Available online at www.sciencedirect.com
ScienceDirect
Physics Procedia 00 (2011) 1–28
Physics Procedia

Entropy and the numerical integration of conservation laws

Gabriella Puppo*

Dipartimento di Matematica, Politecnico di Torino (Italy)

Matteo Semplice*

Dipartimento di Fisica e Matematica, Università dell'Insubria (Como, Italy)

Abstract

In this paper, we review recent results on the role of entropy in the numerical integration of conservation laws. It is well known that weak solutions of systems of conservation laws may not be unique. Physically relevant weak solutions possess a viscous profile and satisfy entropy inequalities. In the discrete case entropy inequalities are used as a tool to prove convergence to entropy dissipating weak solutions. We start with classical results, in which entropy stability is used to prove the convergence of numerical solutions. We continue with entropy stable schemes, which are designed in order to produce entropy dissipation and can be proven to be convergent. Next we consider entropy as an error/regularity indicator permitting a local control on the behavior of the scheme, and thus driving grid and scheme adaptivity.

Keywords:
conservation laws, entropy, high order, error indicator

1. Introduction

We consider systems of m conservation laws in d dimensions:

$$u_t + \sum_{j=1}^{d} \nabla_{x_j} \cdot f_j(u) = 0, \qquad A_j(u) = (f_j)_u(u). \tag{1}$$

We suppose that the system is hyperbolic, i.e. each Jacobian matrix $A_j(u)$ has m real eigenvalues and a complete set of eigenvectors $\forall u$. Many mathematical models describing physical phenomena characterized by signals travelling with finite speed of propagation can be written in this form. Applications include gas dynamics, traffic flow, magnetohydrodynamics, astrophysics, to name just a few.

It is well known that solutions of (1) may exhibit a very rich structure, with discontinuities (shocks) arising in a finite time even from smooth initial data. This possibility obliges to enrich the set of possible solutions of (1)

*Corresponding author
Email addresses: gabriella.puppo@polito.it (Gabriella Puppo), matteo.semplice@uninsubria.it (Matteo Semplice)

introducing the class of weak solutions. The set of weak solutions contains piecewise C^1 functions, with step-like discontinuities whose speed is imposed by the weak form itself, through the so-called Rankine-Hugoniot conditions. The problem is that enlarging the class of admissible solutions, uniqueness is lost, and retrieving uniqueness turns out to be an extremely hard task, which up to now has not been settled in the general case, [1].

In many models of the form (1) coming from physics, the hyperbolic system is obtained disregarding higher order terms, multiplied by a small parameter. In many cases, the higher order term is parabolic, and the small parameter is a diffusive parameter, which, in analogy with fluid dynamics, is called viscosity. In these cases, uniqueness can be retrieved if weak solutions of (1) are required to be limits of viscous solutions as the viscosity tends to zero. On the other hand, uniqueness in physics is often ensured imposing irreversibility on admissible shocks. This results in entropy conditions that a weak solution should satisfy. The link between these two criteria is complex. Weak solutions obtained with the vanishing viscosity method satisfy entropy conditions [1, 2]. The reverse is not true in general, and is linked to the richness of the family of entropies that a system possesses. The best understood case, is the multidimensional scalar case, where entropy conditions are enough to ensure uniqueness, see [3], where Kruzkhov's uniqueness result is clearly described.

The numerical integration of (1) clearly must deal with the same difficulties encountered in the analysis of systems of conservation laws. It is easy to design schemes that compute weak solutions propagating with the speed prescribed by the Rankine Hugoniot conditions. These are the so called conservative schemes, and their introduction has permitted to construct schemes which preserve exactly the correct speeds, with a very easy to implement condition on the numerical fluxes. The issue of ensuring convergence to the correct unique solution has proven to be much harder. Usually, one relies on discrete entropy conditions, which however is satisfactory only in the scalar case.

However, the numerical solution of hyperbolic systems, beside the difficulties arising from the analysis of these equations, introduces new ones, which are linked to the need of dealing with discontinuous solutions. First order schemes have the most favourable theoretical results. The class of monotone schemes produces solutions which are perturbed by a viscous-like error term, and thus they converge to solutions possessing a viscous profile, [3, 4]. Moreover, they satisfy discrete entropy inequalities. However, they converge slowly, thus requiring very fine grids, and they smear discontinuous profiles. Increasing the order is conceptually quite easy, but numerical difficulties increase tremendously. High order schemes tend to produce spurious oscillations on shocks which may interact non linearly with the solution, generating unreliable results. To prevent the onset of spurious oscillations, a huge amount of ideas has been spent, and the literature is very extensive, see [3, 4] for second order schemes and [5] for higher order, and their references. Higher order schemes do not produce numerical solutions with an error term characterized by a neat parabolic structure, as is the case for first order schemes. Thus it is hard to prove that their numerical solutions converge to vanishing viscosity solutions. Initially, the construction of high order schemes has concentrated on the need to avoid spurious oscillations, introducing the concept of TVD (Total Variation Diminuishing) schemes, which are prevented from oscillating by brute force. But this concept cannot be applied in 2D, and it cannot be applied even in 1D for the solution of systems of equations in conservative variables.

For these reasons, exploring the consequences of imposing entropy conditions at the discrete level has been a fruitful tool. This program of work has started at the beginning of the '80s, with a series of works [6], [7], [8], continued with a study of the link between entropy and numerical viscosity in [9], and more recently summarized and further expanded in [10].

In the present work, we describe techniques for the construction of numerical schemes which are linked to the concept of entropy for systems of conservation laws. The paper starts with a brief summary of the basic analytical results involving the entropy, and of the main difficulties of numerical schemes for conservation laws in §2. Then, §3 contains the construction of entropy stable schemes from [10], exploring the link between numerical diffusion and entropy stability. Next the generalization to higher order schemes designed in [11, 12] is described, and we show numerical results illustrating the entropy dissipation of these schemes, especially in the fully discrete case, where the entropic behavior of very common time discretizations is not well known yet. Next in §4 our work [13, 14] is described on the use of entropy residuals to obtain a local estimate of the truncation error of the scheme, based on the computed numerical solution. This tool permits to construct a reliable a posteriori error indicator to drive the construction of a locally adaptive grid, following the unsteady structure of the numerical solution. It is also possible to use the error indicator to modify the method locally (scheme adaptivity), applying high order schemes only on smooth portions of the solution, and/or using the costly non linear devices that prevent the onset of spurious oscillations only where they are actually needed.

Several important results have not found space in this review. A posteriori error control based on entropy residuals can also be found in [15], [16] and recently summarized in [17]. The error indicator proposed and analyzed in these works relies on the family of Kruzkhov's entropies for multidimensional scalar conservation laws. These results are rigorous, but they lack the simplicity and the localization of the indicator proposed in [13, 14] and described below. Moreover, these results depend heavily on the family of entropies for scalar conservation laws, and cannot be generalized easily to hyperbolic systems of equations.

An original approach to the role of entropy in the analysis of systems of conservation laws can be found in [18] and its copious references. Here entropy inequalities are derived minimizing the entropy of kinetic BGK models converging to given systems of conservation laws, as relaxation parameters go to zero. The entropy pairs of systems of conservation laws are naturally obtained studying the convergence towards equilibrium of the BGK system. This approach follows extensive work on the connection between systems of conservation laws and kinetic models, which provides a promising alternative to the idea of characterizing weak solutions through viscous profiles. Note also that through kinetic approximations many numerical schemes for systems of conservation laws have been derived, see [19, 20].

Finally, in [21], a measure of the entropy residual is used to modify the scheme locally. The scheme proposed is a high order linear scheme, without the usual non linear devices preventing the onset of spurious oscillations. Here, the non oscillatory behavior of the solution is preserved adding a degenerate parabolic term which is turned on when the entropy residual surpasses a given threshold.

