Transient L^1 error estimates for well-balanced schemes on non-resonant scalar balance laws

Debora Amadori, Laurent Gosse

Transient L^1 error estimates for well-balanced schemes on non-resonant scalar balance laws

Debora Amadori

DISIM, Università degli Studi dell'Aquila, L'Aquila, Italy

Laurent Gosse *

IAC–CNR "Mauro Picone" (sezione di Roma)
Via dei Taurini, 19 - 00185 Rome, Italy

Abstract

The ability of Well-Balanced (WB) schemes to capture very accurately steady-state regimes of non-resonant hyperbolic systems of balance laws has been thoroughly illustrated since its introduction by Greenberg and LeRoux [15] (see also the anterior WB Glimm scheme in [8]). This paper aims at showing, by means of rigorous $C^0_t(L^1_x)$ estimates, that these schemes deliver an increased accuracy in transient regimes too. Namely, after explaining that for the vast majority of non-resonant scalar balance laws, the $C^0_t(L^1_x)$ error of conventional fractional-step [45] numerical approximations grows **exponentially** in time like $\exp(\max(g')t)\sqrt{\Delta x}$ (as a consequence of the use of Gronwall's lemma), it is shown that WB schemes involving an exact Riemann solver suffer from a much smaller error amplification: thanks to strict hyperbolicity, their error grows at most only **linearly** in time (see also [30]). Numerical results on several test-cases of increasing difficulty (including the classical LeVeque-Yee's benchmark problem [34] in the non-stiff case) confirm the analysis.

Key words: Well-balanced Godunov scheme; Wave-front tracking algorithm; functional equivalent to the L^1 distance; numerical viscosity, non-linear resonance, Gronwall lemma, error estimates.
1991 MSC: 65M06, 35L60

* Corresponding Author
Email addresses: amadori@univaq.it (Debora Amadori), l.gosse@ba.iac.cnr.it (Laurent Gosse).

1 Introduction and main objectives

The main goal of the present text is to emphasize the qualitative difference between Time-Splitting (TS, also called Fractional Step, FS) and Well-Balanced (WB) numerical schemes when it comes to computing the entropy solution [27] of a simple scalar, yet non-resonant, balance law:

$$\partial_t u + \partial_x f(u) = k(x)g(u), \qquad k \in L^1_{loc}(\mathbb{R})\,. \tag{1}$$

We assume that the flux f satisfies the following **non-resonance assumption**, absolutely fundamental for deriving rigorous estimates:

$$\pm f'(u) \geq \nu > 0\,. \tag{2}$$

At this level, no peculiar assumption is made on the source term g except smoothness, $g \in C^1(\mathbb{R})$; especially, g' has no definite sign. The Cauchy problem consists in studying the equation (1)–(2) supplemented by a (possibly discontinuous) initial data,

$$u(t=0,x) = u_0(x) \in L^1 \cap BV(\mathbb{R}), \qquad x \in \mathbb{R}, \tag{3}$$

with $BV(\mathbb{R}) \subset L^\infty(\mathbb{R})$ standing for the space of functions with bounded variation. Moreover assume that the source term is accretive in the following sense:

$$N = \sup\{a'(x)g'(\xi)\ \ x \in \mathbb{R}\,,\ |\xi| \leq \max\{\|u\|_\infty, \|u^{\Delta t}\|_\infty\} > 0\,.$$

The **primary goal** of the paper is to prove that, for $\partial_x a(x) = k(x)$ and w a specific Riemann invariant, a local L^1 error holds for $u^{\Delta t}(t,\cdot)$, a WB approximation of (1): for all $t > 0$,

$$\begin{aligned}
&\int_{x_1}^{x_2} |u^{\Delta t}(t,x) - u(t,x)|dx \leq \\
&\leq\ C \min\left\{ \mathrm{e}^{\kappa_2 \mathrm{TV}\,\{a\}} \left[\Delta x(\mathrm{TV}\,\{u_0\}+1) + \mathrm{TV}\,\{w_0\}\, L\, t\right],\ \sqrt{\Delta x}\,\sqrt{A} + \Delta x\, B \right\},
\end{aligned}$$

where

$$\begin{aligned}
A &= [\mathrm{TV}\,\{w_0\} + \|k\|_{L^1}] \left[\frac{\mathrm{e}^{Nt}-1}{N}\right] \left[\mathrm{e}^{Nt}\,(L+1)\mathrm{TV}\,\{u_0\} + \|k'g\|_\infty \frac{\mathrm{e}^{Nt}-1}{N}\right], \\
B &= \mathrm{e}^{Nt}\,\mathrm{TV}\,\{u_0\} + \|k\|_\infty \frac{\mathrm{e}^{Nt}-1}{N}\ (\mathrm{TV}\,\{w_0\} + \|k\|_{L^1}),
\end{aligned}$$

see (68).

The estimate displays a linear growth in time beyond a certain time. The linear growth in the time variable is classical for homogeneous hyperbolic problems, see for instance [30] and references therein. Here we show that well-balanced Godunov methods allow to extend this property to inhomogeneous problems too.

1.1 A puzzling numerical example

A great deal of effort has been drawn onto deriving error estimates for Cauchy problems of the type (1) during the nineties: see for instance [45,29,42]. The methodology is to adapt the computations appearing in the seminal paper by Kuznetsov [28] to the widely used operator-splitting schemes. This results in the classical "one-half" order of convergence in L^1, which is known to be optimal for Godunov type schemes [43,44]: denoting u, $u^{\Delta t}$ the entropy solution and its numerical approximation, respectively, Tang and Teng state that:

$$\forall t \in [0,T], \qquad \|u^{\Delta t}(t,.) - u(t,.)\|_{L^1(\mathbb{R})} \leq C\sqrt{\Delta t}, \tag{4}$$

where C is a constant independent of Δt (see Theorems 1.1 and 1.2 in [45]). We claim that such a statement, similar to the one in [42], can be misleading [1] because the "constant C" actually depends on the time t (as is suggested in Theorem 2.1 of [29]). A careful inspection of the proofs completed by their own authors reveals that the "constant C" is actually an exponential in time which stems from the application of the Gronwall lemma: see [45] p.116, [29] p.854, [42] p.103 and more recently [38]. A more rigorous statement is:

$$\forall t \in [0,T], \qquad \|u^{\Delta t}(t,.) - u(t,.)\|_{L^1(\mathbb{R})} \leq C \exp(\max[g'(u)]t)\sqrt{\Delta t},$$

where C hinges on the initial data and the "max" is taken on the convex hull of all the values taken by both $u, u^{\Delta t}$. This meets with the a-priori estimate given in [10].

This estimate is disastrous from a computational standpoint because, in order to keep the absolute error below a given tolerance, the computational grid's parameters are meant to decrease exponentially with time (except if $g' \leq 0$, for which $\mathrm{TV}\,(u)(t,.)$ decays exponentially too). One may wonder whether such an exponential amplification of the absolute error can happen in practice, or if it is only a technical discrepancy of the analysis. By considering a particular case of (1), namely,

$$f(u) = u^2/2,\ g(u) = u, \qquad k(x) \equiv 0.2, \qquad u_0(x) = Y(x), \tag{5}$$

Y being the Heaviside function:

$$Y(x) = 0 \quad \text{as } x < 0\,, \qquad Y(x) = 1 \quad \text{as } x > 0\,,$$

one obtains a balance law of which the (smooth) entropy solution is explicitly calculable by the method of characteristics. It is therefore possible to compute accurately the L^1 absolute error of both the TS (involving an exact ODE solver) and WB versions of the Godunov scheme: see Fig. 1. On the left side, the black curve is the L^1 error of the TS scheme, the blue one being the one of the WB one for the numerical solutions appearing on the right side. Moreover, an exponential fitting has been superimposed (in red): the agreement is

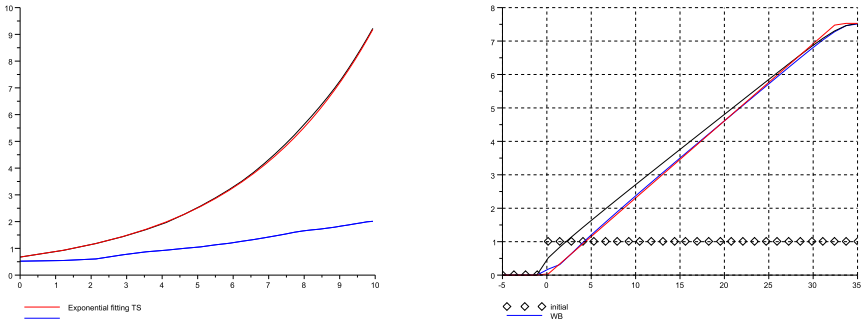

Fig. 1. Time evolution of L^1 error for a rarefaction wave solution of Burgers equation.

very good. For this experiment, 2^7 points in the x variable have been set, the time-step Δt is chosen adaptively in order to maintain a constant CFL of 0.95. The mechanism leading to the exponential amplification is easily discovered by examining the graphic on the right of Fig. 1: for the values very close to zero, the time-step is unreasonably small, and leads to an excessive dissipation through the numerical viscosity [32,46] of the TS scheme. After some time, these artificial values begin to interact with the accretive source term which makes them grow exponentially. The black curve, corresponding to the TS scheme, grows

[1] Observe that it would already be problematic for an homogeneous scalar conservation law in which $k \equiv 0$ because its L^1 error is known to increase in time like $\mathcal{O}(\sqrt{t})$, as explained in *e.g.* [36].

over of the exact solution's red curve in the region $0 \leq x \leq 20$ whereas the WB scheme's blue curve remains very close to it. We stress that such a problem is not meant to stabilize in large times, hence **the gain of the WB scheme has little to do with a steady-state balance between convection flux and source term**. It is the excessive numerical viscosity which is at the origin of the TS scheme's anomalous behavior. In the sequel of the paper, the estimate in [10] will be proved to be quite pessimistic.

Remark 1 *As the crucial issue appears to be the "network viscosity", let us quote words from Oran and Book [41] (p.162) which agree with the test-case of Fig. 1: "The most persistent problem arising in Eulerian representations is numerical diffusion, which moves a small amount of material across cells faster than any physical process. Numerical diffusion may appear as premature mixing throughout a computational cell, when, in fact, the mixing should have just begun at one interface or corner of the cell". Splitting in time between convection treated by means of an Eulerian representation and reaction involving an accretive source term for which g' can be strictly positive is dangerous because of the interplay between the numerical viscosity of the first process being amplified by the second one.*

1.2 Temple class reformulation and uniform BV bounds

The derivation of WB schemes originates with a reformulation of the inhomogeneous balance law (1) under the form of an artificial 2×2 Temple class system [2,11,16,21] by introducing an antiderivative $a(x)$ (defined up to an arbitrary constant),

$$\partial_t u + \partial_x f(u) - g(u)\partial_x a = 0, \qquad \partial_t a = 0, \qquad \partial_x a = k. \tag{6}$$

The non-resonance assumption (2) is equivalent to its strict hyperbolicity and the genuine non-linearity of one of the characteristic fields of (6). The non-conservative product $g(u)\partial_x a$ (see [31]) induces a stationary, trivially linearly degenerate, field which renders locally the effects of the source term. The net gain in considering (6) in a strictly hyperbolic context, is that all the techniques designed for **homogeneous** problems become available because the localized source term is integrated directly inside the self-similar Riemann solver, resulting into a new Rankine-Hugoniot relation.

Two manners of removing the ambiguity linked to the non-conservative product co-exist: the first one consists in considering a sequence of smooth functions a^ϵ converging strongly to a in $L^1_{loc}(\mathbb{R})$. By deriving convenient BV-bounds (see Lemma 7 in [11]), it is possible to study the weak limit of the corresponding sequence $g(u^\epsilon)\partial_x a^\epsilon$, following the general theory of [31]. The second one limits itself to seek a Riemann invariant associated to the linearly degenerate field and its zero eigenvalue (thus justifying the terminology *zero-wave* or *standing wave* used in [1,2,16,21]); here again, the answer is provided in §3.3 of [11], where it is shown that

$$a, \qquad w(u,a) = \phi^{-1}\big(\phi(u) - a\big), \quad \phi' = \frac{f'}{g}, \tag{7}$$

are Riemann invariants of the genuinely non-linear and the linearly degenerate fields, respectively. The expression given in (7) is valid at least when g does not change its sign, so that $\phi' \neq 0$ and then ϕ can be inverted; notice that if $g(u)$ has an isolated zero at some u_o, it is easily checked that w can be extended by continuity through the point u_o by setting $w(u_o, a) = u_o$, and is Lipschitz continuous w.r.t. u, uniformly for a in a bounded set.

The uniform BV-bound on u (solution of both (6) and (1) when $\partial_x a = k \in L^1(\mathbb{R})$ is smooth enough [16]) is therefore an immediate consequence of the absence of any quadratic interaction potential for Temple class systems. The stability and strong compactness of the Godunov scheme for (6) can be quickly established by invoking the results of [33] (as explained in Remark 2 of [14]). Here we show that, going down the general path established by Bressan in [4], the "quasi-decay" of the Lyapunov functional Λ equivalent to the L^1

distance built in [1] yields that,

$$\forall t > 0, \qquad \frac{d\Lambda(t; U^{\Delta t}, U)}{dt} \leq \mathcal{O}(\delta), \qquad U = (a, u)^T,\ U^{\Delta t} = (a^{\Delta x}, u^{\Delta t})^T,$$

thus forbidding any exponential growth in time of the L^1 distance separating the exact solution U from its numerical approximation $U^{\Delta x}$, at least when it is computed by the wave-front tracking algorithm [20]. Bressan's theory henceforth strongly suggests that the introduction of (6) in a strictly hyperbolic context allows for a qualitative improvement of the estimates presented in [29]. It will be shown in the sequel that a similar improvement holds true for the WB Godunov scheme too.

