

Conservation law constrained optimization based upon Front-Tracking

M. Gugat* M. Herty[†] A. Klar[†] G. Leugering*

November 15, 2005

Abstract

We consider models based on conservation laws. For the optimization of such systems, a sensitivity analysis is essential to determine how changes in the decision variables influence the objective function. Here we study the sensitivity with respect to the initial data of objective functions that depend upon the solution of Riemann problems with piecewise linear flux functions. We present representations for the one-sided directional derivatives of the objective functions. The results can be used in the numerical method called Front-Tracking.

1 Introduction

Conservation laws play an important role in physics and engineering. They appear for example as a model of traffic flow [17, 1, 13, 7], in two phase flow problems [16], in water flows [8, 9] and gas dynamics [16, 18, 4]. Usually, certain system parameters can be controlled and therefore, the necessity arises to deal with problems of optimal control. In general the optimal control problem is driven by nonlinear conservation laws and to obtain gradient information we need a sensitivity calculus that incorporates those nonlinear hyperbolic equations.

We consider a sensitivity calculus for scalar nonlinear hyperbolic conservation laws with piece-wise linear flux and piece-wise constant initial data. This setting was introduced by Dafermos in [3]. Later different authors extended the introduced ideas and developed a scheme, called front- [14] or wave-tracking [2] which is at one hand an analytical tool for proving existence and uniqueness of solutions, and at the other hand an efficient numerical method. For further references, see [11, 12, 14, 2].

*Institut für Angewandte Mathematik, FAU Erlangen-Nürnberg, Martensstr. 3, 91058 Erlangen, Germany, {gugat, leugering}@am.uni-erlangen.de

[†]Fachbereich Mathematik, TU Kaiserslautern, Postfach 3049, D-67653 Kaiserslautern, Germany, herty@rhrk.uni-kl.de, klar@mathematik.uni-kl.de

For effective numerical computations we derive a sensitivity calculus for nonlinear hyperbolic equations in the front-tracking setting. The numerical method computing the state **and** a consistent gradient. In terms of the Front-Tracking method, the additional effort for the gradient computation is neglectable. In a forthcoming publication we will present numerical examples using the theoretical results developed here. Also the theoretical results are valid for non-convex and non-concave flux functions.

Related work on sensitivity analysis for networked hyperbolic systems within the framework of continuously differentiable solutions has been given in [6, 5]. A sensitivity analysis for nonlinear hyperbolic conservation laws with source term has been given by S. Ulbrich in [19], [20]. We perform this analysis in the framework of front-tracking, since this yields explicit representations of the derivatives that can be used numerically.

2 Cauchy Problem

We consider Cauchy problems on the real line. Due to Dafermos [3] we have

Lemma 2.1. *Assume f is piece-wise linear, Lipschitz continuous on $[m, M]$ and*

$$u_0(x) = \begin{cases} u_l & x \leq 0 \\ u_r & x > 0 \end{cases} \quad (1)$$

where $u_l, u_r \in [m, M]$ are constants. Then there exists an entropic weak solution of

$$y_t(x, t) + f(y(x, t))_x = 0 \quad (2)$$

and

$$y(x, 0) = u_0(x) \quad (3)$$

which consists of a finite number of constant states separated by shocks centered at the origin.

We consider entropy weak solutions in the sense of Kruzkov, see [15]. An introduction to this type of solutions is given in [14].

The solution can be given explicitly by

$$y(x, t) = \begin{cases} u_l, & -\infty < x/t < \frac{f(y^1) - f(u_l)}{y^1 - u_l}, \\ y^1, & \frac{f(y^1) - f(u_l)}{y^1 - u_l} < x/t < \frac{f(y^2) - f(y^1)}{y^2 - y^1}, \\ \vdots & \\ y^k, & \frac{f(y^k) - f(y^{k-1})}{y^k - y^{k-1}} < x/t < \frac{f(u_r) - f(y^k)}{u_r - y^k}, \\ u_r, & \frac{f(u_r) - f(y^k)}{u_r - y^k} < x/t < \infty \end{cases} \quad (4)$$

where the points y^k are determined as follows. In the case $u_l < u_r$ the boundary of the convex hull of the set

$$\{(y, v) : u_l \leq y \leq u_r, v \geq f(y)\} \quad (5)$$

is a polygonal line with vertices at $(u_l, f(u_l)), (y^1, f(y^1)), \dots, (y^k, f(y^k)), (u_r, f(u_r))$ for $u_l < y^1 < \dots < y^k < u_r$. Herein $(y^i, f(y^i)), i = 1, \dots, k$ are also vertices of the graph of f . The construction is similar for the case $u_r < u_l$.

Let $h = \min\{|y^{i+1} - y^i| : i = 1, \dots, k - 1\}$, $\delta = (\delta_1, \delta_2)$ with $\|\delta\|_\infty < h$. We consider disturbed initial data

$$u_0^\delta(x) = \begin{cases} u_l + \delta_1 & x \leq 0 \\ u_r + \delta_2 & x > 0 \end{cases} \quad (6)$$

and the weak admissible solution y^δ to (2).

3 Main result: Sensitivity calculus for the objective function

We introduce the general optimal control problems for our discussion. We assume the objective functions of the following form, i.e., terminal measurement on a (sub-)set of the real line. Let $x_1 < 0 < x_2$ and $T > 0$ be given. Let g be a continuously differentiable real-valued function on \mathbb{R} with compact support. Define

$$I(u_0) = \int_{x_1}^{x_2} g(y(x, T)) dx \quad (7)$$

where y is the solution of the Cauchy problem (2),(3). We derive a sensitivity calculus for I in Theorem 3.1 below. The corresponding optimization problem is

$$\min_{u_0} I(u_0). \quad (8)$$

For the solution of (8), a sensitivity calculus for the objective function I is essential. In general only the one-sided directional derivatives of I exist.

Definition For $\delta \in \mathbb{R}^2$ with

$$m < u_l + \delta_1 < u_r + \delta_2 < M$$

let f_u^δ denote the lower convex envelope of f in the interval $[u_l + \delta_1, u_r + \delta_2]$. In particular, f_u^0 is the lower convex envelope of f in the interval $[u_l, u_r]$.

Theorem 3.1. Let I be as in (7). Let u_0 be as in (1) and $y_0(x, t)$ be the solution to the Cauchy problem (2),(3). Let u_0^δ be as in (6).

Assume that $u_l, u_r \in (m, M)$, $u_l < u_r$. Assume that $k \geq 1$.

For $\delta \in \mathbb{R}^2$, define

$$h(\delta) = I(u_0^\delta).$$

Then the following expansion holds:

$$I(u_0^\delta) = I(u_0) + \delta_1 \partial_{\delta_1 \pm} h(0) + \delta_2 \partial_{\delta_2 \pm} h(0) + O(\|\delta\|^2) \quad (9)$$

where the right-hand side derivatives ∂_{δ_1+} , ∂_{δ_2+} or the left-hand side derivatives ∂_{δ_1-} , ∂_{δ_2-} are chosen according to the sign of δ_1 , δ_2 respectively. The one-sided partial derivatives of h are given below.

Let

$$D_l = \frac{f(y_1) - f(u_l)}{y_1 - u_l}, \quad D_r = \frac{f(u_r) - f(y_k)}{u_r - y_k}.$$

Let $\chi_{[x_1, x_2]}$ denote the characteristic function of the interval $[x_1, x_2]$ that is

$$\chi_{[x_1, x_2]}(x) = \begin{cases} 0 & \text{if } x \notin [x_1, x_2], \\ 1 & \text{if } x \in [x_1, x_2]. \end{cases} \quad (10)$$

For an interval $[a, b]$, let $\mu([a, b]) = b - a$ denote the length of the interval.

If $D_l - f'_-(u_l) \leq 0$ holds, then

$$\partial_{\delta_1-} h(0) = g'(u_l) \mu([x_1, x_2] \cap (-\infty, TD_l]) - \frac{g(y_1) - g(u_l)}{y_1 - u_l} T [D_l - f'_-(u_l)] \chi_{[x_1, x_2]}(TD_l). \quad (11)$$

If $D_l - f'_-(u_l) > 0$ then

$$\partial_{\delta_1-} h(0) = g'(u_l) \mu([x_1, x_2] \cap (-\infty, T f'_-(u_l)]). \quad (12)$$

If $\{D_l = f'_+(u_l) \text{ or for all } u \in (u_l, y_1) : f(u) > f_u^0(u)\}$ holds, then

$$\partial_{\delta_1+} h(0) = g'(u_l) \mu([x_1, x_2] \cap (-\infty, TD_l]) - \frac{g(y_1) - g(u_l)}{y_1 - u_l} T [D_l - f'_+(u_l)] \chi_{[x_1, x_2]}(TD_l). \quad (13)$$

If $\{D_l \neq f'_+(u_l) \text{ and there exists } u \in (u_l, y_1) : f(u) = f_u^0(u)\}$ then

$$\partial_{\delta_1+} h(0) = g'(u_l) \mu([x_1, x_2] \cap (-\infty, TD_l]) - \frac{g(\bar{u}) - g(u_l)}{\bar{u} - u_l} T [D_l - f'_+(u_l)] \chi_{[x_1, x_2]}(TD_l). \quad (14)$$