2. Conservative schemes and entropy dissipation

In this section, we recall the main properties of the numerical solution of conservation laws. Classical references are [2] and [22] and, more recently, [1] for the analysis of solutions of systems of conservation laws. Several textbooks contain descriptions of the treatment of numerical integration of conservation laws, in particular see [4] and [3].

We start considering the single scalar conservation law with initial condition:

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \qquad u(x, t = 0) = u_0(x). \tag{2}$$

As long as the solution is smooth, there is a unique classical solution, which can be constructed with the method of characteristics. However, it is easy to see that discontinuities may form in finite time, even from smooth initial data. The method of characteristics in fact breaks down if two (or more) characteristics intersect at the same point and a shock develops. In this case, the solutions of the conservation law must be understood in the weak sense. Multiply the conservation law by a smooth test function ϕ with compact support, and integrate in space and time for $t \in [0, \infty) = \mathbb{R}^+$ and $x \in \mathbb{R}$. Integrating by parts, the boundary terms pick up only the initial condition, and the weak form of the conservation law is found:

$$\int_0^\infty \int_{-\infty}^{+\infty} \left[\phi_t u + \phi_x f(u)\right] \mathrm{d}x\, \mathrm{d}t = -\int_{-\infty}^{+\infty} \phi(x, 0) u(x, 0)\, \mathrm{d}x \tag{3}$$

$\forall \phi \in C_0^1(\mathbb{R} \times \mathbb{R}^+)$. The weak solutions obtained from (3) admit discontinuities satisfying the Rankine-Hugoniot jump conditions which dictate the speed at which such solutions can travel. To see on an easy example that the Rankine Hugoniot condition derives from the weak form of the conservation law, we consider the following initial data with a jump:

$$u(x, t = 0) = u_0(x) = \begin{cases} u_L & x < 0 \\ u_R & x \geq 0. \end{cases} \tag{4}$$

The situation is depicted in Fig. 1. Clearly by similarity, this step will travel with constant speed s, so that the solution will be:

$$u(x, t) = \begin{cases} u_L & x < st \\ u_R & x \geq st. \end{cases}$$

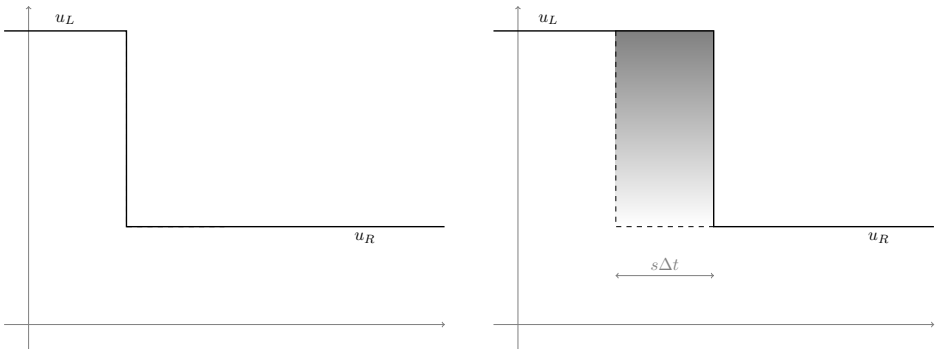


Figure 1: Advection of a step function

Substituting this information in the weak form, one finds:

$$\int_0^\infty \int_{-\infty}^{st} \phi_t u_L \,\mathrm{d}x\,\mathrm{d}t + \int_0^\infty \int_{st}^\infty \phi_t u_R \,\mathrm{d}x\,\mathrm{d}t + \big[f(u_L) - f(u_R)\big] \int_0^\infty \phi(st,t)\,\mathrm{d}t = -\int_{-\infty}^{+\infty} \phi(x,0)u(x,0)\,\mathrm{d}x,$$

which, after some algebra, reduces to:

$$\big[s(u_L - u_R) - (f(u_L) - f(u_R))\big] \int_0^\infty \phi(st,t)\,\mathrm{d}t = 0, \qquad \forall \phi \in C_0^1(\mathbb{R}\times\mathbb{R}^+).$$

Since this equation holds for all test functions, the speed of the jump must satisfy the equation

$$s(u_L - u_R) = f(u_L) - f(u_R). \tag{5}$$

In the scalar case the above relation reduces to

$$s = \frac{f(u_L) - f(u_R)}{u_L - u_R}, \tag{6}$$

which dictates the speed at which a discontinuity will move. The Rankine Hugoniot condition is linked to the conservation of the amount (mass) of u present in the flow. Let $M(t) = \int_a^b u(x,t)\,\mathrm{d}x$ be the amount of u present in an interval (a,b) containing the origin, and consider again the step-like initial condition given above. From Fig. 1, one can see that the mass at time $t + \Delta t$ is given by $M(t+\Delta t) = M(t) + s\Delta t(u_L - u_R)$. On the other hand, the variation of mass is given by the net flux leaving the interval (a,b) in the time Δt, so $M(t+\Delta t) = M(t) + \Delta t(f(u_L) - f(u_R))$, and equating these two expressions the Rankine Hugoniot condition results.

The weak form (3) extends the class of possible solutions of (2), but uniqueness is lost. To recover uniqueness, some further constraints must be imposed. Typically, many systems of conservation laws derive from models of mathematical physics, and are approximations of parabolic systems of equations, with small viscous coefficients. Thus, one is interested in a weak solution possessing a viscous profile, that is a solution that can be obtained as the vanishing viscosity limit of a solution of the parabolic system. It is possible to prove [2, Chapter 20] that vanishing viscosity solutions satisfy entropy inequalities which can be constructed in the following fashion. Suppose the conservation law is endowed with an entropy pair $\eta(u), \psi(u)$ such that η is convex and ψ satisfies $\eta'(u)f'(u) = \psi'(u)$. Note that at least in the scalar case infinite entropies can be found. Multiplying the conservation law by η' one finds that for smooth flows:

$$\frac{\partial \eta}{\partial t} + \frac{\partial \psi(u)}{\partial x} = 0.$$

If shocks develop, the vanishing viscosity weak solution satisfies the inequality:

$$\frac{\partial \eta}{\partial t} + \frac{\partial \psi(u)}{\partial x} \le 0$$

which again should be understood in the weak sense:

$$\int_0^\infty \int_{-\infty}^{+\infty} \left[\phi_t \eta(u) + \phi_x \psi(u)\right] \mathrm{d}x\, \mathrm{d}t \geq -\int_{-\infty}^{+\infty} \phi(x,0)\eta(u(x,0))\, \mathrm{d}x \tag{7}$$

$\forall \phi \in C_0^1(\mathbb{R}\times\mathbb{R}^+), \phi \geq 0$. Note that this time the sign of the test function must be prescribed, so that the sign of the expression being integrated does not depend on the sign of ϕ. Considering again the step like initial condition (4), the weak form of the entropy inequality reduces to:

$$(\eta(u_R) - \eta(u_L))s - (\psi(u_R) - \psi(u_L)) \geq 0, \tag{8}$$

where again s is the speed of the discontinuity. The inequality above selects which initial steps result in entropic shocks, according to the entropy η. See also [23].

The vanishing viscosity solutions satisfy entropy inequalities (7) for all admissible entropies, that is for all entropy pairs satisfying the compatibility condition $\eta' f' = \psi'$, thus the entropy inequality is a necessary condition for uniqueness. In the scalar case, all convex functions are admissible entropies, and Kruzkov in 1970 proved that a weak solution of the conservation law satisfying the entropy inequality for all admissible entropies is unique, [3, chapt 2]. This important result holds also in the multidimensional case. For systems of conservation laws, the existence of an entropy pair is not guaranteed, even though systems deriving from physics usually admit entropy inequalities, as in gas dynamics. However, even for systems which have entropies, the entropy condition is not sufficient to establish uniqueness in the general case, see [1].

