Arch. Rational Mech. Anal. 241 (2021) 357–402
Digital Object Identifier (DOI) https://doi.org/10.1007/s00205-021-01653-4

A Posteriori Error Estimates for Numerical Solutions to Hyperbolic Conservation Laws

Alberto Bressan, Maria Teresa Chiri & Wen Shen

Communicated by Constantine Dafermos

Abstract

This paper is concerned with a posteriori error bounds for a wide class of numerical schemes, for $n \times n$ hyperbolic conservation laws in one space dimension. These estimates are achieved by a "post-processing algorithm", checking that the numerical solution retains small total variation, and computing its oscillation on suitable subdomains. The results apply, in particular, to solutions obtained by the Godunov or the Lax–Friedrichs scheme, backward Euler approximations, and the method of periodic smoothing. Some numerical implementations are presented.

1. Introduction

Consider the Cauchy problem for a strictly hyperbolic system of conservation laws in one space dimension:

$$u_t + f(u)_x = 0, \tag{1.1}$$

$$u(0, x) = \bar{u}(x). \tag{1.2}$$

For initial data with small total variation, it is well known that this problem has a unique entropy-weak solution, depending Lipschitz continuously on the initial data $\bar{u}$ in the $\mathbf{L}^1$ norm [8,9,22,26].

A closely related question is the stability and convergence of various types of approximate solutions. Estimates on the convergence rate for a deterministic version of the Glimm scheme [24,28] were derived in [18], and more recently in [1,6] for a wider class of flux functions. For vanishing viscosity approximations

$$u_t + f(u)_x = \varepsilon\, u_{xx}, \tag{1.3}$$

uniform BV bounds, stability and convergence as $\varepsilon \to 0$ were proved in [5], while convergence rates were later established in [13,19]. Further convergence results

were proved by Bianchini for approximate solutions constructed by the semidiscrete (upwind) Godunov scheme [3], and by the Jin-Xin relaxation model [4].

A major remaining open problem is the convergence of fully discrete approximations, such as the Lax–Friedrichs or the Godunov scheme [25–27]. Indeed, the convergence results known for these numerical algorithms rely on compensated compactness [23]. They apply only to 2×2 systems, and do not yield information about uniqueness or convergence rates.

For a particular class of systems, the convergence of Godunov approximations was proved in [14], relying on uniform bounds on the total variation. For general hyperbolic systems, however, it is known that the Godunov scheme is unstable with respect to the BV norm. In [2] an example was constructed, showing that the total variation of a numerical solution can become arbitrarily large as $t \to +\infty$. Indeed, if the exact solution contains a shock with speed close to a rational multiple of the grid size $\Delta x/\Delta t$, this can cause resonances, producing a large amount of downstream oscillations.

Without an *a priori* bound on the total variation, one cannot compare an approximate solution with trajectories of the semigroup of exact solutions, and all the uniqueness arguments developed in [12,15,16] break down. The counterexample in [2] can thus be regarded as a fundamental obstruction toward the derivation of *a priori* error estimates for fully discrete numerical schemes.

To make progress, in this paper we shift our point of view, focusing on a posteriori error estimates. Namely, we assume that an approximate solution to (1.1), (1.2) has been constructed by some numerical algorithm. Based on some additional information about the approximate solution, we seek an estimate on the difference

$$\|u^{\text{approx}}(T,\cdot) - u^{\text{exact}}(T,\cdot)\|_{\mathbf{L}^1(\mathbb{R})}. \tag{1.4}$$

For any sufficiently small BV initial data $\bar{u}$, it is well known that the unique entropy-admissible BV solution of (1.1), (1.2) has two key properties [8]:

(i) The total variation of $u(t,\cdot)$ remains uniformly small, for all $t \geqq 0$.
(ii) Given a threshold $\rho > 0$, one can identify a finite number of curves in the t-x plane (shocks or contact discontinuities) such that, outside these curves, the solution has local oscillation $< \rho$.

The counterexample in [2] shows that, for an approximation constructed by the Godunov scheme, the property (i) sometimes can fail. Roughly speaking, the result we want to prove in the present paper is the following: let u^{approx} be an approximate solution produced by a conservative scheme which dissipates entropy, and assume that

(i') The total variation of $u^{\text{approx}}(t,\cdot)$ remains small, for all $t \in [0,T]$;
(ii') Outside a finite number of narrow strips in the domain $[0,T] \times \mathbb{R}$, the local oscillation of u^{approx} remains small.

Then the $\mathbf{L}^1$ distance (1.4) is small.

We emphasize that both conditions (i′), (ii′) refer to the output of a numerical computation. In (ii′), we expect that the finitely many strips where the oscillation of u^{approx} is large will have the form

$$\Big\{(t,x);\quad t\in[a_i,b_i],\ x\in[\gamma_i(t)-\delta,\ \gamma_i(t)+\delta]\Big\},$$

where the curve $t\mapsto\gamma_i(t)$ traces the approximate location of a large shock (or a contact discontinuity) in the exact solution. It is also worth noting that our estimates do not require any regularity of the exact solution. In particular, u^{exact} may well have a dense set of discontinuities.

Our goal is to prove error bounds which can be applied to a wide class of approximation schemes. For future reference, we collect the basic assumptions on the system (1.1), and the properties of the approximate solutions that will be used.

(A1) The system (1.1) is strictly hyperbolic, with each characteristic field being either linearly degenerate or genuinely nonlinear. It generates a semigroup of entropy weak solutions $S:\mathcal{D}\times[0,+\infty[\mapsto\mathcal{D}$, where $\mathcal{D}\subset\mathbf{L}^1(\mathbb{R};\,\mathbb{R}^n)$ is a domain containing all functions with sufficiently small total variation, namely

$$\bar u\in\mathbf{L}^1(\mathbb{R};\,\mathbb{R}^n),\quad \text{Tot.Var.}\{\bar u\}\leqq\delta_0\quad\Longrightarrow\quad\bar u\in\mathcal{D}. \tag{1.5}$$

There exist Lipschitz constants C_0, L_0 such that

$$\|S_tu-S_su\|_{\mathbf{L}^1}\leqq C_0\cdot\text{Tot.Var.}\{u\}\cdot|t-s|, \tag{1.6}$$

$$\|S_tu-S_tv\|_{\mathbf{L}^1}\leqq L_0\|u-v\|_{\mathbf{L}^1} \tag{1.7}$$

for all $u,v\in\mathcal{D}$ and $0\leqq s\leqq t$.

(A2) For each genuinely nonlinear field, there exists a strictly convex entropy η, with entropy flux q, which selects the admissible shocks.

We recall that the existence of a semigroup generated by (1.1) was proved in [5,10,11,17], in various degrees of generality. In particular, it is known that the trajectories of the semigroup are the unique limits of vanishing viscosity approximations. To explain the additional assumption **(A2)**, let u^-, u^+ be any two states connected by a genuinely nonlinear shock with speed λ, so that following the Rankine–Hugoniot conditions hold:

$$\lambda\,(u^+-u^-)=f(u^+)-f(u^-).$$

Then, if the shock is NOT admissible, we require that the corresponding entropy should be strictly increasing, namely

$$\lambda\big(\eta(u^-)-\eta(u^+)\big)-\big(q(u^-)-q(u^+)\big)>c_0|u^--u^+|^3, \tag{1.8}$$

for some constant $c_0>0$.

In what follows, we shall use test functions $\varphi\in\mathcal{C}_c(\mathbb{R}^2)$ which are Lipschitz continuous with compact support, with Sobolev norm

$$\|\varphi\|_{W^{1,\infty}}\doteq\max\big\{\|\varphi\|_{\mathbf{L}^\infty},\ \|\varphi_t\|_{\mathbf{L}^\infty},\ \|\varphi_x\|_{\mathbf{L}^\infty}\big\}. \tag{1.9}$$

Given $\varepsilon>0$, we consider approximate solutions $u:[0,T]\mapsto\mathcal{D}$ of the system of conservation laws (1.1), taking values inside the domain of the semigroup S. We assume that these solutions are inductively defined for a discrete set of times $\tau_j=j\varepsilon$, $j=0,1,2\ldots$ For $t\in[\tau_j,\tau_{j+1}[$ one can then define $u(t,\cdot)$ to be the exact solution

to (1.1) which coincides with $u(\tau_j, \cdot)$ at time $t = \tau_j$. In alternative, sometimes it is more convenient to simply define $u(t, \cdot) = u(\tau_j, \cdot)$ for $t \in [\tau_j, \tau_{j+1}[$.

While we do not specify any particular method to construct these approximate solutions, two basic properties will be assumed. The first is the Lipschitz continuity of the map $t \mapsto u(t, \cdot) \in \mathbf{L}^1(\mathbb{R}; \mathbb{R}^n)$, restricted to the discrete set of times τ_j. The second is an approximate weak form of the conservation equations and the entropy conditions. In what follows, L, C denote suitable constants. Moreover, the notation $\varepsilon\mathbb{N} \doteq \{j\varepsilon;\ j = 0, 1, 2, \ldots\}$ will be used.

(AL) *For every* $0 \leqq \tau < \tau' \leqq T$ *with* $\tau, \tau' \in \varepsilon\mathbb{N}$, *one has*

$$\|u(\tau', \cdot) - u(\tau, \cdot)\|_{\mathbf{L}^1} \leqq L\,(\tau' - \tau) \cdot \sup_{t\in[\tau,\tau']} \text{Tot.Var.}\big\{u(t, \cdot)\big\}. \tag{1.10}$$

(P_ε) For every $0 \leqq \tau < \tau' \leqq T$ with $\tau, \tau' \in \varepsilon\mathbb{N}$, and every test function $\varphi \in \mathcal{C}^1_c(\mathbb{R}^2)$, one has

$$\begin{aligned} &\left| \int u(\tau, x)\varphi(\tau, x)\,dx - \int u(\tau', x)\varphi(\tau, x)\,\mathrm{d}x \right.\\ &\qquad \left. + \int_\tau^{\tau'} \int \big\{u\varphi_t + f(u)\varphi_x\big\}\,\mathrm{d}x\,\mathrm{d}t \right| \\ &\qquad \leqq C\varepsilon\|\varphi\|_{W^{1,\infty}} \cdot (\tau' - \tau) \cdot \sup_{t\in[\tau,\tau']} \text{Tot.Var.}\big\{u(t, \cdot)\big\}. \end{aligned} \tag{1.11}$$

Moreover, assuming $\varphi \geqq 0$, one has the entropy inequality

$$\begin{aligned} &\int \eta(u(\tau, x))\varphi(\tau, x)\,\mathrm{d}x - \int \eta(u(\tau', x))\varphi(\tau', x)\,\mathrm{d}x \\ &\quad + \int_\tau^{\tau'} \int \big\{\eta(u)\varphi_t + q(u)\varphi_x\big\}\,\mathrm{d}x\,\mathrm{d}t \\ &\quad \geqq -C\varepsilon\|\varphi\|_{W^{1,\infty}} \cdot (\tau' - \tau) \cdot \sup_{t\in[\tau,\tau']} \text{Tot.Var.}\big\{u(t, \cdot)\big\}. \end{aligned} \tag{1.12}$$

We remark that, for an exact solution, the left hand side of (1.11) would be zero, while the left hand side of (1.12) would be non-negative. Since here we are dealing with ε-approximate solutions, we allow an error that decreases with ε, but increases with the Lipschitz constant of the test function φ.

In the present paper, two main questions will be addressed:

- Given an approximate solution u of (1.1), (1.2) satisfying **(AL)** and ($\mathbf{P}_\varepsilon$), can one estimate the distance between u and the exact solution?
- What kind of approximation schemes satisfy the conditions **(AL)** and (P_ε)?

To answer the first question, using a technique introduced in [7], two types of estimates will be derived:

- On regions where the oscillation is small, the approximate solution u is compared with the solution to a linear hyperbolic problem with constant coefficients.
- Near a point where a large jump occurs, u is compared with the solution to a Riemann problem.

We recall that, for exact solutions, this technique yields the identity $u(t,\cdot) = S_t\bar{u}$, proving that an entropy weak solution is unique and coincides with the corresponding semigroup trajectory [7,8,12,15,16]. In Sections 2 to 4 we develop similar estimates in the case of an approximate solution u, where the right hand side of (1.11), (1.12) is not zero, but vanishes of order $\mathcal{O}(1)\cdot\varepsilon\|\varphi\|_{W^{1,\infty}}$. This will provide a bound on the difference (1.4).

An important aspect must be mentioned here. The uniqueness proofs in [12, 15,16] require some additional regularity condition, such as "Tame Variation" or "Tame Oscillation". These conditions are always satisfied by solutions constructed by front tracking or by the Glimm scheme, but may fail for a numerically approximated solution. To derive rigorous error bounds, we must check that an equivalent condition is satisfied.

For a numerically computed approximation, in Section 5 we introduce a *post-processing algorithm*, which accomplishes three main tasks:

(1) Check that the total variation remains bounded.
(2) Trace the location of a finite number of large shocks.
(3) Check that the oscillation of the solution remains small, on a finite number of polygonal domains, away from the large shocks.

Step (1) is the simplest, yet the crucial one. If the total variation becomes too large, at some time t the approximate solution $u(t,\cdot)$ will fall outside the domain $\mathcal{D}$ of the semigroup. When this happens, the algorithm stops and no error estimate is achieved.

In the favorable case where the total variation remains small, the algorithm can then proceed with steps (2) and (3). To implement these steps, one needs to introduce certain parameters, such as the minimum size of the shocks which will be traced, and the length of the time intervals $[t_j, t_{j+1}]$ used in a new partition of $[0, T]$. For every choice of these parameter values, the algorithm yields an error bound. In practice, the accuracy of this estimate largely depends on the choice of these values. At the end of Section 3, and then again at the end of Section 5, we discuss how to choose these parameters, and the expected order of magnitude of the corresponding error bounds.

To complete our program, in Section 6 we consider various approximation schemes, and prove that they all satisfy the properties **(AL)** and (P_ε). In particular, our analysis applies to: (i) Godunov's scheme, (ii) the Lax–Friedrichs' scheme, (iii) backward Euler approximations, and (iv) approximate solutions obtained by periodic mollifications.

Finally, in Section 7, we discuss details of the post-processing algorithm, and present a numerical simulation. For the "p-system", describing isentropic gas dynamics in Lagrangian coordinates, we consider initial data generating two centered rarefactions, and two shocks that eventually cross each other. After computing an approximate solution by the upwind scheme, we implement the post-processing algorithm. The two shocks are traced (as long as they remain well separated), and the remaining domain is covered by trapezoids where the numerical solution has small oscillation (away from interaction times).

2. Solutions with Small Oscillation

In this section we begin by studying the case where no large shocks are present. Let $u = u(t, x)$ be an approximate solution which satisfies **(AL)** and $(\mathbf{P}_\varepsilon)$. Consider an open interval $]a, b[$, fix a point ξ with $a < \xi < b$ and set

$$A = Df(u(0, \xi)). \tag{2.1}$$

Assuming that all characteristic speeds satisfy

$$\lambda^- < \lambda_i(u) < \lambda^+, \qquad i = 1, \ldots, n, \tag{2.2}$$

fix $\tau \in \varepsilon\mathbb{N}$ and consider the trapezoidal domain

$$\Delta = \Big\{(t, x);\ t \in [0, \tau],\ \ a(t) \doteq a + \lambda^+ t < x < b + \lambda^- t \doteq b(t)\Big\}. \tag{2.3}$$

Following an approach introduced in [7], error estimates will be obtained by comparing u with the solution w of the linear hyperbolic system with constant coefficients

$$w_t + Aw_x = 0, \qquad w(0, x) = u(0, x). \tag{2.4}$$

For this purpose, let $\{\ell_1, \ldots, \ell_n\}$ and $\{r_1, \ldots, r_n\}$ be dual bases of left and right eigenvectors of the matrix A, normalized so that

$$|r_i| = 1, \qquad \ell_i \cdot r_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{2.5}$$

Let $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues of A. For each i, consider the scalar functions

$$u_i(t, x) = \ell_i \cdot u(t, x), \qquad w_i(t, x) = \ell_i \cdot w(t, x).$$

By (2.4), w_i solves the scalar linear equation

$$w_{i,t} + \lambda_i w_{i,x} = 0, \qquad w_i(0, x) = u_i(0, x).$$

For each $i = 1, \ldots, n$, we will estimate the difference $u_i(\tau, \cdot) - w_i(\tau, \cdot)$.

As a preliminary to this, consider a BV function $g : [\alpha, \beta] \mapsto \mathbb{R}$. Since g is regulated, it admits left and right limits $g(x-)$, $g(x+)$ at every point x. By possibly modifying g on the countable set where it has jumps, we can assume that

$$g(x+) \cdot g(x-) \leqq 0 \qquad \Longrightarrow \qquad g(x) = 0. \tag{2.6}$$

We can then select countably many maximal open subintervals $]a_j, b_j[\subset [\alpha, \beta]$ where g has constant sign. Namely,

(G) g has constant sign on each $]a_j, b_j[$, and changes sign on every neighborhood of each endpoint a_j, b_j (unless $a_j = \alpha$ or $b_j = \beta$). Moreover, $g(x) = 0$ for $x \notin \bigcup_j [a_j, b_j]$ (Fig. 1).