1.3 Outline of the paper

A rather original method for deriving new a-priori estimates for scalar balance laws is proposed: specifically, these are able to perceive the WB features of certain discretizations in order to weaken the time-amplification of L^1 truncation bounds. We advocate the idea that an error estimate contains two types of information: the widely recognized dependence on the computational grid's parameter (here Δx or Δt, both related by the classical CFL condition), and the time variable which reveals the characteristic temporal scale inside which the estimate can have a practical significance. The elementary case of a linear advection equation endowed with a space-dependent source term is carried out in §2: in particular, two complementary methods are presented for deriving an error estimate. It is found that conventional "centered source" algorithms lead to bigger errors (even on smooth solutions) displaying a non-linear growth in the time variable. In §3, the scalar law (1) is recast in the framework of the "scalar system" proposed in [11] for which a wavefront tracking approximation is set up, according to [16]: the decay of the corresponding Lyapunov functional Λ is recalled. An hybrid "wavefront tracking/Godunov" scheme is studied in §4, allowing to take advantage of the properties of Λ along two any approximate solutions. As a consequence of Godunov procedure, the jump of the functional at each averaging step must be computed: this point is presented in full detail in §4.2. After having performed the complementary Kuznetsov computation in §4.4, the full estimate, which is the optimum between the two aforementioned ones, is derived in (68). It displays the usual $\sqrt{t}$ growth around $t \simeq 0$, and beyond a certain time, the Lyapunov functional restricts the time-amplification of the error to be linear at most. In §5, more numerical results are displayed, especially on the classical LeVeque-Yee benchmark [34]. After some conclusive remarks, an Appendix A presents a formal computation which explains the main reasons lying behind the functional's decay properties.

A few complementary remarks go as follows:

– the recent book [19] displays the exponential growth in Thm. 5.15 (page 123),

– there are tentatives aiming at reducing artificial viscosity in [18,12],

– different conclusions appear in [9] because it focuses on relaxation processes.

2 Propagation of truncation errors for linearized shallow water equations

Here the goal is to work out an elementary example which, despite its simplicity, contains already a part of the specific features that govern the more complex non-linear cases. By linearizing around a state $\bar{\rho} > 0, \bar{u} = 0$ the usual shallow water equations with topography, one arrives at the following linear system:

$$\partial_t \rho + \partial_x J = 0\,, \qquad \partial_t J + \partial_x \rho = -\partial_x a \tag{8}$$

where the function $a = a(x)$ represents the bottom topography. By linearity, the solutions $\rho(t,\cdot), J(t,\cdot)$ are endowed with the same integrability and smoothness as their initial data. An alternative, customary way to deal with system (8) is to consider $a = a(x)$ as an independent variable and therefore to write:

$$\begin{cases} \partial_t \rho + \partial_x J = 0, \\ \partial_t J + \partial_x(\rho + a) = 0 \\ \partial_t a = 0\,. \end{cases} \tag{9}$$

2.1 Local Error Truncation (LTE) and linear dependence in time

The characteristic speeds of system (8) are ± 1. The diagonal variables $f^{\pm} = \rho \pm J$ satisfy

$$(\partial_t - \partial_x) f^- = \partial_x a\,, \qquad (\partial_t + \partial_x) f^+ = -\partial_x a\,; \tag{10}$$

equivalently, the system (9) is diagonalized as follows:

$$\begin{cases} \partial_t (f^{\pm} + a) \pm \partial_x (f^{\pm} + a) = 0 \\ \partial_t a = 0\,. \end{cases} \tag{11}$$

Correspondingly there exist two approaches for approximating the solutions of (8):

(1) the standard "centered source method" which consists in processing a set of two advection equations with a source term, (10);
(2) the "well-balanced method", which treats two *homogeneous* advection equations from (11): $\partial_t f^{\pm} \pm \partial_x (f^{\pm} + a) = 0$.

Let's assume that we are given a uniform Cartesian computational grid determined by the parameters Δx, $\Delta t = \nu \Delta x$ with $0 < \nu \leq 1$ together with $a \in C^{\infty}(\mathbb{R})$ having compact support, and $f^{\pm}(t = 0, \cdot) \in C^3 \cap W^{3,\infty}(\mathbb{R})$.

Analysis of (1). In each computational cell C_j^n, the residual R_j^n of the "centered source method" is computed by plugging exact solutions into each inhomogeneous advection equation: for instance,

$$\frac{f^+(t^{n+1}, x_j) - f^+(t^n, x_j)}{\Delta t} + \frac{f^+(t^n, x_j) - f^+(t^n, x_{j-1})}{\Delta x} + k(x_j) = R_j^n.$$

By Taylor expansion, one finds that

$$|R_j^n| \leq \frac{\Delta x}{2} (1 - \nu) \, \|\partial_{xx} f^{\pm}(t^n, \cdot)\|_{\infty} + \frac{\Delta t}{2} \|\partial_{xx} a(x)\|_{\infty} + C(\Delta x)^2$$

where the C above depends on $\|\partial_{xxx} f^{\pm}(t^n, \cdot)\|_{\infty}$, $\|\partial_{ttt} f^{\pm}(t^n, \cdot)\|_{\infty}$.

By linearity, the quantity $\partial_{xx} f^{\pm}(t, \cdot)$ is estimated as follows:

$$\|\partial_{xx} f^{\pm}(t, \cdot)\|_{\infty} \leq \|\partial_{xx} f^{\pm}(t = 0)\|_{\infty} + t \|a\|_{C^3}.$$

Hence there are 2 error amplification mechanisms for the "centered source method": the fact that R_j^n contains $\Delta t \|\partial_{xx} a\|_{\infty}$ which doesn't vanish as $\nu = \frac{\Delta t}{\Delta x} = 1$, and the linear growth of $t \mapsto \|\partial_{xx} f^{\pm}(t, \cdot)\|_{\infty}$.

Now, by linearity, the pointwise error,

$$(E^{\pm})_j^n = (f^{\pm} - f^{\pm}_{\Delta x})(t^n, x_j),$$

(where $f^{\pm}_{\Delta x}$ stands for the piecewise-constant approximation of the exact solution $f^{\pm}$) satisfies a slightly modified upwind scheme that easily rewrites under the form of a convex combination plus a source term: for instance,

$$(E^+)^{n+1}_j = (E^+)^n_j \left(1 - \frac{\Delta t}{\Delta x}\right) + \frac{\Delta t}{\Delta x}(E^+)^n_{j-1} + \Delta t R^n_j.$$

It remains to take the modulus, the "sup", and to sum in n all the residuals in order to derive that, for any $t \geq 0$, the following error estimate holds:

$$\begin{aligned} \|E^{\pm}(t,\cdot)\|_{L^\infty} \leq\ & \|E^{\pm}(t=0,\cdot)\|_{L^\infty} \\ & + \Delta x \cdot t\Big[\nu\|a\|_{C^2} + (1-\nu)\left(\|f^{\pm}(t=0,\cdot)\|_{C^2} + t\|a\|_{C^3}\right)\Big] \\ & + O(1)t(\Delta x)^2 \end{aligned} \tag{12}$$

where the $O(1)$ above depends on $\|f^{\pm}(t=0,\cdot)\|_{C^3}$, $\|a\|_{C^4}$ and linearly on t.

Analysis of (2). Oppositely, for the well-balanced method, the same LTE analysis starts with:

$$\frac{f^+(t^{n+1},x_j) - f^+(t^n,x_j)}{\Delta t} + \frac{(f^+ + a)(t^n,x_j) - (f^+ + a)(t^n,x_{j-1})}{\Delta x} = \tilde{R}^n_j.$$

The full upwinding of the topography function a yields a smaller residual:

$$|\tilde{R}^n_j| \leq \frac{\Delta x}{2}(1-\nu)\left\|\partial_{xx}[f^{\pm}(t,\cdot) + a(\cdot)]\right\|_\infty + C(\Delta x)^2,$$

with C depending on the sup norm of $\partial_{xxx}(f^{\pm} + a)$.

Since the diagonal variables $V = (f^{\pm} + a, a)$ are constant along their characteristics, the sup-norm of their derivatives doesn't grow in time. Hence

$$|\tilde{R}^n_j| \leq \frac{\Delta x}{2}(1-\nu)\left\|\partial_{xx}[f^{\pm}(t=0,\cdot) + a(\cdot)]\right\|_\infty + C_0(\Delta x)^2$$

for some constant C_0. Again, by observing that the scheme governing the pointwise error

$$(\tilde{E}^{\pm})^n_j = [(f^{\pm} + a) - (f^{\pm}_{\Delta x} + a)](t^n, x_j)$$

rewrites as a convex combination, one can derive a much better global error estimate:

$$\|\tilde{E}^{\pm}(t,\cdot)\|_{L^\infty} \leq \|\tilde{E}^{\pm}(t=0,\cdot)\|_{L^\infty} + \Delta x \cdot \frac{t}{2}(1-\nu)\|f^{\pm}(t=0,\cdot) + a\|_{C^2} + C_0 t(\Delta x)^2. \tag{13}$$

Let's enumerate the main differences between (12) and (13):

(1) the error (12) of the standard upwind method is a non-linear (quadratic in the present case) function of time whereas the well-balanced estimate (13) is linear,
(2) the source term in (12) always contributes at the same rate regardless to the value of the Courant number ν. In sharp contrast, selecting $\nu = 1$ implies that for the well-balanced method, the initial error (13) stands still, $\|\tilde{E}^{\pm}(t,\cdot)\|_{L^\infty} \simeq \|\tilde{E}^{\pm}(t=0,\cdot)\|_{L^\infty}$.

On Fig. 2, both these theoretical aspects are illustrated: the left graphic shows that the well-balanced pointwise error $\tilde{E}$ grows more strongly when $\nu \to 0$ (the numerical viscosity increases) and the right one reveals that $\tilde{E}$ remains constant in time when $\nu = 1$ but this nice property doesn't hold for a conventional method.

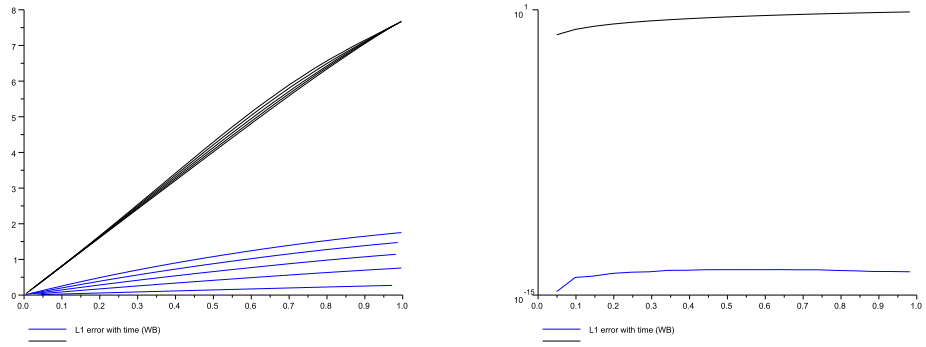

Fig. 2. Time-evolution of the L^∞ errors of well-balanced (blue) and centered source (black) methods for smooth $f^\pm, a$: $\nu = 0.9, 0.7, 0.5, 0.3, 0.1$ (left) and $\nu = 1$ (right)

2.2 Another manner for computing an error estimate

Hereafter we concentrate on the well-balanced method, which on the simple example (11) consists in solving exactly the 2 advection equations, and projecting onto the space of piecewise-constant functions at each time-step. By linearity, the difference between 2 solutions satisfy the same equations (there's no dissipation because no shocks form), thus by calling $W = V - \tilde{V}$ the difference between any 2 solutions, one gets:

$$\|W(t,\cdot)\|_{L^1(\mathbb{R})} = \int_{\mathbb{R}} |a - \tilde{a}| + \left|(f^\pm + a) - (\tilde{f}^\pm + \tilde{a})\right| dx = \|W(t=0,\cdot)\|_{L^1(\mathbb{R})}.$$

Recalling the issue of propagation of truncation errors in a well-balanced Godunov scheme, this equality proves that errors cannot grow during the "exact step". Concerning the projection step, one already knows that the term $\int_{\mathbb{R}} |a - \tilde{a}| dx$ is invariant. Denoting again by $f^\pm_{\Delta x}$ the approximation produced by the Godunov scheme and by $P^{\Delta x}$ the projector on piecewise constant functions, the jump of the L^1 norm reads:

$$\begin{aligned}\int_{\mathbb{R}} \left|(f^\pm + a) - P^{\Delta x}(f^\pm_{\Delta x} + a_{\Delta x})\right| - \left|(f^\pm + a) - (f^\pm_{\Delta x} + a_{\Delta x})\right| dx \\ \leq \int_{\mathbb{R}} \left|P^{\Delta x}(f^\pm_{\Delta x}) - f^\pm_{\Delta x}\right| dx\,,\end{aligned}$$

where we used that $a_{\Delta x} = P^{\Delta x} a$. In this simple case, there are only discontinuities $f^\pm_L, f^\pm_R$ propagating at velocity ± 1 inside every computational cell: let $f^\pm_L$ be in $0 < x < \Delta t$ and $f^\pm_R$ in $\Delta t < x < \Delta x$:

$$\left(\int_0^{\Delta t} + \int_{\Delta t}^{\Delta x}\right) \left|P^{\Delta x}(f^\pm_{\Delta x}) - f^\pm_{\Delta x}\right| dx = 2\Delta x \cdot \nu\,(1-\nu)\,|f^\pm_R - f^\pm_L|\,. \tag{14}$$

Assume that the initial data $f^\pm(t=0,\cdot) \in L^1 \cap BV(\mathbb{R})$ and $a \in BV(\mathbb{R})$. From the equations (11) we notice that $(f^\pm + a)$ is constant across interfaces and its total variation remains constant in time. Recalling that $\Delta t = \nu \Delta x$, we sum up (14) over all cells and deduce that

$$\begin{aligned}\int_{\mathbb{R}} \left|P^{\Delta x}(f^\pm_{\Delta x}) - f^\pm_{\Delta x}\right| dx &\leq 2\Delta t\,(1-\nu)\,TV((f^\pm + a)(t,\cdot)) \\ &= 2\Delta t\,(1-\nu)\,TV((f^\pm + a)(0,\cdot))\,.\end{aligned}$$

By summing on all the time indexes n, another (seemingly less interesting) error estimate in L^1 emerges, which displays again a linear growth in time:

$$\begin{aligned}
&\int_{\mathbb{R}} |f^{\pm}(t,x) - f^{\pm}_{\Delta x}(t,x)|dx \\
&\qquad \leq \int_{\mathbb{R}} |(f^{\pm}+a)(t,x) - (f^{\pm}_{\Delta x} - a_{\Delta x})(t,x)|dx \;+\; \|a - a_{\Delta x}\|_{L^1} \\
&\qquad \leq 2t\;(1-\nu)\,TV(f^{\pm}+a)(0,\cdot) \;+\; \|a - a_{\Delta x}\|_{L^1}
\end{aligned} \tag{15}$$

This bound doesn't vanish as $\Delta x \to 0$, but it is valid for discontinuous solutions and one can see that there's still no time-amplification when the Courant number $\nu = 1$, similarly as (13).