If $f'_+(u_r) \leq D_r$ holds, then

$$\partial_{\delta_2+} h(0) = g'(u_r) \mu([x_1, x_2] \cap [TD_r, \infty)) + \frac{g(u_r) - g(y_k)}{u_r - y_k} T [D_r - f'_+(u_r)] \chi_{[x_1, x_2]}(TD_r). \quad (15)$$

If $f'_+(u_r) > D_r$ then

$$\partial_{\delta_2+} h(0) = g'(u_r) \mu([x_1, x_2] \cap [T f'_+(u_r), \infty)). \quad (16)$$

If $\{D_r = f'_-(u_r) \text{ or for all } u \in (y_k, u_r) : f(u) > f_u^0(u)\}$ holds, then

$$\partial_{\delta_2-} h(0) = g'(u_r) \mu([x_1, x_2] \cap [TD_r, \infty)) + \frac{g(y_k) - g(u_r)}{y_k - u_r} T [D_r - f'_-(u_r)] \chi_{[x_1, x_2]}(TD_r) \quad (17)$$

If $\{D_r \neq f'_-(u_r)\}$ and there exists $u \in (y_k, u_r) : f(u) = f_u^0(u)$, then

$$\partial_{\delta_2^-} h(0) = g'(u_r) \mu([x_1, x_2] \cap [TD_r, \infty)) + \frac{g(\bar{u}) - g(u_r)}{\bar{u} - u_r} T[D_r - f'_-(u_r)] \chi_{[x_1, x_2]}(TD_r) \quad (18)$$

Remark 3.2. *Theorem 3.1 holds true even for a non-convex objective function. In the case $k = 0$ the assertion of the theorem is true, once we replace $y_k = u_l$ and $y_1 = u_r$ at the appropriate places.*

Although our objective function is of integral type, in the derivative integrals do not appear explicitly. However, the terms where measures μ appear can obviously be interpreted as integrals. This is not the case for the additional shift-terms where the characteristic functions and the shock speeds appear.

4 Convex envelope of piecewise linear functions

For our analysis, we need a sensitivity Lemma about the behavior of the lower convex envelope of piecewise linear functions if the interval where the convex envelope is changed. It turns out that if the changes of the boundary points of the interval are small enough, the resulting changes in the convex hull are easy to describe: First, changes on the left-hand side are independent of changes on the right-hand side. Moreover, on each side either the number of breakpoints remains constant (in this case the first breakpoint is moved), or one additional breakpoint appears. As a breakpoint of a function f we understand a vertex of the lower convex envelope of f . To be more precise, we define:

Definition 4.1. *Let $f : \mathbb{R} \rightarrow [m, M]$ be a given piece-wise linear function.*

$u \in \mathbb{R}$ is called a breakpoint of f , if $(u, f(u))$ is a vertex of the lower convex envelope of f .

Lemma 4.2. *Let real numbers m, M be given such that $m < M$. Let $u_l, u_r \in (m, M)$ be given such that $u_l < u_r$. Consider a piecewise linear function f that is defined on $[m, M]$.*

Then f_u^δ is piecewise linear and the set of breakpoints of f_u^δ is contained in the set of breakpoints of f . Assume that f_u^0 has at least one breakpoint in (u_l, u_r) . Let y_1 denote the smallest breakpoint of f_u^0 that is larger than u_l . Let y_k denote the largest breakpoint of f_u^0 that is smaller than u_r .

If $|\delta_1| + |\delta_2|$ is sufficiently small, then between y_1 and y_k the function f_u^δ has the same breakpoints as f_u^0 .

If

$$\delta_1 < 0 \quad \text{and} \quad f'_-(u_l) \geq (f_u^0)'_+(u_l), \quad (19)$$

then $u_l + \delta_1$ is the only breakpoint of f_u^δ that is smaller than y_1 . This is also the case if $\delta_1 > 0$ and

$$\{f'_+(u_l) = (f_u^0)'_+(u_l) \text{ or for all } u \in (u_l, y_1) \text{ we have } f(u) > f_u^0(u)\}. \quad (20)$$

Otherwise f_u^δ has two breakpoints $u_l + \delta_1, \bar{u}$ that are smaller than y_1 . If $\delta_1 < 0$, $\bar{u} = u_l$. If $\delta_1 > 0$, the point \bar{u} is the smallest breakpoint of f in (u_l, y_1) such that on $[\bar{u}, y_1]$ we have $f_u^0 = f$.

If

$$\delta_2 > 0 \text{ and } f'_+(u_r) \leq (f_u^0)'_-(u_r), \quad (21)$$

then $u_r + \delta_2$ is the only breakpoint of f_u^δ that is larger than y_k . This is also the case if $\delta_2 < 0$ and

$$\{f'_-(u_r) = (f_u^0)'_-(u_r) \text{ or for all } u \in (y_k, u_r) \text{ we have } f(u) > f_u^0(u)\}. \quad (22)$$

Otherwise f_u^δ has two breakpoints $\bar{u}, u_r + \delta_2$ that are larger than y_k . If $\delta_2 > 0$, $\bar{u} = u_r$. If $\delta_2 < 0$, the point \bar{u} is the largest breakpoint of f in (y_k, u_r) such that on $[y_k, u_r]$ we have $f_u^0 = f$.

Before we start with the proof of the Lemma we provide two pictures, that should motivate the discussion below. The notation in Figure 1 and 2 are as in the above lemma.

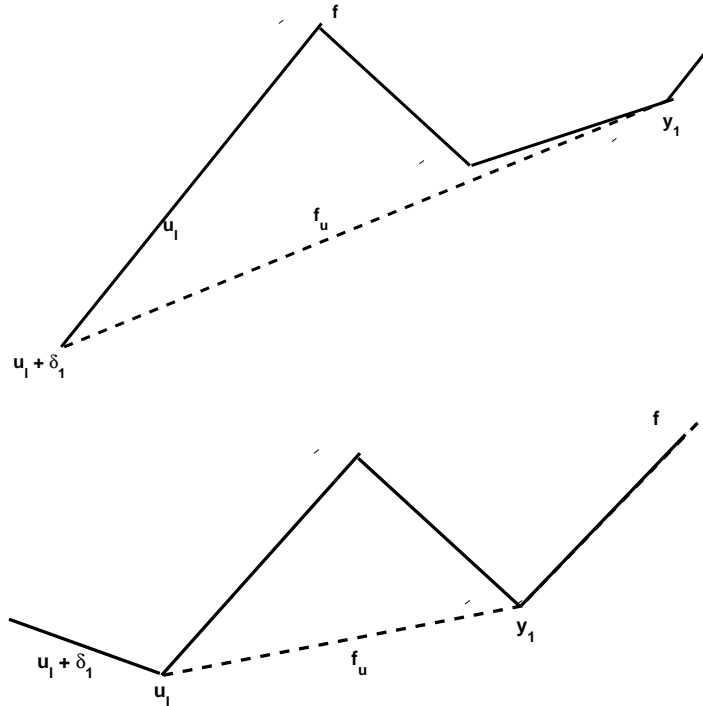


Figure 1: Example for the case $\delta_1 < 0$. Top part corresponds to the case where condition (19) is satisfied. Bottom part shows a case where condition (19) is violated and where we obtain an additional breakpoint.

Proof. The epigraph of f_u^δ is the convex hull of the epigraph of f . Since f is piecewise linear, the convex hull of the epigraph of f has a finite number of vertices. These vertices are breakpoints of f . This implies that f_u^δ is also

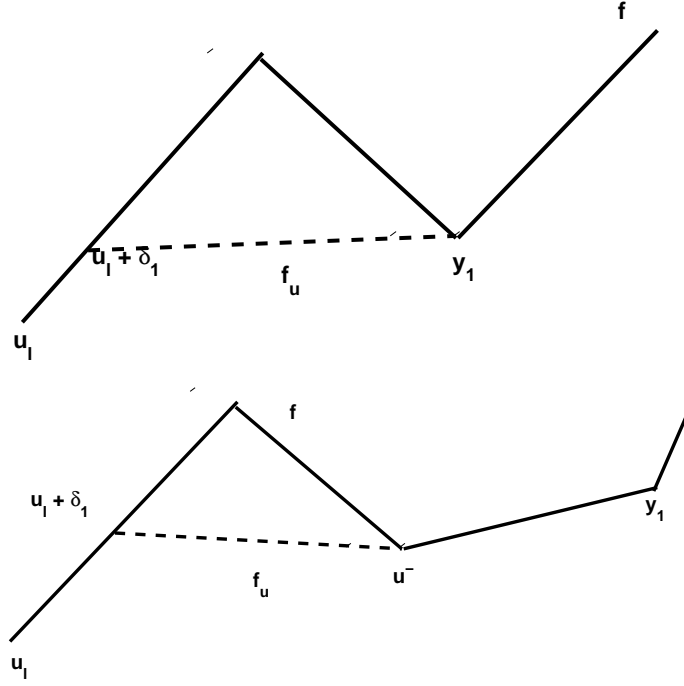


Figure 2: Example for the case $\delta_1 > 0$. Top part of the figure shows a case where (20) holds and the bottom part gives an example for a case where we obtain an additional breakpoint \bar{u} of f .

piecewise linear and the set of breakpoints of f_u^δ is contained in the set of breakpoints of f .