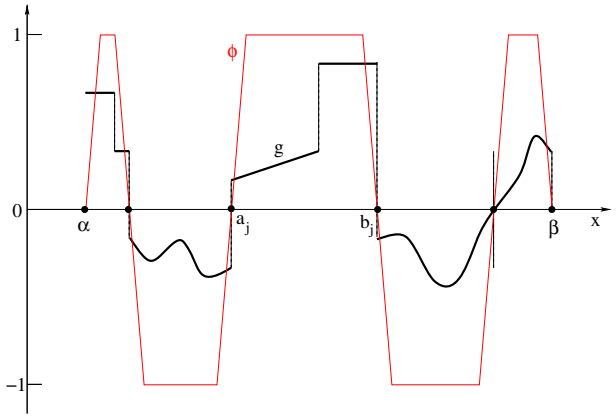

Fig. 1. The test function ϕ defined at (2.7), with Lipschitz constant $\varepsilon^{-2/3}$

For a given $\varepsilon > 0$, consider the test function with Lipschitz constant $\|\phi_x\|_{\mathbf{L}^\infty} = \varepsilon^{-2/3}$

$$\phi(x) \doteq \begin{cases} \min\left\{1,\ \dfrac{x-a_j}{\varepsilon^{2/3}},\ \dfrac{b_j-x}{\varepsilon^{2/3}}\right\}, & \text{if } x \in [a_j, b_j] \text{and} g \text{is positive on }]a_j, b_j[, \\ \max\left\{-1,\ \dfrac{a_j-x}{\varepsilon^{2/3}},\ \dfrac{x-b_j}{\varepsilon^{2/3}}\right\}, & \text{if } x \in [a_j, b_j] \text{ and } g \text{ is negative on }]a_j, b_j[, \\ 0, & \text{if } x \notin \cup_j [a_j, b_j]. \end{cases} \tag{2.7}$$

Lemma 1. *Let $g : [\alpha, \beta] \mapsto \mathbb{R}$ be as above. If g is strictly positive (or strictly negative) for all $x \in [\alpha, \beta]$, then*

$$\int_{\alpha+\varepsilon^{2/3}}^{\beta-\varepsilon^{2/3}} |g(x)|\,\mathrm{d}x \ \leqq\ \int_\alpha^\beta \phi(x)\, g(x)\,\mathrm{d}x. \tag{2.8}$$

On the other hand, if $g(\xi) = 0$ for some $\xi \in [\alpha, \beta]$, then

$$\int_\alpha^\beta |g(x)|\,\mathrm{d}x \ \leqq\ \int_\alpha^\beta \phi(x)\, g(x)\,\mathrm{d}x + 2\varepsilon^{2/3} \cdot \mathrm{Tot.Var.}\big\{g;\ [\alpha, \beta]\big\}. \tag{2.9}$$

Proof. 1. If g has always the same sign, then by construction $\phi(x)\, g(x) \geqq 0$ for all x, while $\phi(x) = \mathrm{sign}(g(x))$ for $x \in [\alpha + \varepsilon^{2/3},\ \beta - \varepsilon^{2/3}]$. Hence the estimate (2.8) is trivially true.

2. If g changes sign, consider the maximal subintervals $[a_j, b_j]$ where g has a constant sign, as in **(G)**. We then have the estimate

$$\begin{aligned} \int_\alpha^\beta |g(x)|\,\mathrm{d}x &= \sum_j \int_{a_j}^{b_j} |g(x)|\,\mathrm{d}x \\ &\leqq \sum_{b_j - a_j \leqq 2\varepsilon^{2/3}} (b_j - a_j) \sup_{a_j < x < b_j} |g(x)| \end{aligned}$$

$$
\begin{aligned}
&+\sum_{b_j-a_j\geqq 2\varepsilon^{2/3}}\left(\int_{a_j+\varepsilon^{2/3}}^{b_j-\varepsilon^{2/3}}+\int_{a_j}^{a_j+\varepsilon^{2/3}}+\int_{b_j-\varepsilon^{2/3}}^{b_j}\right)|g(x)|\,\mathrm{d}x\\
&\leqq\ 2\varepsilon^{2/3}\cdot\sum_{b_j-a_j\leqq 2\varepsilon^{2/3}}\text{Tot.Var.}\{g;\ ]a_j,b_j[\}+\int_\alpha^\beta\phi(x)\,g(x)\,\mathrm{d}x\\
&\quad+2\varepsilon^{2/3}\cdot\sum_{b_j-a_j>2\varepsilon^{2/3}}\ \sup_{a_j<x<b_j}|g(x)|\\
&=\ \int_\alpha^\beta\phi(x)\,g(x)\,\mathrm{d}x+2\varepsilon^{2/3}\cdot\text{Tot.Var.}\big\{g;\ [\alpha,\beta]\big\}.
\end{aligned}
$$

□

Remark 1. Here and in the sequel, one could prove similar results by replacing the exponent $2/3$ with any number $\gamma\in\,]0,1[$, and working with test functions which are Lipschitz continuous with constant $\varepsilon^{-\gamma}$. Our choice of $\gamma=2/3$ is motivated by the heuristic expectation that, in most cases, this should yield the sharpest error bounds. See Remark 2 for further discussion of this point.

We can now state the main result of this section, providing an error estimate on the trapezoidal domain (2.3).

Lemma 2. *There exists a constant C_1 such that the following holds. For a given $\varepsilon>0$, let u be an approximate solution of (1.1) that satisfies the property ($\mathbf{P}_\varepsilon$). Let Δ be the trapezoid in (2.3), and let w be the solution to the linear Cauchy problem (2.4), with A as in (2.1). Then*

$$
\begin{aligned}
&\int_{a+\tau\lambda^+ +\varepsilon^{2/3}}^{b+\tau\lambda^- -\varepsilon^{2/3}}\big|u(\tau,x)-w(\tau,x)\big|\,\mathrm{d}x\\
&\leqq\ C_1\left(\tau\cdot\sup_{(t,x)\in\Delta}\big|u(t,x)-u(0,\xi)\big|+\tau\varepsilon^{1/3}+\varepsilon^{2/3}\right)\\
&\quad\cdot\sup_{t\in[0,\tau]}\text{Tot.Var.}\big\{u(t,\cdot)\big\}.
\end{aligned}
\tag{2.10}
$$

Proof. 1. Fix $i\in\{1,\ldots,n\}$. On the interval $[\alpha,\beta]\doteq[a+\tau\lambda^+,\ b+\tau\lambda^-]$, consider the scalar function

$$
g_i(x)\doteq\ell_i\cdot\big[u(\tau,x)-w(\tau,x)\big]=\ell_i\cdot\big[u(\tau,x)-u(0,x-\lambda_i\tau)\big]. \tag{2.11}
$$

Let $\phi_i:[\alpha,\beta]\mapsto[-1,1]$ be the function with Lipschitz constant $\|\phi_i'\|_{\mathbf{L}^\infty}=\varepsilon^{-2/3}$, defined as in (2.7) with g replaced by g_i. We then extend ϕ_i to the entire real line by setting $\phi_i(x)=0$ if $x\notin[\alpha,\beta]$, and consider a test function $\varphi_i=\varphi_i(t,x)$ such that

$$
\varphi_i(t,x)=\phi_i\big(x-\lambda_i(t-\tau)\big)\,\ell_i\qquad\text{for}\quad t\in[0,\tau],\ x\in\mathbb{R}. \tag{2.12}
$$

2. Observing that $\lambda_i \in [0,1]$ and $\|\varphi_i\|_{W^{1,\infty}} = |\ell_i|\,\varepsilon^{-2/3}$, by (1.11) we now obtain

$$\begin{aligned}
&\left| \int \varphi_i(0,x)u(0,x)\,\mathrm{d}x - \int \varphi_i(\tau,x)u(\tau,x)\,\mathrm{d}x \right.\\
&\qquad \left. + \int_0^\tau \int \big\{u\varphi_{i,t} + f(u)\varphi_{i,x}\big\}\,\mathrm{d}x\,\mathrm{d}t \right|\\
&\qquad \leqq C\varepsilon \|\varphi_i\|_{W^{1,\infty}} \cdot \tau \cdot \sup_{t\in[0,\tau]} \mathrm{Tot.Var.}\big\{u(t,\cdot)\big\}\\
&\qquad \leqq C'\varepsilon^{1/3}\,\tau \cdot \sup_{t\in[0,\tau]} \mathrm{Tot.Var.}\big\{u(t,\cdot)\big\}. \qquad (2.13)
\end{aligned}$$

3. For future use we observe that, if $x \mapsto u(x)$ is Lipschitz and $u^* = u(\xi)$ for some $\xi \in [x_1, x_2]$, then

$$\begin{aligned}
\int_{x_1}^{x_2} \Big|\ell_i\big(f(u)_x - \lambda_i u_x\big)\Big|\,\mathrm{d}x &= \int_{x_1}^{x_2} \Big|\ell_i\big[Df(u) - Df(u^*)\big]u_x\Big|\,\mathrm{d}x\\
&\leqq C_0 \sup_{x_1<x<x_2} |u(x)-u^*| \cdot \int_{x_1}^{x_2} |u_x|\,\mathrm{d}x.
\end{aligned}$$

Here C_0 is a constant depending only on the function f. By an approximation argument, for any BV function $x \mapsto u(x)$ we conclude that

$$\begin{aligned}
&\mathrm{Tot.Var.}\Big\{\ell_i\big(f(u) - \lambda_i u\big);\ [x_1,x_2]\Big\}\\
&\leqq C_0 \left(\sup_{x_1<x<x_2} |u(x)-u^*|\right) \cdot \mathrm{Tot.Var.}\big\{u;\ [x_1,x_2]\big\}. \qquad (2.14)
\end{aligned}$$

4. Since w is a solution to the linear equation (2.4), the choice of the test function φ_i in (2.12) implies

$$\begin{aligned}
\int \varphi_i(0,x)\,u(0,x)\,\mathrm{d}x &= \int \varphi_i(0,x)\,w(0,x)\,\mathrm{d}x\\
&= \int \varphi_i(\tau,x)\,w(\tau,x)\,\mathrm{d}x. \qquad (2.15)
\end{aligned}$$

Moreover, calling $u^* = u(0,\xi)$, integrating by parts and using (2.14) together with the bound $\phi_i(x) \in [-1,1]$, we obtain

$$\begin{aligned}
&\left|\int_0^\tau \int \big\{u\varphi_{i,t} + f(u)\varphi_{i,x}\big\}\,\mathrm{d}x\,\mathrm{d}t\right|\\
&= \left|\int_0^\tau \int \big[f(u) - \lambda_i u\big]\varphi_{i,x}\,\mathrm{d}x\,\mathrm{d}t\right|\\
&\leqq \int_0^\tau \mathrm{Tot.Var.}\Big\{\ell_i(f(u)-\lambda_i u);\ [a+\lambda^+\tau + (t-\tau)\lambda_i,\ b+\lambda^-\tau+(t-\tau)\lambda_i]\Big\}\,\|\varphi_i\|_{\mathbf{L}^\infty}\,\mathrm{d}t\\
&\leqq C_0 \sup_{(t,x)\in\Delta} |u(t,x)-u^*| \cdot \int_0^\tau \mathrm{Tot.Var.}\Big\{u(t,\cdot);\ ]a+t\lambda^+,\ b+t\lambda^-[\Big\}\,\mathrm{d}t. \qquad (2.16)
\end{aligned}$$

5. By (2.15), combining (2.13) with (2.16) we conclude that

$$\begin{aligned}\int_{a+\tau\lambda^+}^{b+\tau\lambda^-} &\phi_i(x)\,\ell_i\big[w(\tau,x)-u(\tau,x)\big]\,\mathrm{d}x\\ &=\int\varphi_i(0,x)u(0,x)\,dx-\int\varphi_i(\tau,x)u(\tau,x)\,\mathrm{d}x\\ &\leqq\left|\int_0^\tau\int\big\{u\varphi_{i,t}+f(u)\varphi_{i,x}\big\}\,\mathrm{d}x\,\mathrm{d}t\right|+\mathcal{E}_i,\end{aligned}\tag{2.17}$$

where

$$\begin{aligned}\mathcal{E}_i\doteq&\left|\int\varphi_i(0,x)u(0,x)\,\mathrm{d}x-\int\varphi_i(\tau,x)u(\tau,x)\,\mathrm{d}x\right.\\&\left.+\int_0^\tau\int\big\{u\varphi_{i,t}+f(u)\varphi_{i,x}\big\}\,\mathrm{d}x\,\mathrm{d}t\right|\\&\leqq C\varepsilon\|\varphi\|_{W^{1,\infty}}\cdot\tau\cdot\sup_{t\in[0,\tau]}\mathrm{Tot.Var.}\big\{u(t,\cdot)\big\}.\end{aligned}\tag{2.18}$$

Notice that the above inequality follows from (1.11). In addition, the first term on the right hand side of (2.17) is estimated by (2.16).

6. If the function $g_i(x)\doteq\ell_i\big[w(\tau,x)-u(\tau,x)\big]$ always keeps the same sign, we now use (2.8). If it changes sign at least once, we use (2.9). Combining the two cases, by (2.16) and (2.17) we deduce that

$$\begin{aligned}&\int_{a+\tau\lambda^+-\varepsilon^{2/3}}^{b+\tau\lambda^-}\Big|\ell_i\big[w(\tau,x)-u(\tau,x)\big]\Big|\,\mathrm{d}x\\&\leqq C_0\sup_{(t,x)\in\Delta}|u(t,x)-u^*|\cdot\int_0^\tau\mathrm{Tot.Var.}\Big\{u(t,\cdot);\ \ ]a+t\lambda^+,\ b+t\lambda^-[\Big\}\,\mathrm{d}t+\mathcal{E}_i\\&\quad+2\varepsilon^{2/3}\cdot\mathrm{Tot.Var.}\Big\{g_i;\ \ ]a+\tau\lambda^+,b+\tau\lambda^-[\Big\}.\end{aligned}\tag{2.19}$$

7. Recalling (2.5), for any vector $v=\sum_i c_i r_i\in\mathbb{R}^n$ one has

$$|v|\ \leqq\ \sum_{i=1}^n|c_i|=\sum_{i=1}^n|\ell_i\cdot v|.$$

We use this inequality with $v=w(\tau,x)-u(\tau,x)$. Using (2.11) to compute the total variation of g_i, summing the inequalities (2.18), (2.19) for $i=1,\ldots,n$, we obtain

$$\begin{aligned}&\int_{a+\tau\lambda^+-\varepsilon^{1/3}}^{b+\tau\lambda^-}\Big|w(\tau,x)-u(\tau,x)\Big|\,\mathrm{d}x\\&\leqq nC_0\sup_{(t,x)\in\Delta}|u(t,x)-u^*|\cdot\tau\cdot\sup_{t\in[0,\tau]}\mathrm{Tot.Var.}\Big\{u(t,\cdot);\ \ ]a+t\lambda^+,\ b+t\lambda^-[\Big\}\\&\quad+nC\varepsilon\|\varphi\|_{W^{1,\infty}}\cdot\tau\cdot\sup_{t\in[0,\tau]}\mathrm{Tot.Var.}\big\{u(t,\cdot)\big\}+2\varepsilon^{2/3}\left(\sum_i|\ell_i|\right)\end{aligned}$$

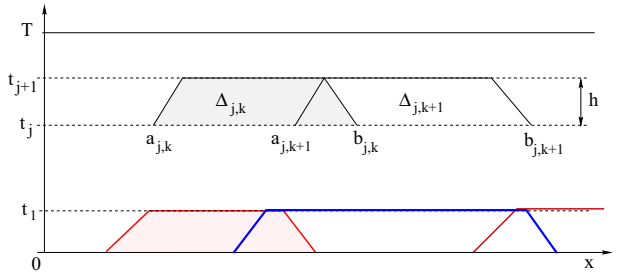

Fig. 2. Covering the strip $[0, T] \times \mathbb{R}$ with finitely many trapezoids Δ_{jk}

$$\cdot\Big(\text{Tot.Var.}\big\{u(0,\cdot);\ [a,b]\big\} + \text{Tot.Var.}\big\{u(\tau,\cdot);\ [a+\tau\lambda^+,\ b+\tau\lambda^-]\big\}\Big).$$

This yields (2.10), for a suitable constant C_1. □

3. Error Bounds for Solutions Without Large Shocks

Consider an approximate solution $u = u(t, x)$ of (1.1), (1.2), constructed by a numerical algorithm with time step $\varepsilon > 0$, which satisfies the properties **(AL)** and $(\mathbf{P}_\varepsilon)$. We fix a new time step $h >> \varepsilon$, and split the interval $[0, T]$ into subintervals $[t_j, t_{j+1}]$ with $t_j = j\,h$. Throughout what follows we choose $h \approx \varepsilon^{1/3}$, say

$$c_0\varepsilon^{1/3} \ \leqq\ h\ \leqq\ \varepsilon^{1/3} \tag{3.1}$$

for some constant $c_0 > 0$, and assume that both h and T are integer multiples of ε. To simplify the discussion, we also assume that $T = \nu h$ for some integer ν. Notice that, in the general case, one can consider the time T' such that

$$T' = \nu h\ \leqq\ T\ <\ (\nu+1)h$$

for some integer ν. By (1.10), the difference can then be estimated by

$$\|u(T,\cdot) - u(T',\cdot)\|_{\mathbf{L}^1}\ \leqq\ Lh\cdot \sup_{t\in[0,T]} \text{Tot.Var.}\{u(t,\cdot)\} = \mathcal{O}(1)\cdot\varepsilon^{1/3}.$$

As shown in Fig. 2, for any given $j = 1, 2, \ldots, \nu - 1$, we cover the real line with finitely many intervals $]a_{jk}, b_{jk}[\,, k = 1, \ldots, N(j)$, so that

$$-\infty = a_{j,1} < a_{j,2} < b_{j,1} < a_{j,3} < \cdots < a_{j,N(j)} < b_{j,N(j)-1} < b_{j,N(j)} = +\infty.$$

We then cover each strip $[t_j, t_{j+1}] \times \mathbb{R}$ with the trapezoids $\Delta_{jk}, k = 1, \ldots, N(j)$. For convenience, these will be expressed as the convex closure of their four vertices:

$$\Delta_{jk} = \text{co}\Big\{(t_j, a_{jk}),\ \ (t_j, b_{jk}),\ \ (t_{j+1}, a_{jk} + h\lambda^+ + \varepsilon^{2/3}),$$
$$(t_{j+1}, b_{jk} + h\lambda^- - \varepsilon^{2/3})\Big\}.$$

Equivalently:

$$\Delta_{jk} = \Big\{(t,x);\ t \in [t_j, t_{j+1}],$$
$$\frac{t_{j+1}-t}{h}a_{jk} + \frac{t-t_j}{h}(a_{jk} + h\lambda^+ + \varepsilon^{2/3}) \leqq x$$
$$\leqq \frac{t_{j+1}-t}{h}b_{jk} + \frac{t-t_j}{h}(b_{jk} + h\lambda^- - \varepsilon^{2/3})\Big\}. \tag{3.2}$$

By suitably choosing the points a_{jk}, b_{jk}, we can assume that the intervals

$$J'_{jk} = [a_{jk} + h\lambda^+ + \varepsilon^{2/3},\ b_{jk} + h\lambda^- - \varepsilon^{2/3}], \qquad k = 1, \ldots, N(j)$$

form a partition of $\mathbb{R}$. Namely

$$b_{jk} + h\lambda^- - \varepsilon^{2/3} = a_{j,k+1} + h\lambda^+ + \varepsilon^{2/3}, \qquad k = 1, \ldots, N(j)-1. \tag{3.3}$$

Furthermore, by choosing the bases of all trapezoids to have length

$$b_{jk} - a_{jk} \ > \ 2h(\lambda^+ - \lambda^-), \tag{3.4}$$

we can assume that each point $(t,x) \in [t_j, t_{j+1}] \times \mathbb{R}$ is contained in at least one and in not more than two of these trapezoids.