In Appendix A we will give a motivation for the choice of a nonlinear functional, which is the counterpart of this analysis for the nonlinear equation (1).

3 Error estimate for non-resonant wave-front tracking algorithm

This section is dedicated to the derivation of an error estimate which is **linear** in both δ and t for the well-balanced wave-front tracking algorithm already introduced in [2,16], where δ represents a mesh parameter for the algorithm (see (20)); this strongly improves the situation depicted in both [29,19].

3.1 Structural assumptions

We consider the Cauchy problem

$$\partial_t u + \partial_x f(u) = k(x)g(u) \tag{16}$$

$$u(t=0,x) = u_0(x) \in L^1 \cap BV(\mathbb{R}) \tag{17}$$

for $x \in \mathbb{R}$, under the assumptions

$$f,\; g \in C^2\,, \qquad \inf_u f'(u) > 0\,, \qquad k(x) \in L^1_{loc}\,.$$

Set $a(x) = \int^x k(s)\,ds$. We also assume that the definition of w, see (7):

$$w(u,a) = \phi^{-1}\big(\phi(u) - a\big), \qquad \phi' = \frac{f'}{g}$$

holds except at a finite number of point, where ϕ is possibly singular. Therefore, for simplicity, we assume g to have a finite number of zeros. As noticed in [11], w can be extended to u with continuity at these points:

$$\forall\, \bar{u} \in \mathbb{R}\,, \quad g(\bar{u}) = 0 \iff \lim_{u \to \bar{u}} w(u,a) = \bar{u}\,.$$

Moreover, with the above extension, the map $u \mapsto w(u;a)$ is Lipschitz continuous, uniformly for a in a bounded set. Indeed, its partial derivative is continuous outside the zeros of g. On the other hand, let $\bar{u}$ satisfy $g(\bar{u}) = 0$. We apply the intermediate value theorem to $\phi(w) = \phi(u) - a$ to obtain

$$w(u;a) - u = -\frac{a}{\phi'(\xi)} = -a\frac{g(\xi)}{f'(\xi)} = -a\frac{g(\xi) - g(\bar{u})}{f'(\xi)}\,,$$

for some ξ intermediate between w and u. Hence $w(u;a) - w(\bar{u};a) = (w(u;a) - u) + (u - \bar{u})$, with

$$|w(u;a) - u| \leq C\|a\|_\infty \frac{\|g'\|_\infty}{\inf f'} |u - \bar{u}|$$

with a constant $C \geq 1$.

We assume also that the initial data are located in an invariant domain for the equation. In terms of the Riemann coordinates (a, w), invariant domains correspond simply to rectangles. We define $w_0(x) = w(u_0(x), a(x))$ and assume that the data are confined into a rectangle:

$$(a(x), w_0(x)) \in K \doteq [\bar{a}_1, \bar{a}_2] \times [\bar{w}_1, \bar{w}_2] \tag{18}$$

for some constants $\bar{a}_1 < \bar{a}_2$, $\bar{w}_1 < \bar{w}_2$. This assumption is quite reasonable in several cases; as an example, it is met for *every* bounded initial data $(a(x), u_0(x))$ with $a \in BV(\mathbb{R})$ if we assume that the trajectories of ordinary differential equation

$$\frac{du(a)}{da} = \frac{g(u)}{f'(u)} \tag{19}$$

do not blow up in finite intervals (see (1.3) in [2] and [39]). In terms of the coordinates (a, u) the invariant domains are of the form

$$\{(a, u) \in \mathbb{R}^2;\ \bar{a}_1 \leq a \leq \bar{a}_2\,,\ \varphi_1(a) \leq u \leq \varphi_2(a)\}$$

where φ_1, φ_2 are solutions to (19) defined on the common interval $[a_1, a_2]$.

3.2 Wave-front tracking approximations

The wave-front tracking technique provides a standard tool for the analysis of solutions to conservation laws. It was introduced in [7] for scalar conservation laws without source terms. It has been later extended to the case of systems (see the reference books [4,20]) and to equations/systems with coefficients depending on x (see for instance [26,23]).

Here we are going to describe a wave-front tracking algorithm for the equation (16), as defined in [16, Section 2.1]. The basic steps are the following.

Step 1. We fix a partition of $[\bar{w}_1, \bar{w}_2]$: $\mathcal{P} = \{\bar{w}_1 = w_0, \ldots, \bar{w}_2 = w_n\}$ and let δ be the corresponding mesh parameter:

$$\delta \doteq \max\{w_i - w_{i-1}\}\,. \tag{20}$$

This will be used to approximate rarefaction fans: see next point (b). Notice that the partition on w induces a partition on the u axis that depends on a: indeed, by defining

$$u = P(w;a)$$

to be the inverse function of $u \mapsto w(a, u)$ (recall that $w_u > 0$), we get the partition $\widetilde{\mathcal{P}}(a) = \{P(w_0;a), \ldots, P(w_n;a)\}$ for the u variable.

Step 2. We define a piecewise constant Riemann solver. Given $U_\ell = (a_\ell, u_\ell)$ and $U_r = (a_r, u_r)$, we consider the usual Riemann problem

$$(a,u)(0,x) = U_\ell \quad \text{for } x < 0\,, \qquad (a,u)(0,x) = U_r \quad \text{for } x > 0\,.$$

The following procedure ensures that, if $w_\ell = w(a_\ell, u_\ell)$ and $w_r = w(a_r, u_r) \in \mathcal{P}$, then the piecewise constant Riemann solver still takes values in $\mathcal{P}$. The solution is composed by:

(i) a single steady wave connecting (a_ℓ, w_ℓ) to (a_r, w_ℓ)

(ii) one or more waves connecting (a_r, w_ℓ) to $(a_r, \underline{w}_r)$. To do this, let $\widetilde{\mathcal{P}}(a_r) = \{u_0, \dots, u_n\}$ the partition on u corresponding to a_r and let $\bar{f}$ be the linear interpolation of f such that $\bar{f}(u_j) = f(u_j)$.

Then, for $x > 0$, the approximate solution u is defined as the *exact* solution of the problem

$$u_t + \bar{f}(u)_x = 0\,, \qquad u(0,x) = \begin{cases} P(w_\ell, a_r) & \text{for } x < 0\,, \\ u_r & \text{for } x > 0\,. \end{cases}$$

Clearly, such solution is piecewise constant, valued in $\mathcal{P}$ and the waves have positive speed.

Step 3. We take a piecewise constant initial data $(a, u_0)(x)$ such that $w(a, u_0)(x) = w_0(x) \in \mathcal{P}$. At each point of discontinuity of $(a(x), w_0(x))$, we solve the corresponding Riemann problem as indicated in (b) . The solution is then defined up to the first time at which an interaction between waves occurs; to prolong the solution after this time, it is enough to consider the new Riemann problem arising at the point of interaction and solve it according to the method described in (b). At the next time where an interaction occurs, the procedure is repeated and so on.

Following [16, Lemma 2.1], the interactions between wave fronts are proved to be finite and therefore the approximate solution is defined for all $t \geq 0$ with values in $\mathcal{P}$. Moreover, the total variation of the Riemann invariant $w(t, \cdot)$ is non-increasing in time. We remark also that the mesh parameter δ concerns only the z, u variables, while $a(x)$ – at this stage – is requested only to be piecewise constant.

3.3 Stability estimates for wave-front tracking approximations

Let $U_1(t,x) = (b,v)(t,x)$ and $U_2(t,x) = (a,u)(t,x)$ be two wave-front tracking approximations, as in Subsect. 3.2. Let $z(t,x)$ be the Riemann coordinate (see (7)) related to $U_1 = (b,v)$. Following [16], we introduce the weight functions

$$W_1(t,x) = \kappa_1 \sum_{y<x} |\Delta z(t,y)| \tag{21}$$

$$W_2(x) = \exp\left(\kappa_2 \sum_{y>x} |\Delta a(y)|\right) \tag{22}$$

where κ_1, κ_2 are constant values to be determined. Then we define the functional

$$\Lambda(t; U_1, U_2) = \int_{x_1+Lt}^{x_2} W_1(t,x) W_2(x) |p(x)| \,+\, W_2(x) |q(t,x)|\, dx\,, \tag{23}$$

where

$$L = \max_{(a,u)\in K} f'(u) \tag{24}$$

and

$$p(x) = a(x) - b(x)\,, \qquad q(t,x) = u(t,x) - \widetilde{v}(t,x) \tag{25}$$

with

$$\widetilde{v}(t,x) = \varphi(a(x); b(x), v(t,x))\,. \tag{26}$$

Here above, φ represents the trajectory of the ODE (19) issued at $(b(x), v(t,x))$. Another way to express $\widetilde{v}$ is by means of the Riemann invariant and reads:

$$\widetilde{v}(t,x) = \phi^{-1}\big(\phi(v(t,x)) + p(x)\big), \qquad \phi' = \frac{f'}{g}\,. \tag{27}$$

The following result was obtained in [16] (Th. 3.1 and Cor. 3.4). Given $x_1 < x_2$, it provides an estimate on the L^1 norm of $U_1 - U_2$ within the domain of determinacy

$$\left\{(t,x):\ (t,x):\quad 0 \le t \le \frac{x_2 - x_1}{L}\,,\ x_1 + Lt < x < x_2\right\}\,.$$

Theorem 1 *Let $U_1 = (b, v)$ and $U_2 = (a, u)$ be two wave-front tracking approximations with values in K. Let $\mathcal{P}_1$, $\mathcal{P}_2$ the corresponding partitions and δ_1, δ_2 be the related mesh parameters (see* (20)*).*

Moreover denote

$$\rho = \mathrm{TV}\left\{a; [x_1, x_2]\right\}, \tag{28}$$

$$r_1 = \mathrm{TV}\left\{z[U_1](0,\cdot); [x_1, x_2]\right\}, \qquad r_2 = \mathrm{TV}\left\{z[U_2](0,\cdot); [x_1, x_2]\right\}. \tag{29}$$

Then there exists a constant $C > 0$ and a choice of κ_1, κ_2 such that the functional $\Lambda(t) := \Lambda(t; U_1, U_2)$ satisfies for all $0 \le s \le t \le (x_2 - x_1)/L$:

$$\frac{\Lambda(t) - \Lambda(s)}{t - s} \le C \cdot \mathrm{e}^{\kappa_2 \rho} \cdot [\delta_1 r_1 + \delta_2 r_2]\,. \tag{30}$$

Concerning the L^1 norm, denote by $\mathcal{I}(t)$ the integral

$$\mathcal{I}(t) = \int_{x_1+Lt}^{x_2} |u(t,x) - v(t,x)|\, dx\,, \qquad 0 \le t \le \frac{x_2 - x_1}{L}\,. \tag{31}$$

Then, estimate (30) *leads to*

$$\begin{aligned}\mathcal{I}(t) \le{}& \mathcal{I}(0) \cdot \mathrm{e}^{\kappa_2\rho} + C'(1 + \kappa_1 r_1)\mathrm{e}^{\kappa_2\rho} \int_{x_1}^{x_2} |a(x) - b(x)|\, dx \\ &+ C \cdot \mathrm{e}^{\kappa_2\rho} \cdot [\delta_1 r_1 + \delta_2 r_2] \cdot t\end{aligned} \tag{32}$$

for a suitable constant C' .

We stress that the quantities ρ, κ_i, δ_i, r_i, C, C' in (30), (32) *do not depend on time*. For convenience of the reader, we report the proof of (32), based on (30).

Proof of (32). We notice that $\Lambda(t; U_1, U_2)$ is equivalent to $\|U_1(t) - U_2(t)\|_{L^1((x_1+Lt, x_2))}$. Indeed, the weight functions are uniformly bounded:

$$0 \le W_1 \le \kappa_1 \mathrm{TV}\left\{z[U_1](0,\cdot); [x_1, x_2]\right\} = \kappa_1 r_1\,, \tag{33}$$

$$1 \le W_2 \le \exp\left(\kappa_2 \mathrm{TV}\left\{a; [x_1, x_2]\right\}\right) = \mathrm{e}^{\kappa_2\rho}\,. \tag{34}$$

Moreover, by defining

$$M \doteq \sup_{u:\ (a,u)\in K} \left|\frac{g(u)}{f'(u)}\right|\,, \tag{35}$$

and recalling (25), (26) we find that

$$|u - v| \ \le\ |q| + M|p|\,, \qquad |q| \ \le\ |u - v| + M|p|\,. \tag{36}$$

From (30) we have

$$\Lambda(t) \le \Lambda(0) + C\,\mathrm{e}^{\kappa_2\rho} \cdot [\delta_1 r_1 + \delta_2 r_2]\, t\,.$$

Therefore we use (33)–(36) to find that

$$\mathcal{I}(t) \le \int_{x_1+Lt}^{x_2} |q|\,dx + M\int_{x_1+Lt}^{x_2} |p|\,dx \tag{37}$$

$$\le \Lambda(t) + M\int_{x_1+Lt}^{x_2} |p|\,dx$$

$$\le \Lambda(0) + M\int_{x_1+Lt}^{x_2} |p|\,dx + C\cdot \mathrm{e}^{\kappa_2\rho}\cdot[\delta_1 r_1 + \delta_2 r_2]\cdot t\,. \tag{38}$$

On the other hand we have

$$\Lambda(0) \le \mathrm{e}^{\kappa_2\rho}\left\{\int_{x_1}^{x_2} |q(0,x)|\,dx + \kappa_1 r_1\int_{x_1}^{x_2} |p(x)|\,dx\right\}$$

$$\le \mathrm{e}^{\kappa_2\rho}\left\{\mathcal{I}(0) + (M+\kappa_1 r_1)\int_{x_1}^{x_2} |p(x)|\,dx\right\}. \tag{39}$$

Therefore, using (39) within (38), we are ready to conclude that

$$\mathcal{I}(t) \le \mathrm{e}^{\kappa_2\rho}\left\{\mathcal{I}(0) + (2M+\kappa_1 r_1)\int_{x_1}^{x_2} |p(x)|\,dx + C[\delta_1 r_1 + \delta_2 r_2]\,t\right\}$$

that leads to (32). □

3.4 Limit $\delta \to 0$ and recovery of Kružkov's entropy solution

Now we assume that U_1, U_2 are wave-front tracking approximations of the same exact solution, associated to the data (a_0, u_0). By letting $\delta_2 \to 0$ in (32), then U_2 approaches the exact solution $(a_0(x), u(t,x))$. Therefore we are able to deduce an error estimate for the wave-front tracking scheme: see next Corollary 1. In order to achieve convergence, we need to specify how the initial data are approximated. Set $\delta = \delta_1$ and

$$b(x) = a(j\delta)\,, \quad v(0,x) = u_0(j\delta) \qquad \text{for } x \in [j\delta, (j+1)\delta)\,.$$

Let $\mathcal{P}$ be any partition of $[\bar w_1, \bar w_2]$ with mesh parameter $\le \delta$ that includes all the points $w\,(a(j\delta), u_0(j\delta))$, as j varies in $\mathbb{Z}$.