Naturally, in the numerical integration of conservation laws, it is desirable to preserve the properties outlined above. The main tools are the Lax-Wendroff theorem, to prove convergence to weak solutions, and the discrete version of the entropy inequality, to prove that a certain scheme produces numerical solutions satisfying the entropy condition.

The computational domain is discretized with a grid in space and time which, in this section, we suppose to be uniform. Let (x_j, t^n) denote the grid points, and suppose that the mesh is characterized by a width h in space and a time step k. Introduce the control volumes $V_j^n = (x_j - h/2, x_j + h/2) \times [t^n, t^n + k)$, and the cell averages $\overline{u}_j^n$ at time t^n defined as:

$$\overline{u}_j^n = \frac{1}{h}\int_{x_j - h/2}^{x_j + h/2} u(x, t^n)\, \mathrm{d}x.$$

Integrating the conservation law in space and time on the control volume V_j^n, the finite volume formulation for the cell averages $\overline{u}$ is obtained:

$$\overline{u}_j^{n+1} = \overline{u}_j^n - \frac{1}{h}\left[\int_0^k f\left(u(x_j + \frac{h}{2}, t^n + t)\right) \mathrm{d}t - \int_0^k f\left(u(x_j - \frac{h}{2}, t^n + t)\right) \mathrm{d}t\right], \tag{9}$$

Note that this relation is exact, and must hold for all control volumes. On the chosen grid, the conservation law reduces to a system of finite volume equations for the cell averages. Note however that this system of equations is not closed, because the equations provide the evolution of the cell averages in terms of fluxes evaluated at the cell interfaces. To obtain the evolution of the cell averages from the finite volume equations, one should extract pointwise information from the cell averages, to compute the integral of the fluxes at the cell interfaces. Numerical schemes in conservation form mimic this structure. Let U denote the numerical solution. Then a numerical scheme in conservation form can be written as:

$$\overline{U}_j^{n+1} = \overline{U}_j^n - \frac{k}{h}\left(F_{j+1/2}^n - F_{j-1/2}^n\right) \tag{10}$$

where $F_{j+1/2}^n$ is the numerical flux function. For the time being, we consider very simple schemes, for which the solution $\overline{U}_j^{n+1}$ at time t^{n+1} depends only on the numerical solution at the previous time step t^n at three grid points. In this case, the numerical flux function will depend on two grid points, namely $F_{j+1/2}^n = F(\overline{U}_j^n, \overline{U}_{j+1}^n)$, and the function F must satisfy the following properties:

- $F(U, U) = f(U)$, (consistency)

- $F(U_j, U_{j+1})$ is at least Lipschitz continuous.

Numerical solutions obtained with conservative schemes are *good* solutions, in the sense that, if they converge, they converge to weak solutions. This is the content of the Lax Wendroff theorem:

Theorem 2.1. Lax Wendroff theorem. *Let $U_h(x,t)$ be a numerical solution obtained on a grid of width h. Suppose that:*

- *U_h is bounded in L^∞ uniformly in h;*
- *$U_h \to U$ as h goes to zero in L^1 on every bounded rectangle in $\mathbb{R}\times\mathbb{R}^+$, with k/h fixed;*
- *U_h is obtained with a conservative scheme.*

Then the limit solution is a weak solution of the conservation law.

The theorem assumes the existence of a sequence of numerical solutions, which are indexed by the mesh width h. The functions in the sequence must enjoy some sort of stability, and the sequence must be convergent. If moreover all functions in the sequence are obtained with a conservative scheme, then the limit solution is guaranteed to satisfy the weak form of the conservation law (3), and thus to contain shocks satisfying the Rankine Hugoniot jump conditions.

The proof is interesting, see [3], and is based on the following ingredients:

- multiply the conservative form of the scheme by a test function ϕ, summing over all grid points in space and time;
- sum by parts, discharging the differences from U and $F_{j+1/2}$ on ϕ;
- pass to the limit for $h \to 0$, using the boundedness of U_h, the convergence of U_h, and the consistency of the numerical flux F, and get the weak form of the conservation law.

Note that when summing by parts all interior terms can be grouped together in pairs, because, thanks to the conservative form, the fluxes are evaluated at the interfaces, and therefore they always appear twice: once as flux exiting from the cell V_j^n, and the second time as flux entering the cell V_{j+1}^n. This point is important, and we illustrate it using again the initial condition (4). Suppose the computational domain is the interval $[a,b]$ which is large enough to contain the initial step up to a final time T. Let $x_{j_a} = a + \frac{h}{2}$ and $x_{j_b} = b - \frac{h}{2}$ be the grid points at the end points of the interval. Then the total mass contained in the interval $[a,b]$ at time $t^n < T$ is given by:

$$M(t^n) = \sum_{j=j_a}^{j_b} \overline{U}_j^n \, h$$

Applying the conservative scheme (10), the total mass at time $t^{n+1} < T$ is given by:

$$\begin{aligned} M(t^n + k) &= \sum_{j=j_a}^{j_b} \overline{U}_j^{n+1} \, h \\ &= \sum_{j=j_a}^{j_b} \left[\overline{U}_j^n - \frac{k}{h}\left(F_{j+1/2}^n - F_{j-1/2}^n \right) \right] h \\ &= M(t^n) - k\left(F_{j_b+1/2}^n - F_{j_a-1/2}^n \right), \end{aligned}$$

where all interior terms cancel out because the conservative structure of the scheme ensures that whatever flows out a cell, flows into the following cell. At the boundary the consistency of the scheme can be applied. Here by hypothesis the solution is constant, so:

$$F_{j_b+1/2}^n = F(\overline{U}_{j_b}^n, \overline{U}_{j_b+1}^n) = F(u_R, u_R) = f(u_R) \qquad F_{j_a-1/2}^n = F(\overline{U}_{j_a-1}^n, \overline{U}_{j_a}^n) = F(u_L, u_L) = f(u_L)$$

Substituting this information above, it can be seen that the scheme provides the exact evolution of the total mass of the system, and thus the speed of propagation of shocks is also exact. In more general cases, the propagation speed of shocks will not be exact, because the end points of the jump may depend on time and they will not be evolved exactly, but we wish to stress the fact that no error in the propagation speed will occur if the end points of the jump are evolved correctly, even though the shock may be spread on several cells. This in turn is guaranteed by the fact that evaluating the total mass, the numerical fluxes cancel out for interior cell boundaries.

In other words, to preserve global properties, such as convergence, the numerical flux functions do not need to be sophisticated, as long as they are consistent and they are evaluated at cell boundaries, as in (10). Convergence means that the numerical solution approximates the true solution *as* $h \to 0$, but how small should h be before the error between the numerical and the true solution lies below a given threshold remains unknown. Naturally, one is interested in the behavior of the solution for a finite mesh size. But for a finite value of h convergence theorems are not helpful. In this case, error estimates are needed, as in [24] or [17], and a posteriori error estimators [17, 25, 14]. As a matter of fact, a convergent scheme may be quite lowsy locally, for a finite mesh width.

To transfer the entropy condition to the discrete level, integrate the entropy inequality over the control volumes V_j^n defined above as in (9), obtaining:

$$\frac{1}{h}\int_{x_j-h/2}^{x_j+h/2}\left[\eta(u(x,t^{n+1}))-\eta(u(x,t^n))\right]\mathrm{d}x+\frac{1}{h}\int_0^k\left[\psi(u(x_{j+1/2},t^n+t))-\psi(u(x_{j-1/2},t^n+t))\right]\mathrm{d}t\le 0. \tag{11}$$

This inequality involves the unknown exact solution, and it should be extended to the numerical solution. The finite volume conservative scheme (10) constructs the solution as a sequence of point values which coincide with the cell averages of the numerical solution. These point values are used to define a piece-wise constant function $U(x,t)$ as follows:

$$U(x,t)=\overline{U}_j^n \quad (x,t)\in[x_j-\tfrac{h}{2},x_j+\tfrac{h}{2})\times[t^n,t^{n+1}). \tag{12}$$

Note that this reconstruction is conservative, in the sense that it preserves the amount of mass U:

$$\int_{x_j-h/2}^{x_j+h/2}U(x,t)\,\mathrm{d}x=\overline{U}_j^n\,h \qquad t\in[t^n,t^{n+1}).$$

With this assumption, one can evaluate the space integral in the entropy inequality. For the integral in time, which involves the entropy fluxes, numerical entropy fluxes are constructed. For a 3-point scheme with numerical flux $F^n_{j+1/2}=F(\overline{U}^n_j,\overline{U}^n_{j+1})$, the numerical entropy flux will be a function $\Psi^n_{j+1/2}=\Psi(\overline{U}^n_j,\overline{U}^n_{j+1})$, satisfying the following properties:

- $\Psi(U,U)=\psi(U)$, (consistency)
- $\Psi(U_j,U_{j+1})$ is at least Lipschitz continuous.