Next, we recall that the oscillation of u over a set Δ is defined as

$$\text{Osc.}\big\{u;\ \Delta\big\} \doteq \sup_{(t,x),(s,y)\in\Delta} |u(t,x) - u(t,y)|.$$

For each fixed $j \in \{1, \ldots, \nu\}$, the maximum oscillation of u over all trapezoids Δ_{jk} will be denoted by

$$\kappa_j \doteq \max_{1\leqq k\leqq N(j)} \text{Osc.}\big\{u;\ \Delta_{jk}\big\}. \tag{3.5}$$

Let now $S : [0, +\infty[\, \times \mathcal{D} \mapsto \mathcal{D}$ be the Lipschitz semigroup generated by the hyperbolic system (1.1), as in (1.6), (1.7). In particular, $t \mapsto S_t\bar{u}$ yields the exact solution to the Cauchy problem (1.1), (1.2). As proved in [7,8], for any approximate solution u one has the error estimate

$$\|u(T,\cdot) - S_T\bar{u}\|_{\mathbf{L}^1} \ \leqq \ L_0 \cdot \sum_{j=0}^{\nu-1} \Big\|u(t_{j+1},\cdot) - S_h u(t_j,\cdot)\Big\|_{\mathbf{L}^1}. \tag{3.6}$$

For each j, we will show that the corresponding term on the right hand side of (3.6) can be estimated using (2.10).

Consider the covering of the strip $[t_j, t_{j+1}] \times \mathbb{R}$ in terms of the trapezoids Δ_{jk}, introduced at (3.2). As in (2.4), for $k = 1, \ldots, N(j)$ we shall denote by $w^{(k)}$ the solution to the linearized problem with constant coefficients

$$w_t + Aw_x = 0, \qquad w(t_j,\cdot) = u(t_j,\cdot), \qquad A \doteq Df\big(u(t_j,\xi_k)\big) \tag{3.7}$$

for some given points $\xi_k \in]a_{jk},\ b_{jk}\,[$.

Let $\ell_i^{(k)}$ be the ith left eigenvector of the above matrix A, normalized as in (2.5). Using (2.19) on each trapezoid Δ_{jk} we obtain

$$\begin{aligned}
&\int_{a_{jk}+h_j\lambda^+ +\varepsilon^{2/3}}^{b_{jk}+h_j\lambda^- -\varepsilon^{2/3}} \left|\ell_i^{(k)} \cdot \big[u(t_{j+1},x) - w^{(k)}(t_{j+1},x)\big]\right| \mathrm{d}x \\
&\leqq C_0 \cdot \mathrm{Osc.}\{u;\ \Delta_{jk}\} \\
&\quad \cdot \int_{t_j}^{t_{j+1}} \mathrm{Tot.Var.}\Big\{u(t,\cdot);\ ]a_{jk}+(t-t_j)\lambda^+, b_{jk}+(t-t_j)\lambda^-[\Big\}\,\mathrm{d}t \\
&\quad + \left| \int \varphi_i^{(k)}(t_j,x)u(t_j,x)\,\mathrm{d}x - \int \varphi_i^{(k)}(t_{j+1},x)u(t_{j+1},x)\,\mathrm{d}x \right. \\
&\quad \left. + \iint_{\Delta_{jk}} \Big\{u\varphi_{i,t}^{(k)} + f(u)\varphi_{i,x}^{(k)}\Big\}\,\mathrm{d}x\,\mathrm{d}t \right| \\
&\quad + 2\varepsilon^{2/3} \cdot \mathrm{Tot.Var.}\Big\{\ell_i \cdot u(t_j,\cdot);\ [a_{jk}, b_{jk}]\Big\} \\
&\quad + 2\varepsilon^{2/3} \cdot \mathrm{Tot.Var.}\Big\{\ell_i \cdot u(t_{j+1},\cdot);\ [a_{jk}+h_j\lambda^+,\ b_{jk}+h_j\lambda^-]\Big\} \\
&\doteq A_{ik} + B_{ik} + C_{ik} + D_{ik}. \qquad (3.8)
\end{aligned}$$

For notational convenience, call χ_{jk} the characteristic function of the interval $[a_{jk}+h_j\lambda^+ +\varepsilon^{2/3},\ b_{jk}+h_j\lambda^- -\varepsilon^{2/3}]$. Our next goal is to estimate the quantity

$$E_j \doteq \int_{-\infty}^{+\infty} \left| u(t_{j+1},x) - \sum_k w^{(k)}(t_{j+1},x)\cdot\chi_{jk}(x)\right| \mathrm{d}x. \qquad (3.9)$$

This can of course be achieved by summing the terms on the right hand side of (3.8) over all $i = 1,\dots,n$ and $k = 1,\dots,N(j)$. In working towards this goal, we recall the key assumption that every point $(t,x) \in [t_j, t_{j+1}] \times \mathbb{R}$ belongs to one and no more than two of the trapezoids Δ_{jk}. More precisely, we have the implication

$$|k-k'| \geqq 2 \quad \Longrightarrow \quad \Delta_{jk} \cap \Delta_{jk'} = \emptyset. \qquad (3.10)$$

Recalling (3.5), for a fixed i we thus obtain

$$\sum_{k=1}^{N(j)} A_{ik} \leqq C_0\,\kappa_j \cdot \int_{t_j}^{t_{j+1}} 2\,\mathrm{Tot.Var.}\{u(t,\cdot)\}\,\mathrm{d}t, \qquad (3.11)$$

and

$$\begin{aligned}
\sum_{k=1}^{N(j)} C_{ik} &\leqq 4\varepsilon^{2/3}\,\mathrm{Tot.Var.}\{u(t_j,\cdot)\}, \\
\sum_{k=1}^{N(j)} D_{ik} &\leqq 4\varepsilon^{2/3}\,\mathrm{Tot.Var.}\{u(t_{j+1},\cdot)\}. \qquad (3.12)
\end{aligned}$$

The estimate for $\sum_k B_{jk}$ is a bit more delicate, because if we use (1.11) separately on each subdomain Δ_{jk}, the error term on the right side would be multiplied by $N(j)$, which can be a very large number.

For this reason, we argue as follows. For each $i \in \{1, \ldots, n\}$, we consider test functions $\varphi, \widetilde{\varphi}_i$ which satisfy, for $t \in [t_j, t_{j+1}]$,

$$\varphi_i(t,x) = \begin{cases} \varphi_i^{(k)}(t,x) & \text{if } (t,x) \in \Delta_{jk}, \quad k \text{ even}, \\ 0 & \text{otherwise.} \end{cases}$$

$$\widetilde{\varphi}_i(t,x) = \begin{cases} \varphi_i^{(k)}(t,x) & \text{if } (t,x) \in \Delta_{jk}, \quad k \text{ odd}, \\ 0 & \text{otherwise.} \end{cases}$$

For convenience, we denote by $\ell_{\max}$ an upper bound for the norm of all left eigenvectors $\ell_i = \ell_i(u)$ of all matrices $A(u) = Df(u)$, normalized as in (2.5). With this notation we have

$$\|\varphi_i\|_{W^{1\infty}} \leqq \ell_{\max} \cdot \varepsilon^{-2/3}, \qquad \|\widetilde{\varphi}_i\|_{W^{1\infty}} \leqq \ell_{\max} \cdot \varepsilon^{-2/3}. \tag{3.13}$$

Applying (1.11) to the test function φ_i, then to $\widetilde{\varphi}_i$, we obtain

$$\sum_{k \text{ even}} B_{ik} \leqq C\varepsilon h\, \varepsilon^{-2/3} \ell_{\max} \cdot \sup_{t \in [t_j, t_{j+1}]} \text{Tot.Var.}\big\{u(t,\cdot)\big\},$$

$$\sum_{k \text{ odd}} B_{ik} \leqq C\varepsilon h\, \varepsilon^{-2/3} \ell_{\max} \cdot \sup_{t \in [t_j, t_{j+1}]} \text{Tot.Var.}\big\{u(t,\cdot)\big\}.$$

Summing over k, we thus obtain

$$\sum_{k=1}^{N(j)} B_{ik} \leqq 2Ch\varepsilon^{1/3} \ell_{\max} \cdot \sup_{t \in [t_j, t_{j+1}]} \text{Tot.Var.}\big\{u(t,\cdot)\big\}. \tag{3.14}$$

All together, the inequalities (3.11), (3.12), and (3.14) yield

$$\begin{aligned} &\int_{-\infty}^{+\infty} \left| u(t_{j+1}, x) - \sum_k w^{(k)}(t_{j+1}, x) \cdot \chi_{jk}(x) \right| \mathrm{d}x \\ &\leqq C_0\, 2n\kappa_j \int_{t_j}^{t_{j+1}} \text{Tot.Var.}\big\{u(t,\cdot)\big\}\, \mathrm{d}t + 4n\varepsilon^{2/3} \text{Tot.Var.}\big\{u(t_j,\cdot)\big\} \\ &\quad + 4n\varepsilon^{2/3} \text{Tot.Var.}\big\{u(t_{j+1},\cdot)\big\} + 2Ch\, \varepsilon^{1/3} \ell_{\max} \\ &\quad \cdot \sup_{t \in [t_j, t_{j+1}]} \text{Tot.Var.}\big\{u(t,\cdot)\big\}. \end{aligned} \tag{3.15}$$

Next, we replace the approximate solution u with the exact solution $u^{\text{exact}}(t_j + s, \cdot) = S_s u(t_j, \cdot)$ of (1.1) having the same data at $t = t_j$. As proved in [7], with the same notation used in (3.9), as long as $u(t_j, \cdot) \in \mathcal{D}$ remains in the domain of the semigroup, one has

$$\int_{-\infty}^{+\infty} \left| u^{\text{exact}}(t_{j+1}, x) - \sum_k w^{(k)}(t_{j+1}, x) \cdot \chi_{jk}(x) \right| \mathrm{d}x$$

$$\leqq\ C_2\, h \left(\max_{1\leqq k\leqq N(j)} \text{Osc.}\left\{u(t_j,\cdot);\ \Delta_{jk}\right\} \right) \cdot \text{Tot.Var.}\big\{u(t_j,\cdot)\big\}, \quad (3.16)$$

for a suitable constant C_2.

Combining (3.15) with (3.16) and recalling (3.5), we obtain

$$\begin{aligned} &\int_{-\infty}^{+\infty} \Big| u(t_{j+1},x) - \big(S_{t_{j+1}-t_j} u(t_j,\cdot)\big)(x) \Big|\, \mathrm{d}x \\ &\leqq\ C_3 \left(\kappa_j\, h + \varepsilon^{2/3} + h\varepsilon^{1/3}\right) \cdot \sup_{t\in[t_j,t_{j+1}]} \text{Tot.Var.}\big\{u(t,\cdot)\big\}. \end{aligned} \quad (3.17)$$

Recalling that $h \approx \varepsilon^{1/3}$ and $T = \nu\varepsilon^{1/3}$, from the above analysis we obtain

Theorem 1. *Let the basic assumptions* **(A1), (A2)** *hold. Let $t \mapsto u(t,\cdot) \in \mathcal{D}$ be an approximate solution to the Cauchy problem* (1.1), (1.2)*, taking values in the domain $\mathcal{D}$ of the semigroup and satisfying* **(AL)** *and* **($\mathbf{P}_\varepsilon$)**. *Then, for some constant C_4, the following holds: assume that the strip $[0,T]\times\mathbb{R}$ can be covered by trapezoids Δ_{jk}, $j = 0,\ldots,\nu-1$, $k = 1,\ldots N(j)$ as in* (3.2)*, so that* (3.3)–(3.5) *hold. Then the difference between $u(T,\cdot)$ and the exact solution $S_T\bar u$ is bounded by*

$$\|u(T,\cdot) - S_T\bar u\|_{\mathbf{L}^1} \ \leqq\ C_4 \left(2T + \sum_{j=0}^{\nu-1} \kappa_j\right) \varepsilon^{1/3} \cdot \sup_{t\in[0,T]} \text{Tot.Var.}\big\{u(t,\cdot)\big\} \quad (3.18)$$

Proof. Let L_0 be the Lipschitz constant of the semigroup in (1.7). From (3.6) and (3.17) it now follows that

$$\begin{aligned} \|u(T,\cdot) - S_T\bar u\|_{\mathbf{L}^1} &\leqq L_0 \cdot \sum_{j=0}^{\nu-1} \Big\| u(t_{j+1},\cdot) - S_h u(t_j,\cdot) \Big\|_{\mathbf{L}^1} \\ &\leqq L_0 \cdot \sum_{j=0}^{\nu-1} C_3 \left(\kappa_j\, h + \varepsilon^{2/3} + h\varepsilon^{1/3}\right) \cdot \sup_{t\in[t_j,t_{j+1}]} \text{Tot.Var.}\big\{u(t,\cdot)\big\} \\ &\leqq L_0\, C_3 \cdot \left(2T + \sum_{j=0}^{\nu-1} \kappa_j\right) \varepsilon^{1/3} \cdot \sup_{t\in[0,T]} \text{Tot.Var.}\big\{u(t,\cdot)\big\}. \end{aligned} \quad (3.19)$$

This yields (3.18), with $C_4 = L_0 C_3$. □

Remark 2. Based on the estimate (3.19), we seek to understand at which rate the error in the approximate solution may approach zero, as $\varepsilon \to 0$.

Having chosen $h \approx \varepsilon^{1/3}$, we can choose all bounded trapezoids Δ_{jk}, $1 < k < N(j)$, to be of diameter $\mathcal{O}(1)\cdot\varepsilon^{1/3}$. Moreover, by choosing every b_{j1} suitably large and negative, and $a_{j,N(j)}$ large and positive, we can assume that the solution is nearly constant on the unbounded trapezoids $\Delta_{j,1}$ and $\Delta_{j,N(j)}$. Here and in the sequel, the Landau symbol $\mathcal{O}(1)$ denotes a uniformly bounded quantity.

If the exact solution is Lipschitz continuous, we expect that the maximum oscillation (3.5) will be of size $\kappa_j = \mathcal{O}(1)\cdot\varepsilon^{1/3}$ for every $j \in \{0,1,\ldots,\nu-1\}$. In

this case, as $\varepsilon \to 0$ the quantity $2T + \sum_{j=1}^{\nu} \kappa_j$ remains uniformly bounded, and the estimate (3.19) indicates that the error vanishes of order $\mathcal{O}(1) \cdot \varepsilon^{1/3}$.

Next, assume that the initial data $\bar{u}$ contains a jump, generating a centered rarefaction wave of strength σ. In this case, taking into account the decay caused by genuine nonlinearity, we expect that the oscillation of u over a trapezoid Δ_{jk} of diameter $\mathcal{O}(1) \cdot \varepsilon^{1/3}$ will satisfy a bound of the form

$$\text{Osc.}\left\{u(t,\cdot);\ \Delta_{jk}\right\} \ = \ \mathcal{O}(1) \cdot \min\left\{\sigma,\ \frac{\varepsilon^{1/3}}{t_j}\right\}. \tag{3.20}$$

Recalling that $t_j = j\varepsilon^{1/3}$ and $\nu = T\varepsilon^{-1/3}$, this leads to

$$\sum_{j=1}^{\nu} \kappa_j = \sum_{j=1}^{\nu} \min\left\{\sigma,\ \frac{C\varepsilon^{1/3}}{j\varepsilon^{1/3}}\right\} = \mathcal{O}(1) \cdot \log \nu = \mathcal{O}(1) \cdot |\log \varepsilon|. \tag{3.21}$$

In this case, the estimate (3.19) would indicate that the error vanishes of order $\mathcal{O}(1) \cdot \varepsilon^{1/3} |\log \varepsilon|$. The same should hold if the exact solution contains finitely many centered rarefaction waves.

We emphasize, however, that this is only a heuristic expectation. For a numerically computed solution, it needs to be confirmed by a post-processing algorithm, which can actually provide a bound on the oscillations κ_j in (3.18).

4. Solutions with an Isolated Large Shock

The error estimates developed in the previous section are not effective for solutions containing large shocks. Indeed, around a shock, the oscillation will be large. As a consequence, even when the diameters of the trapezoids Δ_{jk} in (3.2) approach zero, the maximum oscillation κ_j in (3.5) will remain uniformly large. For this reason, we do not expect that the right hand side of the error bound (3.18) will approach zero as $\varepsilon \to 0$. To cope with this problem, in this section we develop additional tools to estimate the numerical error in a neighborhood of a shock.