Choose $|\delta_1| > 0$ so small that f is linear between u_l and $u_l + \delta_1$. Choose $|\delta_2| > 0$ so small that f is linear between u_r and $u_r + \delta_2$.

We have

$$\frac{f(y_1) - f(u_l)}{y_1 - u_l} = (f_u^0)'_+(u_l) = (f_u^0)'_-(y_1) < (f_u^0)'_+(y_1).$$

Thus we can choose $|\delta_1|$ so small that we have

$$\frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1} < (f_u^0)'_+(y_1).$$

We have

$$\frac{f(u_r) - f(y_k)}{u_r - y_k} = (f_u^0)'_-(u_r) = (f_u^0)'_+(y_k) > (f_u^0)'_-(y_k).$$

Thus we can choose $|\delta_2|$ so small that we have

$$\frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k} > (f_u^0)'_-(y_k).$$

Since $(f_u^0)'_+(u_l) < (f_u^0)'_-(u_r)$, we can choose $|\delta_1| + |\delta_2|$ so small that

$$\frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1} < \frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k}. \quad (23)$$

4.1 The case $\delta_1 < 0, \delta_2 > 0$

Now we consider the case $\delta_1 < 0, \delta_2 > 0$ where $[u_l, u_r] \subset [u_l + \delta_1, u_r + \delta_2]$.

Consider the polygonal line through the points

$$(u_l + \delta_1, f(u_l + \delta_1)), (u_l, f(u_l)), (y_1, f(y_1)). \quad (24)$$

If it is the graph of a convex function, then the polygonal line through

$$(u_l + \delta_1, f(u_l + \delta_1)), (u_l, f(u_l)), (y_1, f(y_1)), \dots, (y_k, f(y_k))$$

is also the graph of a convex function g_1 . Moreover, $g_1 \leq f$ on $[u_l + \delta_1, y_k]$.

If the polygonal line through the points (24) is *not* the graph of a convex function (this is the case if (19) holds), then the polygonal line through the points

$$(u_l + \delta_1, f(u_l + \delta_1)), (y_1, f(y_1)), \dots, (y_k, f(y_k))$$

is the graph of a convex function g_1 . Moreover, $g_1 \leq f$ on $[u_l + \delta_1, y_k]$.

Now consider the polygonal line through

$$(y_k, f(y_k)), (u_r, f(u_r)), (u_r + \delta_2, f(u_r + \delta_2)). \quad (25)$$

If it is the graph of a convex function, then also the polygonal line through

$$(y_1, f(y_1)), \dots, (y_k, f(y_k)), (u_r, f(u_r)), (u_r + \delta_2, f(u_r + \delta_2))$$

is the graph of a convex function g_2 . Moreover, $g_2 \leq f$ on $[y_1, u_r + \delta_2]$.

If the polygonal line through the points (25) is *not* the graph of a convex function (this is the case if (21) holds), then the polygonal line through the points

$$(y_1, f(y_1)), \dots, (y_k, f(y_k)), (u_r + \delta_2, f(u_r + \delta_2))$$

is the graph of a convex function g_2 with $g_2 \leq f$ on $[y_1, u_r + \delta_2]$.

Thus we have constructed convex functions g_1, g_2 with $g_1 = g_2$ on $[y_1, y_k]$.

Define the function g on $[u_l + \delta_1, u_r + \delta_2]$ by $g(u) = g_1(u)$ for $u \in [u_l + \delta_1, y_1]$ and $g(u) = g_2(u)$ for $u > y_1$.

Then g is convex on $[u_l + \delta_1, u_r + \delta_2]$. (If $y_1 = y_k$, this is due to (23).) Moreover, $g \leq f$ on $[u_l + \delta_1, u_r + \delta_2]$. By our construction, at the breakpoints of g we have $f(u) = g(u)$. This implies that $g = f_u^\delta$.

4.2 The case $\delta_1 > 0, \delta_2 > 0$

Now we consider the case $\delta_1 > 0, \delta_2 > 0$.

Assume that (20) holds. If $f'_+(u_l) = (f_u^0)'_+(u_l)$, we define the convex function $g_3(u) = f_u^0$ on $[u + \delta_1, \dots, y_k]$. Then $g_3(u) \leq f$ on $[u + \delta_1, \dots, y_k]$ and $g_3(u + \delta_1) = f(u + \delta_1)$.

If $f'_+(u_l) > (f_u^0)'_+(u_l)$, condition (20) implies that for all $u \in (u_l, y_1)$, we have $f(u) > f_u^0(u)$. Hence if δ_1 is sufficiently small, the line through $(u_l + \delta_1, f(u_l + \delta_1))$, $(y_1, f(y_1))$ is below the graph of f . Thus for the convex function g_3 whose graph is the polygonal line through $(u_l + \delta_1, f(u_l + \delta_1))$, $(y_1, f(y_1))$, $\dots, (y_k, f(y_k))$ we have $g_3(u) \leq f$ on $[u + \delta_1, \dots, y_k]$.

If (20) is violated, then

$$f'_+(u_l) > (f_u^0)'_+(u_l) \quad (26)$$

and there exist breakpoints of f in (u_l, y_1) such that on $[\bar{u}, y_1]$ we have $f_u^0 = f$. Let \bar{u} be the smallest breakpoint of this type. Consider the polygonal line through the points

$$(u_l + \delta_1, f(u_l + \delta_1)), (\bar{u}, f(\bar{u})), (y_1, f(y_1)). \quad (27)$$

Since f is linear on the interval $[u_l, u_l + \delta_1]$, condition (26) implies that we have

$$\frac{f(\bar{u}) - f(u_l + \delta_1)}{\bar{u} - u_l - \delta_1} < \frac{f(\bar{u}) - f(u_l)}{\bar{u} - u_l} = (f_u^0)'_+(u_l). \quad (28)$$

Since

$$(f_u^0)'_+(u_l) = \frac{f(y_1) - f(\bar{u})}{y_1 - \bar{u}},$$

(28) shows that the polygonal line through the points (27) is the graph of a convex function h . Moreover, we can choose $|\delta_1|$ so small that on the interval $[u_l + \delta_1, y_1]$ we have $h \leq f$. Thus we can define the convex function g_3 on $[u_l + \delta_1, y_k]$ as the function whose graph is the polygonal line through the points

$$(u_l + \delta_1, f(u_l + \delta_1)), (\bar{u}, f(\bar{u})), (y_1, f(y_1)), \dots, (y_k, f(y_k)).$$

Then we have $g_3(u) \leq f$ on $[u + \delta_1, \dots, y_k]$.

Now define the function g_2 as in Section 4.1. Then we define the function g on $[u_l + \delta_1, u_r + \delta_2]$ by $g(u) = g_3(u)$ for $u \in [u_l + \delta_1, y_1]$ and $g(u) = g_2(u)$ for $u > y_1$. As in Section 4.1, we see that g is a convex function and $g = f_u^\delta$.

4.3 Other cases

The other two cases ($\delta_1 > 0, \delta_2 < 0$) and ($\delta_1 < 0, \delta_2 < 0$) are treated analogously by constructing convex function on $[u_l + \delta_1, y_k]$ and $[y_1, u_r + \delta_2]$ that are the lower convex envelope of f on these intervals. In fact, these two functions are equal to the lower convex envelope of f on $[u_l + \delta_1, u_r + \delta_2]$ on their respective domains.

5 Second order expansion of piecewise linear functions

We also need the following Lemma about piece-wise differentiable functions, is a statement about their first derivative.

Lemma 5.1. *Let f be a piecewise linear function. Let $u, y \in \mathbb{R}$, $u \neq y$ be given. Define*

$$D = \frac{f(y) - f(u)}{y - u}.$$

Assume that $|\delta|$ is such that f is linear between $u + \delta$ and u .

If $\delta < 0$, we have

$$\frac{f(y) - f(u + \delta)}{y - u - \delta} = D + \delta \frac{D - f'_-(u)}{y - u} + \frac{\delta^2}{(y - u)(y - u - \delta)} [D - f'_-(u)]. \quad (29)$$

If $\delta > 0$, we have

$$\frac{f(y) - f(u + \delta)}{y - u - \delta} = D + \delta \frac{D - f'_+(u)}{y - u} + \frac{\delta^2}{(y - u)(y - u - \delta)} [D - f'_+(u)]. \quad (30)$$

Proof. Let $\delta < 0$.

$$\begin{aligned} & \frac{f(y) - f(u + \delta)}{y - u - \delta} - D \\ &= \frac{f(y) - f(u + \delta)}{y - u - \delta} - \frac{f(y) - f(u)}{y - u} \\ &= \frac{[f(y) - f(u + \delta)][y - u] - [f(y) - f(u)][y - u - \delta]}{[y - u - \delta][y - u]} \\ &= \frac{[f(y) - f(u)]\delta + [f(u) - f(u + \delta)][y - u]}{[y - u - \delta][y - u]} \\ &= \delta \frac{D}{y - u - \delta} - \delta \frac{f(u) - f(u + \delta)}{-\delta} \frac{1}{y - u - \delta} \\ &= \delta [D - f'_-(u)] \frac{1}{y - u - \delta} \\ &= \delta \frac{D - f'_-(u)}{y - u} \left[1 + \frac{\delta}{y - u - \delta} \right]. \end{aligned}$$

The proof for $\delta > 0$ works analogously.

6 Approximation of the derivative of the state

Here we derive an approximation of the derivative of y with respect to u_0 .