Substituting this information in the finite volume formulation of the entropy inequality (11), it is found that the numerical solution should satisfy the *discrete entropy inequality*:

$$\eta\left(\overline{U}_j^{n+1}\right)-\eta\left(\overline{U}_j^n\right)+\frac{k}{h}\left[\Psi^n_{j+1/2}-\Psi^n_{j-1/2}\right]\le 0. \tag{13}$$

A scheme satisfying the discrete entropy condition and the hypotheses of the Lax-Wendroff theorem converges to an entropy satisfying weak solution. In the scalar the following theorem holds [3]:

Theorem 2.2. Convergence to the entropic solution. *Let $U_h(x,t)$ be a numerical solution obtained on a grid of width h. Suppose that:*

- *U_h is bounded in L^∞ uniformly in h;*
- *$U_h \to U$ as h goes to zero in L^1 on every bounded rectangle in $\mathbb{R}\times\mathbb{R}^+$, with k/h fixed;*
- *U_h is obtained with a conservative scheme.*

- *U_h satisfies the cell entropy inequality* (13) *with a consistent entropic flux* Ψ *constructed for all admissible entropy pairs* η, ψ *and for all* h.

Then the limit solution is the unique entropic weak solution of the conservation law.

The proof is very similar to the proof of the Lax Wendroff theorem, and shows that the limit solution U satisfies the weak form of the entropy inequality (7) for all admissible entropies. Again, the proof relies on the fact that the numerical entropy fluxes are evaluated at the cell interfaces, see [3].

3. Numerical diffusion and entropy stable schemes

In this section we will start our review of numerical schemes obtained using the entropy as a tool to design numerical schemes. This section starts with the notion of entropy conservative and entropy stable schemes introduced in [9] and expanded in [10]. Next, we show a few typical results in the one-dimensional scalar case. The section ends with a description of the high order entropy stable schemes of [12].

The easiest conservative scheme one can think of has the simple numerical flux $F^n_{j+1/2} = \frac{1}{2}\left(f(\overline{U}^n_j) + f(\overline{U}^n_{j+1})\right)$, leading to

$$\overline{U}^{n+1}_j = \overline{U}^n_j - \frac{k}{2h}\left(f(\overline{U}^n_{j+1}) - f(\overline{U}^n_{j-1})\right).$$

This scheme is however disastrous, since it is unstable for all k. Using discrete Fourier analysis for the equation $u_t + au_x = 0$, it is easy to see that this scheme actually introduces a *negative numerical diffusion*, since its modified equation has the form:

$$u_t + au_x = -\frac{1}{2}ka^2u_{xx}.$$

The viscosity in this equation is $\nu = -\frac{1}{2}ka^2$, and the numerical solution will blow up exponentially for every choice of the timetep k. Thus, the scheme must be modified adding a diffusive term. The simplest choice is:

$$\overline{U}^{n+1}_j = \overline{U}^n_j - \frac{k}{2h}\left(f(\overline{U}^n_{j+1}) - f(\overline{U}^n_{j-1})\right) + \frac{k}{2h}Q\left(\overline{U}^n_{j+1} - 2\overline{U}^n_j + \overline{U}^n_{j-1}\right)$$

Let $\lambda = \frac{k}{h}$ denote the mesh ratio, which, in this section, will be constant. The parameter Q must be determined in order to ensure that the scheme preserves the monotonicity of the solution, in order to avoid spurious oscillations. Schemes satisfying this request are called monotone, and they can be characterized by the constraints

$$\frac{\partial \overline{U}^{n+1}_j}{\partial \overline{U}^n_{j+l}} \geq 0 \qquad l = -1, 0, 1.$$

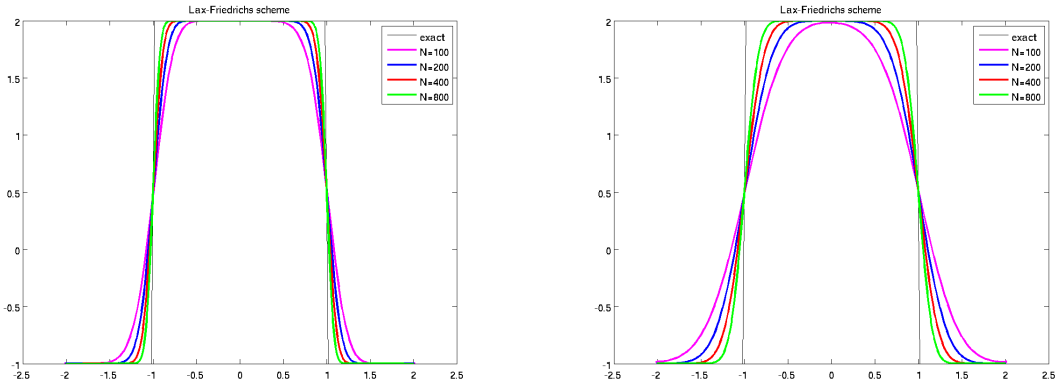
It is easy to see that the scheme above is monotone, provided that $Q \geq \max|f'(u)|$, with $\lambda Q \leq 1$, which implies also the CFL condition. These schemes can be proven to satisfy all entropy inequalities, and thus they converge to the unique entropy solution in the scalar case, [3, 26]. More generally, the following modified numerical flux is introduced:

$$F_{j+1/2} = \frac{1}{2}\left(f(\overline{U}^n_j) + f(\overline{U}^n_{j+1})\right) - \frac{1}{2}Q_{j+1/2}\left(\overline{U}^n_{j+1} - \overline{U}^n_j\right). \tag{14}$$

This scheme is monotone if $Q_{j+1/2} \geq f'(\overline{U}_{j+1}), Q_{j-1/2} \geq -f'(\overline{U}_{j-1})$ and $\frac{1}{2}\lambda(Q_{j+1/2} + Q_{j-1/2}) \leq 1$. Most first order numerical schemes can be written in this form. The Local Lax Friedrichs scheme for instance has a coefficient of numerical viscosity given by

$$Q_{j+1/2} = \max(|f'(\overline{U}_{j+1})|, |f'(\overline{U}_j)|).$$

Monotone schemes are nice, because they converge to the exact solution, but they are limited to first order accuracy [4]. This means that to get an accurate solution very fine grids are needed, and jumps in the numerical solution, especially for contact discontinuities, are smeared on several grid points. We illustrate this behavior on the simple advection equation $u_t + u_x = 0$ with periodic boundary conditions, considering the evolution of a square wave using



(a) (b)

Figure 2: Linear advection of a square wave with the Lax Friedrichs scheme, $N = 100, N = 200, N = 400, N = 800$ cells. (a) $T = 8$, (b) $T = 32$.

the Lax Friedrichs scheme, with $Q_{j+1/2} = 1$ which gives the minimum amount of numerical dissipation in this case, and $\lambda = 0.9$, see Fig. 2.