Consider an approximate solution u, which satisfies **(AL)** and $(\mathbf{P}_\varepsilon)$. We seek a sharper error bound, assuming that the oscillation of u is concentrated in a narrow region of the form

$$\Gamma \doteq \Big\{(t,x);\ t \in [0,h],\ |x - \gamma(t)| < \delta\Big\}, \qquad \gamma(t) = x_0 + \lambda t. \tag{4.1}$$

Of course, we expect that $\gamma(\cdot)$ will trace the position of a large shock in the exact solution. Here $\rho, \delta > 0$ are suitable parameters. Different choices of these values will lead to different estimates. As a rule of thumb, it will be useful for the reader to keep in mind their order of magnitude:

$$h \approx \varepsilon^{1/3}, \qquad \rho \approx \varepsilon^{1/3}, \qquad \delta \approx \frac{\varepsilon}{\rho} = \varepsilon^{2/3}. \tag{4.2}$$

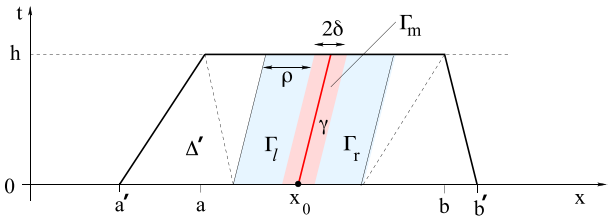

Fig. 3. The regions Γ_l, Γ_m, Γ_r introduced at (4.13) to trace a large shock, and the trapezoid Δ' at (4.4)

Referring to Fig. 3, we introduce the points

$$\begin{cases} a = x_0 + \lambda^- h - \delta - \rho, \\ b = x_0 + \lambda^+ h + \delta + \rho, \end{cases} \qquad \begin{cases} a' \doteq a - \lambda^+ h, \\ b' \doteq b - \lambda^- h, \end{cases} \tag{4.3}$$

and consider the trapezoidal domain

$$\begin{aligned} \Delta' &\doteq \text{co}\Big\{(0,a'),\ (0,b'),\ (h,a),\ (h,b)\Big\} \\ &= \Big\{(t,x);\ t\in[0,h],\ \ a'+\lambda^+ t \leqq x \leqq b'+\lambda^- t\Big\}. \end{aligned} \tag{4.4}$$

Our basic assumption is that, outside the narrow strip Γ, the oscillation of u remains small. More precisely, consider the left and right domains

$$\begin{aligned} \Delta'_l &= \Big\{(t,x);\ t\in[0,h],\ \ a'+\lambda^+ t \leqq x \leqq x_0-\delta+\lambda t\Big\}, \\ \Delta'_r &= \Big\{(t,x);\ t\in[0,h],\ \ x_0+\delta+\lambda t \leqq x \leqq b'+\lambda^- t\Big\}, \end{aligned} \tag{4.5}$$

and define

$$\kappa' \doteq \max\Big\{\text{Osc.}\{u;\ \Delta'_l\},\ \text{Osc.}\{u;\ \Delta'_r\}\Big\}. \tag{4.6}$$

Saying that

$$u^- \doteq u\big(0, x_0-\delta\big), \qquad u^+ \doteq u\big(0, x_0+\delta\big), \tag{4.7}$$

the above definition of κ' implies

$$\begin{aligned} |u(t,x)-u^-| &\leqq \kappa' \ \text{ for } \ \ (t,x)\in\Delta'_l, \\ |u(t,x)-u^+| &\leqq \kappa' \ \text{ for } \ \ (t,x)\in\Delta'_r. \end{aligned} \tag{4.8}$$

Assuming that κ' is small, the following result provides a bound on the distance between u and the exact solution, $\|u(h,\cdot)-S_h\bar{u}\|_{\mathbf{L}^1([a,b])}$, restricted to the interval $[a,b]$.

Theorem 2. *Let $t \mapsto u(t,\cdot) \in \mathcal{D}$ be an approximate solution to the Cauchy problem (1.1), (1.2), taking values in the domain $\mathcal{D}$ of the semigroup, and satisfying* **(AL)** *and* ($\mathbf{P}_\varepsilon$)*. Then, for some constant C_5, in the above setting we have the error bound*

$$\int_a^b \big|u(h,x) - (S_h\bar u)(x)\big|\,\mathrm{d}x \;\leqq\; C_5\cdot h\left(\frac{\varepsilon}{\rho}+\kappa'+\frac{\rho\kappa'+\delta}{h}\right)^{2/3} + C_5\Big(\rho\kappa'+h\kappa'+\delta\Big). \tag{4.9}$$

Moreover, there exists a constant K_1 such that, if

$$|u^+-u^-| \;\geqq\; K_1\cdot\left(\frac{\varepsilon}{\delta}+\kappa'+\frac{\rho\kappa'+\delta}{h}\right)^{1/3}, \tag{4.10}$$

then the estimate (4.9) can be improved to

$$\int_a^b \big|u(h,x) - (S_h\bar u)(x)\big|\,\mathrm{d}x \;\leqq\; C_5\cdot h\left(\frac{\varepsilon}{\rho}+\kappa'+\frac{\rho\kappa'+\delta}{h}\right). \tag{4.11}$$

Remark 3. It may seem surprising that the error bound (4.11), valid for large jumps, is actually better than (4.9), which applies to small jumps. To understand what is involved here, is useful to observe that if the strength $\sigma = |u^+-u^-|$ is small, it could be that this jump is tracing a centered rarefaction wave within the exact solution, which gets approximated by a single jump by the numerical algorithm (indeed, this is a common feature of front tracking approximations). If σ is small enough, the entropy produced by the jump is small, and the assumptions (1.11), (1.12) can still be satisfied. This is a "worst-case scenario": as shown in Fig. 4, the corresponding $\mathbf{L}^1$ error is $\mathcal{O}(1)\cdot h\sigma^2$. On the other hand, if the strength σ of the jump is large, the entropy dissipation assumption (1.12) rules out this possibility. Therefore, the jump must trace an entropic shock in the exact solution.

Proof of Theorem 2. **1.** As a first step, using ($\mathbf{P}_\varepsilon$) we will provide a bound for the error

$$\Big|f(u^+)-f(u^-)-\lambda\big(u^+-u^-\big)\Big|. \tag{4.12}$$

As shown in Fig. 3, denote by Γ_l, Γ_m, and Γ_r the left, middle, and right domains

$$\begin{aligned}
\Gamma_l &\doteq \big\{(t,x);\ t\in[0,h],\ \ x\in[\gamma(t)-\delta-\rho,\ \gamma(t)-\delta]\big\},\\
\Gamma_m &\doteq \big\{(t,x);\ t\in[0,h],\ \ x\in[\gamma(t)-\delta,\ \gamma(t)+\delta]\big\},\\
\Gamma_r &\doteq \big\{(t,x);\ t\in[0,h],\ \ x\in[\gamma(t)+\delta,\ \gamma(t)+\delta+\rho]\big\}.
\end{aligned} \tag{4.13}$$

Recalling (4.1), (4.3) and (4.4), we observe that the above definitions imply

$$\Gamma_m=\Gamma, \qquad\qquad \Gamma_l\cup\Gamma_m\cup\Gamma_r \;\subset\; \Delta'.$$

Given $\rho > 0$, consider a Lipschitz test function φ such that, for $t \in [0, h]$, one has

$$\phi(t,x) = \begin{cases} 0 & \text{if} \quad |x-\gamma(t)| \geqq \delta+\rho, \\ 1 & \text{if} \quad |x-\gamma(t)| \leqq \delta, \\ \dfrac{\delta+\rho-|x-\gamma(t)|}{\rho} & \text{if} \quad \delta < |x-\gamma(t)| < \delta+\rho. \end{cases} \tag{4.14}$$

Then choose any unit vector $\mathbf{w} \in \mathbb{R}^n$ and set $\varphi(t,x) = \phi(t,x)\,\mathbf{w}$. By construction, for $t \in [0,h]$ the test function φ vanishes outside the union $\Gamma_l \cup \Gamma_m \cup \Gamma_r$. Notice that

$$\|\varphi_x\|_{\mathbf{L}^\infty} = \frac{1}{\rho}, \qquad \|\varphi_t\|_{\mathbf{L}^\infty} = \frac{|\lambda|}{\rho}.$$

Assuming that the approximate solution u satisfies (P_ε), by (1.11), it follows that

$$\begin{aligned} &\left| \int \varphi(0,x)u(0,x)\,\mathrm{d}x - \int \varphi(h,x)u(h,x)\,\mathrm{d}x \right. \\ &\quad \left. + \iint_{\Gamma_l \cup \Gamma_m \cup \Gamma_r} u\varphi_t + f(u)\varphi_x \,\mathrm{d}x\,\mathrm{d}t \right| \\ &\quad \leqq C\,\varepsilon\, h \cdot \frac{\max\{1,\ |\lambda|\}}{\rho} \cdot \sup_{t\in[0,h]} \mathrm{Tot.Var.}\big\{u(t,\cdot)\big\}. \end{aligned} \tag{4.15}$$

Using (4.8), we now estimate

$$\left(\iint_{\Gamma_l} + \iint_{\Gamma_m} + \iint_{\Gamma_r} \right) \big\{ u\varphi_t + f(u)\varphi_x \big\}\,\mathrm{d}x\,\mathrm{d}t \doteq \mathbf{I}_l + \mathbf{I}_m + \mathbf{I}_r. \tag{4.16}$$

Trivially, $\mathbf{I}_m = 0$ because $\varphi_t = \varphi_x = 0$ on Γ_m. By (4.8) it follows that

$$\begin{aligned} \mathbf{I}_r + \mathbf{I}_l &= \iint_{\Gamma_r \cup \Gamma_l} \big\{ u\varphi_t + f(u)\varphi_x \big\}\,\mathrm{d}x\,\mathrm{d}t \\ &= \int_0^h \Big\{ \big[\lambda u^+ - f(u^+)\big] - \big[\lambda u^- - f(u^-)\big] \Big\} \mathbf{w}\,\mathrm{d}t + \mathcal{O}(1)\cdot h\,\kappa'. \end{aligned} \tag{4.17}$$

Next, by (4.6), one obtains

$$\begin{aligned} &\left| \int \varphi(0,x)u(0,x)\,\mathrm{d}x - \int \varphi(h,x)u(h,x)\,\mathrm{d}x \right| \\ &\quad \leqq \left(\int_{x_0-\delta-\rho}^{x_0-\delta} + \int_{x_0+\delta}^{x_0+\delta+\rho} \right) \Big| u(0,x) - u(h, x+\lambda h) \Big|\,\mathrm{d}x \\ &\qquad + \int_{x_0-\delta}^{x_0+\delta} \Big(|u(0,x)| + |u(h,x+\lambda h)| \Big)\,\mathrm{d}x \\ &\quad \leqq \rho \cdot \mathrm{Osc.}\{u;\ \Delta_l'\} + \rho \cdot \mathrm{Osc.}\{u;\ \Delta_r'\} + 2\delta \cdot 2\|u\|_{\mathbf{L}^\infty} \\ &\quad \leqq 2\rho\kappa' + 4\delta\,\|u\|_{\mathbf{L}^\infty}. \end{aligned} \tag{4.18}$$

From (4.15), by (4.17) and (4.18) it follows that

$$\begin{aligned}&\left|\int_0^h \Big\{\lambda(u^+-u^-)-\big[f(u^+)-f(u^-)\big]\Big\}\mathbf{w}\,\mathrm{d}t\right|\\&\quad=\mathcal{O}(1)\cdot\left(\frac{\varepsilon\,h}{\rho}+h\kappa'+\rho\kappa'+\delta\right),\end{aligned}\tag{4.19}$$

where the factor $\mathcal{O}(1)$ already accounts for the uniform bound on the total variation. Choosing the unit vector

$$\mathbf{w}=\frac{\lambda(u^+-u^-)-\big[f(u^+)-f(u^-)\big]}{\Big|\lambda(u^+-u^-)-\big[f(u^+)-f(u^-)\big]\Big|},$$

by (4.19) we conclude that the error in the Rankine–Hugoniot equations has the size

$$\Big|\lambda(u^+-u^-)-\big[f(u^+)-f(u^-)\big]\Big|=\mathcal{O}(1)\cdot\left(\frac{\varepsilon h}{\rho}+h\kappa'+\frac{\rho\kappa'+\delta}{h}\right).\tag{4.20}$$

2. Next, consider the piecewise constant function

$$w(t,x)\doteq\begin{cases}u^- & \text{if}\quad x<x_0+\lambda t,\\ u^+ & \text{if}\quad x>x_0+\lambda t.\end{cases}\tag{4.21}$$

The aim of the next two steps is to prove that the difference between w and an exact solution having the same initial data is bounded by

$$\|w(h,\cdot)-S_h w(0,\cdot)\|_{\mathbf{L}^1(\mathbb{R})}=\mathcal{O}(1)\cdot h\,\Big|\lambda(u^+-u^-)-\big[f(u^+)-f(u^-)\big]\Big|.\tag{4.22}$$

With this goal in mind, define the averaged Jacobian matrix

$$A=\int_0^1 Df\big(su^+ +(1-s)u^-\big)\,\mathrm{d}s.$$

Call $\lambda_1<\cdots<\lambda_n$ the eigenvalues of A. Let $\{r_1,\ldots,r_n\}$ and $\{\ell_1,\ldots,\ell_n\}$ be dual bases of right and left eigenvectors of A, normalized as (2.5). Moreover, let c_i, ℓ_{max} be such that

$$u^+-u^-=\sum_{i=1}^n c_i r_i,\qquad\qquad \ell_{max}\doteq\max\{|\ell_1|,\ldots,|\ell_n|\}.\tag{4.23}$$

For every $i=1,\ldots,n$, we then have

$$\begin{aligned}&\Big|\ell_i\cdot(\lambda I-A)(u^+-u^-)\Big|=|c_i|\,|\lambda-\lambda_i|\\&\quad\leqq\ell_{max}\Big|\lambda(u^+-u^-)-\big[f(u^+)-f(u^-)\big]\Big|.\end{aligned}\tag{4.24}$$

Let $i^* \in \{1, \dots, n\}$ be a characteristic family such that $|\lambda - \lambda_{i^*}| = \min_i |\lambda - \lambda_i|$. Since the eigenvalues of A are strictly separated, by (4.24), it follows that

$$|c_i| = \mathcal{O}(1) \cdot \Big|\lambda(u^+ - u^-) - \big[f(u^+) - f(u^-)\big]\Big| \qquad \text{for all } i \neq i^*. \quad (4.25)$$

We now consider the solution to the Riemann problem with left and right states u^-, u^+. Let $\sigma_1, \dots, \sigma_n$ be the sizes of the waves in this solution. As usual, if the ith field is genuinely nonlinear, we choose the sign so that $\sigma_i > 0$ corresponds to a rarefaction wave, while $\sigma_i < 0$ yields an entropy admissible shock. For future use, we denote by

$$u^- = u_0,\ u_1,\ \dots,\ u_n = u^+ \quad (4.26)$$

the intermediate states. If the i^*th characteristic field is linearly degenerate, standard estimates on the strength of these waves yield the bound

$$\sum_{i=1}^{n} |\sigma_i - c_i| = \mathcal{O}(1) \cdot \sum_{i \neq i^*} |c_i|. \quad (4.27)$$

Indeed, (4.27) is trivially true when the right hand side is zero. The general case is obtained by an application of the implicit function theorem. The same estimate (4.27) is achieved when the i^*th field is genuinely nonlinear and $c_{i^*} < 0$. By (4.27) it follows that

$$|u_{i^*} - u^+| + |u_{i^*-1} - u^-| + \sum_{i \neq i^*} |u_i - u_{i-1}| = \mathcal{O}(1) \cdot \sum_{i \neq i^*} |c_i|. \quad (4.28)$$

In both of the above cases, combining (4.24), (4.25), and (4.28), the distance between $w(h, \cdot)$ and an exact solution can be estimated as

$$\begin{aligned}
&\frac{1}{h} \|w(h, \cdot) - S_h w(0, \cdot)\|_{\mathbf{L}^1(\mathbb{R})} \\
&= \mathcal{O}(1) \cdot \sum_{i \neq i^*} |c_i| + \mathcal{O}(1) \cdot |\lambda_{i^*} - \lambda|\, |c_{i^*}| \\
&\quad + \mathcal{O}(1) \cdot \big(|u_{i^*} - u^+| + |u_{i^*-1} - u^-|\big)\, |\lambda| \\
&= \mathcal{O}(1) \cdot \Big|\lambda(u^+ - u^-) - \big[f(u^+) - f(u^-)\big]\Big| \\
&= \mathcal{O}(1) \cdot \left(\frac{\varepsilon}{\rho} + \kappa' + \frac{\rho\kappa' + \delta}{h}\right). \qquad (4.29)
\end{aligned}$$

Notice that the last estimate was obtained from (4.20). □

3. It remains to study the case where the i^*th field is genuinely nonlinear, but $c_{i^*} > 0$. For this purpose, call $\widetilde{w} = \widetilde{w}(t, x)$ the solution to the Riemann problem with initial data $\widetilde{w}(0, \cdot) = w(0, \cdot)$, which contains a non-entropic i^*-shock of size

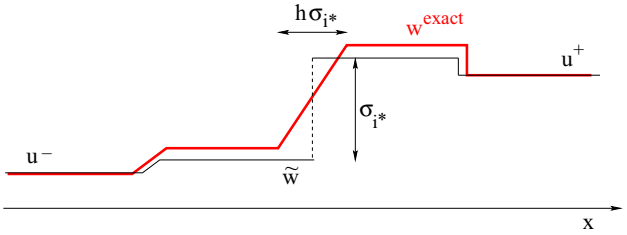

Fig. 4. Comparing the entropic solution w^{exact} to the Riemann problem with left and right states u^-, u^+ with another weak solution $\widetilde{w}$ containing a non-admissible i^*-shock of strength $\sigma_{i^*} > 0$. Taking into account the presence of a centered rarefaction wave in w^{exact}, The difference between the two solutions can be bounded as $\|\widetilde{w}(h,\cdot) - w^{\text{exact}}(h,\cdot)\|_{\mathbf{L}^1} = \mathcal{O}(1)\cdot h\sigma_{i^*}^2$

σ^*, while all other waves are entropy admissible. We observe that all the above estimates still apply to $\widetilde{w}$. In particular,

$$\frac{1}{h}\|\widetilde{w}(h,\cdot) - S_h w(0,\cdot)\|_{\mathbf{L}^1(\mathbb{R})} = \mathcal{O}(1)\cdot\sum_{i\neq i^*}|c_i| + \mathcal{O}(1)\cdot|\lambda_{i^*}-\lambda|\,|c_{i^*}|. \tag{4.30}$$

It remains to estimate the difference between $\widetilde{w}$ and the entropic solution to the same Riemann problem. Call $u^- = \widetilde{u}_0, \widetilde{u}_1, \ldots, \widetilde{u}_n = u^+$ the intermediate states for the non-entropic solution $\widetilde{w}$. Since shock and rarefaction curves have a second order tangency, comparing with the intermediate states (4.26) of the entropic solution, we find

$$|\widetilde{u}_i - u_i| = \mathcal{O}(1)\cdot\sigma_{i^*}^3 \qquad i = 0, 1, \ldots, n. \tag{4.31}$$

Taking into account that the wave connecting the states u_{i^*-1} and u_{i^*} is a centered rarefaction instead of a single jump, we obtain the bound

$$\frac{1}{h}\|w(h,\cdot) - \widetilde{w}(h,\cdot)\|_{\mathbf{L}^1(\mathbb{R})} = \mathcal{O}(1)\cdot\sigma_{i^*}^2. \tag{4.32}$$

Combining (4.30) with (4.32) we conclude that

$$\begin{aligned}\frac{1}{h}\|w(h,\cdot) - S_h w(0,\cdot)\|_{\mathbf{L}^1(\mathbb{R})} &= \mathcal{O}(1)\\ &\cdot\sum_{i\neq i^*}|c_i| + \mathcal{O}(1)\cdot|\lambda_{i^*}-\lambda|\,|c_{i^*}| + \mathcal{O}(1)\cdot c_{i^*}^2.\end{aligned} \tag{4.33}$$

