability of solutions of nonlinear hyperbolic systems of conservation laws*, Math. Comput. Simul. **50** (1999) 77–95.
25. L. Gosse, *A Donoho-Stark criterion for stable signal recovery in discrete wavelet subspaces*, J. Comput. & Applied Math. **235** (2011) 5024–5039.
26. L. Gosse & F. James, *Numerical approximation of linear one-dimensional conservation equations with discontinuous coefficients*, Math. Comput. **69** (2000) 987–1015.
27. J.A. Hogan, J.D. Lakey, **Time-frequency and time-scale methods**, Birkhauser, 2005.
28. H. Holden and N.H. Risebro, **Front Tracking for Hyperbolic Conservation Laws**. Applied Mathematical Sciences **152** (Springer-Verlag, New York, 2002)
29. E. Hopf, The partial differential equation $u_t + uu_x = u_{xx}$, Comm. Pure Appl. Math. 3 (1950) 20–23.
30. R. Lecaros, E. Zuazua, Tracking control of the 1D scalar conservation laws in the presence of shocks, Trends in Contemporary Mathematics, Volume 8, Springer INdAM Series 195–219.
31. Y. Li, S. Osher, R. Tsai, Heat source identification based on ℓ^1 constrained minimization, Inverse Problems and Imaging 8 (2014) 199–221.
32. T.-P- Liu, M. Pierre, Source-Solutions and Asymptotic Behavior in Conservation Laws, J. Differ. Equations 51 (1984) 419-441.
33. S. Mallat, **A wavelet tour of signal processing**, Academic Press, 1998.
34. P. Martinez, J. Vancostenoble, Carleman estimates for one-dimensional degenerate heat equations, J. Evol. Equ. 6, no. 2, 325-362 (2006)
35. Y. Meyer, *Ondelettes sur l'intervalle*, Riv. Matem. Iberoamericana **7** (1991) 115–133.
36. Negreanu, M.; Matache, A.M.; Schwab, C., Wavelet Filtering for Exact Controllability of the wave equation, SIAM J. Scient. Comput. 28 (2006)
37. K.J. Plotkin, State of the art of sonic boom modeling, J. Acoust. Soc. Am. 111 (1), Pt. 2, Jan. 2002.
38. K.J. Plotkin, S.K. Rallabhandi, W. Li, Generalized Formulation and Extension of Sonic Boom Minimization Theory for Front and Aft Shaping, 47th AIAA Aerospace Sciences Meeting (Including The New Horizons Forum and Aerospace Exposition) 5–8 January 2009, Orlando, Florida.
39. A. Pozo, L.I. Ignat, E. Zuazua, *Large-time asymptotics, vanishing viscosity and numerics for 1-D scalar conservation laws*, Math. Comput. **84** (2015) 1633–1662.
40. S.K. Rallabhandi, Advanced sonic boom prediction using augmented Burgers equation, Journal of Aircraft 48 (2011), 1245-1253.
41. D. Serre, Systems of Conservation Laws I: Hyperbolicity, Entropies, Shock Waves, Cambridge University Press (1999).
42. D. Slepian, *Some comments on Fourier analysis, uncertainty and modeling*, SIAM Rev. **25** (1983) 379–393.
43. J. Tropp, On the linear independence of spikes and sines. Journal of Fourier Analysis and Applications, 14 (2008) 838–858.
44. J. Vancostenoble, E. Zuazua, Null controllability for the heat equation with singular inverse-square potential, J. Functional Analysis 254: 1864–1902 (2008).
45. Gilbert G Walter, Xiaoping Shen , *Wavelet based on prolete spheroidal wave functions*, J. Fourier Anal. Appl. 10 (2004), 1–26
46. H. Xiao, V. Rokhlin, N. Yarvin, *Prolate spheroidal wave functions, quadrature and interpolation*, Inverse Problems **17** (2001) 805–838.
47. G.B. Whitham, The Flow Pattern of a Supersonic Projectile, Comm. Pure Applied Math. V, 301-348 (1952).