Corollary 1 *Let $u(t,x)$ be a solution to* (1), (3) *and let $a(x) = \int_{-\infty}^{x} k(s)\,ds$. Denote $w_0 = w(u_0, a)$ and let (v, b) be the approximate solution with parameter δ corresponding to the data above. Then the following inequality holds:*

$$\int_{x_1+Lt}^{x_2} |u(t,x) - v(t,x)|\,dx \le \delta\, C\,\mathrm{e}^{\kappa_2\rho}\, r\cdot t \tag{40}$$

$$+ \delta\,\mathrm{e}^{\kappa_2\rho}\big[\mathrm{TV}\,\{u_0; [x_1, x_2]\} + C'(1+\kappa_1 r)\rho + \mathcal{O}(1)\delta\big],$$

where $\rho = \mathrm{TV}\,\{a; [x_1, x_2]\}$, $r = \mathrm{TV}\,\{w_0; [x_1, x_2]\}$.

Proof. We first consider a sequence $\delta_{2,k} \to 0$ and perform the limit in (32), similarly as done in [16, Th. 4.1].

Given a sequence of partitions $\mathcal{P}_k$, with corresponding $\delta_{2,k} \to 0$ (see (20)) as $k \to \infty$, we choose $a_{0,k}$, $u_{0,k}$ piecewise constant and such that:

$$a_{0,k} \to a_0\,, \quad u_{0,k} \to u_0 \quad \text{in } L^1_{loc}\,, \qquad w_{0,k} = w(a_{0,k}, u_{0,k}) \in \mathcal{P}_k\,,$$

and, for some R independent on k:

$$\begin{aligned}
&\mathrm{TV}\,\{a_{0,k}\} \le \mathrm{TV}\,\{a_0\}\,, \qquad \mathrm{TV}\,(a_{0,k}, u_{0,k}) \le R\,,\\
&\limsup_{k\to\infty} \mathrm{TV}\, w(a_{0,k}, u_{0,k}) \le \mathrm{TV}\, w(a_0, u_0)\,.
\end{aligned}$$

Choosing the approximation of (a_0, u_0) as above, we find that $\rho_k \le \rho$ and that $r_{2,k}$ is uniformly bounded (see (28) and (29)). Moreover, the constant values C, C' are uniform in k. Therefore, computing the limit in (32) we find that

$$\begin{aligned}
&\int_{x_1+Lt}^{x_2} |u(t,x) - v(t,x)|\,dx \qquad (41)\\
&\le \mathrm{e}^{\kappa_2\rho}\left[\int_{x_1}^{x_2} |u_0(x) - v(0,x)|\,dx + C'(1+\kappa_1 r)\int_{x_1}^{x_2} |a(x)-b(x)|\,dx + C\,\delta_1\,r\,t\right].
\end{aligned}$$

Now we notice that

$$\begin{aligned}
&\int_{x_1}^{x_2} |u(0,x) - v(0,x)|\,dx \le \delta\,\left\{\mathrm{TV}\,\{u_0; [x_1,x_2)\} + O(1)\delta\right\}\,,\\
&\int_{x_1}^{x_2} |a(x)-b(x)|\,dx \le \delta\left\{\int_{x_1}^{x_2} |a'(x)|\,dx + \delta\|a'\|_\infty\right\}. \qquad (42)
\end{aligned}$$

We then substitute in (41) and obtain

$$\begin{aligned}
&\int_{x_1+Lt}^{x_2} |u(t,x) - v(t,x)|\,dx\\
&\qquad \le \delta\,\mathrm{e}^{\kappa_2\rho}\big[\mathrm{TV}\,\{u_0; [x_1,x_2)\} + C'(1+\kappa_1 r)\rho + O(1)\delta + C\,r\,t\big]\,.
\end{aligned}$$

□

4 Error estimate for the non-resonant Godunov scheme

By construction, the WB Godunov scheme has zero numerical viscosity at steady-state: here it is rigorously shown that, thanks to its Temple reformulation, it suffers from a less harmful error amplification as time grows, one thing probably leading to the other.

4.1 Design of a "wave-front tracking/Godunov scheme"

Hereafter, a uniform Cartesian computational grid is considered, with a mesh-width and time-step denoted by Δx and Δt respectively, always supposed to satisfy the classical CFL stability restriction $L\Delta t = \Delta x$, where L is given at (24). For all $j \in \mathbb{Z}$, the typical computational cell is $C_j = ((j-\frac{1}{2})\Delta x, (j+\frac{1}{2})\Delta x)$. Given a parameter $\delta > 0$, a numerical approximation $u = u^{\Delta t,\delta}$ is built as follows.

We are interested in a local-in-space estimate, on the domain of dependence established by a certain interval $[x_1, x_2]$.

(i) Initial data $a(x)$, u_0 are approximated by

$$a^{\Delta x}(x) = a(j\Delta x)\,, \qquad u^{\Delta t,\delta}(x,0) = u_0(j\Delta x)\,, \qquad x \in C_j\,; \qquad (43)$$

this choice preserves steady solutions.

A partition $\mathcal{P}_0$ is introduced, with mesh parameter $\leq \delta$ and that contains all the values of $w(a^{\Delta x}, u^{\Delta t,\delta})(x)$, $x \in \cup C_j : C_j \cap [x_1, x_2] \neq \emptyset$; the partition is finite.

(ii) On the time interval $(0, \Delta t)$, $u^{\Delta t,\delta}(t,x)$ is defined according to the WFT procedure; here it simply corresponds to solving the Riemann problems at the points $(j+\frac{1}{2})\Delta x$ with the piecewise constant Riemann solver (rarefaction waves are partitioned, see (b) of the WFT procedure). The solution turns out to be piecewise constant; thanks to CFL condition, wave fronts do not interact.

(iii) At time $t = \Delta t$, the projection step is performed:

$$u^{\Delta t,\delta}(\Delta t+, x) = \frac{1}{\Delta x} \sum_j \chi_{C_j}(x) \int_{C_j} u^{\Delta t,\delta}(\Delta t-, x)\, dx \tag{44}$$

$$\doteq P\left(u^{\Delta t,\delta}(\Delta t-, \cdot)\right) . \tag{45}$$

This procedure may introduce values $w(a^{\Delta x}, u^{\Delta t,\delta})(\Delta t+, \cdot)$ that do not belong to $\mathcal{P}_0$. If this is the case, these new values (that are a finite number) are added to the partition, leading to a new partition $\mathcal{P}_1$ whose mesh parameter will be still $\leq \delta$; otherwise we simply set $\mathcal{P}_1 = \mathcal{P}_0$.

(iv) Step (ii) is repeated with initial data $u^{\Delta t,\delta}(n\Delta t+, \cdot)$ and partition $\mathcal{P}_n$, followed by the projection step (iii). It induces a time-marching process which goes on arbitrarily.

We recall that $\mathrm{TV}\, w(a^{\Delta x}, u^{\Delta t,\delta})$ does not increase across the averaging step, (44). In order to extend the error estimate (40) to the aforementioned WFT-Godunov scheme, we need to compare the exact solution (a, u) with the approximate solution $(a^{\Delta x}, u^{\Delta t,\delta})$. To do so we employ the functional (23), that compares $(a^{\Delta x}, u^{\Delta t,\delta})$ with any WFT approximation of the exact solution, having an arbitrarily small parameter δ_1; the limiting process $\delta_1 \to 0$ (keeping Δx, Δt, δ fixed) leads to the desired estimate. The key point is to understand how the functional (23) jumps at each averaging step (44); in the remaining part of the layer, it is governed by (30).

4.2 Control of the functional's jump at each averaging step

Let $b(x)$ be a piecewise constant approximation of $a(x)$ and let $v(t,x)$ be a WFT approximation of the exact solution u with the same mesh parameter δ. We assume that b satisfies (42). Set $U_1 = (b, v)$ and $U_2 = (a^{\Delta x}, u^{\Delta t,\delta})$; the functional $\Lambda(t) := \Lambda(t; U_1, U_2)$, see (23), is defined out of time steps $t_n = n\Delta t$, while at each time t_n it changes due to the Godunov projection. Since it affects only the term $u^{\Delta t,\delta}$, while the terms $a^{\Delta x}$, b and v do not change, the terms

$$W_1(t,x)\,, \qquad W_2(x)\,, \qquad p(x)$$

(see (21), (22)) do not change either. Hence

$$\Delta\Lambda(t_n) = \int_{x_1+Lt}^{x_2} W_2(x)\left(|Pu(t_n) - \widetilde{v}(t_n,x)| - |u(t_n-,x) - \widetilde{v}(t_n,x)|\right) dx \tag{46}$$

(we drop superscripts on u for simplicity). The variation of the above term is estimated in the following proposition.

Proposition 1 *Let $\mathcal{C} = \{\cup_j C_j :\ C_j \cap (x_1 + Lt, x_2) \neq \emptyset\}$. One has*

$$\int_{x_1+Lt}^{x_2} |Pu - \widetilde{v}(x)| - |u(x) - \widetilde{v}(x)|\, dx \leq 2\Delta x\, \mathrm{TV}\left\{\widetilde{v}; \mathcal{C}\right\}. \tag{47}$$

In addition, if either u or v contains only zero-waves, then (47) *improves to*

$$\int_{x_1+Lt}^{x_2} |Pu - \widetilde{v}(x)| - |u(x) - \widetilde{v}(x)|\, dx \leq 0\,. \tag{48}$$

Proof. For any constant $c \in \mathbb{R}$, we apply triangle inequality and get

$$|Pu - \widetilde{v}(x)| - |u(x) - \widetilde{v}(x)| \leq |Pu - c| - |u(x) - c| + 2|c - \widetilde{v}(x)|\,.$$

On each set C_j we choose a constant c_j belonging to the range of $\widetilde{v}$ on C_j. Therefore we find that

$$\begin{aligned} &\int_{C_j} |Pu - c_j| - |u(x) - c_j|\,dx \leq \sup_a \int_{C_j} |Pu - a| - |u(x) - a|\,dx \leq 0\,, \\ &\int_{C_j} |c_j - \widetilde{v}(x)|\,dx \leq \Delta x\,\mathrm{TV}\,\{\widetilde{v}; C_j\}\,. \end{aligned} \tag{49}$$

Putting together the last two estimates, we end up with (47):

$$\begin{aligned} &\int_{x_1+Lt}^{x_2} |Pu - \widetilde{v}(x)| - |u(x) - \widetilde{v}(x)|\,dx \\ &\leq \sum_j \int_{C_j} |Pu - \widetilde{v}(x)| - |u(x) - \widetilde{v}(x)|\,dx \;\leq\; 2\Delta x\,\mathrm{TV}\,\{\widetilde{v}; \cup_j C_j\}\,. \end{aligned}$$

Concerning (48), assume first that u contains only zero-waves. They are located on the grid $\{(j + 1/2)\Delta x\}$, so that u is constant on each cell C_j. Therefore $Pu(t_n) = u(t_n)$ and the integral in (48) is equal to 0.

On the other hand, assume that v contains only zero-waves, that is, the corresponding Riemann invariant $\phi^{-1}(\phi(v) - b)$ is constant. Then $\widetilde{v} = \phi^{-1}(\phi(v) + a^{\Delta x} - b)$ may change only due to the presence of $a^{\Delta x}$: in other words, it is constant on each cell C_j. By setting $\widetilde{v}(x) = \widetilde{v}_j$ on C_j, and arguing as in (49), we have

$$\int_{x_1+Lt}^{x_2} |Pu - \widetilde{v}(x)| - |u(x) - \widetilde{v}(x)|\,dx \;\leq\; \sum_j \int_{C_j} |Pu - \widetilde{v}_j| - |u(x) - \widetilde{v}_j|\,dx \;\leq\; 0\,.$$

□

Remark 2 The inequality (48) quantifies accurately the well-balanced character of our approximations. One way to rephrase it can be: the averaging step of the Godunov procedure ceases to increase the functional (23) as soon as either u or v reaches steady-state (one 0-Riemann invariant becomes a constant).

We now ready to quantify the L^1 distance between $u^{\Delta t,\delta}$ and v.

Theorem 2 *For U_1, U_2 as above and having set $I(t) = \int_{x_1+Lt}^{x_2} |u^{\Delta t,\delta}(t,x) - v(t,x)|\,dx$, one has*

$$I(t) \leq \mathrm{e}^{\kappa_2\rho}\,\{I(0) \,+\, \mathcal{O}(1)\,(\Delta x + \delta)\} \,+\, \mathrm{e}^{\kappa_2\rho}\,\mathcal{O}(1)\,\{L \,+\, \delta\}\,r\,t \tag{50}$$

where $\rho = \mathrm{TV}\,\{a; [x_1, x_2]\}$, $r = \mathrm{TV}\,\{w_0; [x_1, x_2]\}$.