We claim that the jump $c_{i^*} > 0$ must be small, otherwise the approximate entropy inequality (1.12) would fail. Intuitively, this means that the approximate solution cannot contain a large, non-admissible shock. Indeed, let η be a convex entropy, with entropy flux q, such that (1.8) holds for every non-admissible shock of the i^* family. Let ϕ be the test function in (4.14). Arguing as in (4.15), by (1.12) we now obtain

$$\int \phi(0,x)\eta(u(0,x))\,\mathrm{d}x - \int \phi(h,x)\eta(u(h,x))\,\mathrm{d}x$$

$$+\iint_{\Gamma_l\cup\Gamma_m\cup\Gamma_r}\eta(u)\phi_t+q(u)\phi_x\,\mathrm{d}x\,\mathrm{d}t\ \geqq\ -\mathcal{O}(1)\cdot\frac{\varepsilon}{\rho}h. \tag{4.34}$$

As in (4.18), we have

$$\begin{aligned}&\left|\int\phi(0,x)\eta(u(0,x))\,\mathrm{d}x-\int\phi(h,x)\eta(u(h,x))\,\mathrm{d}x\right|\\&\quad=\mathcal{O}(1)\cdot(\rho\kappa'+\delta).\end{aligned} \tag{4.35}$$

Repeating the argument at (4.16), (4.17) we obtain

$$\begin{aligned}\int_0^h\int\big\{\eta(u)\phi_t+q(u)\phi_x\big\}\,\mathrm{d}x\,\mathrm{d}t=h\Big\{&\big[\lambda\eta(u^+)-q(u^+)\big]\\&-\big[\lambda\eta(u^-)-q(u^-)\big]\Big\}+\mathcal{O}(1)h\kappa'.\end{aligned} \tag{4.36}$$

Next, consider the state $\widetilde{u}^+$, connected to u^- by a (not entropy admissible) i^*-shock of size $c_{i^*}>0$. By the implicit function theorem, one has the bound

$$|\widetilde{u}^+-u^+|=\mathcal{O}(1)\cdot\sum_{i\neq i^*}|c_i|. \tag{4.37}$$

Recalling (1.8), from (4.36) we obtain that

$$\begin{aligned}&\frac{1}{h}\int_0^h\int\big\{\eta(u)\phi_t+q(u)\phi_x\big\}\,\mathrm{d}x\,\mathrm{d}t\\&=\Big\{\big[\lambda\eta(\widetilde{u}^+)-q(\widetilde{u}^+)\big]-\big[\lambda\eta(u^-)-q(u^-)\big]\Big\}\\&\quad+\mathcal{O}(1)\cdot|\widetilde{u}^+-u^+|+\mathcal{O}(1)\cdot\kappa'\\&\leqq\ -c_0|\widetilde{u}^+-u^-|^3+\mathcal{O}(1)\cdot|\widetilde{u}^+-u^+|+\mathcal{O}(1)\cdot\kappa'\\&\leqq\ -c_0c_{i^*}^3+\mathcal{O}(1)\cdot\sum_{i\neq i^*}|c_i|\\&\leqq\ -c_0c_{i^*}^3+\mathcal{O}(1)\cdot\Big|\lambda(u^+-u^-)-\big[f(u^+)-f(u^-)\big]\Big|,\end{aligned} \tag{4.38}$$

where (4.25) was used in the last inequality. Combining (4.38) with (4.34)–(4.36), and using (4.20) to bound the last term in (4.38), we obtain

$$c_{i^*}=\mathcal{O}(1)\cdot\left(\frac{\varepsilon}{\rho}+\kappa'+\frac{\rho}{h}\kappa'+\frac{\delta}{h}\right)^{1/3}. \tag{4.39}$$

Starting from (4.33), and using (4.24), (4.25), (4.20), and (4.39), we obtain

$$\frac{1}{h}\|w(h,\cdot)-S_hw(0,\cdot)\|_{\mathbf{L}^1(\mathbb{R})}=\ \mathcal{O}(1)\cdot\left(\frac{\varepsilon}{\delta}+\kappa'+\frac{\rho\kappa'+\delta}{h}\right)^{2/3}. \tag{4.40}$$

4. Notice that the estimate (4.40) is somewhat weaker, compared with (4.29). In this step we show that, if the jump $|u^+-u^-|$ is sufficiently large, then in the

genuinely nonlinear case we must have $c_{i^*} < 0$, hence the stronger estimate (4.29) holds. Recalling (4.23), notice that

$$|c_{i^*}| \geqq |u^+ - u^-| - \sum_{i \neq i^*} |c_i| \geqq |u^+ - u^-| - \mathcal{O}(1) \cdot \left(\frac{\varepsilon}{\delta} + \kappa' + \frac{\rho\kappa' + \delta}{h} \right). \tag{4.41}$$

Therefore, there exists a constant K_1 large enough so that, if (4.10) holds, then (4.41) provides a contradiction with (4.39). Since (4.39) was obtained by assuming that $c_{i^*} > 0$, we conclude that (4.10) is a sufficient condition to guarantee that $c_{i^*} \leqq 0$. In this case, the stronger estimate (4.29) holds.

5. Restricted to the interval $[a, b]$, by (4.8) and (4.40), the difference between u and the exact solution having initial data $u(0, x) = \bar{u}(x)$ can now be estimated by

$$\begin{aligned}
&\int_a^b \big|u(h,x) - (S_h\bar{u})(x)\big|\,\mathrm{d}x \\
&\leqq \left(\int_a^{x_0+\lambda h-\delta} + \int_{x_0+\lambda h-\delta}^{x_0+\lambda h+\delta} + \int_{x_0+\lambda h+\delta}^{b} \right) \big|u(h,x) - w(h,x)\big|\,\mathrm{d}x \\
&\quad + \|w(h,\cdot) - S_h w(0,\cdot)\|_{\mathbf{L}^1} \\
&\quad + L_0 \left(\int_{a'}^{x_0-\delta} + \int_{x_0-\delta}^{x_0+\delta} + \int_{x_0+\delta}^{b'} \right) \big|w(0,x) - \bar{u}(x)\big|\,\mathrm{d}x \\
&\leqq \Big[(b-a-2\delta)\,\kappa' + 4\delta\|u\|_{\mathbf{L}^\infty} \Big] + \mathcal{O}(1)\cdot h \left(\frac{\varepsilon}{\rho} + \kappa' + \frac{\rho\kappa' + \delta}{h} \right)^{2/3} \\
&\quad + \Big[(b'-a'-2\delta)\,\kappa' + 4\delta\|u\|_{\mathbf{L}^\infty} \Big] \\
&\leqq C_5 \cdot h \left(\frac{\varepsilon}{\rho} + \kappa' + \frac{\rho\kappa' + \delta}{h} \right)^{2/3} + C_5(\rho + h)\kappa' + C_5\delta,
\end{aligned} \tag{4.42}$$

for a suitable constant C_5. Indeed, from (4.3) it follows that

$$\begin{aligned}
b - a &= 2\rho + 2\delta + (\lambda^+ - \lambda^-)h, \\
b' - a' &= 2\rho + 2\delta + 2(\lambda^+ - \lambda^-)h.
\end{aligned} \tag{4.43}$$

On the other hand, if (4.10) holds, then we can use (4.29) instead of (4.40). The same argument used in (4.42) now yields that

$$\begin{aligned}
\int_a^b \big|u(h,x) - (S_h\bar{u})(x)\big|\,\mathrm{d}x &\leqq C_5 \cdot h \left(\frac{\varepsilon}{\rho} + \kappa' + \frac{\rho\kappa' + \delta}{h} \right) \\
&\quad + C_5(\rho + \delta + h)\kappa' + C_5\delta.
\end{aligned} \tag{4.44}$$

5. A Post-processing Algorithm

There are various ways to use the estimates developed in Sections 3 and 4, to obtain a posteriori error bounds. The underlying idea is to isolate a finite number of thin regions enclosing the large jumps, where the estimates (4.9) or (4.11) can be used. Then use the bounds (3.18) on the remaining portion of the domain.

The algorithm described below can be applied to any BV solution of (1.1), but it is designed in order to be most effective when the exact solution is piecewise Lipschitz with finitely many shocks (or contact discontinuities) and centered rarefaction waves.

Let $u : [0, T] \times \mathbb{R} \mapsto \mathbb{R}^n$ be an approximate solution of (1.1), (1.2), which satisfies the properties **(AL)** and $(\mathbf{P}_\varepsilon)$. In this section we introduce an algorithm which checks its total variation, identifies the location of large shocks, and constructs trapezoidal subdomains where the oscillation remains small. In view of our previous analysis, this will yield an error bound on the $\mathbf{L}^1$ distance (1.4) between u and an exact solution.

The algorithm includes three steps.

STEP 1. For each $t \in [0, T]$, we compute the total variation of $u(t, \cdot)$. Let $\delta_0 > 0$ be the constant in (1.5). If

$$\sup_{t\in[0,T]} \text{Tot.Var.}\big\{u(t,\cdot)\big\} \leqq \delta_0, \tag{5.1}$$

then the algorithm can proceed. On the other hand, if (5.1) fails, the approximate solution may lie outside the domain of the semigroup and no error estimate can be provided. In this case, the algorithm stops.

STEP 2. We now split the interval $[0, T]$ into equal subintervals of size $h = \varepsilon^{1/3}$, inserting the times $t_j = j\,h$, $j = 0, 1, \ldots, \nu = T/h$. The next goal is to identify the location of the large shocks, on each strip $[t_j, t_{j+1}] \times \mathbb{R}$. For this purpose, we set $\rho = h = \varepsilon^{1/3}$, $\delta = \varepsilon^{2/3}$, and choose two additional parameters:

- A lower bound σ_{min} for the size of the jump to be traced;
- An upper bound κ' for the oscillation of u on a region to the right and to the left of the jump.

In view of (4.10), it will be convenient to choose these values so that

$$\sigma_{min} \geqq K_1 \cdot \left(2\varepsilon^{1/3} + 2\kappa'\right)^{1/3}. \tag{5.2}$$

In this way, the sharper estimate (4.11) in Theorem 2 will be available.

Recalling the construction at (4.1)–(4.5), we introduce

Definition 1. Given an interval $[x_0 - \delta, x_0 + \delta]$ and a speed $\lambda \in [\lambda^-, \lambda^+]$, consider the polygonal regions

$$\begin{aligned}
\Gamma &\doteq \Big\{(t,x);\ \ t \in [t_j, t_{j+1}],\ \ x_0 - \delta + \lambda(t - t_j) \leqq x \leqq x_0 - \delta + \lambda(t - t_j)\Big\},\\
\Delta'_l &= \Big\{(t,x);\ t \in [t_j, t_{j+1}],\ \ a' + \lambda^+(t - t_j) \leqq x \leqq x_0 - \delta + \lambda(t - t_j)\Big\},\\
\Delta'_r &= \Big\{(t,x);\ t \in [t_j, t_{j+1}],\ \ x_0 + \delta + \lambda(t - t_j) \leqq x \leqq b' + \lambda^-(t - t_j)\Big\},
\end{aligned} \tag{5.3}$$

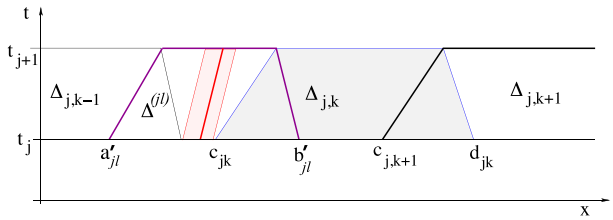

Fig. 5. Implementing a post-processing algorithm, each strip $[t_j, t_{j+1}] \times \mathbb{R}$ is covered with trapezoids Δ_{jk} where the oscillation remains small (as far as possible), and trapezoids $\Delta^{(j\ell)}$ containing a large traced shock

with a', b' as in (4.3). We say that Γ **traces a shock** during the time interval $[t_i, t_{i+1}]$ if

$$\max\Big\{\text{Osc.}\{u;\ \Delta'_l\},\ \text{Osc.}\{u;\ \Delta'_r\}\Big\} \ \leqq\ \kappa', \tag{5.4}$$

$$\big|u(t_j, x_0+\delta) - u(t_j, x_0+\delta)\big| \ \geqq\ \sigma_{min}. \tag{5.5}$$

In what follows, we shall denote by

$$\Delta^{(j\ell)} = \Big\{(t,x);\ t\in[t_j,t_{j+1}],\ \ a'_{j\ell}+\lambda^+(t-t_j) \ \leqq\ x\ \leqq\ b'_{j\ell}+\lambda^-(t-t_j)\Big\},$$
$$\ell = 1,\ldots,N'(j) \tag{5.6}$$

the trapezoids containing the traced shocks (see Fig. 5).

STEP 3. We cover the remaining region $[t_j, t_{j+1}] \setminus \bigcup_{\ell=1}^{N'(j)} \Delta^{(j\ell)}$ with finitely many trapezoids of the same form as in (3.2)

$$\Delta_{jk} = \Big\{(t,x);\ t\in[t_j,t_{j+1}],\ \frac{t_{j+1}-t}{h}c_{jk} + \frac{t-t_j}{h}(c_{jk}+h\lambda^+ + \varepsilon^{2/3})$$
$$\leqq\ x\ \leqq\ \frac{t_{j+1}-t}{h}d_{jk} + \frac{t-t_j}{h}(d_{jk}+h\lambda^- - \varepsilon^{2/3})\Big\}, \tag{5.7}$$

in such a way that each point $(t,x) \in [t_j, t_{j+1}] \times \mathbb{R}$ is contained in at most two of these trapezoids (see Fig. 5). More precisely, we can assume that (3.10) holds, for all $k, k' \in \{1,\ldots,N(j)\}$. Within each time interval $[t_j, t_{j+1}]$, we compute the maximum oscillation of u over these trapezoids:

$$\kappa_j \doteq \max_{1\leqq k\leqq N(j)} \text{Osc.}\{u;\ \Delta_{jk}\}. \tag{5.8}$$

The next result provides an a posteriori estimate on the $\mathbf{L}^1$ error in the approximate solution. Here the estimate refers to the outcome of a post-processing algorithm, depending on the choice of κ' and σ_{min} in (5.4), (5.5). We remark that any choice of such parameters, determining which shocks are actually traced, leads to some error bound. However, the sharpness of the estimate heavily depends on a suitable choice of these parameter values.

Theorem 3. *Consider a system of conservation laws satisfying the basic assumptions* **(A1), (A2)**. *Then there exist constants C', C'' such that the following holds.*

Let $u : [0, T] \times \mathbb{R} \mapsto \mathbb{R}^n$ be an approximate solution to the Cauchy problem (1.1), (1.2), satisfying the conditions **(AL)** *and* **($\mathbf{P}_\varepsilon$)**, *together with (5.1). Let κ_j, κ' be the oscillation bounds in (3.5) and (5.4), for a covering with trapezoids Δ_{jk}, $\Delta^{(j\ell)}$ produced by a post-processing algorithm. Then the difference between $u(T, \cdot)$ and the exact solution is bounded by*

$$\begin{aligned}\big\|u(T,\cdot) - S_T\bar u\big\|_{\mathbf{L}^1(\mathbb{R})} \leqq\; & C'\left(T + \sum_{j=0}^{\nu-1}\kappa_j\right)\varepsilon^{1/3} \\ & + C''\left(\varepsilon^{1/3}\kappa' + \varepsilon^{2/3}\right)\cdot\sum_{j=0}^{\nu-1} N'(j). \qquad (5.9)\end{aligned}$$

Proof. 1. Denoting L_0 as the Lipschitz constant of the semigroup at (1.7), we have

$$\|u(T,\cdot) - S_T\bar u\|_{\mathbf{L}^1} \;\leqq\; L_0\cdot\sum_{j=0}^{\nu-1}\Big\|u(t_{j+1},\cdot) - S_h u(t_j,\cdot)\Big\|_{\mathbf{L}^1}. \qquad (5.10)$$

For each j, in order to estimate the difference $u(t_{j+1}, \cdot) - S_h u(t_j, \cdot)$, we consider a covering of the strip $[t_j, t_{j+1}] \times \mathbb{R}$ by trapezoids $\Delta^{(j\ell)}$, $\ell = 1, \ldots N'(j)$ as in (5.6), and Δ_{jk}, $k = 1, \ldots, N(j)$, as in (5.7).

2. Recalling (4.3), we denote by $\{t_{j+1}\} \times [a_{j\ell}, b_{j\ell}]$ the upper boundaries of the trapezoids $\Delta^{(j\ell)}$. These are the trapezoids which contain one large traced shock. Moreover, we call $\{t_{j+1}\} \times [\widehat{c}_{jk}, \widehat{d}_{jk}]$ the upper boundaries of the remaining trapezoids Δ_{jk}. According to (5.7), this means that

$$[\widehat{c}_{jk},\, \widehat{d}_{jk}] = \Big[c_{jk} + h\lambda^+ + \varepsilon^{2/3}\;,\; d_{jk} + h\lambda^- - \varepsilon^{2/3}\Big].$$

3. The same argument used at (3.17) now yields an error bound on the set

$$V_j \doteq \bigcup_{k=1}^{N(j)} [\widehat{c}_{jk},\, \widehat{d}_{jk}].$$

Indeed, recalling the uniform bound (5.1) on the total variation, one obtains

$$\int_{V_j}\Big|u(t_{j+1},x) - \big(S_h u(t_j,\cdot)\big)(x)\Big|\,\mathrm{d}x \;\leqq\; C_3\left(\kappa_j\, h + \varepsilon^{2/3} + h\varepsilon^{1/3}\right)\cdot\delta_0. \qquad (5.11)$$

On the other hand, for each $\ell \in \{1, \ldots, N'(j)\}$, applying the estimate (4.11) on the interval $[a_{j\ell}, b_{j\ell}]$ we obtain the error bound

$$\int_{a_{j\ell}}^{b_{j\ell}}\big|u(h,x) - (S_h\bar u)(x)\big|\,\mathrm{d}x \;\leqq\; C_5\cdot h\left(\frac{\varepsilon}{\rho} + \kappa' + \frac{\rho\kappa' + \delta}{h}\right)$$

$$+C_5\,\big(\rho\kappa' + h\kappa' + \delta\big)\,. \tag{5.12}$$

4. Recalling our choices

$$\rho = h = \varepsilon^{1/3}, \qquad \delta = \varepsilon^{2/3}, \qquad \nu h = \nu\varepsilon^{1/3} = T,$$

summing the terms in (5.11) over all $j \in \{0, \ldots, \nu - 1\}$, and summing the terms in (5.12) over all j and all $\ell \in \{1, \ldots, N'(j)\}$, we obtain (5.9). □

Remark 4. It is interesting to speculate about the rate at which the error bound on the right hand side of (5.9) will approach zero as $\varepsilon \to 0$. We begin by assuming that the exact solution we are trying to compute is piecewise Lipschitz, with a finite number of centered rarefaction waves, and finitely many non-interacting shocks.