We want to prove that a first-order expansion of u^δ with respect to δ exists.

Theorem 6.1. *Assume that $u_l, u_r \in (m, M)$, $u_l < u_r$. Assume that $k \geq 1$. There exists a function $H^\delta(x, t)$ of δ, x, t and such that for $\|\delta\|$ sufficiently small*

$$\|y^\delta - (y + H^\delta)\|_{L^1((0,T) \times \mathbb{R})} = O(\|\delta\|^2) \quad (31)$$

and H^δ can be computed in the following way:

$$H^\delta = H_{l-}^\delta + H_{r+}^\delta \text{ if } \delta_1 \leq 0 \text{ and } \delta_2 \geq 0,$$

$$H^\delta = H_{l+}^\delta + H_{r+}^\delta \text{ if } \delta_1 \geq 0 \text{ and } \delta_2 \geq 0,$$

etc. with $H_{l-}^\delta, H_{l+}^\delta, H_{r-}^\delta, H_{r+}^\delta$ defined in the following way:

$$\text{If } f'_-(u_l) \geq (f_u^0)'_+(u_l), \quad (32)$$

we have

$$H_{l-}^\delta(x, t) = \begin{cases} \delta_1, & -\infty < x/t < \frac{f(y_1) - f(u_l)}{y_1 - u_l}, \\ u_l - y_1, & \frac{f(y_1) - f(u_l)}{y_1 - u_l} < x/t < \frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1}, \\ 0, & \frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1} < x/t < \infty. \end{cases}$$

If (32) is violated, we have

$$H_{l-}^\delta(x, t) = \begin{cases} \delta_1, & -\infty < x/t < \frac{f(u_l) - f(u_l + \delta_1)}{-\delta_1}, \\ 0, & \frac{f(u_l) - f(u_l + \delta_1)}{-\delta_1} < x/t < \frac{f(y^1) - f(u^l)}{y^1 - u^l}, \\ 0, & \frac{f(y^1) - f(u^l)}{y^1 - u^l} < x/t < \infty. \end{cases}$$

$$\text{If } \{f'_+(u_l) = (f_u^0)'_+(u_l) \text{ or for all } u \in (u_l, y_1) \text{ we have } f(u) > f_u^0(u)\} \quad (33)$$

then

$$H_{l+}^\delta(x, t) = \begin{cases} \delta_1, & -\infty < x/t < \frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1}, \\ y_1 - u_l, & \frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1} < x/t < \frac{f(y_1) - f(u_l)}{y_1 - u_l}, \\ 0, & \frac{f(y_1) - f(u_l)}{y_1 - u_l} < x/t < \infty. \end{cases}$$

If (33) is violated, we have

$$H_{l+}^\delta(x, t) = \begin{cases} \delta_1, & -\infty < x/t < \frac{f(\bar{u}) - f(u_l + \delta_1)}{\bar{u} - u_l - \delta_1}, \\ \bar{u} - u_l, & \frac{f(\bar{u}) - f(u_l + \delta_1)}{\bar{u} - u_l - \delta_1} < x/t < \frac{f(y_1) - f(u_l)}{y_1 - u_l}, \\ 0, & \frac{f(y_1) - f(u_l)}{y_1 - u_l} < x/t < \infty. \end{cases}$$

$$\text{If } f'_+(u_r) \leq (f_u^0)'_-(u_r), \quad (34)$$

we have

$$H_{r+}^{\delta}(x, t) = \begin{cases} 0, & -\infty < x/t < \frac{f(u_r+\delta_2)-f(y_k)}{u_r+\delta_2-y_k}, \\ u_r - y_k, & \frac{f(u_r+\delta_2)-f(y_k)}{u_r+\delta_2-y_k} < x/t < \frac{f(u_r)-f(y_k)}{u_r-y_k}, \\ \delta_2, & \frac{f(u_r)-f(y_k)}{u_r-y_k} < x/t < \infty. \end{cases}$$

If (34) is violated, we have

$$H_{r+}^{\delta}(x, t) = \begin{cases} 0, & -\infty < x/t < \frac{f(u_r)-f(y_k)}{u_r-y_k}, \\ 0, & \frac{f(u_r)-f(y_k)}{u_r-y_k} < x/t < \frac{f(u_r+\delta_2)-f(u_r)}{\delta_2}, \\ \delta_2, & \frac{f(u_r+\delta_2)-f(u_r)}{\delta_2} < x/t < \infty. \end{cases}$$

$$\text{If } \{f'_-(u_r) = (f_u^0)'_-(u_r) \text{ or for all } u \in (y_k, u_r) \text{ we have } f(u) > f_u^0(u)\} \quad (35)$$

we have

$$H_{r-}^{\delta}(x, t) = \begin{cases} 0, & -\infty < x/t < \frac{f(u_r)-f(y_k)}{u_r-y_k}, \\ y_k - u_r, & \frac{f(u_r)-f(y_k)}{u_r-y_k} < x/t < \frac{f(u_r+\delta_2)-f(y_k)}{u_r+\delta_2-y_k}, \\ \delta_2, & \frac{f(u_r+\delta_2)-f(y_k)}{u_r+\delta_2-y_k} < x/t < \infty. \end{cases}$$

If (35) is violated, we have

$$H_{r-}^{\delta}(x, t) = \begin{cases} 0, & -\infty < x/t < \frac{f(u_r)-f(y_k)}{u_r-y_k}, \\ \bar{u} - u_r, & \frac{f(u_r)-f(y_k)}{u_r-y_k} < x/t < \frac{f(u_r+\delta_2)-f(\bar{u})}{u_r+\delta_2-\bar{u}}, \\ \delta_2, & \frac{f(u_r+\delta_2)-f(\bar{u})}{u_r+\delta_2-\bar{u}} < x/t < \infty. \end{cases}$$

Proof.

First we compute y^{δ} . The change from u_l to $u_l + \delta_1$ causes a shift of the first vertex in the polygonal line that is the boundary of the convex hull of the set

$$\{(y, v) : u_l + \delta_1 \leq y \leq u_r + \delta_2, v \geq f(y)\}. \quad (36)$$

Moreover, if $|\delta_1|$ is sufficiently small, at most one additional vertex compared with the set (5) appears between the first and the second vertex.

The change from u_r to $u_r + \delta_2$ causes a shift of the last vertex. Moreover, if $|\delta_2|$ is sufficiently small, at most one additional vertex compared with the set (5) appears between the last and the second to last vertex.

6.0.1 Case 1: Assume that $\delta_1 < 0$.

Subcase 1a) Lemma 4.2 implies that if $|\delta_1| + |\delta_2|$ is sufficiently small, and (32) holds, the first vertex in the convex hull of the set (5) is shifted to $(u_l + \delta_1, f(u_l + \delta_1))$ and no additional vertices are generated.

In this case we have

$$y^{\delta}(x, t) = \begin{cases} u_l + \delta_1, & -\infty < x/t < \frac{f(y^1)-f(u_l+\delta_1)}{y^1-u_l-\delta_1}, \\ y^1, & \frac{f(y^1)-f(u_l+\delta_1)}{y^1-u_l-\delta_1} < x/t < \frac{f(y^2)-f(y^1)}{y^2-y^1}, \\ \vdots & \end{cases} \quad (37)$$

Subcase 1b) Lemma 4.2 implies that if $|\delta_1| + |\delta_2|$ is sufficiently small, and (32) is violated, the first vertex in the convex hull of the set (5) is shifted to $(u_l + \delta_1, f(u_l + \delta_1))$ and the point $(u_l, f(u_l))$ appears as an additional vertex.

In this case we have

$$y^\delta(x, t) = \begin{cases} u_l + \delta_1, & -\infty < x/t < \frac{f(u_l) - f(u_l + \delta_1)}{-\delta_1}, \\ u^l, & \frac{f(u_l) - f(u_l + \delta_1)}{-\delta_1} < x/t < \frac{f(y^1) - f(u^l)}{y^1 - u^l}, \\ y^1, & \frac{f(y^1) - f(u^l)}{y^1 - u^l} < x/t < \frac{f(y^2) - f(y^1)}{y^2 - y^1}, \\ \vdots & \end{cases} \quad (38)$$

Detection of Subcase 1a)/ Subcase 1b) How can we detect the Subcase 1a)/ Subcase 1b) respectively numerically?

Let p be the largest breakpoint of f that is less than u_l if such a breakpoint exists, and $p = u_l - 1$ otherwise.

If

$$f'_-(u_l) = \frac{f(u_l) - f(p)}{u_l - p} \geq \frac{f(y_1) - f(u_l)}{y_1 - u_l} \quad (39)$$

we are in Subcase 1a); otherwise we are in Subcase 1b). Note that the last differential quotient in (39) gives the right-hand side derivative in the point u_l of the function that is defined by the boundary of the convex hull of the epigraph of f .

6.0.2 Case 2: Assume that $\delta_1 > 0$.

Subcase 2a)

Lemma 4.2 implies that if $|\delta_1| + |\delta_2|$ is sufficiently small, and (33) holds, the first vertex in the convex hull of the set (5) is shifted to $(u_l + \delta_1, f(u_l + \delta_1))$ and no additional vertices are generated.