Proof. Set $\Lambda(t) := \Lambda(t; U_1, U_2)$. In view of (30), we find that

$$\Lambda(t) \leq \Lambda(0) + \sum_n \Delta\Lambda(t_n) + C \cdot \mathrm{e}^{\kappa_2\rho} \cdot \delta\,r\,t\,.$$

Recalling (46) and (47), we find

$$\sum_n \Delta\Lambda(t_n) \leq 2\Delta x\,\mathrm{e}^{\kappa_2\rho} \sum_n \mathrm{TV}\,\{\widetilde{v}(t_n); \mathcal{C}(t_n)\}\,.$$

Recalling (27) and that $z[U_1] = w = \phi^{-1}\big(\phi(v) - b\big)$, we have

$$\widetilde{v} = \phi^{-1}\big(\phi(v) + a^{\Delta x} - b\big) = \phi^{-1}\big(\phi(w) + a^{\Delta x}\big)\,.$$

Since $a^{\Delta x}$ changes only at the points x_j and the map $w \mapsto \phi^{-1}\left(\phi(w) + a\right)$ is Lipschitz continuous, uniformly w.r.t. a, we deduce that

$$\mathrm{TV}\left\{\widetilde{v}(t_n); C_j\right\} \leq L_1 \mathrm{TV}\left\{w(t_n); C_j\right\}$$

for a suitable constant L_1. Henceforth:

$$\sum_n \Delta\Lambda(t_n) \leq 2\Delta x\, \mathrm{e}^{\kappa_2\rho}\, \frac{t}{\Delta t}\, L_1\left(r + \mathcal{O}(1)\Delta x\right) \,=\, 2\,\mathrm{e}^{\kappa_2\rho}\, L\, L_1\left(r + \mathcal{O}(1)\Delta x\right) t\,.$$

In conclusion we find that

$$\Lambda(t) \leq \Lambda(0) + \,\mathrm{e}^{\kappa_2\rho}\,\left\{2L\, L_1\left(r + \mathcal{O}(1)\Delta x\right) + C\cdot\delta\, r\right\}\, t\,.$$

By estimating $\Lambda(t)$, $\Lambda(0)$ as in (37), (39) respectively,

$$I(t) \leq \mathrm{e}^{\kappa_2\rho}\,\left\{I(0) \;+\; (2M + \kappa_1 r)\int_{x_1+Lt}^{x_2} |p(x)|\,dx \;+\; (2L\, L_1 + C\cdot\delta)\; r\, t\right\}\,.$$

About $\int |p|$, we proceed as in (42) and write

$$\int |p(x)|\,dx \leq \int |a - a^{\Delta x}|\,dx + \int |b - a|\,dx = \mathcal{O}(1)\left(\Delta x + \delta\right)\,.$$

In conclusion we get (50). □

Remark 3 In the case of $a'(x)g(u) \leq 0$ for all x, u the estimate (50) can be improved by replacing $\mathrm{e}^{\kappa_2\rho}$ with 1 on the right hand side. This is because the weight W_2 is not needed anymore and can be replaced by 1 in the functional (see [16]).

4.3 Sending $\delta \to 0$: the Well-Balanced Godunov scheme

As $\delta \to 0$, the wave-front tracking/Godunov scheme reduces to a classical Well-Balanced Godunov scheme, that reads as follows. We set Δx, Δt as in Subsec. 4.1 and $t_n = n\Delta t$, $x_j = (j + \frac{1}{2})\Delta x$.

(i) At time $t = 0$ we set the initial data as in (43): $u_j^0 = u_0(j\Delta x)$.

(ii) On each time strip (t_n, t_{n+1}) and each cell $C_j = (x_{j-1}, x_j)$ the approximate solution $u^{\Delta t}$ is defined as follows. At x_{j-1} a stationary wave is introduced by solving

$$\phi(u^n_{j-\frac{1}{2}}) = \phi(u^n_{j-1}) + \int_{x_{j-1}}^{x_j} k(y)\,dy\,, \qquad \phi' = \frac{f'}{g}\,; \tag{51}$$

the intermediate state $u^n_{j-\frac{1}{2}}$ is well-defined by such equation. Now, in the set $(t_n, t_{n+1})\times C_j$, $u^{\Delta t}$ is given by the solution of the problem

$$\partial_t v + \partial_x f(v) = 0\,, \qquad v(t, x_{j-1}+) = u^n_{j-\frac{1}{2}}\,, \qquad v(t_n, x) = u^n_j\,,$$

given by a single wave of positive speed (either shock or rarefaction) between $u^n_{j-\frac{1}{2}}$ and u^n_j.

(iii) At time $t = t_{n+1}$ the projection step is performed, with P as in (45):

$$u^{\Delta t}(t_{n+1}, x) = P\left(u^{\Delta t}(t_{n+1}-, \cdot)\right) .$$

Steps **(ii)** and **(iii)** are repeated inductively, therefore defining $u^{\Delta t}$ for all positive times.

Passing to the limit as $\delta \to 0$ in (50), we obtain an estimate on the distance between the exact solution (a, u) and its WB Godunov approximation $(a^{\Delta x}, u^{\Delta t})$, defined as above:

$$\begin{aligned}\|u^{\Delta t}(t,\cdot) - u(t,\cdot)\|_{L^1([x_1+Lt,x_2])} &\le \mathrm{e}^{\kappa_2 \mathrm{TV}\{a;[x_1,x_2]\}} \|u^{\Delta t}(0,\cdot) - u_0\|_{L^1([x_1,x_2])} \\ &+ \mathrm{e}^{\kappa_2 \mathrm{TV}\{a;[x_1,x_2]\}} \mathcal{O}(1)\left(\Delta x + \mathrm{TV}\{w_0;[x_1,x_2]\}\, L\, t\right) . \end{aligned} \tag{52}$$

4.4 Kuznetsov's exponential estimate: $O(\sqrt{t.\Delta x})$ for $t \simeq 0$

In this section we are going to present a more standard approach to derive an error estimate for the WB Godunov scheme; it is based on a classical Kuznetsov argument ([28]). Let us define

$$N = \sup\{a'(x)g'(\xi) \;\; x \in \mathbb{R}\,, \;\; |\xi| \le \max\{\|u\|_\infty, \|u^{\Delta t}\|_\infty\}\,. \tag{53}$$

In the following we assume that $N > 0$, that corresponds to a source term which is **not dissipative**. As we will see below, this is the case in which an exponential amplification of the error generically occurs, following the classical approach by Kuznetsov. The case of $N \le 0$ is much easier, thanks to the L^1 contractivity property related to (16) (see also Remark 3).

Theorem 3 *For $x_1 < x_2$ there exists a suitable constant C such that*

$$\int_{x_1}^{x_2} |u^{\Delta t}(t,x) - u(t,x)|dx \;\le\; C\,\sqrt{\Delta x}\,\sqrt{A} \,+\, C\,\Delta x\, B\,, \tag{54}$$

where

$$\begin{aligned} A &= [\mathrm{TV}\{w_0\} + \|k\|_{L^1}]\left[\frac{\mathrm{e}^{Nt}-1}{N}\right]\left[\mathrm{e}^{Nt}\,(L+1)\mathrm{TV}\{u_0\} + \|k'\|_\infty\,\|g\|_\infty \frac{\mathrm{e}^{Nt}-1}{N}\right] , \\ B &= \mathrm{e}^{Nt}\,\mathrm{TV}\{u_0\} + \|k\|_\infty \frac{\mathrm{e}^{Nt}-1}{N}\,(\mathrm{TV}\{w_0\} + \|k\|_{L^1})\,. \end{aligned}$$

Proof. Let η be smooth and convex, and q such that $q' = \eta' f'$. Set $R_{j,n} = (t_n, t_{n+1}) \times C_j$. For the numerical approximation $v = u^{\Delta t}$ and a test function $0 \le \varphi \in \mathcal{D}((0,+\infty)\times\mathbb{R})$, the entropy inequality reads as:

$$\begin{aligned} &-\int_{\mathbb{R}\times\mathbb{R}^+} \eta(v)\varphi_t + q(v)\varphi_x\, dxdt = -\sum_{j,n}\int_{R_{j,n}} \ldots\, dxdt \\ &\le \sum_{j,n}\int_{x_{j-1}}^{x_j} [\eta(u_j^n) - \eta(v(t_n-,x))]\varphi(t_n,x)\,dx + \int_{t_n}^{t_{n+1}} [q(u_j^n) - q(u_{j+1/2}^n)]\varphi(t,x_{j+1})\,dt \\ &\doteq \sum_{j,n}\left[I^1_{j,n} + I^2_{j,n}\right] . \end{aligned}$$

As customary (see for instance [2, p.258]) the first term $I^1_{j,n}$ is treated by Jensen's inequality,

that gives

$$\eta(u_j^n) \le (\Delta x)^{-1} \int_{x_{j-1}}^{x_j} \eta(v(t_n-,x))\,dx\,.$$

Therefore

$$\begin{aligned}\sum_{n,j} I_{j,n}^1 &\le \sum_{n,j} \int_{x_{j-1}}^{x_j} \left[\eta(u_j^n) - \eta(v(t_n-,x))\right]\left[\varphi(t_n,x) - \varphi(t_n,x_j)\right]\,dx \\ &\le \Delta x \|\eta'\|_\infty \sum_{n,j} \mathrm{TV}\left\{v(t_n-,\cdot); C_j\right\} \int_{C_j} |\varphi_x(t_n,x)|\,dx\,.\end{aligned}$$

Concerning $I_{j,n}^2$, we recall (51) and that $q' = \eta' f' = \eta' g \phi'$. Then

$$\begin{aligned}q(u_j^n) - q(u_{j+1/2}^n) &= \int_{u_{j+1/2}^n}^{u_j^n} q'(u)\,du = \int_{\phi(u_{j+1/2}^n)}^{\phi(u_j^n)} (\eta' g)(\phi^{-1}(\alpha))\,d\alpha \\ &= (\eta' g)(\zeta_j^n) \cdot \int_{x_j}^{x_{j+1}} k(y)\,dy\end{aligned}$$

with ζ_j^n that belongs to the interval with extrema u_j^n and $u_{j+1/2}^n$. Hence

$$\begin{aligned}I_{j,n}^2 = &\int_{R_{j,n}} (\eta' g)(v)\,k(x)\varphi(t,x)\,dtdx \\ &+ \int_{R_{j,n}} \underbrace{k(x)\left[(\eta' g)(\zeta_j^n) - (\eta' g)(v)\right]}_{=\alpha(t,x)} \varphi(t,x_{j+1})\,dtdx \\ &+ \int_{R_{j,n}} \underbrace{k(x)(\eta' g)(v)}_{=\beta(t,x)} \left[\varphi(t,x_{j+1}) - \varphi(t,x)\right]\,dtdx\,.\end{aligned}$$

Therefore

$$\begin{aligned}&-\int_{\mathbb{R}\times\mathbb{R}^+} \eta(v)\partial_t\varphi + q(v)\partial_x\varphi + k(x)(\eta' g)(v)\varphi(t,x)\,dxdt \qquad (55)\\ &\le \Delta x \|\eta'\|_\infty \sum_{n,j} \mathrm{TV}\left\{v(t_n-,\cdot); C_j\right\} \int_{C_j} |\varphi_x(t_n,x)|\,dxdt \\ &\quad + \sum_{n,j} \int_{R_{j,n}} \alpha(t,x)\varphi(t,x_{j+1})\,dtdx \\ &\quad + \sum_{n,j} \int_{R_{j,n}} \beta(t,x)\left[\varphi(t,x_{j+1}) - \varphi(t,x)\right]\,dtdx\,. \qquad (56)\end{aligned}$$

For a given $\ell \in \mathbb{R}$, we approximate the Kružkov entropy $|u-\ell|$ as follows: let $E \in C^2(\mathbb{R})$ be such that $E'' \ge 0$, $E(v) = |v|$ for $|v| \ge 1$, $E'(0) = 0$. Then define $\eta_\delta(v) = \delta E(\frac{v-\ell}{\delta})$. To pass to the limit as $\delta \to 0$ in (55)–(56), we need to estimate α and β. Recalling that, on the cell $R_{j,n}$, $v(t,x)$ takes values between $u_{j-1/2}^n$ and u_j^n, we obtain:

$$
\begin{aligned}
|\alpha(t,x)| &\le |k(x)| Lip(\eta' g) \left|\zeta_j^n - v(t,x)\right| \\
&\le |k(x)| Lip(\eta' g) \mathrm{TV}\left\{v(t_n,\cdot); (x_{j-1}+, x_j+)\right\}, \\
|\beta(t,x)| &\le |k(x)| \|\eta' g\|_\infty .
\end{aligned}
$$

Using the dominated convergence theorem as $\delta \to 0$ in (55)–(56), we obtain

$$
\begin{aligned}
&-\int_{\mathbb{R}\times\mathbb{R}^+} |v-\ell|\partial_t\varphi + |f(v)-f(\ell)|\partial_x\varphi + k(x)\mathrm{sgn}(v-\ell)g(v)\varphi(t,x)\,dxdt \\
&\le \Delta x \sum_{n,j} \mathrm{TV}\left\{v(t_n-,\cdot); C_j\right\} \int_{C_j} |\varphi_x(t_n,x)|\,dxdt \\
&\quad + \widetilde{C}_2 \sum_{n,j} \mathrm{TV}\left\{v(t_n,\cdot); (x_{j-1}+, x_j+)\right\} \int_{R_{j,n}} |k(x)|\varphi(t,x_{j+1})\,dtdx \\
&\quad + \widetilde{C}_3 \sum_{n,j} \int_{R_{j,n}} |k(x)| \left[\varphi(t,x_{j+1}) - \varphi(t,x)\right]\,dtdx
\end{aligned}
$$

where the constants $\widetilde{C}_j$ depend on L, $\|u\|_\infty$, $\|g\|_{C^1}$. On the other hand, since u is an exact (entropy) solution, we have:

$$
-\int_{\mathbb{R}\times\mathbb{R}^+} |u-\ell|\partial_s\varphi + |f(u)-f(\ell)|\partial_y\varphi + \mathrm{sgn}(u-\ell)k(y)g(u)\varphi(s,y)\,dsdy \le 0 .
$$

Following [3,10], we introduce a test function ϕ of the form

$$
\varphi(t,x,s,y) = \Phi(t,x)\zeta(t-s,x-y)\exp(-Nt)
$$

with $\Phi \ge 0$, $\zeta \ge 0$; $\Phi \in \mathcal{D}((0,+\infty)\times\mathbb{R})$, $\zeta \in \mathcal{D}((-\infty,0)\times\mathbb{R})$ chosen as in (2.10), (2.14) of [3], with parameters $\nu = 0$, $\delta = \Delta$ and $\theta = \Delta/4$. Then

$$
\begin{aligned}
0 \le &\iiiint |v(t,x)-u(s,y)|\Phi_t(t,x)\,\zeta(t-s,x-y)\mathrm{e}^{-Nt}\,dsdydtdx && (57)\\
&+ \iiiint |f(v(t,x))-f(u(s,y))|\Phi_x(t,x)\,\zeta(t-s,x-y)\mathrm{e}^{-Nt}\,dsdydtdx && (58)\\
&-N \iiiint |v(t,x)-u(s,y)|\varphi(t,x,s,y)\,dsdydtdx && (59)\\
&+ \iiiint \left[k(x)\mathrm{sgn}(v(t,x)-\ell)g(v) - k(y)\mathrm{sgn}(u(s,y)-\ell)g(u)\right]\varphi\,dsdydtdx && (60)\\
&+\Delta x \sum_{n,j} \mathrm{TV}\left\{v(t_n-,\cdot); C_j\right\} \iint dyds \int_{C_j} |\varphi_x(t_n,x,s,y)|\,dx && (61)\\
&+\widetilde{C}_2 \sum_{n,j} \mathrm{TV}\left\{v(t_n,\cdot); (x_{j-1},x_j]\right\} \iint dyds \int_{R_{j,n}} |k(x)|\varphi(t,x_{j+1},s,y)\,dtdx && (62)\\
&+\widetilde{C}_3 \iint dyds \sum_{n,j} \int_{R_{j,n}} |k(x)| \left[\varphi(t,x_{j+1},s,y) - \varphi(t,x,s,y)\right]\,dtdx . && (63)
\end{aligned}
$$

The first two lines, (57) and (58), are treated as in [3] by a suitable choice of the function Φ. After approximation of characteristic functions, they lead to the term

$$\int_{x_1+LT-\Delta}^{x_2+\Delta} |u(0,x)-v(0,x)|\,dx - \mathrm{e}^{-NT}\int_{x_1-\Delta/2}^{x_2+\Delta/2} |u(T,x)-v(T,x)|\,dx$$
$$+2(L+1)\mathrm{TV}\,\{u_0\}\Delta\,, \tag{64}$$

see (2.13), (2.16), (2.18), (2.23)-(2.24) and finally (2.9) in [3].