As is Remark 2, the first term on the right hand side of (5.9) is expected to approach zero as $\varepsilon^{1/3}|\log\varepsilon|$. Concerning the second term, we can fix a constant C_0 and choose $\kappa' = C_0\varepsilon^{1/3}$. If the exact solution contains N' shocks, we expect that, for all $\varepsilon > 0$ sufficiently small, each of these shocks will be traced, satisfying the inequality (5.2). The second term on the right hand side of (5.9) will thus have the form

$$C''\left(\varepsilon^{1/3}C_0\varepsilon^{1/3} + \varepsilon^{2/3}\right)\cdot \nu N' = \mathcal{O}(1)\cdot\varepsilon^{1/3}.$$

In this case, (5.9) would yield

$$\big\|u(T,\cdot) - S_T\bar{u}\big\|_{\mathbf{L}^1(\mathbb{R})} = \mathcal{O}(1)\cdot\varepsilon^{1/3}|\log\varepsilon|. \tag{5.13}$$

More generally, let us now assume that some of the shocks in the solution interact with each other. Let $\tau \in [0, T]$ be one of the (finitely many) interaction times. During a time interval $[\tau^-, \tau^+]$ around τ, of size $\tau^+ - \tau^- = \mathcal{O}(1)\cdot\rho$, we shall not be able to trace the interacting shocks. As a consequence, for $[t_j, t_{j+1}]\cap[\tau^-, \tau^+] \neq \emptyset$, the oscillation on one of the trapezoids Δ_{jk} in (5.7) (the one which contains a non-traced shock) will be large. This will force κ_j to be large. However, we expect that the total length of all intervals $[t_j, t_{j+1}]$, where some large shock cannot be traced, will have size

$$\mathcal{O}(1)\cdot\rho\cdot[\text{total number of shock interactions}] = \mathcal{O}(1)\cdot\varepsilon^{1/3}.$$

In conclusion, the presence of finitely many shock interactions will contribute an additional error term $\mathcal{O}(1)\cdot\varepsilon^{1/3}$ to the right hand side of (5.9). This will not change its overall order of magnitude.

One could also argue that, if the solution contains a finite number of compression waves, from which new shocks are formed, these (non-traced) waves would contribute an error term of the same nature as a centered rarefaction wave. Therefore, a bound of the order (5.13) would still be obtained.

Once again, we emphasize that the bounds (5.13) represent only a heuristic expectation. For a numerically computed solution, they needs to be confirmed by a post-processing algorithm, computing a bound on the oscillations κ_j in (3.5).

6. Properties of Approximation Schemes

In this section we analyze various approximation methods, and check that they all satisfy the assumptions **(AL)** and $(\mathbf{P}_\varepsilon)$.

6.1. The Godunov Scheme

To simplify our discussion, without loss of generality we assume that all characteristic speeds (that is all eigenvalues of the Jacobian matrices $Df(u)$) lie in the interior of the interval $[0, 1]$. This can be achieved by a linear rescaling and a shift of the coordinates t, x. In this case, the Godunov scheme reduces to an upwind scheme. Given a mesh size $\varepsilon > 0$, consider the grid points

$$P_{mj} = (t_m, x_j) = (\varepsilon m, \varepsilon j).$$

As shown in Fig. 6, left, we consider approximate solutions $u = u(t, x)$ with the following properties:

(i) At each time $t = t_m$, the function $u(t_m, \cdot)$ is piecewise constant, namely

$$u(t_m, x) = U_{mj} \qquad \text{for} \quad x_j < x < x_{j+1},$$

(ii) For $t \in [t_m, t_{m+1}[$, the function $u(t, \cdot)$ yields the exact solution to (1.1) with initial data $u(t_m, \cdot)$. This is obtained by solving the Riemann problems at each node x_j.

(iii) At time t_{m+1}, we take the average of $u(t_{m+1}-, \cdot)$ on each interval $[x_j, x_{j+1}]$. Namely

$$u(t_{m+1}, x) = U_{m+1,j} \doteq \frac{1}{\varepsilon} \int_{x_j}^{x_{j+1}} u(t_{m+1}, y)\, \mathrm{d}y \quad \text{for } x_j < x < x_{j+1}. \tag{6.1}$$

Since we are assuming that all wave speeds are contained in the interval $[0, 1]$, using the conservation equations these average values $U_{m+1,j}$ can be computed by

$$U_{m+1,\,j} = U_{m,j} + \Big(f(U_{m,j-1}) - f(U_{m,j}) \Big). \tag{6.2}$$

We check that an approximate solution u produced by the Godunov scheme with mesh size $\varepsilon > 0$ satisfies the Lipschitz condition **(AL)**. Indeed, for every $0 \leqq \tau < \tau' \leqq T$ with $\tau, \tau' \in \varepsilon\mathbb{N}$,

$$\begin{aligned} \big\| u(\tau', \cdot) - u(\tau, \cdot) \big\|_{\mathbf{L}^1(\mathbb{R})} &\leqq \sum_{\tau < t_m \leqq \tau'} \sum_j \int_{x_{j-1}}^{x_j} \big| u(t_{m+1}, x) - u(t_m, x) \big|\, \mathrm{d}x \\ &= \sum_{\tau < t_m \leqq \tau'} \sum_j \varepsilon \Big| f(U_{m,j-1}) - f(U_{m,j}) \Big| \\ &\leqq \varepsilon \sum_{\tau < t_m \leqq \tau'} \mathrm{Tot.Var.} \Big\{ f(u(t_m, \cdot)) \Big\} \end{aligned}$$

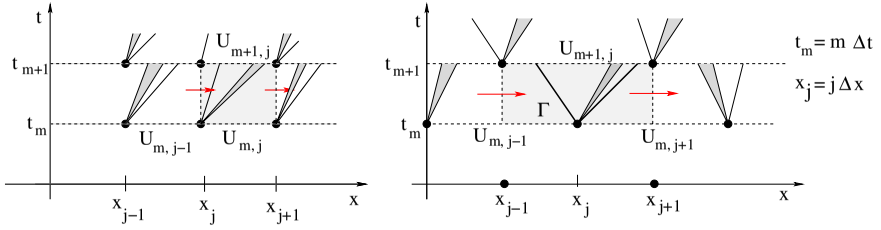

Fig. 6. Left: the upwind Godunov scheme is obtained by solving the Riemann problems at each node $P_{mj} = (t_m, x_j)$, then by replacing each solution by its average on each of the intervals $[x_j, x_{j+1}]$. By the conservation equation, these averages can be explicitly computed by (6.2). Right: a similar construction, on a staggered grid, leads to the Lax–Friedrichs scheme (6.12)

$$\leqq\ (\tau' - \tau) \cdot \mathrm{Lip}(f) \cdot \sup_{t\in[\tau,\tau']} \mathrm{Tot.Var.}\big\{u(t,\cdot)\big\}. \tag{6.3}$$

Here and in the sequel, $\mathrm{Lip}(f)$ denotes the Lipschitz constant of the function f.

To prove that the property $(\mathbf{P}_\varepsilon)$ also holds, we shall use

Lemma 3. *Let $w : [0, \varepsilon] \mapsto \mathbb{R}$ be any function with bounded variation, and assume $\varphi \in \mathcal{C}^1$. Consider the average value*

$$\overline{w} \doteq \frac{1}{\varepsilon}\int_0^\varepsilon w(y)\,\mathrm{d}y.$$

Then

$$\left|\int_0^\varepsilon [w(x) - \overline{w}]\,\varphi(x)\,\mathrm{d}x\right| \ \leqq\ \mathrm{Tot.Var.}\big\{w;\ ]0,\varepsilon[\,\big\}\cdot\varepsilon^2\|\varphi_x\|_{\mathbf{L}^\infty}. \tag{6.4}$$

Proof. Call $\overline{\varphi}$ the average value of φ over $[0, \varepsilon]$. Then

$$\begin{aligned}\left|\int_0^\varepsilon [w(x) - \overline{w}]\,\varphi(x)\,\mathrm{d}x\right| \ &=\ \left|\int_0^\varepsilon [w(x) - \overline{w}]\,(\varphi(x) - \overline{\varphi})\,\mathrm{d}x\right| \\ &\leqq \int_0^\varepsilon \|w(\cdot) - \overline{w}\|_{\mathbf{L}^\infty}\cdot\ \|\varphi(\cdot) - \overline{\varphi}\|_{\mathbf{L}^\infty}\,\mathrm{d}x \\ &\leqq \varepsilon\cdot\mathrm{Tot.Var.}\big\{w;\ ]0,\varepsilon[\,\big\}\cdot\varepsilon\|\varphi_x\|_{\mathbf{L}^\infty}. \end{aligned}\tag{6.5}$$

□

Next, fix $\varepsilon > 0$ and consider any test function $\varphi \in \mathcal{C}^1_c(\mathbb{R}^2)$. Since the Godunov approximations coincide with exact solutions on each of the half-open intervals $[t_m, t_{m+1}[$, we have

$$\begin{aligned}&\left|\int u(\tau,x)\varphi(\tau,x)\,\mathrm{d}x - \int u(\tau',x)\varphi(\tau',x)\,\mathrm{d}x\right. \\ &\qquad\left. + \int_\tau^{\tau'}\int \big\{u\varphi_t + f(u)\varphi_x\big\}\ \mathrm{d}x\,\mathrm{d}t\right| \\ &= \left|\sum_{\tau<t_m\leqq\tau'}\int [u(t_m,x) - u(t_m-,x)]\varphi(t_m,x)\,\mathrm{d}x\right|. \end{aligned}\tag{6.6}$$

Using Lemma 3 we obtain

$$\begin{aligned}&\sum_{\tau<t_m\leqq\tau'}\left|\int [u(t_m,x) - u(t_m-,x)]\,\varphi(t_m,x)\,\mathrm{d}x\right| \\ &\leqq \sum_{\tau<t_m\leqq\tau'}\sum_j \varepsilon^2\,\mathrm{Tot.Var.}\Big\{u(t_m,\cdot);\ ]j\varepsilon,(j+1)\varepsilon[\,\Big\}\cdot\|\varphi_x\|_{\mathbf{L}^\infty} \\ &= \varepsilon\,(\tau'-\tau)\|\varphi_x\|_{\mathbf{L}^\infty}\cdot\sup_{t\in[\tau,\tau']}\mathrm{Tot.Var.}\big\{u(t,\cdot)\big\}. \end{aligned}\tag{6.7}$$

This yields (1.11). In order to prove ($\mathbf{P}_\varepsilon$), given a convex entropy with entropy flux q, it remains to check that (1.12) is satisfied as well. Let $\varphi \geqq 0$ be a test function in $\mathcal{C}^1_c(\mathbb{R}^2)$. Integration by parts yields

$$\begin{aligned}
&\int \eta(u(\tau,x))\varphi(\tau,x)\,\mathrm{d}x - \int \eta(u(\tau',x))\varphi(\tau',x)\,\mathrm{d}x + \int_\tau^{\tau'}\int \eta(u)\varphi_t + q(u)\varphi_x\,\mathrm{d}t\,\mathrm{d}x\\
&= -\int_\tau^{\tau'}\int \{\eta(u)_t + q(u)_x\}\varphi\,\mathrm{d}x\,\mathrm{d}t\\
&\quad + \sum_{\tau<t_m\leqq\tau'}\int (\eta(u(t_m-,x)) - \eta(u(t_m,x)))\varphi(t_m,x)\,\mathrm{d}x. \qquad (6.8)
\end{aligned}$$

By construction, the approximation u is an entropy weak solution of the hyperbolic system of conservation law in every strip $[t_{m-1}, t_m[\times\mathbb{R}$, therefore the first term on the right hand side of (6.8) is non negative.

By the convexity of η, we can apply Jensen's inequality and obtain

$$\eta(u(t_m,x)) \;=\; \eta\Big(\frac{1}{\varepsilon}\int_{x_j}^{x_{j+1}} u(t_m-,y)\,\mathrm{d}y\Big) \;\leqq\; \frac{1}{\varepsilon}\int_{x_j}^{x_{j+1}} \eta(u(t_m-,y))\,\mathrm{d}y \qquad (6.9)$$

for $x_j < x < x_{j+1}$. In turn, this yields

$$\begin{aligned}
&\sum_{\tau<t_m\leqq\tau'}\int_{\mathbb{R}} (\eta(u(t_m-,x)) - \eta(u(t_m,x)))\varphi(t_m,x)\,\mathrm{d}x\\
&= \sum_{\tau<t_m\leqq\tau'}\sum_j \int_{x_j}^{x_{j+1}} (\eta(u(t_m-,x)) - \eta(u(t_m,x)))\varphi(t_m,x)\,\mathrm{d}x\\
&\geqq \sum_{\tau<t_m\leqq\tau'}\sum_j \int_{x_j}^{x_{j+1}} \Big(\eta(u(t_m-,x)) - \frac{1}{\varepsilon}\int_{x_j}^{x_{j+1}} \eta(u(t_m-,y))\,\mathrm{d}y\Big)\varphi(t_m,x)\,\mathrm{d}x\\
&\geqq -\operatorname{Lip}(\eta)\,\varepsilon\cdot(\tau'-\tau)\|\varphi_x\|_{\mathbf{L}^\infty}\cdot \sup_{t\in[\tau,\tau']} \operatorname{Tot.Var.}\big\{u(t,\cdot)\big\}. \qquad (6.10)
\end{aligned}$$

6.2. The Lax–Friedrichs Scheme

Consider step sizes $\Delta t, \Delta x > 0$ so that all characteristic speeds satisfy the CFL condition

$$|\lambda_i| \;<\; \frac{\Delta x}{\Delta t}. \qquad (6.11)$$

As shown in Fig. 6, right, we then construct a staggered grid with nodes at the points

$$P_{mj} = (m\,\Delta t,\; j\,\Delta x), \qquad\qquad m+j \text{ even}.$$

The Lax–Friedrichs approximations are defined inductively as follows. Given a piecewise constant function $u(t_m,\cdot)$, with jumps at the points P_{mj} with $m+j$ even,

for $t \in [t_m, t_{m+1}[$ we let $u(t, \cdot)$ be the exact solution of the system of conservation laws (1.1) with the given data at $t = t_m$. We then define $u(t_{m+1}, \cdot)$ to be the piecewise constant function obtained by taking the average of $u(t_{m+1}-, \cdot)$ over every interval $[x_{j-1}, x_{j+1}]$ with $m + j$ even. By the conservation equations, if all characteristic speeds satisfy $|\lambda_i| < \Delta x/\Delta t$, these average values

$$U_{m+1,j} = \frac{1}{2\Delta x} \int_{x_{j-1}}^{x_{j+1}} u(t_{m+1}-, x)\, \mathrm{d}x, \qquad m + j \text{ even},$$

are inductively computed by the Lax–Friedrichs scheme

$$U_{m+1,j} = \frac{1}{2}(U_{m,j+1} + U_{m,j-1}) - \frac{\Delta t}{2\Delta x}\big[f(U_{m,j+1}) - f(U_{m,j-1})\big]. \quad (6.12)$$

Setting $\varepsilon = \Delta t$, both the approximate Lipschitz condition **(AL)** and the property $(\mathbf{P}_\varepsilon)$ can be proved in the same way as for the Godunov scheme. We thus omit details.

Remark 5. If the stability condition (6.11) is violated, it is well known that the numerical algorithm becomes unstable. As a consequence, the numerical solution will develop a large amount of oscillations. In this case, the total variation becomes too large, the assumption (5.1) fails, and the error bound (5.9) does not apply.

6.3. Backward Euler Approximations

We now discretize time but keep space continuous. We assume that all characteristic speeds are strictly positive. Calling $\varepsilon = \Delta t$ the time step, the backward Euler approximations are defined in terms of the implicit equations

$$u(t + \varepsilon, x) = u(t, x) - \varepsilon f(u(t + \varepsilon, x))_x.$$

Equivalently, setting $U_m(x) = u(m\varepsilon, x)$, one needs to solve the sequence of ODEs

$$U_{m+1}(x) = U_m(x) - \varepsilon\, f(U_{m+1}(x))_x. \quad (6.13)$$

The convergence of these approximations for general $n \times n$ hyperbolic systems has not been studied. Complete results are available in the scalar case [20], which is covered by the general theory of nonlinear contractive semigroups [21].