In this case we have

$$y^\delta(x, t) = \begin{cases} u_l + \delta_1, & -\infty < x/t < \frac{f(y^1) - f(u_l + \delta_1)}{y^1 - u_l - \delta_1}, \\ y^1, & \frac{f(y^1) - f(u_l + \delta_1)}{y^1 - u_l - \delta_1} < x/t < \frac{f(y^2) - f(y^1)}{y^2 - y^1}, \\ \vdots & \end{cases} \quad (40)$$

Note that $y^\delta(x, t)$ has the same form as in Subcase 1a).

Subcase 2b)

Lemma 4.2 implies that if $|\delta_1| + |\delta_2|$ is sufficiently small, and (33) is violated, the first vertex in the convex hull of the set (5) is shifted to $(u_l + \delta_1, f(u_l + \delta_1))$ and the point $(\bar{u}, f(\bar{u}))$ appears as an additional vertex.

In this case we have

$$y^\delta(x, t) = \begin{cases} u_l + \delta_1, & -\infty < x/t < \frac{f(\bar{u}) - f(u_l + \delta_1)}{\bar{u} - u_l - \delta_1}, \\ \bar{u}, & \frac{f(\bar{u}) - f(u_l + \delta_1)}{\bar{u} - u_l - \delta_1} < x/t < \frac{f(y^1) - f(\bar{u})}{y^1 - \bar{u}}, \\ y^1, & \frac{f(y^1) - f(\bar{u})}{y^1 - \bar{u}} < x/t < \frac{f(y^2) - f(y^1)}{y^2 - y^1}, \\ \vdots & \end{cases} \quad (41)$$

Detection of Subcase 2a)/ Subcase 2b) In order to check whether we are in Subcase 2b), we have to verify whether the following condition holds:

$$f'_+(u_l) > \frac{f(y_1) - f(u_l)}{y_1 - u_l} \quad (42)$$

and there exists a breakpoint \bar{u} of f in the open interval (u_l, y_1) such that

$$\bar{u} = \bar{\lambda}u_l + (1 - \bar{\lambda})y_1, \quad f(\bar{u}) = \bar{\lambda}f(u_l) + (1 - \bar{\lambda})f(y_1). \quad (43)$$

If (42) and (43) hold, we are in Subcase 2b), otherwise we are in Subcase 2a).

In Subcase 2b) we have to determine the smallest \bar{u} satisfying (43). This is in fact a finite optimization problem since we only have to compare the values of f and the lower convex envelope of f on (u_l, y_1) at the breakpoints of f in (u_l, y_1) .

Now we analyze analogously the changes that can occur as a result of the shift from u_r to $u_r + \delta_2$ if $|\delta_2|$ is sufficiently small. Again we have overall four cases.

6.0.3 Case 1: Assume that $\delta_2 > 0$.

Subcase 1a)

Lemma 4.2 implies that if $|\delta_1| + |\delta_2|$ is sufficiently small, and (34) holds, the last vertex in the convex hull of the set (5) is shifted to $(u_r + \delta_2, f(u_r + \delta_2))$ and no additional vertices are generated.

In this case we have

$$y^\delta(x, t) = \begin{cases} \vdots \\ y_k, & \frac{f(y^k) - f(y^{k-1})}{y^k - y^{k-1}} < x/t < \frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k}, \\ u_r + \delta_2, & \frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k} < x/t < \infty. \end{cases} \quad (44)$$

Subcase 1b) Lemma 4.2 implies that if $|\delta_1| + |\delta_2|$ is sufficiently small, and (34) is violated, the last vertex in the convex hull of the set (5) is shifted to $(u_r + \delta_2, f(u_r + \delta_2))$ and the point $(u_r, f(u_r))$ appears as an additional vertex.

In this case we have

$$y^\delta(x, t) = \begin{cases} \vdots \\ y_k, & \frac{f(y^k) - f(y^{k-1})}{y^k - y^{k-1}} < x/t < \frac{f(u_r) - f(y_k)}{u_r - y_k}, \\ u_r, & \frac{f(u_r) - f(y_k)}{u_r - y_k} < x/t < \frac{f(u_r + \delta_2) - f(u_r)}{\delta_2}, \\ u_r + \delta_2, & \frac{f(u_r + \delta_2) - f(u_r)}{\delta_2} < x/t < \infty. \end{cases} \quad (45)$$

Detection of Subcase 1a)/ Subcase 1b) If

$$f'_+(u_r) = \frac{f(u_r + \delta_2) - f(u_r)}{\delta_2} \leq \frac{f(u_r) - f(y_k)}{u_r - y_k} \quad (46)$$

we are in Subcase 1a), otherwise we are in Subcase 1b). Note that the last differential quotient in (46) gives the left-hand side derivative in the point u_r of the lower convex envelope of f on $[u_r, u_l]$.

6.0.4 Case 2: Assume that $\delta_2 < 0$.

Subcase 2a) Lemma 4.2 implies that if $|\delta_1| + |\delta_2|$ is sufficiently small, and (35) holds, the last vertex in the convex hull of the set (5) is shifted to $(u_r + \delta_2, f(u_r + \delta_2))$ and no additional vertices are generated.

In this case we have

$$y^\delta(x, t) = \begin{cases} \vdots \\ y_k, & \frac{f(y^k) - f(y^{k-1})}{y^k - y^{k-1}} < x/t < \frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k}, \\ u_r + \delta_2, & \frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k} < x/t < \infty. \end{cases} \quad (47)$$

Note that $y^\delta(x, t)$ has the same form as in Subcase 1a).

Subcase 2b) Lemma 4.2 implies that if $|\delta_1| + |\delta_2|$ is sufficiently small, and (35) is violated, the last vertex in the convex hull of the set (5) is shifted to $(u_r + \delta_2, f(u_r + \delta_2))$ and a point $(\bar{u}, f(\bar{u}))$ appears as an additional vertex; This point \bar{u} is a breakpoint of f that can be found as follows:

There exists $\bar{\lambda} \in (0, 1)$ such that

$$f(\bar{\lambda}y_k + (1 - \bar{\lambda})u_r) = \bar{\lambda}f(y_k) + (1 - \bar{\lambda})f(u_r).$$

Choose the largest $\bar{\lambda}$ of this type; in fact this $\bar{\lambda}$ is less than one. Define

$$\bar{u} = \bar{\lambda}y_k + (1 - \bar{\lambda})u_r.$$

In this case we have

$$y^\delta(x, t) = \begin{cases} \vdots \\ y_k, & \frac{f(y^k) - f(y^{k-1})}{y^k - y^{k-1}} < x/t < \frac{f(\bar{u}) - f(y_k)}{\bar{u} - y_k}, \\ \bar{u}, & \frac{f(\bar{u}) - f(y_k)}{\bar{u} - y_k} < x/t < \frac{f(u_r + \delta_2) - f(\bar{u})}{u_r + \delta_2 - \bar{u}}, \\ u_r + \delta_2, & \frac{f(u_r + \delta_2) - f(\bar{u})}{u_r + \delta_2 - \bar{u}} < x/t < \infty. \end{cases}$$

Detection of Subcase 2a)/ Subcase 2b) If

$$f'_-(u_r) < \frac{f(u_r) - f(y_k)}{u_r - y_k} \quad (48)$$

and there exists a breakpoint \bar{u} of f in the open interval (y_k, u_r) such that

$$\bar{u} = \bar{\lambda}y_k + (1 - \bar{\lambda})u_r, \text{ and } f(\bar{u}) = \bar{\lambda}f(y_k) + (1 - \bar{\lambda})f(u_r). \quad (49)$$

then we are in Subcase 2b).

If (48) or (49) is violated, we are in Subcase 2a).

6.0.5 The expansion of y^δ

The above considerations show that in the expansion

$$y^\delta = y + H^\delta$$

we can decompose H^δ in the form

$$H^\delta = H_l^\delta + H_r^\delta,$$

where H_l^δ gives the sensitivity with respect to u_l and H_r^δ contains the sensitivity information for u_r . Moreover, we have to consider one-sided sensitivities H_{l+}^δ , H_{l-}^δ , H_{r+}^δ , H_{r-}^δ corresponding to the cases $\delta_1 > 0$, $\delta_1 < 0$, $\delta_2 > 0$, $\delta_2 < 0$ respectively.

Define

$$G^\delta(x, t) = y^\delta(x, t) - y(x, t).$$

Variations of δ_1 : First we consider the effect of the variations of δ_1 . In Subcase 1a) we have

$$G_{l-}^\delta(x, t) = \begin{cases} \delta_1, & -\infty < x/t < \frac{f(y_1) - f(u_l)}{y_1 - u_l}, \\ u_l + \delta_1 - y_1, & \frac{f(y_1) - f(u_l)}{y_1 - u_l} < x/t < \frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1}, \\ 0, & \frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1} < x/t < \infty. \end{cases} \quad (50)$$

Note that

$$\frac{f(y_1) - f(u_l)}{y_1 - u_l} \leq \frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1}$$

since $u_l \in [u_l + \delta_1, y_1]$ is not a breakpoint of the lower convex envelope of f on $[u_l + \delta_1, y_1]$.