Let us evaluate the terms (59)–(63). The term in [...] in (60) is bounded by:

$$|\,[\ldots]\,| \le |k(y)-k(x)||g(v(t,x))| + |k(y)||g(v(t,x))-g(u(s,y))|\,.$$

The last term is compensated by the term in (59), thanks to the definition (53) of N. The remaining term from (60) is estimated as follows:

$$\iiiint |k(y)-k(x)||g(v(t,x))|\varphi\,dsdydtdx \le \|k'\|_\infty \frac{\Delta}{2}\,\|g\|_\infty \iiiint \varphi(t,x,s,y)\,dsdydtdx\,.$$

The last integral is bounded by:

$$\begin{aligned}\iiiint \varphi\,dsdydtdx &= \iiint \Phi(t,x)\frac{1}{\Delta}\zeta_1^x\left(\frac{x-y}{\Delta}\right)\mathrm{e}^{-Nt}\,dydtdx\\ &= \iint \Phi(t,x)\,\mathrm{e}^{-Nt}\,dtdx\\ &\le m(I)\frac{1-\mathrm{e}^{-NT}}{N}\end{aligned}$$

where we used that $0\le\Phi\le 1$; here I is some bounded interval such that $\Phi(t,x)=0$ for $x\notin I$. It remains to consider (61)–(63). Let us first estimate (61). By the specific form of Φ, that satisfies $|\Phi_x|\le\mathcal{O}(1)/\Delta$ (see (2.26) in [3] with $\theta=\Delta/4$), and by definition of ζ, we find that

$$\exp(Nt)\partial_x\varphi(t,x,s,y) = \partial_x\Phi(t,x)\zeta(t-s,x-y) \;+\; \Phi(t,x)\zeta^t(t-s)\partial_x\left\{\frac{1}{\Delta}\zeta_1^x\left(\frac{x-y}{\Delta}\right)\right\}$$

so that

$$\iint dyds\int_{C_j}|\varphi_x(t_n,x,s,y)|\,dx \le \mathrm{e}^{-Nt_n}C\,\frac{\Delta x}{\Delta}$$

and therefore

$$\begin{aligned}(61) &\le C\frac{(\Delta x)^2}{\Delta}\sum_{n,j}\mathrm{TV}\,\{v(t_n-,\cdot);C_j\}\mathrm{e}^{-Nt_n}\\ &\le C'\frac{\Delta x}{\Delta}\left(\mathrm{TV}\,\{w_0\}+\|k\|_{L^1}\right)\frac{1}{N}\left(1-\mathrm{e}^{-NT}\right)\end{aligned} \tag{65}$$

for suitable constants C, C'.

About (62) one has, using that $\iint\varphi\,dyds=\Phi(t,x)\mathrm{e}^{-Nt}\le\mathrm{e}^{-Nt}$:

$$\begin{aligned}(62) &\le \widetilde{C}_2\sum_{n,j}\mathrm{TV}\,\{v(t_n,\cdot);(x_{j-1}+,x_j+)\}\int_{R_{j,n}}|k(x)|\mathrm{e}^{-Nt}\,dtdx\\ &\le \widetilde{C}_2\|k\|_\infty\Delta x\sum_{n,j}\mathrm{TV}\,\{v(t_n,\cdot);(x_{j-1}+,x_j+)\}\int_{t_n}^{t_{n+1}}\mathrm{e}^{-Nt}\,dt\\ &\le C'\widetilde{C}_2\|k\|_\infty\,\Delta x\,(\mathrm{TV}\,\{w_0\}+\|k\|_{L^1})\frac{1-\mathrm{e}^{-NT}}{N}\,.\end{aligned} \tag{66}$$

Finally, let us estimate (63). Using that $|\Phi_x| \le C/\Delta$ as above, we have:

$$\begin{aligned}
&\iint \left[\sum_{n,j} \int_{R_{j,n}} |k(x)| \left[\varphi(t, x_{j+1}, s, y) - \varphi(t, x, s, y) \right] dtdx \right] dyds \\
&= \sum_{n,j} \int_{R_{j,n}} |k(x)| \left[\Phi(t, x_{j+1}) - \Phi(t, x) \right] \mathrm{e}^{-Nt} \, dtdx \\
&\le C \frac{\Delta x}{\Delta} \sum_{n,j} \int_{R_{j,n}} |k(x)| \mathrm{e}^{-Nt} \, dtdx \\
&\le C \frac{\Delta x}{\Delta} \|k\|_{L^1} \frac{1 - \mathrm{e}^{-NT}}{N} \, . \qquad (67)
\end{aligned}$$

Now we sum up (64)–(67) and get

$$\begin{aligned}
\mathrm{e}^{-NT} \int_{x_1 - \Delta/2}^{x_2 + \Delta/2} |u(T,x) - v(T,x)| \, dx &\le \int_{x_1 + LT - \Delta}^{x_2 + \Delta} |u(0,x) - v(0,x)| \, dx \\
&+ \Delta \left[2(L+1) \mathrm{TV}\left\{u_0\right\} + \|k'\|_\infty \|g\|_\infty m(I) \frac{1 - \mathrm{e}^{-NT}}{N} \right] \\
&+ \frac{\Delta x}{\Delta} C'' \frac{1 - \mathrm{e}^{-NT}}{N} \left(\mathrm{TV}\left\{w_0\right\} + \|k\|_{L^1} \right) \\
&+ \Delta x \, C' \widetilde{C}_2 \, \frac{1 - \mathrm{e}^{-NT}}{N} \|k\|_\infty \left(\mathrm{TV}\left\{w_0\right\} + \|k\|_{L^1} \right)
\end{aligned}$$

for a suitable constant C''. The integral at $t = 0$ is simply bounded by $\Delta x \, \mathrm{TV}\left\{u_0\right\}$. Now we can choose Δ in order to minimize the above quantity: by writing $ax + \frac{b}{x}$, one seeks the zero of its derivative, being $x = \sqrt{b/a}$ that gives the minimum value $2\sqrt{ab}$. Therefore we obtain the estimate (54). □

4.5 A threshold effect and behavior in large time

Having obtained the two estimates (52) and (54), one can compare them and take the more convenient one:

$$\begin{aligned}
&\int_{x_1}^{x_2} |u^{\Delta t}(t,x) - v(t,x)| dx \le \qquad (68) \\
&\le C \min \left\{ \mathrm{e}^{\kappa_2 \mathrm{TV}\left\{a\right\}} \left[\Delta x (\mathrm{TV}\left\{u_0\right\} + 1) + \mathrm{TV}\left\{w_0\right\} L \, t \right] , \, \sqrt{\Delta x} \sqrt{A} + \Delta x \, B \right\}
\end{aligned}$$

with A and B as in (54), that grow exponentially in time. This formula highlights a critical threshold phenomenon: initially, for $t \simeq 0$, the exponential growth obtained through Kuznetsov's method takes place, but by Taylor's expansion, it remains close to a $\frac{1}{2}$-power growth in t:

$$\sqrt{A} \simeq \sqrt{t} \, \left[\mathrm{TV}\left\{w_0\right\} + \|k\|_{L^1} \right]^{1/2} \left[(L+1) \mathrm{TV}\left\{u_0\right\} + \|k'\|_\infty \|g\|_\infty t \right]^{1/2} ,$$

$$B \simeq \mathrm{TV}\left\{u_0\right\} + \|k\|_\infty \left(\mathrm{TV}\left\{w_0\right\} + \|k\|_{L^1} \right) t \, .$$

Once a critical value is reached, the L^1 error is bounded by the estimate (52) which displays a rigorous linear growth in t. The presence of the constant term $\mathcal{O}(1) \, \mathrm{TV}\left\{w_0; [x_1, x_2]\right\}$ which

doesn't tend to zero when $\Delta x \to 0$ means that for very fine grids, the Kuznetsov estimate dominates. Thus the new bound (68) is significant mainly for coarse grids (which are the most interesting in terms of CPU cost): it is illustrated on various test-cases hereafter.

Another interesting feature in the error estimate (52) is that it decouples the effects of the grid parameter Δx and the ones of the time. More precisely, the mesh width Δx affects the error at time $t \simeq 0$, but thanks to the fact that none of the terms acting on t depends on Δx, its overall influence shrinks as time grows. This is easily checked numerically by setting up the Burgers equation corresponding to (5) in the domain $x \in (-4, 60)$ for several grid parameters. On Fig. 3, a comparison of the time evolution of L^1 errors is displayed for both

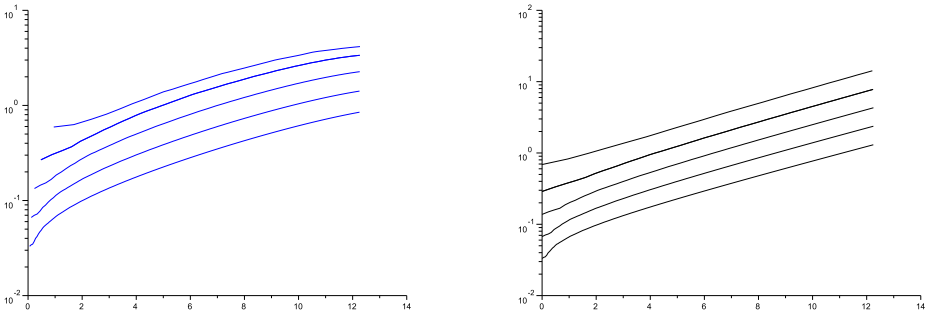

Fig. 3. Time evolution of the measured L^1 error for (5) with $\Delta x = 2^{-n}$, $n = 0, 1, 2, 3, 4$. The WB scheme (left) shows a weaker dependence on the grid compared to the TS one (right) which displays a neat exponential growth.

the WB and the TS scheme: besides the exponential growth, a stronger dependence of the TS scheme with respect to Δx clearly appears.

Finally we remark that, thanks to the non-resonance assumption (2), the estimate (68) can be significantly improved whenever initial data approach a stationary wave as $x \to -\infty$.

Indeed, assume that $w_0(x) = \bar{w}_0$ for $x < x_0$. The exact solution will be stationary with $w(t, x) = \bar{w}_0$ for $x < x_0 + \nu t$, where $\nu = \inf f' > 0$. Thanks to the choice (43) of the approximated initial data, the approximate solution $(a^{\Delta x}, u^{\Delta t, \delta})$ will be stationary itself, containing only 0-waves, with $w(a^{\Delta x}, u^{\Delta t, \delta}) = \bar{w}_0$ on $x < x_0 + \nu t$; same for $\delta \to 0$.

Now consider $x_2 > x_1 > x_0$. After time $t^* = (x_2 - x_0)/\nu + \Delta t$, the approximated solution ($\delta > 0$ and $\delta = 0$) will be steady on $[x_1, x_2]$ and the functional does not longer increase across fractional steps: see (48). Therefore, for $t > t^*$ and $\mathcal{I} = (x_1, x_2)$, one has that $u(t, x) = \phi^{-1}(\phi(w_0) + a(x))$ and (52) is replaced by

$$\begin{aligned} \|u^{\Delta t}(t, \cdot) - u(t, \cdot)\|_{L^1(\mathcal{I})} &= \|u^{\Delta t}(t^*, \cdot) - u(t^*, \cdot)\|_{L^1(\mathcal{I})} \\ &\leq C \|a^{\Delta x}(\cdot) - a(\cdot)\|_{L^1(\mathcal{I})} = \mathcal{O}(1)\Delta x \,, \end{aligned}$$

meaning that only the error of the projector onto piecewise constant functions remains.

5 More numerical results in transient regime

One difficulty lies in finding meaningful examples of inhomogeneous scalar balance laws which admit explicit solutions thus allowing for the computation of the L^1 error of the numerical scheme at each time-step.