Let $U_m(\cdot)$ be a sequence of solutions to (6.13) with $m = 0, 1, \ldots$ and define the approximate solution u by setting

$$u(t, x) = U_m(x) \qquad \text{for} \quad m\varepsilon \leqq t < (m + 1)\varepsilon.$$

Then **(AL)** follows by

$$\begin{aligned} \big\|u(\tau, \cdot) - u(\tau', \cdot)\big\|_{\mathbf{L}^1(\mathbb{R})} &\leqq \sum_{\tau < t_m \leqq \tau'} \big\|u(t_m, \cdot) - u(t_{m-1}, \cdot)\big\|_{\mathbf{L}^1(\mathbb{R})} \\ &= \varepsilon \sum_{\tau < t_m \leqq \tau'} \int \big|f(u(t_m, x))_x\big|\, \mathrm{d}x \leqq \mathrm{Lip}(f)\,(\tau' - \tau) \cdot \Big(\sup_{t \in [\tau, \tau']} \mathrm{Tot.Var.}\{u(t, \cdot)\}\Big). \end{aligned}$$

Next, we verify ($\mathbf{P}_\varepsilon$). As before, fix $0 \leqq \tau < \tau' \leqq T$ with $\tau, \tau' \in \varepsilon\mathbb{N}$. Given a test function $\varphi \in \mathcal{C}^1_c(\mathbb{R}^2)$, we compute that

$$
\begin{aligned}
&\left| \int u(\tau, x)\varphi(\tau, x)\,\mathrm{d}x - \int u(\tau', x)\varphi(\tau', x)\,\mathrm{d}x + \int_\tau^{\tau'} \int \big\{u\varphi_t + f(u)\varphi_x\big\}\,\mathrm{d}x\,\mathrm{d}t \right| \\
&= \left| -\int_\tau^{\tau'} \int f(u)_x \varphi\,\mathrm{d}x\,\mathrm{d}t + \int \sum_{\tau < t_m \leqq \tau'} \Big(u(t_{m-1}, x) - u(t_m, x)\Big)\varphi(t_m, x)\,\mathrm{d}x \right| \\
&= \left| -\int \left(\sum_{\tau < t_m \leqq \tau'} \frac{u(t_{m-1}, x) - u(t_m, x)}{\varepsilon} \int_{t_{m-1}}^{t_m} \varphi(t, x)\,\mathrm{d}t \right) dx \right. \\
&\quad \left. + \int \sum_{\tau < t_m \leqq \tau'} \Big(u(t_{m-1}, x) - u(t_m, x)\Big)\varphi(t_m, x)\,\mathrm{d}x \right| \\
&\leqq \int \sum_{\tau < t_m \leqq \tau'} \big|u(t_{m-1}, x) - u(t_m, x)\big| \left| \varphi(t_m, x) - \frac{1}{\varepsilon} \int_{t_{m-1}}^{t_m} \varphi(t, x)\,\mathrm{d}t \right| \mathrm{d}x \\
&\leqq \varepsilon\,(\tau' - \tau)\,\|\varphi_t\|_{\mathbf{L}^\infty}\,\mathrm{Lip}(f) \cdot \Big(\sup_{t \in [\tau, \tau']} \mathrm{Tot.Var.}\{u(t, \cdot)\} \Big).
\end{aligned}
$$

Finally, let η be a convex entropy with entropy flux q. If $\varphi \geqq 0$ is a test function in $\mathcal{C}^1_c(\mathbb{R}^2)$, we have

$$
\begin{aligned}
&\int \eta(u(\tau, x))\varphi(\tau, x)\,\mathrm{d}x - \int \eta(u(\tau', x))\varphi(\tau', x)\,\mathrm{d}x + \int_\tau^{\tau'} \int \{\eta(u)\varphi_t + q(u)\varphi_x\}\,\mathrm{d}x\,\mathrm{d}t \\
&= -\int_\tau^{\tau'} \int q(u)_x \varphi\,\mathrm{d}t\,\mathrm{d}x + \int \sum_{\tau < t_m \leqq \tau'} \Big(\eta(u(t_{m-1}, x)) - \eta(u(t_m, x))\Big)\varphi(t_m, x)\,\mathrm{d}x \\
&= -\sum_{\tau < t_m \leqq \tau'} \int \Big(D\eta(u(t_m, x)) \cdot \frac{u(t_{m-1}, x) - u(t_m, x)}{\varepsilon} \int_{t_{m-1}}^{t_m} \varphi(t, x)\,\mathrm{d}t \Big)\,\mathrm{d}x \\
&\quad + \int \sum_{\tau < t_m \leqq \tau'} \Big(\eta(u(t_{m-1}, x)) - \eta(u(t_m, x))\Big)\varphi(t_m, x)\,\mathrm{d}x \\
&\geqq \sum_{\tau < t_m \leqq \tau'} \int \Big(\eta(u(t_{m-1}, x)) - \eta(u(t_m, x))\Big)\Big(\varphi(t_m, x) - \frac{1}{\varepsilon} \int_{t_{m-1}}^{t_m} \varphi(t, x)\,\mathrm{d}t\Big)\,\mathrm{d}x \\
&\geqq -\varepsilon\,(\tau' - \tau)\,\|\varphi_t\|_{\mathbf{L}^\infty}\,\mathrm{Lip}(\eta)\mathrm{Lip}(f) \cdot \Big(\sup_{t \in [\tau, \tau']} \mathrm{Tot.Var.}\{u(t, \cdot)\} \Big).
\end{aligned}
$$

Notice that the convexity of η was here used in the inequality

$$
D\eta(u(a, x)) \cdot \big(u(a, x) - u(b, x)\big) \;\geqq\; \eta(u(a, x)) - \eta(u(b, x)). \tag{6.14}
$$

6.4. The Smoothing Method

Next, we consider an approximate solution to (1.1), (1.2) obtained by periodic mollifications, taking the convolution with a smoothing kernel $K \in \mathcal{C}_c^\infty(\mathbb{R})$. We assume that

$$\begin{cases} K(x) \in [0,1] \text{ for } |x| < 1, \\ K(x) = 0 \quad \text{for } |x| \geqq 1, \end{cases} \qquad K(x) = K(-x), \qquad \int K(x)\,dx = 1, \tag{6.15}$$

and set $K_\delta(x) \doteq \delta^{-1} K(\delta^{-1}x)$. We fix a time step $\varepsilon > 0$ and define an ε-approximate solution u by setting

$$t_m = m\,\varepsilon, \qquad u(t_m, \cdot) = K_\delta * u(t_m-, \cdot),$$

and letting u be a classical solution to (1.1) on each half-open interval $[t_m, t_{m+1}[$.

As in the scalar case (see [26]), the method is well-defined provided that the ratio ε/δ is suitably small. To see this, in connection with the quasilinear system

$$u_t + A(u)u_x = 0, \qquad A(u) = Df(u), \tag{6.16}$$

we choose bases $\{r_1(u), \ldots, r_n(u)\}$ and $\{l_1(u), \ldots, l_n(u)\}$ of right and left eigenvectors of $A(u)$, normalized so that

$$|r_i(u)| = 0, \qquad l_i(u) \cdot r_i(u) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{6.17}$$

We denote by $u_x^i = l_i \cdot u_x$ the ith component of the gradient vector u_x with respect to this basis. From (6.17) and (6.16) it follows that

$$u_x = \sum_{i=1}^n u_x^i r_i(u) \qquad u_t = -\sum_{i=1}^n \lambda_i(u) u_x^i r_i(u).$$

Differentiating the first equation with respect to t and the second one with respect to x, then equating the results, one obtains a semilinear system of evolution equations for the scalar components u_x^i, having the form

$$(u_x^i)_t + \lambda_i (u_x^i)_x = \sum_{j,k=1}^n g_{jk}^i(u) u_x^j u_x^k \tag{6.18}$$

(see, for example, Section 1.6 in [9] for details). Assume that

$$\left|g_{jk}^i(u)\right| < M_g$$

for all i, j, k, and all u in the domain were the solution is defined. Let $t \mapsto Z(t)$ be the solution to the ODE

$$\frac{d}{dt} Z(t) = n^2 M_g\, Z^2(t), \qquad Z(0) = Z_0. \tag{6.19}$$

Assume that, at time $t = 0$, it holds that

$$\left|u_x^i(0,x)\right| \leqq Z_0 \qquad \text{for all } x \in \mathbb{R}, \quad i = 1, \ldots, n. \tag{6.20}$$

A comparison argument now yields

$$\left|u^i_x(t,x)\right| \leqq Z(t) \qquad \text{for all } x \in \mathbb{R},\ t \in [0, T_0[,$$

where

$$T_0 = \frac{1}{n^2 M_g\, Z_0}$$

is the time where the solution to (6.19) blows up.
It remains to give an upper bound for the gradient components after each mollification. This is achieved observing that

$$\begin{aligned}\|u^i_x(t_m,\cdot)\|_{\mathbf{L}^\infty} &\leqq \|u_x(t_m,\cdot)\|_{\mathbf{L}^\infty}\cdot \sup_u |l_i(u)|\\ &\leqq \frac{\text{Tot.Var.}\{u(t_m-,\cdot)\}}{\delta}\,\|K\|_{\mathbf{L}^\infty}\cdot \sup_u |l_i(u)|.\end{aligned}$$

Therefore, if we choose

$$0<\varepsilon<\frac{\delta}{n^2\, M_g\,\|K\|_{\mathbf{L}^\infty}}\Big(\sup_m \text{Tot.Var.}\{u(t_m-,\cdot)\}\Big)^{-1}\Big(\sup_i \sup_u |l_i(u)|\Big)^{-1}, \tag{6.21}$$

all the components u^i_x remain bounded on each strip $[t_m, t_{m+1}[\times\mathbb{R}$, and the approximate solution is well-defined.

We now check that the assumption **(AL)** holds:

$$\begin{aligned}&\big\|u(\tau',\cdot)-u(\tau,\cdot)\big\|_{\mathbf{L}^1(\mathbb{R})}\\ &\leqq \sum_{\tau<t_m\leqq\tau'}\int\big|\,K_\delta * u(t_m-,x)-u(t_m-,x)\big|\,\mathrm{d}x+\int\big|u(t_m-,x)-u(t_{m-1},x)\big|\,\mathrm{d}x\\ &\leqq 2\delta\Big(\frac{\tau'-\tau}{\varepsilon}\Big)\cdot\sup_{t\in[\tau,\tau']}\text{Tot.Var.}\big\{u(t,\cdot)\big\}+L\,(\tau'-\tau)\cdot\sup_{t\in[\tau,\tau']}\text{Tot.Var.}\big\{u(t,\cdot)\big\}\\ &= C\,(\tau'-\tau)\cdot\sup_{t\in[\tau,\tau']}\text{Tot.Var.}\big\{u(t,\cdot)\big\}.\end{aligned}$$

Here we are using the estimate

$$\begin{aligned}&\int\big|\,K_\delta * u(t_m-,x)-u(t_m-,x)\big|\,\mathrm{d}x\\ &\leqq \int\left(\int_{-\delta}^{\delta}K_\delta(y)\sup_{|y|<\delta}\Big|u(t_m-,x)-u(t_m-,x-y)\Big|\,\mathrm{d}y\right)\mathrm{d}x\\ &\leqq \int\text{Tot.Var.}\Big\{u(t_m-,\cdot);\ ]x-\delta,x+\delta[\,\Big\}\,\mathrm{d}x\\ &\leqq 2\delta\cdot\text{Tot.Var.}\{u(t_m-,\cdot)\}.\end{aligned}\tag{6.22}$$

To prove $(\mathbf{P}_\varepsilon)$ we shall need the following result:

Lemma 4. *Let $w : \mathbb{R} \mapsto \mathbb{R}$ be any function with bounded variation, and assume $\varphi \in \mathcal{C}^1$. Let $K \in \mathcal{C}_c^\infty$ be a smoothing kernel as in (6.15), and define*

$$\widetilde{w} \doteq K_\delta * w.$$

Then

$$\left| \int [\widetilde{w}(x) - w(x)]\, \varphi(x)\, \mathrm{d}x \right| \leqq \delta^2 \, \|\varphi_x\|_{\mathbf{L}^\infty} \cdot \text{Tot.Var.}\{w\}. \tag{6.23}$$

Proof. 1. Assuming that w is Lipschitz continuous, we rewrite the left hand side of (6.23) in a more suitable way:

$$\begin{aligned} &\left| \int [\widetilde{w}(x) - w(x)]\, \varphi(x)\, \mathrm{d}x \right| \\ &= \left| \iint K_\delta(x-y)[w(y) - w(x)]\, \varphi(x)\, \mathrm{d}y\, \mathrm{d}x \right| \\ &= \left| \iint K_\delta(x-y)[\varphi(y) - \varphi(x)]\, w(x)\, \mathrm{d}y\, \mathrm{d}x \right| \\ &= \left| \int \left(\int_{-\infty}^{x} \int K_\delta(z-y)[\varphi(y) - \varphi(z)]\, \mathrm{d}y\, \mathrm{d}z \right) w'(x)\, \mathrm{d}x \right|. \end{aligned} \tag{6.24}$$

2. Next we prove the estimate

$$\left| \int_{-\infty}^{x} \int K_\delta(z-y)[\varphi(y) - \varphi(z)]\, \mathrm{d}y\, \mathrm{d}z \right| \leqq \delta^2 \, \|\varphi_x\|_{\mathbf{L}^\infty}. \tag{6.25}$$

Indeed, since the integrand vanishes for $|z - y| > \delta$, the domain where it can be nonzero can be split as $\Sigma_0 \cup \Sigma$, where (see Fig. 7)

$$\Sigma_0 = \Big\{ (z, y);\ \ |z - y| \leqq \delta, \quad y + z < x - \delta \Big\},$$

$$\Sigma = \Big\{ (z, y);\ \ z \leqq x, \quad y - z \leqq \delta, \quad y + z \geqq x - \delta \Big\}.$$

Since K is an even function, by symmetry (switching the variables z and y), we immediately obtain

$$\iint_{\Sigma_0} K_\delta(z-y)[\varphi(y) - \varphi(z)]\, \mathrm{d}y\, \mathrm{d}z = 0. \tag{6.26}$$

On the other hand, the integral over the triangle Σ can be estimated by

$$\begin{aligned} &\left| \iint_{\Sigma} K_\delta(z-y)[\varphi(y) - \varphi(z)]\, \mathrm{d}y\, \mathrm{d}z \right| \\ &\quad \leqq \text{meas}(\Sigma) \cdot \|K_\delta\|_{\mathbf{L}^\infty} \cdot \sup_{(z,y)\in\Sigma} |\varphi(y) - \varphi(z)| \\ &\quad \leqq \delta^2 \cdot \delta^{-1} \|K\|_{\mathbf{L}^\infty} \cdot \delta \|\varphi_x\|_{\mathbf{L}^\infty}. \end{aligned} \tag{6.27}$$

Putting together (6.26) and (6.27), and recalling that $\|K\|_{\mathbf{L}^\infty} \leqq 1$, we obtain (6.25).

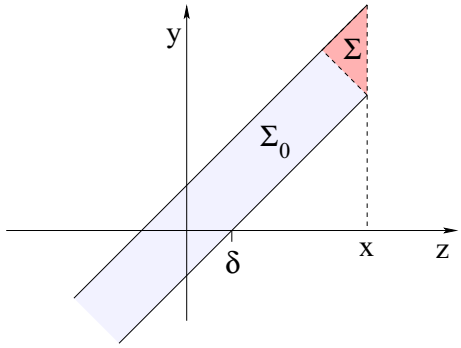

Fig. 7. Splitting the domain of integration $\Sigma_0 \cup \Sigma$, in the proof of (6.25)

3. Using (6.25) in (6.24), we now obtain the bound (6.23), for every Lipschitz function w. Since the previous arguments do not depend on the Lipschitz constant of w, by an approximation argument we conclude that (6.23) remains valid for every function w with bounded variation. □

We are now ready to prove $(\mathbf{P}_\varepsilon)$. Consider a test function $\varphi \in \mathcal{C}^1_c(\mathbb{R}^2)$. Using Lemma 4, we obtain

$$\begin{aligned}
&\left| \int u(\tau,x)\varphi(\tau,x)\,\mathrm{d}x - \int u(\tau',x)\varphi(\tau',x)\,\mathrm{d}x + \int_\tau^{\tau'} \int \big\{u\varphi_t + f(u)\varphi_x\big\}\,\mathrm{d}x\,\mathrm{d}t \right| \\
&= \left| \int \sum_{\tau < t_j \leqq \tau'} \big(u(t_j,x) - u(t_j-,x)\big)\,\varphi(t_j,x)\mathrm{d}x \right| \\
&\leqq \delta^2\Big(\frac{\tau'-\tau}{\varepsilon}\Big)\|\varphi_x\|_{\mathbf{L}^\infty} \cdot \sup_{t\in[\tau,\tau']} \text{Tot.Var.}\{u(t,\cdot)\} \\
&= C_1\,\varepsilon\,(\tau'-\tau)\|\varphi_x\|_{\mathbf{L}^\infty} \cdot \sup_{t\in[\tau,\tau']} \text{Tot.Var.}\{u(t,\cdot)\}.
\end{aligned}$$

In view of (6.21), here we can choose the constant $C_1 = \delta^2/\varepsilon^2$.

Finally, let η be a convex entropy with entropy flux q. For any non-negative test function $\varphi \in \mathcal{C}^1_c(\mathbb{R}^2)$ one has

$$\begin{aligned}
&\int \eta(u(\tau,x))\varphi(\tau,x)\,\mathrm{d}x - \int \eta(u(\tau',x))\varphi(\tau',x)\,\mathrm{d}x + \int_\tau^{\tau'}\int \big\{\eta(u)\varphi_t + q(u)\varphi_x\big\}\,\mathrm{d}x\,\mathrm{d}t \\
&= -\int \sum_{\tau < t_m \leqq \tau'} \Big(\eta(u(t_m,x)) - \eta(u(t_m-,x))\Big)\,\varphi(t_m,x)\,\mathrm{d}x \\
&\geqq -\int \sum_{\tau < t_m \leqq \tau'} D\eta(u(t_m,x))\big(u(t_m,x) - u(t_m-,x)\big)\,\varphi(t_m,x)\,\mathrm{d}x \\
&\geqq -\delta^2\Big(\frac{\tau'-\tau}{\varepsilon}\Big)\|\varphi_x\|_{\mathbf{L}^\infty}\,\mathrm{Lip}(\eta) \cdot \sup_{t\in[\tau,\tau']} \text{Tot.Var.}\{u(t,\cdot)\} \\
&= -C_1\,\varepsilon\,(\tau'-\tau)\|\varphi_x\|_{\mathbf{L}^\infty}\,\mathrm{Lip}(\eta) \cdot \sup_{t\in[\tau,\tau']} \text{Tot.Var.}\{u(t,\cdot)\},
\end{aligned}$$

where the first inequality follows from the convexity of η, by (6.14).

7. Numerical Implementation

In this last section we discuss details of the post-processing algorithm, and present a numerical simulation.

STEP 1 of the algorithm, computing the total variation of the numerical solution $u(t, \cdot)$, is entirely straightforward.

STEP 2, identifying the location of the large shocks, requires more attention. Given a pair of constants $M > 0$ and $\sigma >> \varepsilon > 0$, we first identify regions where the total variation of u is large. For this purpose, we introduce

Definition 2. For a given function $u : [0, T] \times \mathbb{R} \mapsto \mathbb{R}^n$, the points (t, x) such that

$$\begin{aligned} \min \Big\{ &\text{Tot.Var.}\big\{u(t, \cdot);\ [x - \sigma, x + \varepsilon]\big\}, \\ &\text{Tot.Var.}\big\{u(t, \cdot);\ [x - \varepsilon,\ x + \sigma]\big\}\Big\} > M\sigma \end{aligned} \tag{7.1}$$

will be called **flagged points**.

Notice that, by definition, the variation of $u(t, \cdot)$ on a small interval to the left or to the right of a flagged point must be large. Roughly speaking, the following result shows that, outside flagged points, solutions are approximately Lipschitz continuous with constant $2M$.