To obtain a first order approximation of G_{l-}^δ we define

$$H_{l-}^\delta(x, t) = \begin{cases} \delta_1, & -\infty < x/t < \frac{f(y_1) - f(u_l)}{y_1 - u_l}, \\ u_l - y_1, & \frac{f(y_1) - f(u_l)}{y_1 - u_l} < x/t < \frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1}, \\ 0, & \frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1} < x/t < \infty. \end{cases} \quad (51)$$

In Subcase 1b) we have

$$G_{l-}^\delta(x, t) = \begin{cases} \delta_1, & -\infty < x/t < \frac{f(u_l) - f(u_l + \delta_1)}{-\delta_1}, \\ 0, & \frac{f(u_l) - f(u_l + \delta_1)}{-\delta_1} < x/t < \frac{f(y^1) - f(u^l)}{y^1 - u^l}, \\ 0, & \frac{f(y^1) - f(u^l)}{y^1 - u^l} < x/t < \infty. \end{cases} \quad (52)$$

To obtain a first order approximation of G_{l-}^δ we define

$$H_{l-}^\delta(x, t) = G_{l-}^\delta(x, t). \quad (53)$$

In Subcase 2a) we have

$$G_{l+}^{\delta}(x, t) = \begin{cases} \delta_1, & -\infty < x/t < \frac{f(y_1)-f(u_l+\delta_1)}{y_1-u_l-\delta_1} \\ y_1 - u_l, & \frac{f(y_1)-f(u_l+\delta_1)}{y_1-u_l-\delta_1} < x/t < \frac{f(y_1)-f(u_l)}{y_1-u_l} \\ 0, & \frac{f(y_1)-f(u_l)}{y_1-u_l} < x/t < \infty. \end{cases} \quad (54)$$

To obtain a first order approximation of G_{l+}^{δ} we define

$$H_{l+}^{\delta}(x, t) = G_{l+}^{\delta}(x, t). \quad (55)$$

In Subcase 2b) we have

$$G_{l+}^{\delta}(x, t) = \begin{cases} \delta_1, & -\infty < x/t < \frac{f(\bar{u})-f(u_l+\delta_1)}{\bar{u}-u_l-\delta_1} \\ \bar{u} - u_l, & \frac{f(\bar{u})-f(u_l+\delta_1)}{\bar{u}-u_l-\delta_1} < x/t < \frac{f(y_1)-f(u_l)}{y_1-u_l} \\ 0, & \frac{f(y_1)-f(u_l)}{y_1-u_l} < x/t < \infty. \end{cases} \quad (56)$$

To obtain a first order approximation of G_{l+}^{δ} we define

$$H_{l+}^{\delta}(x, t) = G_{l+}^{\delta}(x, t). \quad (57)$$

Variations of δ_2 : Now we consider the effect of the variations of δ_2 .

In Subcase 1a) we have

$$G_{r+}^{\delta}(x, t) = \begin{cases} 0, & -\infty < x/t < \frac{f(u_r+\delta_2)-f(y_k)}{u_r+\delta_2-y_k}, \\ u_r + \delta_2 - y_k, & \frac{f(u_r+\delta_2)-f(y_k)}{u_r+\delta_2-y_k} < x/t < \frac{f(u_r)-f(y_k)}{u_r-y_k} \\ \delta_2, & \frac{f(u_r)-f(y_k)}{u_r-y_k} < x/t < \infty. \end{cases} \quad (58)$$

To obtain a first order approximation of G_{r+}^{δ} , we define

$$H_{r+}^{\delta}(x, t) = \begin{cases} 0, & -\infty < x/t < \frac{f(u_r+\delta_2)-f(y_k)}{u_r+\delta_2-y_k}, \\ u_r - y_k, & \frac{f(u_r+\delta_2)-f(y_k)}{u_r+\delta_2-y_k} < x/t < \frac{f(u_r)-f(y_k)}{u_r-y_k} \\ \delta_2, & \frac{f(u_r)-f(y_k)}{u_r-y_k} < x/t < \infty. \end{cases} \quad (59)$$

In Subcase 1b) we have

$$G_{r+}^{\delta}(x, t) = \begin{cases} 0, & -\infty < x/t < \frac{f(u_r)-f(y_k)}{u_r-y_k}, \\ 0, & \frac{f(u_r)-f(y_k)}{u_r-y_k} < x/t < \frac{f(u_r+\delta_2)-f(u_r)}{\delta_2} \\ \delta_2, & \frac{f(u_r+\delta_2)-f(u_r)}{\delta_2} < x/t < \infty. \end{cases} \quad (60)$$

To obtain a first order approximation of G_{r+}^{δ} , we define

$$H_{r+}^{\delta}(x, t) = G_{r+}^{\delta}(x, t). \quad (61)$$

In Subcase 2a) we have

$$G_{r-}^{\delta}(x, t) = \begin{cases} 0, & -\infty < x/t < \frac{f(u_r)-f(y_k)}{u_r-y_k}, \\ y_k - u_r, & \frac{f(u_r)-f(y_k)}{u_r-y_k} < x/t < \frac{f(u_r+\delta_2)-f(y_k)}{u_r+\delta_2-y_k} \\ \delta_2, & \frac{f(u_r+\delta_2)-f(y_k)}{u_r+\delta_2-y_k} < x/t < \infty. \end{cases} \quad (62)$$

To obtain a first order approximation of G_{r-}^δ , we define

$$H_{r-}^\delta(x, t) = G_{r-}^\delta(x, t). \quad (63)$$

In Subcase 2b) we have

$$G_{r-}^\delta(x, t) = \begin{cases} 0, & -\infty < x/t < \frac{f(u_r) - f(y_k)}{u_r - y_k}, \\ \bar{u} - u_r, & \frac{f(u_r) - f(y_k)}{u_r - y_k} < x/t < \frac{f(u_r + \delta_2) - f(\bar{u})}{u_r + \delta_2 - \bar{u}}, \\ \delta_2, & \frac{f(u_r + \delta_2) - f(\bar{u})}{u_r + \delta_2 - \bar{u}} < x/t < \infty. \end{cases} \quad (64)$$

To obtain a first order approximation of G_{r-}^δ , we define

$$H_{r-}^\delta(x, t) = G_{r-}^\delta(x, t). \quad (65)$$

6.0.6 Convergence of the expansion of y^δ

Let

$$l(\delta) = \max\left\{\frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1}, \frac{f(y_1) - f(u_l)}{y_1 - u_l}\right\} \text{ and}$$

$$r(\delta) = \min\left\{\frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k}, \frac{f(u_r) - f(y_k)}{u_r - y_k}\right\}.$$

Then we have

$$\begin{aligned} & \|y^\delta - (y + H^\delta)\|_{L^1((0, T) \times \mathbb{R})} \\ &= \int_0^T \int_{\mathbb{R}} |y^\delta(x, t) - y(x, t) - H^\delta(x, t)| dx dt \\ &= \int_0^T \int_{-\infty}^{l(\delta)t} |G^\delta(x, t) - H^\delta(x, t)| dx dt \\ &\quad + \int_0^T \int_{r(\delta)t}^{\infty} |G^\delta(x, t) - H^\delta(x, t)| dx dt \\ &\leq \int_0^T \delta_1 t \left| \frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1} - \frac{f(y_1) - f(u_l)}{y_1 - u_l} \right| dt \\ &\quad + \int_0^T \delta_2 t \left| \frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k} - \frac{f(u_r) - f(y_k)}{u_r - y_k} \right| dt \\ &\leq \frac{T^2}{2} [\delta_1^2 h_1(u_l, y_1, f) + \delta_2^2 h_2(y_k, u_r, f)] \end{aligned}$$

with real numbers $h_1(u_l, y_1, f)$, $h_2(y_k, u_r, f)$. Note that for δ_1 only in Subcase 1a) we have a contribution for $x/t \leq l(\delta)$ and for δ_2 only in Subcase 1a) we have a contribution for $x/t \geq r(\delta)$.

This finishes the proof.

Note that the shock structure of H^δ is similar to y and y^δ , respectively, except for at most two additional intervals. This allows an effective computation of

H^δ . In Figure (3) we give an example for the shock structure of H^δ in the $x - t$ -plane. The value of the state ξ_1 depends whether the conditions (19) for $\delta_1 < 0$ and (20) for $\delta_1 > 0$ are satisfied or not. Similarly the state ξ_2 depends on the conditions (21) and (22) for $\delta_2 > 0$ and $\delta_2 < 0$, respectively.

By the above theorem we obtain that H^δ can be seen as surrogate for the first derivative of y^δ and hence can be used in optimization methods to compute the gradient.

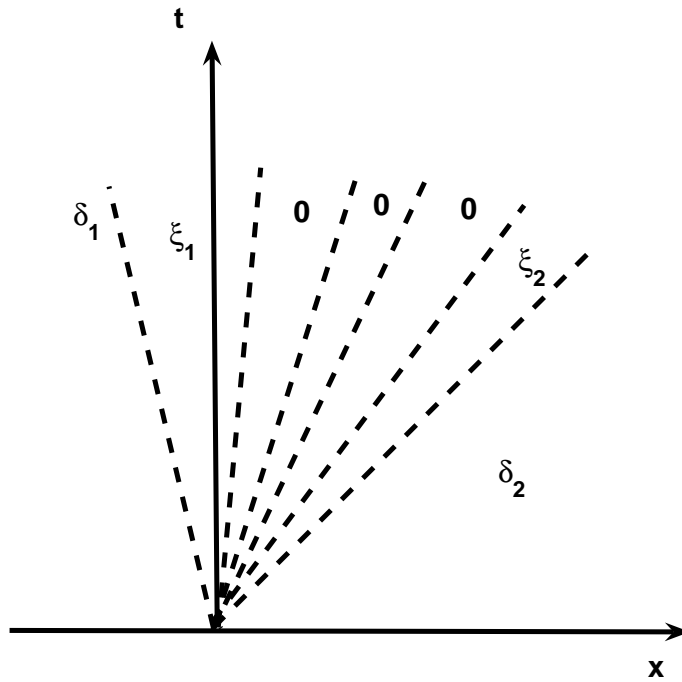


Figure 3: Structure of H^δ in $x - t$ -plane. The states $\xi_1 = 0$ if (19) is satisfied and the value of the state ξ_2 depends on the condition (21).