It is clear that the profile is smeared as time increases, even for very fine grids. Clearly this scheme is too diffusive. In this case, the same results would be obtained with the Godunov scheme which, in the linear case, has a coefficient of numerical viscosity $Q_{j+1/2} = |a|$, where a is the propagation speed. In order to reduce drastically the artificial diffusion, a possibility is to increase the order of the scheme. Fig. 3 shows the effect of increasing the order of accuracy, on a smooth solution for long integration times (here the initial hump completes its period 8 times). While the first order scheme is still far away from the exact solution with $N = 800$ grid points, the second order solution is almost coinciding with the exact solution with only 200 grid points. However, using a standard second order scheme on a discontintinuity, the solution starts oscillating. The Lax Wendroff scheme for the linear advection equation $u_t + au_x = 0$ can be written in viscosity form (14) with a coefficient $Q_{j+1/2} = \lambda a^2$. Since for stability $\lambda < 1/|a|$, the Lax Wendroff scheme indeed has a smaller viscosity than the Godunov scheme, but the problem is that it has too little viscosity. The problem than is: *is it possible to say what is the minimum amount of viscosity that a numerical scheme should have?*

Entropy conservative schemes address exactly this issue. Since it is known that for convergence a numerical scheme should dissipate entropy, at least on shocks, a starting point is to construct schemes which dissipate no entropy. Then the numerical diffusion inherent in these schemes can be used as a benchmark to set the minimum numerical diffusion that a scheme should have. This idea has been developed by Tadmor in a series of papers [8], [9] and the review [10]. Here we illustrate the behavior of the schemes for scalar conservation laws, but the main setting will be described for systems of equations.

In this section, we consider one-dimensional systems of m conservation laws:

$$u_t + f_x(u) = 0, \qquad A(u) = f_u(u). \tag{15}$$

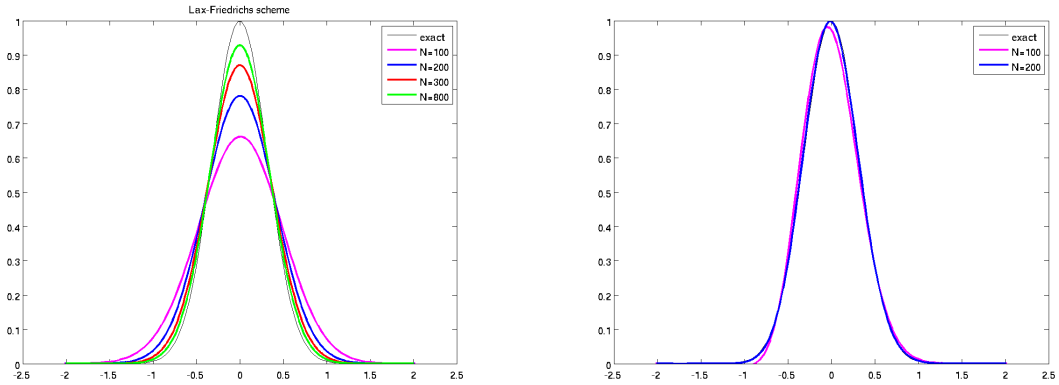
Again, the system is hyperbolic, i.e. the Jacobian matrix $A(u)$ has m real eigenvalues and a complete set of eigenvectors $\forall u$.

Recall that an entropy pair (η, ψ) for the system (15) is a couple of functions $\eta(u)$ and $\psi(u)$ such that η is convex (i.e. the Hessian η_{uu} is a positive definite symmetric matrix) and ψ satisfies the compatibility relation

$$(\eta_u)^T A(u) = (\psi_u)^T. \tag{16}$$

For example gas dynamic equations have a family of entropy pairs of the form

$$\eta(u) = -\rho h(S), \quad \psi(u) = -mh(S), \qquad u = [\rho, m, E]^T, \qquad S = \ln(p\rho^{-\gamma}) \tag{17}$$



(a) (b)

Figure 3: Linear advection of a smooth function with the Lax Friedrichs scheme, and with the Lax Wendroff scheme $N = 100, N = 200, N = 400, N = 800$ cells, up tp $T = 32$. (a) Lax-Friedrichs, (b) Lax-Wendroff.

where ρ is the density, m is momentum, E is the density of total energy, S is the specific entropy, p is the pressure and γ is the ratio of specific heats. The scalar function h must satisfy the properties $h' - \gamma h'' > 0$, $h' > 0$, [6]. In particular for $h(S) = S$ we recover the standard entropy used in gas dynamics.

Introduce the *entropy variables* $v(u) = \eta_u$. Since the Jacobian of the transformation is η_{uu}, which is invertible, the map associating the conservative variables u to the entropy variables is invertible. Entropy variables are used to rewrite the sytem of conservation laws in symmetric form:

$$\partial_t u(v) + \partial_x g(v) = 0, \qquad g(v) = f(u(v)). \tag{18}$$

In fact, carrying out differentiation, the system becomes

$$u_v \partial_t v + f_u u_v \partial_x v = (\eta_{uu})^{-1} \partial_t v + A(u)(\eta_{uu})^{-1} \partial_x v = 0,$$

in which the matrix $A(u)(\eta_{uu})^{-1}$ is symmetric, because of the compatibility condition (16). However, here the interesting point is that entropy variables permit to introduce the *entropy potentials* ϕ and θ:

$$\phi(v) = v \cdot u(v) - \eta(u(v)) \tag{19}$$
$$\theta(v) = v \cdot g(v) - \psi(u(v)), \tag{20}$$

where $u \cdot v$ denotes the inner product between the vectors u and v. Note that $\phi'(v) = u(v)$ while $\theta'(v) = g(v)$.

3.1. Semidiscrete entropy stable schemes

Entropy stable schemes are first defined in the semidiscrete case, in which only the space discretization is considered, while time is left as a continuous variable. The original system (15) is re-written in terms of entropy variables, and on this system a semidiscrete conservative scheme is applied:

$$\frac{\mathrm{d}}{\mathrm{d}t} u(v)_j + \frac{1}{h}\left(G_{j+1/2} - G_{j-1/2}\right) = 0 \tag{21}$$

where $G_{j+1/2}$ denotes a numerical flux consistent with the flux function $g(u(v))$. The aim is to compute a numerical flux $G^*_{j+1/2}$ such that the scheme above is entropy conservative, in the sense that a local cell entropy *equality* is satisfied in each cell:

$$\frac{\mathrm{d}}{\mathrm{d}t} \eta(u)_j + \frac{1}{h}\left(\Psi_{j+1/2} - \Psi_{j-1/2}\right) = 0 \tag{22}$$

for some consistent numerical entropy flux Ψ. In particular, let [10]:

$$\Psi_{j+1/2} = \Psi(v_j, v_{j+1}) = \frac{1}{2}(v_j + v_{j+1}) \cdot G_{j+1/2} - \frac{1}{2}\left(\theta(v_j) + \theta(v_{j+1})\right). \tag{23}$$

Note that, on grid values, $u_j = u(v_j)$, and $f(u_j) = g(v_j)$. It is easy to check that the expression above defines a consistent entropy flux, provided $G_{j+1/2}$ is also consistent with the flux $g(v)$. Now the entropy dissipation due to the scheme can be computed:

$$\frac{\mathrm{d}}{\mathrm{d}t}\eta(u)_j + \frac{1}{h}\left(\Psi_{j+1/2} - \Psi_{j-1/2}\right) = v_j \frac{\mathrm{d}}{\mathrm{d}t}u_j + \frac{1}{h}\left(\Psi_{j+1/2} - \Psi_{j-1/2}\right).$$

Substituting the numerical entropy flux (23) with $G^*_{j+1/2}$, and the expression for $\frac{\mathrm{d}}{\mathrm{d}t}u(v)_j$ using (21), the entropy dissipation becomes

$$-\frac{1}{h}v_j\left(G^*_{j+1/2} - G^*_{j-1/2}\right) + \frac{1}{2h}\left((v_j + v_{j+1}) \cdot G^*_{j+1/2} - \left(\theta(v_j) + \theta(v_{j+1})\right) - (v_j + v_{j-1}) \cdot G^*_{j-1/2} + \left(\theta(v_j) + \theta(v_{j-1})\right)\right) = 0$$

Rearranging terms

$$\frac{1}{2h}\left[(v_{j+1} - v_j) \cdot G^*_{j+1/2} - \left(\theta(v_{j+1}) - \theta(v_j)\right) + (v_j - v_{j-1}) \cdot G^*_{j-1/2} - \left(\theta(v_j) - \theta(v_{j-1})\right)\right] = 0$$

and clearly the equality holds, provided that

$$G^*_{j+1/2} \cdot (v_{j+1} - v_j) = \theta(v_{j+1}) - \theta(v_j). \tag{24}$$

So, with this choice the entropy dissipation in (22) is zero. On the other hand, if the numerical flux $G_{j+1/2}$ satisfies

$$G_{j+1/2} \cdot (v_{j+1} - v_j) \le \theta(v_{j+1}) - \theta(v_j), \tag{25}$$

the left hand side of (22) is negative and entropy dissipation occurs within the cell.