Lemma 5. *Assume that all points (t, x) with $x \in [a, b]$ are not flagged. Then*

$$|u(t, a) - u(t, b)| \leqq \left(1 + \frac{b - a}{\sigma + \varepsilon}\right) 2\sigma M. \tag{7.2}$$

Proof. 1. By assumption, we can split the interval as $[a, b] = I^+ \cup I^-$, where I^+, I^- are two disjoint sets with the following property. Setting

$$J_x = \begin{cases} [x - \varepsilon, x + \sigma] \text{ if } \ x \in I^+, \\ [x - \sigma, x + \varepsilon] \text{ if } \ x \in I^-, \end{cases}$$

one has

$$\text{Tot.Var.}\{u(t, \cdot);\ J_x\} \leqq \sigma M \qquad \text{for all } x \in [a, b].$$

2. We claim that every subinterval $[c, d] \subseteq [a, b]$ with length $d - c \leqq \sigma + \varepsilon$ can be covered by two of the intervals J_x. Indeed, three cases can arise:

CASE 1: $c + \varepsilon \in I^+$. Then $[c, d] \subseteq J_{c+\varepsilon}$.

CASE 2: $d - \varepsilon \in I^-$. Then $[c, d] \subseteq J_{d-\varepsilon}$.

CASE 3: $c + \varepsilon \in I^-$ and $d - \varepsilon \in I^+$. Then we can find to points $c + \varepsilon \leqq x < y \leqq d - \varepsilon$, such that

$$y - x < \varepsilon, \qquad x \in I^-,\ y \in I^+.$$

In this case, $[c, d] \subseteq J_x \cup J_y$, proving our claim.

3. To complete the proof, we cover $[a, b]$ with finitely many intervals $[c_i, d_i]$, $i = 1, \ldots, N$, such that

$$N \leqq 1 + \frac{b-a}{\sigma+\varepsilon}, \qquad d_i - c_i \leqq \sigma + \varepsilon \qquad \text{for all } i = 1, \ldots, N,$$

By the previous construction, for every i we have that

$$\text{Tot.Var.}\big\{u(t,\cdot);\ [c_i, d_i]\big\} \leqq 2M\sigma.$$

Therefore, the total variation of $u(t, \cdot)$ over $[a, b]$ is bounded by $2N\sigma M$. This yields (7.2). □

Having defined the set $\mathcal{F} \subseteq [0, T] \times \mathbb{R}$ of all flagged points, for each $t_j = j\varepsilon \in [0, T]$, $j = 0, 1, \ldots, \nu$, we denote by

$$\mathcal{F}_j \doteq \big\{x \in \mathbb{R};\ \ (t_j, x) \in \mathcal{F}\big\}$$

the set of flagged points at time t_j.

For every time t_j, we identify intervals $[a, b]$ such that $b - a \leqq \delta \doteq \varepsilon^{2/3}$ and moreover

$$a \in \mathcal{F}_j, \qquad b \in \mathcal{F}_j, \qquad \mathcal{F}_j \cap [a-\rho, a[= \mathcal{F}_j \cap]b, b+\rho] = \emptyset.$$

In other words, the points (t_j, a) and (t_j, b) are flagged, but points to the left of a and to the right of b are not flagged.

Each such interval $[a, b]$ locates a possible isolated shock at time t_j. To check if this shock can be traced over the entire interval $[t_j, t_{j+1}]$, we check if there exists an interval $[c, d]$ satisfying the same properties at time $t = t_{j+1}$, and moreover

$$[c, d] \ \subset\ [a + \lambda^- h\ ,\ b + \lambda^+ h].$$

In the positive case, we approximate the shock location with a straight line as follows:

$$\gamma(t) = x_0 + \lambda(t - t_j), \qquad t \in [t_j, t_{j+1}].$$

Here we choose x_0, λ so that

$$x_0 = \gamma(t_j) = \frac{a+b}{2}, \qquad \gamma(t_{j+1}) = \frac{c+d}{2}.$$

We then consider the polygonal regions Γ, Δ_l, Δ'_r defined as in (5.3). If the two inequalities (5.4), (5.5) are both satisfied, we say that the parallelogram

$$\Gamma = \Big\{(t, x);\ t \in [t_j, t_{j+1}],\ \ x_0 - \delta + (t - t_j)\lambda \leqq x \leqq x_0 + \delta + (t - t_j)\lambda\Big\},$$

traces the shock. The trapezoid

$$\Delta' = \Big\{(t, x);\ t \in [t_j, t_{j+1}],\ \ a' + \lambda^+(t - t_j) \leqq x \leqq b' + \lambda^-(t - t_j)\Big\},$$

where

$$a' \doteq x_0 - \rho - \delta - (\lambda^+ - \lambda^-)h, \qquad b' \doteq x_0 + \rho + \delta + (\lambda^+ - \lambda^-)h,$$

is then inserted within the list of trapezoids $\Delta^{(j\ell)}$ in (5.6), containing a traced shock. On the other hand, if one of the inequalities (5.4), (5.5) fails, the shock is not traced.

STEP 3 of the algorithm provides a covering of each domain

$$\Big([t_j, t_{j+1}] \times \mathbb{R}\Big) \setminus \bigcup_{\ell=1}^{N'(j)} \Delta^{(j\ell)}, \qquad j = 0, 1, \ldots, \nu - 1,$$

with finitely many trapezoids Δ_{jk} as in (5.7). This step is straightforward. The algorithm terminates by computing the constants κ_j in (5.8), which provide an upper bound on the oscillation of u on each Δ_{jk}, $k = 1, \ldots, N(j)$.

Example. We consider a model of isentropic gas dynamics in Lagrangian coordinates. Using a shifted system of coordinates, this can be written as

$$\begin{cases} v_t - u_x + v_x = 0, \\ u_t + \left(\dfrac{1}{2v^2}\right)_x + u_x = 0. \end{cases} \tag{7.3}$$

Here u is the velocity of the gas, while v denotes the specific volume. By the choice of coordinates, the characteristic speeds are

$$\lambda^{\pm} = 1 \pm v^{-3/2}.$$

In particular, when $v \geqq 1$, one has $\lambda(v) \in [\lambda^-, \lambda^+] \doteq [0, 2]$. We consider the Cauchy problem with piecewise constant initial data

$$v(0, x) = \begin{cases} 2 \text{ if } x < 0, \\ 3 \text{ if } 0 < x < 1/2, \\ 1 \text{ if } x > 1/2, \end{cases} \qquad u(0, x) = 0. \tag{7.4}$$

The exact solution is shown in Fig. 8.

We compute an approximate solution using the upwind scheme with mesh sizes

$$\Delta x = \varepsilon = 0.0005, \qquad \Delta t = \frac{\varepsilon}{2} = 0.00025.$$

The profiles of the two components of the solution, at the final time $T = 1.5$, are shown in Fig. 9.

To illustrate how the post-processing algorithm works, in Fig. 10, left, we plot the set of flagged points. These are computed according to Definition 2, choosing $\sigma = 0.0063$ and $M = 25$. In Fig. 10, right, we identify the shocks that can be traced on each time interval $[t_j, t_{j+1}]$. Here $\kappa' = 0.1$, $\sigma_{min} = 0.4$. Notice that, according to our previous construction, each trapezoid $\Delta^{(j\ell)}$ around a traced shock will have the form

$$\Delta^{(j\ell)} = \Big\{(t, x);\ \ t \in [t_j,\, t_{j+1}],\ \ x_{j\ell} - \delta - \rho - 2h$$

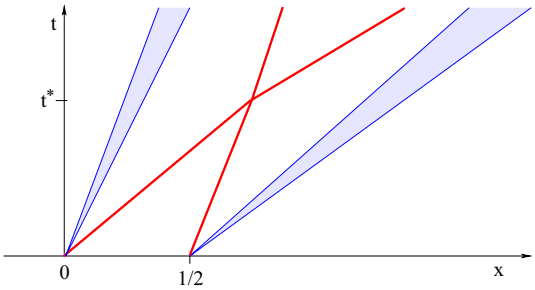

Fig. 8. A sketch of the exact solution to (7.3), (7.4), containing two centered rarefaction waves and two shocks, interacting at time t^*

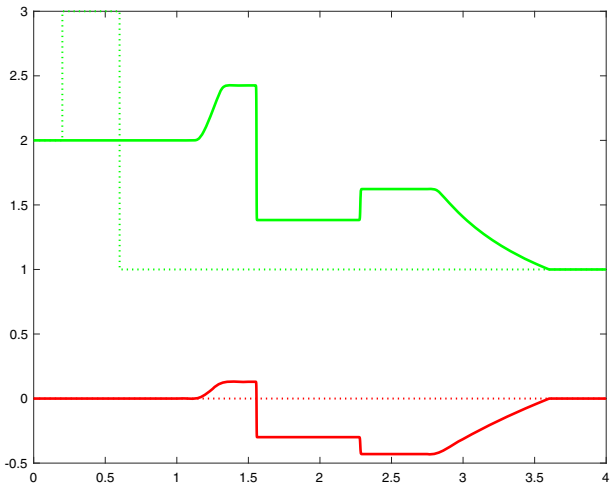

Fig. 9. The components of the solution at the terminal time $T = 1.5$, computed by the upwind scheme. Above: the specific volume $v(T, \cdot)$. Below: the velocity $u(T, \cdot)$. The dotted lines show the initial data

$$+2(t - t_j) \;\leqq\; x \;\leqq\; x_{j\ell} + \delta + \rho + 2h\Big\}.$$

Indeed, this is obtained from (4.3), (4.4), with

$$a' = x_{j\ell} - \delta - \rho - 2h, \qquad b' = x_{j\ell} + \delta + \rho + 2h, \qquad \lambda^- = 0, \qquad \lambda^+ = 2.$$

Finally, in Fig. 11 we plot an approximate graph of the function

$$\kappa(t) = \kappa_j \;\doteq\; \max_{1\leqq k\leqq N(j)} \text{Osc.}\{u;\ \Delta_{jk}\} \qquad \text{if } t \in [t_j, t_{j+1}[. \tag{7.5}$$

One can think of $\kappa(t)$ as the maximum oscillation of the numerical solution $u(t, \cdot)$ on domains of diameter $\mathcal{O}(1) \cdot \varepsilon^{1/3}$, outside the large traced shocks. In view of (5.9), this function $\kappa(\cdot)$ determines the rate at which the distance $\|u(t, \cdot) - S_t\bar{u}\|_{\mathbf{L}^1}$ between the approximate and the exact solution increases in time. Notice that $\kappa(t)$ is large for $t \approx 0$, when the main contribution to the error comes from the two

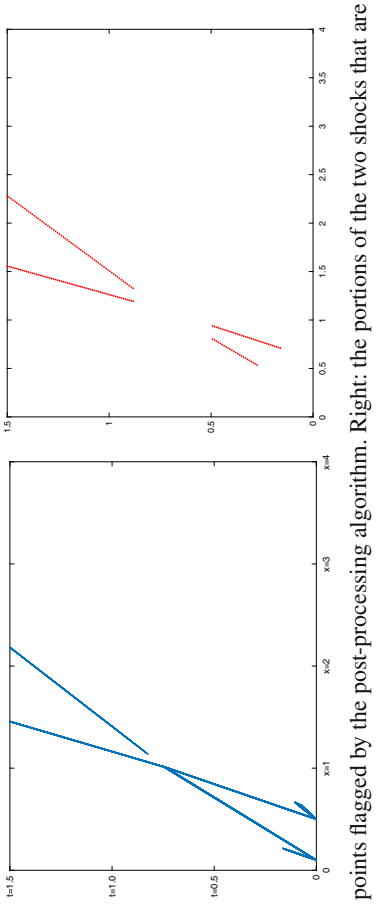

Fig. 10. Left: the points flagged by the post-processing algorithm. Right: the portions of the two shocks that are actually traced

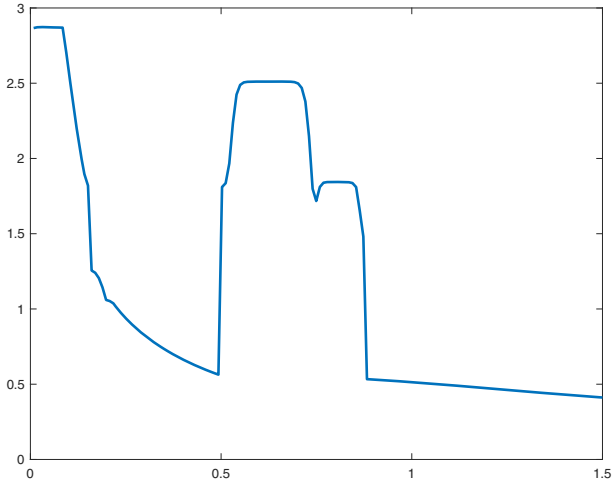

Fig. 11. An approximate computation of the function $\kappa(t)$ defined at (7.5), determining the error rate

centered rarefactions. As time increases, the rarefactions decay, and the value of $\kappa(t)$ decays as well. As t approaches the interaction time t^*, the two shocks cannot be individually traced. As a consequence, the value of $\kappa(t)$ suddenly becomes very large. Finally, when the shocks move away from each other and can be traced once again, we see that $\kappa(t)$ reverts to a small value.

Acknowledgements. This research was partially supported by NSF with Grant DMS-2006884, "Singularities and error bounds for hyperbolic equations".

References

1. ANCONA, F., MARSON, A.: Sharp convergence rate of the Glimm scheme for general nonlinear hyperbolic systems. *Commun. Math. Phys*. **302**, 581–630, 2011
2. BAITI, P., BRESSAN, A., JENSSEN, H.K.: BV instability of the Godunov scheme. *Commun. Pure Appl. Math*. **59**, 1604–1638, 2006
3. BIANCHINI, S.: BV solutions of the semidiscrete upwind scheme. *Arch. Ration. Mech. Anal*. **167**, 1–81, 2003
4. BIANCHINI, S.: Hyperbolic limit of the Jin–Xin relaxation model. *Commun. Pure Appl. Math*. **59**, 688–753, 2006
5. BIANCHINI, S., BRESSAN, A.: Vanishing viscosity solutions of nonlinear hyperbolic systems. *Ann. Math*. **161**, 223–342, 2005
6. BIANCHINI, S., MODENA, S.: Convergence rate of the Glimm scheme. *Bull. Inst. Math. Acad. Sin*. **11**, 235–300, 2016
7. BRESSAN, A.: The unique limit of the Glimm scheme. *Arch. Ration. Mech. Anal*. **130**, 205–230, 1995

8. Bressan, A.: *Hyperbolic Systems of Conservation Laws. The One-Dimensional Cauchy Problem*. Oxford University Press, Oxford 2000
9. Bressan, A.: Hyperbolic conservation laws: an illustrated tutorial. In: Ambrosio, L., Bressan, A., Helbing, D., Klar, A., Zuazua, E. (eds.) *Modelling and Optimisation of Flows on Networks. Springer Lecture Notes in Mathematics* vol. 2062, pp. 157–245, 2012
10. Bressan, A., Colombo, R.M.: The semigroup generated by 2×2 conservation laws. *Arch. Ration. Mech. Anal*. **113**, 1–75, 1995
11. Bressan, A., Crasta, G., Piccoli, B.: Well posedness of the Cauchy problem for $n \times n$ systems of conservation laws. *Am. Math. Soc. Mem.* **694**, 2000
12. Bressan, A., Goatin, P.: Oleinik type estimates and uniqueness for $n \times n$ conservation laws. *J. Differ. Equ*. **156**, 26–49, 1999
13. Bressan, A., Huang, F., Wang, Y., Yang, T.: On the convergence rate of vanishing viscosity approximations for nonlinear hyperbolic systems. *SIAM J. Math. Anal*. **44**, 3537–3563, 2012
14. Bressan, A., Jenssen, H.K.: On the convergence of Godunov scheme for nonlinear hyperbolic systems. *Chin. Ann. Math. B* **21**, 1–16, 2000
15. Bressan, A., LeFloch, P.: Uniqueness of weak solutions to systems of conservation laws. *Arch. Ration. Mech. Anal*. **140**, 301–317, 1997
16. Bressan, A., Lewicka, M.: A uniqueness condition for hyperbolic systems of conservation laws. *Discrete Contin. Dyn. Syst*. **6**, 673–682, 2000
17. Bressan, A., Liu, T.P., Yang, T.: L^1 stability estimates for $n \times n$ conservation laws. *Arch. Ration. Mech. Anal*. **149**, 1–22, 1999
18. Bressan, A., Marson, A.: Error bounds for a deterministic version of the Glimm scheme. *Arch. Ration. Mech. Anal*. **142**, 155–176, 1998
19. Bressan, A., Yang, T.: On the rate of convergence of vanishing viscosity approximations. *Commun. Pure Appl. Math*. **57**, 1075–1109, 2004
20. Crandall, M.G.: The semigroup approach to first order quasilinear equations in several space variables. *Isr. J. Math*. **12**, 108–132, 1972
21. Crandall, M.G., Liggett, T.M.: Generation of semigroups of nonlinear transformations on general Banach spaces. *Am. J. Math*. **93**, 265–298, 1971
22. Dafermos, C.: *Hyperbolic Conservation Laws in Continuum Physics*, 4th edn. Springer, Berlin 2016
23. Ding, X., Chen, G.Q., Luo, P.: Convergence of the fractional step Lax–Friedrichs scheme and Godunov scheme for the isentropic system of gas dynamics. *Commun. Math. Phys*. **121**, 63–84, 1989
24. Glimm, J.: Solutions in the large for nonlinear hyperbolic systems of equations. *Commun. Pure Appl. Math*. **18**, 697–715, 1965
25. Godunov, S.K.: A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Mat. Sb. (N.S.)* **47**(89), 271–306, 1959. **(Russian)**
26. Holden, H., Risebro, N.: *Front Tracking for Hyperbolic Conservation Laws*. Springer, Berlin 2002
27. LeVeque, R.J.: *Numerical Methods for Conservation Laws*. Birkhäuser, Basel 1990
28. Liu, T.P.: The deterministic version of the Glimm scheme. *Commun. Math. Phys*. **57**, 135–148, 1975

Alberto Bressan, Maria Teresa Chiri and Wen Shen
Department of Mathematics,
Penn State University,
University Park
PA
16802
USA.
e-mail: axb62@psu.edu

and

Maria Teresa Chiri
e-mail: mxc6028@psu.edu

and

Wen Shen
e-mail: wxs27@psu.edu

(*Received October 5, 2020 / Accepted March 29, 2021*)
Published online April 11, 2021