7 Proof of Theorem 3.1

In this section, we present the proof of Theorem 3.1.

Proof. If $D_l - f'_-(u_l) \leq 0$, since $D_l = (f_u^0)'_+(u_l)$, condition (32) holds. Hence we have for $\delta = (\delta_1, 0)$ with $\delta_1 < 0$

$$h(\delta) - h(0)$$

$$\begin{aligned}
&= \int_{-\infty}^{D_l T} g'(u_l) \delta_1 \chi_{[x_1, x_2]}(x) dx \\
&+ \int_{D_l T}^{\frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1} T} [g(u_l + \delta_1) - g(y_1)] \chi_{[x_1, x_2]}(x) dx + O(\delta_1^2) \\
&= \int_{-\infty}^{D_l T} g'(u_l) \delta_1 \chi_{[x_1, x_2]}(x) dx \\
&+ \int_{D_l T}^{\frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1} T} [g(u_l) + g'(u_l) \delta_1 - g(y_1)] \chi_{[x_1, x_2]}(x) dx + O(\delta_1^2) \\
&= g'(u_l) \delta_1 \mu([x_1, x_2] \cap (-\infty, T D_l]) \\
&+ [g(u_l) - g(y_1) + g'(u_l) \delta_1] \mu([x_1, x_2] \cap [D_l T, \frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1} T]) + O(\delta_1^2).
\end{aligned}$$

On account of (29) we have

$$\begin{aligned}
&\mu([x_1, x_2] \cap [D_l T, \frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1} T]) \\
&= \mu([x_1, x_2] \cap [D_l T, D_l T + \delta_1 \frac{D_l - f'_-(u_l)}{y_1 - u_l} T + O(\delta_1^2)]) \\
&= \delta_1 \frac{D_l - f'_-(u_l)}{y_1 - u_l} T \chi_{[x_1, x_2]}(D_l T) + O(\delta_1^2)
\end{aligned}$$

and this implies (11).

If (32) is violated, we have for $\delta = (\delta_1, 0)$ with $\delta_1 < 0$

$$\begin{aligned}
&h(\delta) - h(0) \\
&= \int_{-\infty}^{\frac{f(u_l) - f(u_l + \delta_1)}{-\delta_1} T} g'(u_l) \delta_1 \chi_{[x_1, x_2]}(x) dx \\
&= g'(u_l) \delta_1 \mu([x_1, x_2] \cap (-\infty, \frac{f(u_l) - f(u_l + \delta_1)}{-\delta_1} T]) \\
&= g'(u_l) \delta_1 \mu([x_1, x_2] \cap (-\infty, f'_-(u_l) T]) + O(\delta_1^2).
\end{aligned}$$

This implies (12).

If $\{D_l = f'_+(u_l)$ or for all $u \in (u_l, y_1) : f(u) > f_u^0(u)\}$, then (33) holds and we have for $\delta = (\delta_1, 0)$ with $\delta_1 > 0$

$$\begin{aligned}
&h(\delta) - h(0) \\
&= \int_{-\infty}^{\frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1} T} g'(u_l) \delta_1 \chi_{[x_1, x_2]}(x) dx \\
&+ \int_{\frac{f(y_1) - f(u_l + \delta_1)}{y_1 - u_l - \delta_1} T}^{D_l T} [g(y_1) - g(u_l)] \chi_{[x_1, x_2]}(x) dx + O(\delta_1^2) \\
&= g'(u_l) \delta_1 \mu([x_1, x_2] \cap (-\infty, T D_l]) \\
&+ [g(y_1) - g(u_l)] \mu([x_1, x_2] \cap [D_l T + \delta_1 \frac{D_l - f'_+(u_l)}{y_1 - u_l} T, D_l T]) + O(\delta_1^2).
\end{aligned}$$

where for the last step we have used (30). This implies (13).

If (33) is violated, then

$$\frac{f(\bar{u}) - f(u_l)}{\bar{u} - u_l} = D_l,$$

and for $\delta = (\delta_1, 0)$ with $\delta_1 > 0$ we have

$$\begin{aligned} & h(\delta) - h(0) \\ &= \int_{-\infty}^{\frac{f(\bar{u}) - f(u_l + \delta_1)}{\bar{u} - u_l - \delta_1} T} g'(u_l) \delta_1 \chi_{[x_1, x_2]}(x) dx \\ &+ \int_{\frac{f(\bar{u}) - f(u_l + \delta_1)}{\bar{u} - u_l - \delta_1} T}^{D_l T} [g(\bar{u}) - g(u_l)] \chi_{[x_1, x_2]}(x) dx + O(\delta_1^2) \\ &= g'(u_l) \delta_1 \mu([x_1, x_2] \cap (-\infty, T D_l]) \\ &+ [g(\bar{u}) - g(u_l)] \mu([x_1, x_2] \cap [D_l T + \delta_1 \frac{D_l - f'_+(u_l)}{\bar{u} - u_l} T, D_l T]) + O(\delta_1^2). \end{aligned}$$

where for the last step we have used (30). This implies (13).

If $f'_+(u_r) \leq D_r$ since $D_r = (f_u^0)'_-(u_r)$ condition (34) holds. Hence for $\delta = (0, \delta_2)$ with $\delta_2 > 0$ we have

$$\begin{aligned} & h(\delta) - h(0) \\ &= \int_{\frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k} T}^{D_r T} [g(u_r + \delta_2) - g(y_k)] \chi_{[x_1, x_2]}(x) dx \\ &+ \int_{D_r T}^{\infty} g'(u_r) \delta_2 \chi_{[x_1, x_2]}(x) dx + O(\delta_2^2) \\ &= \int_{\frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k} T}^{D_r T} [g(u_r) + g'(u_r) \delta_2 - g(y_k)] \chi_{[x_1, x_2]}(x) dx \\ &+ \int_{D_r T}^{\infty} g'(u_r) \delta_2 \chi_{[x_1, x_2]}(x) dx + O(\delta_2^2) \\ &= g'(u_r) \delta_2 \mu([x_1, x_2] \cap (T D_r, \infty]) \\ &+ [g(u_r) - g(y_k) + g'(u_r) \delta_2] \mu([x_1, x_2] \cap [\frac{f(y_k) - f(u_r + \delta_2)}{y_k - u_r - \delta_2} T, D_r T]) + O(\delta_2^2). \end{aligned}$$

On account of (30) we have

$$\begin{aligned} & \mu([x_1, x_2] \cap [\frac{f(y_k) - f(u_r + \delta_2)}{y_k - u_r - \delta_2} T, D_r T]) \\ &= \mu([x_1, x_2] \cap [D_r T + \delta_2 \frac{D_r - f'_+(u_r)}{y_k - u_r} T + O(\delta_2^2), D_r T]) \\ &= -\delta_2 \frac{D_r - f'_+(u_r)}{y_k - u_r} T \chi_{(x_1, x_2]}(D_r T) + O(\delta_2^2) \end{aligned}$$

and this implies (15).

If (34) is violated, we have for $\delta = (0, \delta_2)$ with $\delta_2 > 0$

$$\begin{aligned}
& h(\delta) - h(0) \\
&= \int_{\frac{f(u_r + \delta_2) - f(u_r)}{\delta_2} T}^{\infty} g'(u_r) \delta_2 \chi_{[x_1, x_2]}(x) dx + O(\delta_2^2) \\
&= g'(u_r) \delta_2 \mu([x_1, x_2] \cap [\frac{f(u_r + \delta_2) - f(u_r)}{\delta_2} T, \infty)) \\
&= g'(u_r) \delta_2 \mu([x_1, x_2] \cap [f'_+(u_r) T, \infty)) + O(\delta_2^2).
\end{aligned}$$

This implies (16).

If $\{D_r = f'_-(u_r)$ or for all $u \in (y_k, u_r) : f(u) > f_u^0(u)\}$, condition (35) holds and we have for $\delta = (0, \delta_2)$ with $\delta_2 < 0$

$$\begin{aligned}
& h(\delta) - h(0) \\
&= \int_{D_r T}^{\frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k} T} [g(y_k) - g(u_r)] \chi_{[x_1, x_2]}(x) dx \\
&+ \int_{\frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k} T}^{\infty} g'(u_r) \delta_2 \chi_{[x_1, x_2]}(x) dx + O(\delta_2^2) \\
&= g'(u_r) \delta_2 \mu([x_1, x_2] \cap [\frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k} T, \infty)) \\
&+ [g(y_k) - g(u_r)] \mu([x_1, x_2] \cap [D_r T, \frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k} T]) + O(\delta_2^2).
\end{aligned}$$

On account of (29) we have

$$\frac{f(u_r + \delta_2) - f(y_k)}{u_r + \delta_2 - y_k} = D_r + \delta_2 \frac{D_r - f'_-(u_r)}{y_k - u_r} + O(\delta_2^2).$$

Thus $h(\delta) - h(0)$

$$= g'(u_r) \delta_2 \mu([x_1, x_2] \cap [D_r T, \infty)) + [g(y_k) - g(u_r)] \delta_2 \chi_{[x_1, x_2]}(T D_r) (D_r - f'_-(u_r)) + O(\delta_2^2).$$

This implies (17).