Definition 3.1. Entropy conservative and entropy stable schemes
A numerical scheme is entropy conservative if its numerical flux $G_{j+1/2}$ satisfies (24)*. It is entropy stable if its numerical flux satisfies* (25)*.*

Note that a numerical flux $G_{j+1/2}$ in the entropy variables induces a corresponding numerical flux in the conservative variables: $F_{j+1/2} = F(U_j, U_{j+1}) = G(v_j, v_{j+1}) = G_{j+1/2}$. In the scalar case, the entropy conservative flux is unique:

$$G^*_{j+1/2} = \frac{\theta(v_{j+1}) - \theta(v_j)}{v_{j+1} - v_j}.$$

For systems of equations, several choices are possible, since (24) is a single scalar constraint.

Example 3.2. Burgers' equation
For Burgers' equation $f(u) = \frac{1}{2}u^2$. Choosing the entropy $\eta(u) = \frac{1}{2}u^2$, the entropy flux is $\psi(u) = \frac{1}{3}u^3$. The entropy variable in this case is simply $v(u) = u$ and the entropy potential is $\theta(u) = \frac{1}{6}u^3$. Thus the entropy conservative flux is

$$F^*_{j+1/2} = F^*(U_j, U_{j+1}) = \frac{1}{6}\left(U^2_{j+1} + U_{j+1}U_j + U^2_j\right).$$

As the example shows, the entropy conservative flux is linked to the particular entropy chosen. This is not a problem in the scalar case for convex flux functions, since in this case entropy stability for the quadratic entropy is enough to select the unique entropic solution, see [10, 27].

Now, it is possible to rewrite the entropy conservative scheme in viscous form, to emphasize the numerical viscosity of these schemes. It is convenient to rewrite the viscous form of the scheme in terms of entropy variables:

$$G_{j+1/2} = \frac{1}{2}\left(f(\overline{U}^n_j) + f(\overline{U}^n_{j+1})\right) - \frac{1}{2}\tilde{Q}_{j+1/2}\left(v^n_{j+1} - v^n_j\right).$$

Substituting this expression in (25), and using (24), one finds that the viscosity of an entropy stable scheme must satisfy

$$-\frac{1}{2}\tilde{Q}_{j+1/2}\|v_{j+1}-v_j\|^2 \le -\frac{1}{2}\tilde{Q}^*_{j+1/2}\|v_{j+1}-v_j\|^2,$$

which gives

$$\tilde{Q}^*_{j+1/2} \le \tilde{Q}_{j+1/2}, \tag{26}$$

that is any entropy stable scheme must contain more numerical viscosity than an entropy conservative one. Note that in the scalar case

$$\tilde{Q}_{j+1/2} = Q_{j+1/2}\frac{U_{j+1}-U_j}{v_{j+1}-v_j},$$

and since $v(u)$ is a monotone increasing function of u, the fraction on the right hand side is always non-negative. Thus the viscosity coefficients of entropy stable schemes must satisfy

$$Q^*_{j+1/2} \le Q_{j+1/2}. \tag{27}$$

The entropy conservative coefficient of numerical viscosity $Q^*_{j+1/2}$ depends on the particular entropy chosen. It is interesting to see what is the maximum of $Q^*_{j+1/2}$ over all possible entropies. To compute this value, rewrite $Q^*_{j+1/2}$ as

$$Q^*_{j+1/2} = \frac{f(U_j)+f(U_{j+1})-2F^*_{j+1/2}}{U_{j+1}-U_j}$$

and compute the supremum over all possible entropies. It has been proven [8, 7] that the supremum is achieved by Godunov flux. Thus an entropy stable scheme, which is uniformly stable for all entropies, has a coefficient of numerical viscosity $Q_{j+1/2}$ which is larger than the coefficient of viscosity of the Godunov scheme. These are the so called E-schemes [7], which are again first order schemes.

However, it is easy to see, at least in the scalar case, that entropy conservative schemes are second order accurate. In fact, the entropy conservative flux $G^*_{j+1/2}$ can be written as

$$G^*_{j+1/2} = \frac{1}{v_{j+1}-v_j}\int_{v_j}^{v_{j+1}}\theta'(v)\,\mathrm{d}v = \frac{1}{v_{j+1}-v_j}\int_{v_j}^{v_{j+1}} g(v)\,\mathrm{d}v,$$

where the definition of the entropy potential (19) has been applied. Introduce the change of variables $v_{j+1/2}(\xi) = \frac{1}{2}(v_j+v_{j+1})+\xi(v_{j+1}-v_j)$. The expression above becomes

$$G^*_{j+1/2} = \int_{-1/2}^{1/2} g(v_{j+1/2}(\xi))\,\mathrm{d}\xi.$$

Integrating by parts, one obtains

$$G^*_{j+1/2} = \frac{1}{2}(g(v_{j+1})+g(v_j)) - \int_{-1/2}^{1/2}\xi g'(v_{j+1/2}(\xi))(v_{j+1}-v_j)\,\mathrm{d}\xi.$$

Substituting the viscous coefficient $Q_{j+1/2}$, this expression gives

$$Q^*_{j+1/2} = \int_{-1/2}^{1/2} 2\xi g'(v_{j+1/2}(\xi))(v_{j+1}-v_j)\,\mathrm{d}\xi.$$

Integrating by parts again, the following expression for Q^* is finally found:

$$Q^*_{j+1/2} = (v_{j+1}-v_j)\left(\int_{-1/2}^{1/2}\left(\frac{1}{4}-\xi^2\right)g''(v_{j+1/2}(\xi))\,\mathrm{d}\xi\right),$$

which shows that $Q^*_{j+1/2} = O(v_{j+1} - v_j) = O(h)$. Therefore the resulting scheme

$$\frac{\mathrm{d}U_j}{\mathrm{d}t} = -\frac{\lambda}{2}\left(f(U_{j+1}) - f(U_{j-1})\right) + \frac{\lambda}{2}\left(Q^*_{j+1/2}(v_{j+1} - v_j) - Q^*_{j-1/2}(v_j - v_{j-1})\right)$$

is second order accurate in space and it conserves entropy. To obtain a second order entropy stable flux therefore one possibility is to increase the artificial diffusion of an entropy conservative scheme with a term of order $v_{j+1} - v_j$. Thus the numerical flux [9] is modified as

$$G_{j+1/2} = G^*_{j+1/2} - \frac{1}{2}D_{j+1/2}, \tag{28}$$

where

$$D_{j+1/2} = \alpha_{j+1/2}(v_{j+1} - v_j) \tag{29}$$

and $\alpha_{j+1/2}$ is a positive coefficient. For instance one can take $\alpha_{j+1/2} = \max(|f(U_j)|, |f(U_{j+1})|)$. The entropy dissipation due to this scheme can be computed defining the numerical entropy flux as in (23) and it is given by:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\eta(U)_j + \frac{1}{h}\left(\Psi_{j+1/2} - \Psi_{j-1/2}\right) &= -\frac{1}{4h}\left[D_{j+1/2}(v_{j+1} - v_j) + D_{j-1/2}(v_j - v_{j-1})\right] \\ &= -\frac{1}{4h}\left[\alpha_{j+1/2}(v_{j+1} - v_j)^2 + \alpha_{j-1/2}(v_j - v_{j-1})^2\right].\end{aligned} \tag{30}$$