5.1 An inhomogeneous N-wave

This transient example is taken from [17]; it stems from a similarity-solution analysis of the accretive balance law,

$$\partial_t u + \partial_x f(u) = u, \quad f(u) = u^4/4, \quad u_0(x) = \mathrm{sgn}(x)(3|x|)^{\frac{1}{3}} \chi_{|x|<\frac{1}{2}}, \tag{69}$$

where χ_A stands for the characteristic function of a set A. According to [17], the entropy solution of (69) simply reads:

$$\forall t > 0, \qquad u(t,x) = \mathrm{sgn}(x)(3|x|)^{\frac{1}{3}} \chi_{|x|<\frac{1}{2}\exp(3t/4)}.$$

It is easy to set up both the WB and the TS Godunov scheme for (69) with an adaptive

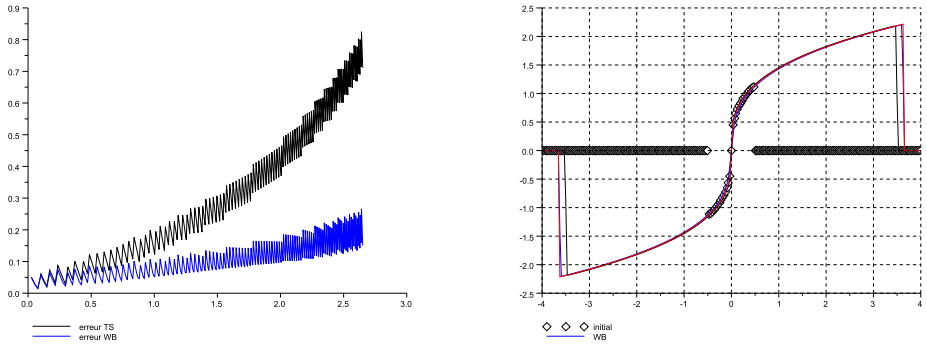

Fig. 4. Time evolution of L^1 error for a N-wave solving (69).

time-step selection in order to keep the CFL number at 0.95. 2^8 points in the x variable have been used to grid the interval $[-4, 4]$ and the marching schemes have been iterated up to $T = 2.65$ to produce the results displayed in Fig. 4. On the left side of the figure, one observes again an exponential-type amplification of the absolute L^1 error for the TS scheme (black curve) whereas the blue curve shows that the WB discretization performs better. By observing the right side of Fig. 4, it is easy to understand that the excessive amplification of the TS scheme's (black curve) error manifests itself through discontinuities moving with a wrong speed. Instead, the blue curve, corresponding to the WB scheme, remains very close to the red curve which illustrates the exact similarity-solution. On Fig. 5, we show that the

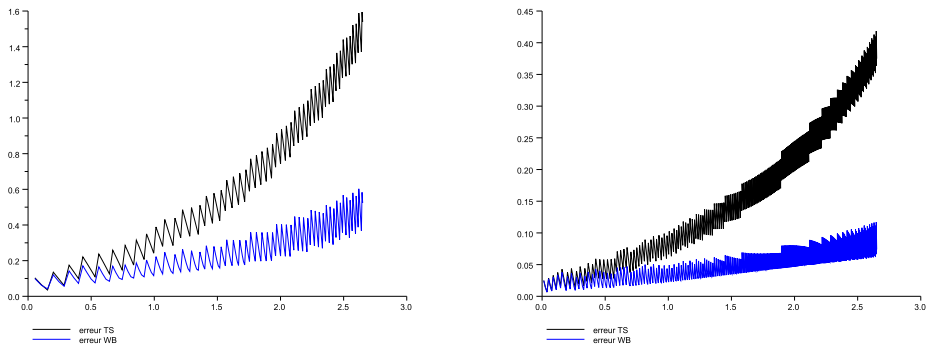

Fig. 5. Time evolution of L^1 error for a N-wave solving (69) with different grids.

qualitative behavior of the measured L^1 error as a function of time doesn't depend on Δx: 2^7 (resp., 2^9) points are used for the left (resp., right) side of the Figure.

5.2 LeVeque-Yee's effect for Riccati source term

This test-case is inspired by the long-standing benchmark proposed in [34] (see also [25,48]):

$$\partial_t u + \partial_x \frac{u^2}{2} = \pm k(1+u)(2-u), k(x) = 2\left(1+\sin(\frac{\pi x}{10})\right), x \in [-0.1, 49.9]. \qquad (70)$$

By prescribing the initial data $u_0(x) = 2 - 3Y(x)$, where Y is the Heaviside function

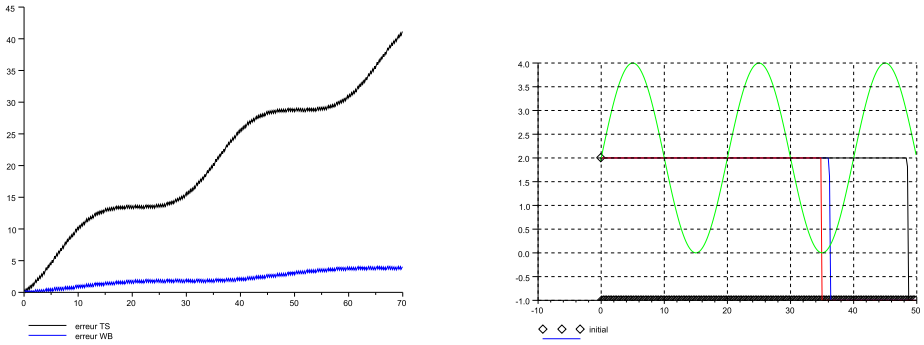

Fig. 6. Time evolution of L^1 error for a "speedup" Riccati source term.

(see (5)), one obtains an entropic shock traveling at constant speed $\sigma = \frac{1}{2}$. However, as a consequence of numerical viscosity, it is quite difficult to reproduce it numerically in a correct manner (even in the non-stiff case) because, according to the sign placed beside k:

- 2 is stable, -1 is unstable with the "+ sign": the values created by numerical viscosity are increased by the source term thus speeding up the shock. This case is depicted on Fig. 6.
- 2 is unstable, -1 is stable with the "− sign": the values created by numerical viscosity are decreased by the source term thus slowing down the shock. This case is depicted on Fig. 7.

On the right side of Fig. 6, one can see that, at time $T = 70$ the black curve generated

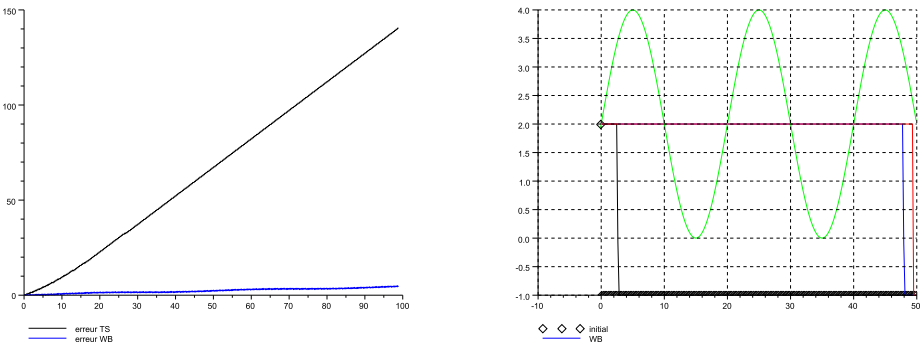

Fig. 7. Time evolution of L^1 error for a "slowdown" Riccati source term.

by the TS scheme is much ahead of both the blue (WB scheme) and red (exact solution) curves. The green curve corresponds to $k(x)$. The WB scheme is only slightly ahead of the exact solution, thanks to its lower numerical dissipation of the original discontinuity. This is revealed by the time-evolution of the L^1 absolute error: the TS error (in black) is 10 times bigger than the WB one (in blue). When the minus sign is selected, the time-amplification of the L^1 error up to $T = 99$ for the TS scheme is even more dramatic because the shock

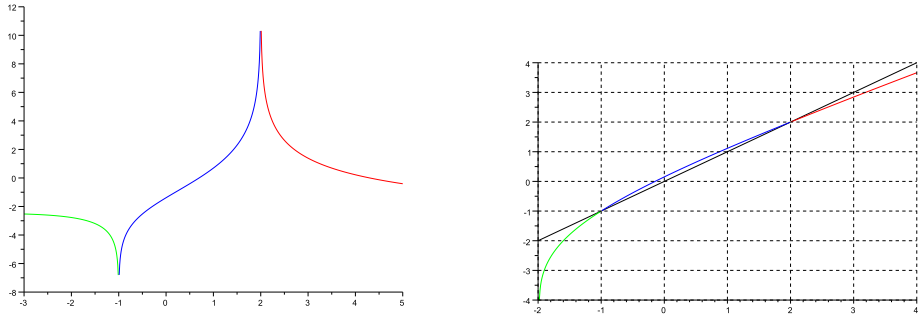

Fig. 8. Illustration of the functions $\phi(u)$ and $w(u,a)$ for the equation (70).

decelerates so much that it becomes nearly static (see the right side of Fig. 7, black curve). The blue curve (WB scheme) remains quite close to the red one (exact solution) and the L^1 error of the WB scheme seems to be independent of the sign put in front of k in (70). For this experiment, 2^8 points have been set up in the x variable, the time-step is chosen $\Delta t = 0.95\Delta x/2$, and the ODE solver involved in the TS scheme is a second order Runge-Kutta. In order to set up the WB scheme, one has to perform logarithmic integrals in order to calculate the function $\phi(u) = \log|1+u| - 2\log|2-u|$: it is displayed (on the left), together with the Riemann invariant (on the right) $w(u, a = 0.3)$ in Fig. 8.

5.3 A stationary roll-wave

For the sake of completeness, a stationary example has also been taken from [17,22]: it consists in simulating a transonic roll-wave which is another similarity-solution of the quartic balance law (69). For this test-case, it is important to be careful in choosing the computational grid in such a way it contains a point in $x = 0$. By doing so, one obtains the results displayed in Fig. 9 with 2^7 points uniformly in space and a constant time-step $\Delta t = 0.95\Delta x/1.9$. The L^1 error of the WB scheme is of the order of the machine precision

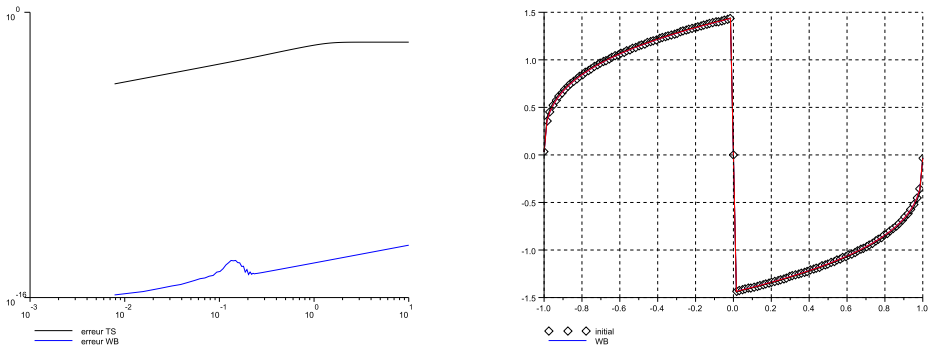

Fig. 9. Time evolution of L^1 error for the static roll-wave.

(blue curve): this isn't surprising since the discrete initial data contains only zero-waves thus is supposed to remain unchanged for all times. A tiny amplification appears though on the left of Fig. 9: this is quite similar to the small destabilization process already encountered in a completely different context in [13], Fig. 5.3.

6 Conclusion and outlook

Several types of anomalous/spurious behavior of time-splitting numerical approximations to scalar balance laws in the non-stiff regimes can be related to the interaction between the numerical viscosity inherent to the homogeneous evolution step and an non-dissipative source term for which $g' > 0$ locally. A similar mechanism can occur for the wave-front tracking algorithm for which one has to deal with the interaction of small but artificial "rarefaction fronts" with a destabilizing source term. It has been presently shown both theoretically and numerically that these drawbacks can be strongly reduced (if not suppressed) by setting up a well-balanced strategy involving a supplementary linearly degenerate field associated to the source term (which is thus rendered by means of a Rankine-Hugoniot relation inside an overall self-similar Riemann solver). Such a formulation displays at least 3 advantages: (already evoked in [13])

(1) lower numerical viscosity leading to significantly smaller error estimates,
(2) stiffness doesn't constitute an important problem,
(3) preservation of stationary regimes (initial data containing only zero waves) because zero numerical dissipation remains at steady-state [37].

Hence, besides Property 3. which is well-known, Properties 1. and 2., mostly concerned with transient regimes, express interesting and perhaps less well-known features of WB schemes. Future directions of investigation would address typically the cases of initial-boundary value problems, space-dependent fluxes and source terms, and (hopefully) $n \times n$ inhomogeneous non-resonant Temple class systems. A first step in such direction could be the derivation of error estimates for the WB/AP schemes proposed in [14] for the approximation of quasi-monotone 2-velocity kinetic models.

A A decaying functional

In this appendix we show formally that the Lyapunov functional Λ is non-increasing along two solutions, provided that κ_1, κ_2 are chosen large enough. We recall (21)–(27):

$$\Lambda(t; U_1, U_2) = \int_{x_1+Lt}^{x_2} W_2(x)|u(t,x) - \widetilde{v}(t,x)| \, + \, W_1(t,x)W_2(x)|a(x) - b(x)| \, dx \, ,$$

where

$$W_1(t,x) = \kappa_1 \int_{-\infty}^{x} |z_x(t,y)| \, dy \, , \qquad W_2(x) = \exp\left\{ \kappa_2 \int_x^{\infty} |a'(y)| \, dy \right\} ,$$
$$\widetilde{v} = \phi^{-1}\big(\phi(v) + a - b\big), \qquad z = \phi^{-1}\big(\phi(v) - b\big) \, , \qquad \phi' = f'/g \, .$$

For simplicity here we assume that ϕ' is bounded, that is, g is bounded away from zero. We first analyze the time evolution of $|u - \widetilde{v}|$. We claim that

$$\partial_t |u - \widetilde{v}| + \partial_x |f(u) - f(\widetilde{v})| \leq |g(u) - g(\widetilde{v})| \, |\partial_x a| \, + \, C|a - b| \, |\partial_x z| \tag{A.1}$$

for a suitable constant $C > 0$.