If (35) is violated, we have for $\delta = (0, \delta_2)$ with $\delta_2 < 0$

$$\begin{aligned}
& h(\delta) - h(0) \\
&= \int_{D_r T}^{\frac{f(u_r + \delta_2) - f(\bar{u})}{u_r + \delta_2 - \bar{u}} T} [g(\bar{u}) - g(u_r)] \chi_{[x_1, x_2]}(x) dx \\
&+ \int_{\frac{f(u_r + \delta_2) - f(\bar{u})}{u_r + \delta_2 - \bar{u}} T}^{\infty} g'(u_r) \delta_2 \chi_{[x_1, x_2]}(x) dx + O(\delta_2^2) \\
&= g'(u_r) \delta_2 \mu([x_1, x_2] \cap [\frac{f(u_r + \delta_2) - f(\bar{u})}{u_r + \delta_2 - \bar{u}} T, \infty)) \\
&+ [g(y_k) - g(u_r)] \mu([x_1, x_2] \cap [D_r T, \frac{f(u_r + \delta_2) - f(\bar{u})}{u_r + \delta_2 - \bar{u}} T]) + O(\delta_2^2).
\end{aligned}$$

On account of (29) and since $D_r = [f(\bar{u}) - f(u_r)]/[\bar{u} - u_r]$ and we have

$$\frac{f(u_r + \delta_2) - f(\bar{u})}{u_r + \delta_2 - \bar{u}} = D_r + \delta_2 \frac{D_r - f'_-(u_r)}{y_k - u_r} + O(\delta_2^2).$$

Thus $h(\delta) - h(0)$

$$= g'(u_r)\delta_2\mu([x_1, x_2] \cap [D_r T, \infty)) + [g(\bar{u}) - g(u_r)]\delta_2\chi_{[x_1, x_2]}(T D_r)(D_r - f'_-(u_r)) + O(\delta_2^2).$$

This implies (18).

8 Example: Application to traffic flow problems

The example is academic and illustrates the possible application of the derived calculus only. The main advantage of the calculus is the application as a numerical method and as already mentioned we will present numerical results in a forthcoming publication.

We consider an (simplified) example related to traffic flow. A typical flux-function $f(\rho)$ in traffic flow applications is concave with a single maximum [10]. We normalize the maximal density to $\rho_{\max} = 1$ and use in the following a piece-wise linear approximation on $[0, 1]$ of

$$f(\rho) = \rho(1 - \rho).$$

It is well-known that for Riemann initial data such that $u_l \leq \operatorname{argmax} f(\cdot) = \frac{1}{2}$ and $u_r \geq 1 - u_l$, the solution to (2,3) is a shock wave travelling with non-positive velocity. We show that this information is also obtained by considering an optimal control problem and by discussing the necessary and sufficient optimality conditions derived by Theorem (3.1).

We fix $u_l < \frac{1}{2}$ and denote by y the solution to (2),(3) and (1). We consider the objective function given by

$$I(u_0) = \frac{1}{2} \int_{-1}^0 (y(x, T) - a)^2 dx$$

for T suitable large and a constant $a > 0$. We can control u_r , i.e., a part of the initial data for u_0 for $x \geq 0$. To be in the setting of the Theorem (3.1) we assume $u_r \geq u_l$.

A necessary condition for u_r to be a minimizer, is, that the left-hand sided derivative $\partial_{\delta_2^-} h(0) \leq 0$ and the right-hand sided derivative $\partial_{\delta_2^+} h(0) \geq 0$.

Then, $\partial_{\delta_2^+} h(0) \geq 0$ implies $u_r \geq 1 - u_l$. Indeed, we distinguish the cases *A* and *B*.

$$(A) \quad f'_+(u_r) = 1 - 2u_r \leq D_r = 1 - (u_r + y_k).$$

This implies $u_r \geq y_k$ and due to the concavity of f we obtain $y_k \equiv u_l$. For T sufficiently large we obtain by (15) and $\partial_{\delta_2^+} h(0) \geq 0$, $D_r \leq 0$ and $u_r \geq a$.

$$(B) f'_+(u_r) > D_r.$$

Then $u_r < y_k$ and by construction $u_r < y_k \cdots \leq u_l < \frac{1}{2}$, which contradicts the assumption $u_r \geq u_l$.

By $0 \geq \partial_{\delta_2^-} h(0)$ we obtain similarly $0 \geq u_r - a$ by (17) and $D_r < 0$, which finally yields the expected, necessary optimality conditions for $u_r : u_r = a$ and $u_r > 1 - u_l$. This also shows, that the above is controllable, if and only if $a > 1 - u_l$ and T sufficiently large.

9 Summary

We derived a sensitivity calculus for the optimal control problem governed by nonlinear conservation laws with initial Riemann data. The calculus is valid also for non-convex, piecewise linear flux functions and general objective functionals. The generalization to piecewise constant control functions is straightforward and will be considered in a forthcoming paper containing also numerical results. The introduced calculus is easy to implement as numerical method for gradient computation due to the explicit formulas. Further it can be incorporated in existing Front-Tracking codes due to the similar structure of solutions.

Acknowledgments

This work has been partly supported by the Kaiserslautern Excellence Cluster Dependable Adaptive Systems and Mathematical Modelling.

References

- [1] A. Aw and M. Rascle. Resurrection of second order models of traffic flow? *SIAM J. Appl. Math.*, 60:916, 2000.
- [2] A. Bressan. *Hyperbolic Systems of Conservation Laws*. Oxford University Press, Oxford, 2000.
- [3] C. M. Dafermos. Polygonal approximations of solutions of the initial value problem for a conservation law. *J. Math. Anal. Appl.*, 38:33, 1972.
- [4] K. Ehrhardt and M. Steinbach. Nonlinear optimization in gas networks. *Modeling, Simulation and Optimization of Complex Processes (eds. H. G. Bock, E. Kostina, H.X. Phu, R. Ranacher)*, page 139, 2005.
- [5] M. Gugat. Nodal control of networked hyperbolic systems. *AMO Advanced Modeling and Optimization*, 7:1, 2005.

- [6] M. Gugat. Optimal nodal control of networked hyperbolic systems: Evaluation of derivatives. *AMO Advanced Modeling and Optimization*, 7:1, 2005.
- [7] M. Gugat, M. Herty, A. Klar, and G. Leugering. Optimal control for traffic flow networks. *Journal of Optimization Theory and Applications*, 126:589, 2005.
- [8] M. Gugat, G. Leugering, K. Schittkowski, and E. J. P. G. Schmidt. Modelling, stabilization and control of flow in networks of open channels. In Rambau Grötschel, Krumke, editor, *Online optimization of large scale systems*, page 250. Springer, 2001.
- [9] M. Gugat, G. Leugering, and E. J. P. G. Schmidt. Global controllability between steady supercritical flows in channel networks. *Mathematical Methods in the Applied Sciences*, 2003.
- [10] D. Helbing. *Verkehrsdynamik*. Springer-Verlag, Berlin, Heidelberg, New York, 1997.
- [11] H. Holden and L. Holden. On scalar conservation laws in one-dimension. In Albeverio, editor, *Ideas and Methods in Mathematical Analysis, Stochastics and Applications*, page 480, 1992.
- [12] H. Holden, L. Holden, and R. Hoegh-Krohn. A numerical method for first order nonlinear scalar conservation laws in one-dimension. *Comput. Math. Anal.*, 15:595, 1988.
- [13] H. Holden and N. H. Risebro. A mathematical model of traffic flow on a network of unidirectional roads. *SIAM J. Math. Anal.*, 26:999, 1995.
- [14] H. Holden and N. H. Risebro. *Front tracking for hyperbolic conservation laws*. Springer, New York, Berlin, Heidelberg, 2002.
- [15] S. N. Kruzkov. First order quasi linear equations in several independent variables. *Math. USSR Sbornik*, 10:217, 1970.
- [16] R. J. LeVeque. *Numerical methods for conservation laws*. Birkhäuser Verlag, Basel, Boston, Berlin, 1990.
- [17] M.J. Lighthill and J.B. Whitham. On kinematic waves. *Proc. Royal Soc. Edinburgh*, A229:281, 1955.
- [18] J. Smoller. *Shock waves and reaction diffusion equations*. Springer, New York Heidelberg Berlin, 1994.
- [19] S. Ulbrich. A sensitivity and adjoint calculus for discontinuous solutions of hyperbolic conservation laws with source terms. *SIAM J. Control Optim.*, 41:740, 2002.
- [20] S. Ulbrich. Adjoint-based derivative computations for the optimal control of discontinuous solutions of hyperbolic conservation laws. *Systems & Control Letters*, 3:309, 2003.