Note that the entropy dissipation has the correct sign if $\alpha_{j+1/2} \geq 0$, thus yielding an entropy stable scheme. This construction can be easily extended to systems of equations. Now the space contribution to the cell entropy dissipation can be computed as

$$S^{(x)}_j = -\frac{1}{4h}\left[(v_{j+1} - v_j)^T\alpha_{j+1/2}(v_{j+1} - v_j) + (v_j - v_{j-1})^T\alpha_{j-1/2}(v_j - v_{j-1})\right], \tag{31}$$

where $\alpha_{j+1/2}$ will be positive definite matrices. In the scalar case, the entropy stable scheme constructed above has an artificial diffusion coefficient

$$Q_{j+1/2} = Q^*_{j+1/2} + \frac{D_{j+1/2}}{v_{j+1} - v_j}.$$

Therefore, if the numerical diffusion term is defined as in (29), $Q_{j+1/2} = O(1)$ and thus the scheme is only first order accurate. To obtain higher order schemes the diffusion term defined in (29) must be modified, following [12]. See §3.3.

3.2. The fully discrete problem

Numerical solutions are obtained with fully discrete schemes, in which the time derivative of u is discretized usually with Runge-Kutta schemes or multistep methods. Therefore, in order to ensure entropy stability of the final solution, the entropy production due to time discretizations must be computed. This problem is addressed in [10], where it is proven that the Explicit Euler scheme creates spurious entropy, while the second order implicit Crank Nicolson scheme is entropy conservative, and the Implicit Euler scheme dissipates entropy, thus improving the stability of entropy stable space discretizations. A more general approach can be found in [28, 11] where implicit high order entropy conservative schemes are constructed.

Most schemes for hyperbolic systems of conservation laws are explicit. The theoretical results imply that in order to achieve entropy stability it is necessary to counterbalance the entropy production due to the explicit time integration with the entropy dissipation of the space part. It is possible to follow the bounds on [10] for the Explicit Euler scheme to obtain entropy stable schemes for a first order in time explicit integration. However the entropy behavior of higher order explicit time discretizations is less known.

The analysis in [10] considers a few prototype cases. We start from the Backward Euler scheme. In this case, the cell entropy dissipation of the fully discrete scheme can be defined as

$$S_j = \frac{1}{k}\left(\eta(v)^{n+1}_j - \eta(v)^n_j\right) + \frac{1}{h}\left(\Psi^{n+1}_{j+1/2} - \Psi^{n+1}_{j-1/2}\right). \tag{32}$$

Adding and subtracting the quantity $\frac{1}{k}\, v_j^{n+1} \cdot (U_j^{n+1} - U_j^n)$, S_j can be broken up in the two contributions $S_j = S_j^{(x)} + S_j^{(t)}$, defined below:

$$S_j^{(x)} = \frac{1}{k}\, v_j^{n+1} \cdot (U_j^{n+1} - U_j^n) + \frac{1}{h}\left(\Psi_{j+1/2}^{n+1} - \Psi_{j-1/2}^{n+1}\right) \tag{33}$$

$$= -\frac{1}{h}\, v_j^{n+1} \cdot (G_{j+1/2}^{n+1} - G_{j-1/2}^{n+1}) + \frac{1}{h}\left(\Psi_{j+1/2}^{n+1} - \Psi_{j-1/2}^{n+1}\right)$$

$$S_j^{(t)} = \frac{1}{k}\left(\eta(v_j^{n+1}) - \eta(v_j^n)\right) - \frac{1}{k}\, v_j^{n+1} \cdot (U_j^{n+1} - U_j^n). \tag{34}$$

$S_j^{(x)}$ has already been computed in (31). Note that the diffusion terms will also be evaluated implicitly. For $S_j^{(t)}$, the first term, dropping the index j, can be rewritten as:

$$\eta(v^{n+1}) - \eta(v^n) = \int_{v^n}^{v^{n+1}} \eta_v(v)\, \mathrm{d}v = \int_{v^n}^{v^{n+1}} \eta_u(u) u_v\, \mathrm{d}v.$$

Let $H(v) = (\eta_{uu})^{-1} = u_v$, which is a symmetric positive definite matrix. Recall that $v = \eta_u(u)$, and change variables as $v^{n+1/2}(\xi) = \frac{1}{2}(v^{n+1} + v^n) + \xi(v^{n+1} - v^n)$:

$$\eta(v^{n+1}) - \eta(v^n) = \int_{-1/2}^{1/2} v^{n+1/2} H(v^{n+1/2})(v^{n+1} - v^n)\, \mathrm{d}\xi. \tag{35}$$

The second term, dropping again the index j, can be written as

$$v^{n+1} \cdot (U^{n+1} - U^n) = v^{n+1} \int_{v^n}^{v^{n+1}} u_v\, \mathrm{d}v = v^{n+1} \int_{-1/2}^{1/2} H(v^{n+1/2})(v^{n+1} - v^n)\, \mathrm{d}\xi.$$

Subtracting this expression from the one above, the entropy dissipation due to the time discretization can be written as

$$S_j^{(t)} = -\frac{1}{k} \int_{-1/2}^{1/2} \left(\frac{1}{2} - \xi\right)(v^{n+1} - v^n)^T H(v^{n+1/2})(v^{n+1} - v^n)\, \mathrm{d}\xi \le 0. \tag{36}$$

Therefore, the Backward Euler differentiation introduces a *negative* term to the cell entropy dissipation, thus enhancing entropy stability. For the Forward Euler scheme, the cell entropy dissipation can be defined as

$$S_j = \frac{1}{k}\left(\eta(v)_j^{n+1} - \eta(v)_j^n\right) + \frac{1}{h}\left(\Psi_{j+1/2}^n - \Psi_{j-1/2}^n\right). \tag{37}$$

This time, add and subtract the quantity $\frac{1}{k}\, v_j^n \cdot (U_j^{n+1} - U_j^n)$. The total entropy dissipation S_j will now be given by the two contributions $S_j = S_j^{(x)} + S_j^{(t)}$ defined as

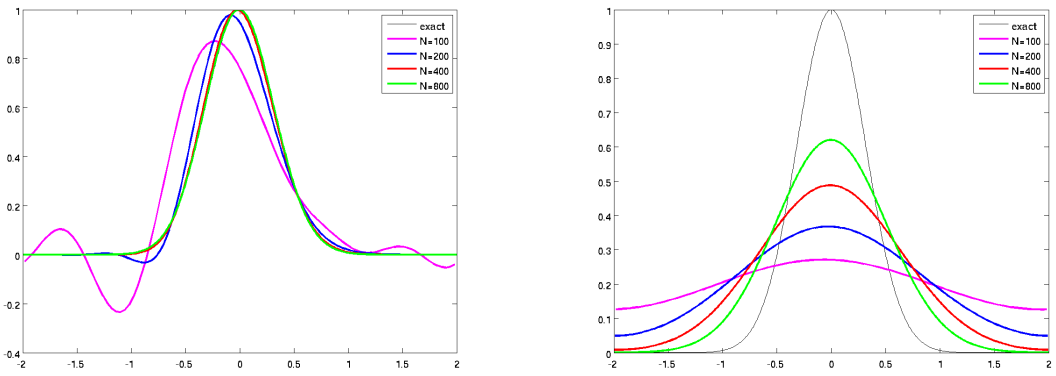
$$S_j^{(x)} = -\frac{1}{h}\, v_j^n \cdot (G_{j+1/2}^n - G_{j-1/2}^n) + \frac{1}{h}\left(\Psi_{j+1/2}^n - \Psi_{j-1/2}^n\right) \tag{38}$$

$$S_j^{(t)} = \frac{1}{k}\left(\eta(v)_j^{n+1} - \eta(v)_j^n\right) - \frac{1}{k}\, v_j^n \cdot (U_j^{n+1} - U_j^n). \tag{39}$$