With the help of (A.1) we can show that, for κ_1 and κ_2 large enough, $\partial_t \Lambda \leq 0$. The weights W_2 and W_1 are used to integrate the two terms appearing in (A.1), respectively. More specifically,

$$\partial_x W_2 = -\kappa_2 W_2 |\partial_x a| \, , \tag{A.2}$$
$$\partial_t W_1 \, = \, -\kappa_1 f'(v) |\partial_x z| \, . \tag{A.3}$$

Notice that (A.2) follows from the definition of W_2; while (A.3) will be proved later on. Using (A.2) we obtain:

$$\begin{aligned}
&\partial_t\{W_2|u-\widetilde{v}|\}+\partial_x\left\{W_2|f(u)-f(\widetilde{v})|\right\}\\
&\qquad\le W_2\left[\left(|g(u)-g(\widetilde{v})|-\kappa_2|f(u)-f(\widetilde{v})|\right)|\partial_x a|+C|a-b||\partial_x z|\right]. \qquad \text{(A.4)}
\end{aligned}$$

By choosing κ_2 large enough, the first term in (A.4) is negative. Hence we get

$$\partial_t\{W_2|u-\widetilde{v}|\}+\partial_x\left\{W_2|f(u)-f(\widetilde{v})|\right\}\ \le\ CW_2|a-b||\partial_x z|\,. \qquad \text{(A.5)}$$

To deal with the r.h.s. of (A.5), we use (A.3) and find that

$$\begin{aligned}
&\partial_t\{W_2|u-\widetilde{v}|+W_1W_2|a-b|\}+\partial_x\left\{W_2|f(u)-f(\widetilde{v})|\right\}\\
&\qquad\le W_2|a-b||\partial_x z|\left[C-\kappa_1 f'(v)\right]\\
&\qquad\le 0
\end{aligned}$$

for κ_1 sufficiently large. Now, using the previous inequality and the definition of L, see (24), we find that

$$\begin{aligned}
&\partial_t\Lambda(t;U_1,U_2)\\
&\le -W_2|f(u)-f(\widetilde{v})||_{x=x_2}+W_2|f(u)-f(\widetilde{v})||_{x=x_1+Lt}\ -\ LW_2|u-\widetilde{v}||_{x=x_1+Lt}\\
&\le 0+W_2\left\{|f(u)-f(\widetilde{v})|\ -\ L|u-\widetilde{v}|\right\}|_{x=x_1+Lt}\\
&\le 0\,.
\end{aligned}$$

This concludes the (formal) proof.

Proof of (A.1). We first obtain an equation satisfied by $\widetilde{v}$. By the definition of $\widetilde{v}$ and z, we have

$$\phi(\widetilde{v})=\phi(v)+a-b=\phi(z)+a\,,\qquad \phi(z)=\phi(v)-b\,,\quad \phi'=f'/g\,.$$

Notice that $\phi(z)$ is constant along v-characteristics. Indeed,

$$\begin{aligned}
(\partial_t+f'(v)\partial_x)\,\phi(z)&=\phi'(v)\,(\partial_t+f'(v)\partial_x)\,(v)-f'(v)\partial_x b\\
&=\phi'(v)\left[\partial_t v+f'(v)\partial_x v-g(v)\partial_x b\right]\ =\ 0\,.
\end{aligned}$$

As a consequence we find that

$$\begin{aligned}
(\partial_t+f'(\widetilde{v})\partial_x)\,\phi(\widetilde{v})&=(\partial_t+f'(\widetilde{v})\partial_x)\,\phi(z)+f'(\widetilde{v})\partial_x a\\
&=(f'(\widetilde{v})-f'(v))\,\partial_x\phi(z)+f'(\widetilde{v})\partial_x a
\end{aligned}$$

and therefore

$$\begin{aligned}
(\partial_t+f'(\widetilde{v})\partial_x)\,(\widetilde{v})&=\frac{1}{\phi'(\widetilde{v})}\,(\partial_t+f'(\widetilde{v})\partial_x)\,\phi(\widetilde{v})\\
&=\frac{1}{\phi'(\widetilde{v})}\,(f'(\widetilde{v})-f'(v))\,\partial_x\phi(z)+g(\widetilde{v})\partial_x a\,.
\end{aligned}$$

In summary: we have to evaluate $\partial_t|u-\widetilde{v}|$, where

$$\partial_t u + \partial_x f(u) = g(u)\partial_x a\,,$$
$$\partial_t \widetilde{v} + \partial_x f(\widetilde{v}) = g(\widetilde{v})\partial_x a + \frac{\phi'(z)}{\phi'(\widetilde{v})}\left(f'(\widetilde{v}) - f'(v)\right)\partial_x z\,.$$

Using that $f' > 0$ we get

$$\partial_t|u-\widetilde{v}| + \partial_x|f(u)-f(\widetilde{v})| \le \operatorname{sgn}(u-\widetilde{v})\left(g(u)-g(\widetilde{v})\right)\partial_x a + \left|\frac{\phi'(z)}{\phi'(\widetilde{v})}\left(f'(\widetilde{v}) - f'(v)\right)\partial_x z\right| \tag{A.6}$$

Notice that, if $a'(x)g'(u) \le 0$ for all x, u, then the last term in (A.6) is ≤ 0. In general this is not true; this source contribution is balanced by the weight W_2 (see also Remark 3). About the last term, we notice that

$$|a-b| = |\phi(\widetilde{v}) - \phi(v)| \ge \inf|\phi'||\widetilde{v}-v| = c|\widetilde{v}-v|$$

with $c > 0$, since $\phi' = f'/g$ is bounded away from zero. Hence we deduce the estimate

$$\left|\frac{\phi'(z)}{\phi'(\widetilde{v})}\right||f'(\widetilde{v}) - f'(v)| \le \left|\frac{\phi'(z)}{\phi'(\widetilde{v})}\right|\sup|f''|\,|\widetilde{v}-v| \;\le\; C|a-b|$$

for a suitable constant $C > 0$. This completes the proof of (A.1). □

Proof of (A.3). Recalling that z is constant along v-characteristics, deriving by x the equation $z_t + f'(v)z_x = 0$ and by setting $q = z_x$, we find

$$q_t + [f'(v)q]_x = 0\,.$$

Multiplying by $\operatorname{sgn}(q)$ we formally get $|q|_t + \left[f'(v)|q|\right]_x = 0\,.$

We can now evaluate $\partial_t W_1$:

$$\partial_t W_1(t,x) = \kappa_1 \int_{-\infty}^{x} \partial_t|z_x(t,y)|\,dy = -\kappa_1 f'(v(x,t))|z_x(x,t)|\,.$$

Hence (A.3) is proved. □

Acknowledgments. Thanks are due to Prof. Albert Cohen for having given us an unpublished version of a manuscript [6] about Kuznetsov's method and to François Bouchut for having driven our attention onto this problem.

References

[1] D. Amadori, L. Gosse and G. Guerra, *Global BV entropy solutions and uniqueness for hyperbolic systems of balance laws*, Arch. Ration. Mech. Anal. **162** (2002), 327–366

[2] D. Amadori, L. Gosse and G. Guerra, *Godunov-type approximation for a general resonant balance law with large data*, J. Differential Equations **198** (2004), 233–274

[3] F. Bouchut and B. Perthame, *Kružkov's estimates for scalar conservation laws revisited*, Trans. Amer. Math. Soc. **350** (1998), no. 7, 2847–2870

[4] A. Bressan, **Hyperbolic Systems of Conservation Laws – The one-dimensional Cauchy problem**, Oxford Lecture Series in Mathematics and its Applications **20** (Oxford University Press, Oxford, 2000)

[5] B. Cockburn, *Continuous dependence and error estimation for viscosity methods*, Acta Numerica **12** (2003), 127–180

[6] A. Cohen, W. Dahmen and R. DeVore, *Some Comments on Kutznetsov's Error Estimates*, unpublished manuscript

[7] C.M. Dafermos, *Polygonal approximations of solutions of the initial value problem for a conservation law*, J. Math. Anal. Appl. **38** (1972), 33–41

[8] W. E, *Homogenization of scalar conservation laws with oscillatory forcing terms*, SIAM J. Appl. Math. **52** (1992), 959–972

[9] T. Gallouet, J.-M. Hérard, O. Hurisse and A.-Y. LeRoux, *Well-balanced schemes versus fractional step method for hyperbolic systems with source terms*, Calcolo **43** (2006), 217–251

[10] L. Gosse, *A priori error estimate for a well-balanced scheme designed for inhomogeneous scalar conservation laws*, C.R. Acad. Sc. Paris Série I **327** (1998), 467–472

[11] L. Gosse, *Localization effects and measure source terms in numerical schemes for balance laws*, Math. Comp. **71** (2002), 553–582

[12] L. Gosse, *Time-splitting schemes and measure source terms for a quasilinear relaxing system*, M3AS **13**(8) (2003), 1081–1101

[13] L. Gosse, *Maxwellian decay for well-balanced approximations of a super-characteristic chemotaxis model*, SIAM J. Scient. Comput. **34** (2012), A520–A545

[14] L. Gosse and G. Toscani, *Space localization and well-balanced scheme for discrete kinetic models in diffusive regimes*, SIAM J. Numer. Anal. **41** (2003), 641–658

[15] J. Greenberg and A.-Y. LeRoux, *A well balanced scheme for the numerical processing of source terms in hyperbolic equations*, SIAM J. Numer. Anal. **33** (1996), 1–16

[16] G. Guerra, *Well-posedness for a scalar conservation law with singular nonconservative source*, J. Differential Equations **206** (2004), 438–469

[17] Y. Ha and Y.J. Kim, *Explicit solutions to a convection-reaction equation and defects of numerical schemes*, J. Comp. Phys. **220** (2006), 511–531

[18] C. Helzel, R.J. LeVeque and G. Warnecke, *A modified fractional step method for the accurate approximation of detonation waves*, SIAM J. Sci. Comput. **22** (2000), 1489–1510

[19] H. Holden, K.H. Karlsen, K.-A. Lie and N.H. Risebro, **Splitting Methods for Partial Differential Equations with Rough Solutions. Analysis and MATLAB programs**. EMS Series of Lectures in Mathematics. European Mathematical Society (EMS), Zürich, 2010

[20] H. Holden and N.H. Risebro, **Front Tracking for Hyperbolic Conservation Laws**. Applied Mathematical Sciences **152** (Springer-Verlag, New York, 2002)

[21] E. Isaacson and B. Temple, *Convergence of the* 2×2 *Godunov method for a general resonant nonlinear balance law*, SIAM J. Appl. Math. **55** (1995), 625–640

[22] S. Jin and Y.J. Kim, *On the computation of roll waves*, Math. Mod. Numer. Anal. **35** (2001), 463–480

[23] K.H. Karlsen, N.H. Risebro and J.D. Towers, *Front tracking for scalar balance equations*, J. Hyperbolic Differ. Equ. **1** (2004), 115–148

[24] M.A. Katsoulakis, G. Kossioris and Ch. Makridakis, *Convergence and error estimates of relaxation schemes for multidimensional conservation laws*, Comm. Partial Differential Equations **24** (1999), no. 3-4, 395–424

[25] P. Klingenstein, *Hyperbolic Conservation Laws with Source Terms: Errors of the Shock Location*, Research report 94-07 (1994) ETH Zürich

[26] C. Klingenberg and N.H. Risebro, *Convex conservation laws with discontinuous coefficients. Existence, uniqueness and asymptotic behavior*, Comm. Partial Differential Equations **20** (1995), no. 11-12, 1959–1990

[27] S.N. Kružkov, *First order quasilinear equations in several independant space variables*, Mat. USSR Sbornik **81** (1970), 228–255

[28] N.N. Kuznetsov, *Accuracy of some approximate methods for computing the weak solutions of a first-order quasilinear equation*, Zh. Vychisl. Mat. i Mat. Fiz., 16 (1976), pp. 1489–1502; English transl. in USSR Comp. Math. and Math. Phys., 16 (1976), pp. 105–119

[29] J.O. Langseth, A. Tveito and R. Winther, *On the convergence of operator splitting applied to conservation laws with source terms*, SIAM J. Numer. Anal. **33** (1996), 843–863

[30] W.J. Layton, *Error Estimates for Finite Difference Approximations to Hyperbolic Equations for Large Time*, Proc. Amer. Math. Soc. **92** (1984), 425–431

[31] Ph. LeFloch and A.E. Tzavaras, *Representation of weak limits and definition of nonconservative products*, SIAM J. Math. Anal. **30** (1999), 1309–1342

[32] A.-Y. Le Roux, *A Numerical Conception of Entropy for Quasi-Linear Equations*, Math. Comp. **31** (1977), 848–872

[33] R.J. LeVeque and B. Temple, *Stability of Godunov's method for a class of* 2×2 *systems of conservation laws*, Trans. A.M.S. **288** (1985), 115–123

[34] R.J. LeVeque and H.C. Yee, *A study of numerical methods for hyperbolic conservation laws with stiff source terms*, J. Comp. Phys. **86** (1990), 187–210

[35] T.-P. Liu and T. Yang, *A New Entropy Functional for a Scalar Conservation Law*, Comm. Pure Applied Math. **52** (1999), 1427–1442

[36] B.J. Lucier, *Error bounds for the methods of Glimm, Godunov, and LeVeque*, SIAM J. Numer. Anal., 22 (1985), 1074–1081

[37] C. Mascia and A. Terracina, *Long-time behavior for conservation laws with source in a bounded domain*, J. Differential Equations **159** (1999), 485–514

[38] X. Meng, C.-W. Shu, Q. Zhang and B. Wu, *Superconvergence of discontinuous Galerkin method for scalar nonlinear conservation laws in one space dimension*, SIAM Journal on Numerical Analysis **50** (2012), 2336–2356

[39] R. Natalini. A. Tesei, *Blow-up of solutions for a class of balance laws*, Comm. Part. Diff. Eqns. **19** (1994), 417–453

[40] H. Nessyahu, E. Tadmor and T. Tassa, The convergence rate of Godunov type schemes, SIAM J. Numer. Anal. **31** (1994), 1–16

[41] Elaine S. Oran. Jay P. Boris, **Numerical Simulation of Reactive Flow**, Second Edition, Cambridge Univ. Press (2005)

[42] F. Peyroutet, *Splitting Method Applied to Hyperbolic Problem with Source Term*, Appl. Math. Lett. **14** (2001), 99–104

[43] F. Sabac, *The optimal convergence rate of monotone finite difference methods for hyperbolic conservation laws*, SIAM J. Numer. Anal. **34** (1997), 2306–2318

[44] T. Tang and Z.-H. Teng, *The sharpness of Kuznetsov's* $O(\sqrt{\Delta x})$ L^1*-error estimate for monotone difference schemes*, Math. Comp. **64** (1995), 581–589

[45] T. Tang and Z.-H. Teng, *Error bounds for fractional step methods for conservation laws with source terms*, SIAM J. Numer. Anal. **32** (1995), 110–127

[46] E. Tadmor, *Numerical viscosity and the entropy condition for conservative difference schemes*, Math. Comp. **43** (1984), 369–381

[47] W.-C. Wang, *On* L^1 *convergence rate of viscous and numerical approximate solutions of genuinely nonlinear scalar conservation laws*, SIAM J. Math. Anal. **30** (1998), 38–52

[48] W. Wang, C.-W. Shu, H. C. Yee and B. Sjögreen, *High order finite difference methods with subcell resolution for advection equations with stiff source terms*, J. Comp. Phys. **231**(1) (2012), 190–214