Again, $S_j^{(x)}$ is already known from (31), except that now the diffusion terms will be computed explicitly. For $S_j^{(t)}$, the first term has already been computed in (35), and the following expression results:

$$S_j^{(t)} = \frac{1}{k} \int_{-1/2}^{1/2} \left(\frac{1}{2} + \xi\right)(v^{n+1} - v^n)^T H(v^{n+1/2})(v^{n+1} - v^n)\, \mathrm{d}\xi \ge 0. \tag{40}$$

Thus, for the explicit Euler differentiation the entropy dissipation is positive, and therefore to achieve entropy stability, the time step k must be reduced, in order to ensure that $S_j^{(t)} < |S_j^{(x)}|$. This is found in [10] at the price of a suboptimal



(a) (b)

Figure 4: Linear advection of a smooth function with entropy conservative space differencing, $N = 100, N = 200, N = 400, N = 800$ cells, up to $T = 32$. (a) Crank Nicolson, (b) Backward Euler.

CFL condition. Finally, consider the Crank Nicolson scheme, which, unlike [10], we use in its standard form. The cell entropy production will be given by:

$$S_j = \frac{1}{k}\left(\eta(v_j^{n+1}) - \eta(v_j^n)\right) + \frac{1}{2h}\left(\Psi_{j+1/2}^{n+1} + \Psi_{j+1/2}^{n} - \Psi_{j-1/2}^{n+1} - \Psi_{j-1/2}^{n}\right) \tag{41}$$

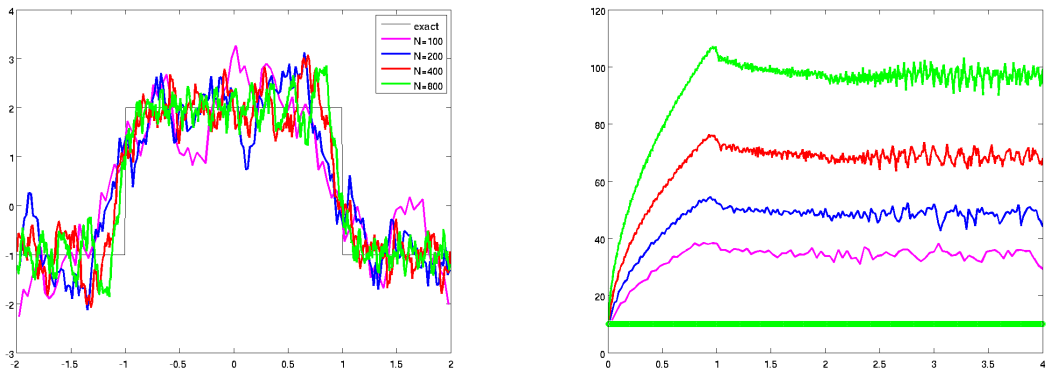
Using the same techniques illustrated above, and (31), the space and time entropy dissipations will be given by:

$$\begin{aligned} S_j^{(x)} &= -\frac{1}{8h}\Big[(v_{j+1}^{n+1} - v_j^{n+1})^T \alpha_{j+1/2}(v_{j+1}^{n+1} - v_j^{n+1}) + (v_{j+1}^{n} - v_j^{n})^T \alpha_{j+1/2}(v_{j+1}^{n} - v_j^{n}) \\ &\quad + (v_j^{n+1} - v_{j-1}^{n+1})^T \alpha_{j-1/2}(v_j^{n+1} - v_{j-1}^{n+1}) + (v_j^{n} - v_{j-1}^{n})^T \alpha_{j-1/2}(v_j^{n} - v_{j-1}^{n})\Big], \\ S_j^{(t)} &= \frac{1}{k}\int_{-1/2}^{1/2} \xi\, (v_j^{n+1} - v_j^{n})^T H(v_j^{n+1/2})(v_j^{n+1} - v_j^{n})\, \mathrm{d}\xi. \end{aligned}$$

For the scalar case, one can choose the quadratic entropy $\eta(u) = \frac{1}{2}u^2$, which gives $H(v) \equiv 1$, and $S_j^{(t)} = 0$. This shows, that for the scalar case, the Crank Nicolson scheme is entropy conservative.

Now, we want to show a few numerical results illustrating the behavior of entropy conservative schemes in the scalar case. We consider the linear advection problem $u_t + au_x = 0$, with the quadratic entropy $\eta(u) = \frac{1}{2}u^2$. As already noted, in this case $v(u) = u$. The entropy flux is $\psi(u) = \frac{1}{2}au^2$ and the entropy potential is $\theta(v) = \frac{1}{2}av^2$. Consequently the entropy conservative numerical flux is $G^*_{j+1/2} = \frac{1}{2}(v_{j+1} - v_j)$, which gives the centered scheme. Clearly, this space discretization is unstable, when coupled with the Forward Euler scheme. In fact the entropy production due to the explicit time differencing is not counterbalanced by entropy dissipation from the space discretization. When the entropy conservative scheme is coupled with Crank Nicolson time differencing, a fully discrete entropy conservative scheme is obtained. The results obtained in this case on a smooth problem are compared with an entropy stable in time scheme, as the Backward Euler time differencing in Fig. 4. It is clear that the fully discrete entropy conservative scheme is much more accurate than Backward Euler. The error for the entropy conservative scheme is due mainly to a dispersion error characteristic of second order schemes, while the error of the entropy stable implicit scheme is due to a very high numerical diffusion. In both cases, the grid parameter $\lambda = \frac{k}{h} = 1$, although both schemes are unconditionally stable.

On the other hand, the fully discrete entropy conservative scheme introduces spurious oscillations on non smooth solutions, see Fig. 5. Note that these are contact discontinuities, so that no entropy dissipation occurs in the exact solution. Still, entropy dissipation is clearly necessary to obtain a solution without oscillations. However, the scheme does have some weak form of stability, which is illustrated by the fact that the Total Variation of the numerical solution remains bounded in time.



(a) (b)

Figure 5: Linear advection of a non smooth function with a fully discrete entropy conservative scheme, $N = 100, N = 200, N = 400, N = 800$ cells, up to $T = 32$. On the right, we show the behavior of the total variation and the total entropy as functions of time (bottom curve), which remains perfectly constant for all grids. (a) Solution up to $T = 32$, (b) Total variation and entropy in time.

3.3. High order schemes

The construction of high order entropy stable schemes follows [11] and [12]. First, high order entropy conservative space discretizations are constructed, which are later modified adding diffusion terms to preserve high order and ensure entropy stability. The construction starts from the basic entropy conservative flux computed in (24). This numerical flux is based on two points, and higher order entropy conservative schemes are obtained enlarging the stencil and considering linear combinations of the basic two point entropy conservative flux:

$$G^{p,*}_{j+1/2}(v_{j-p+1},\dots,v_{j+p}) = \sum_{k,r=0}^{p-1} \beta_{r,k} G^*(v_{j-k}, v_{j+1+r}). \tag{42}$$

Note that for $p = 1$ the basic entropy conservative flux is recovered, which is second order accurate. In [12], it is proven that, choosing appropriately the coefficients $\beta_{r,k}$, it is possible to obtain numerical fluxes which are $2p$ order accurate. For example:

$$G^{2,*}_{j+1/2} = \frac{4}{3}G^*(v_j, v_{j+1}) - \frac{1}{6}G^*(v_{j-1}, v_{j+1}) - \frac{1}{6}G^*(v_j, v_{j+2}), \tag{43}$$

$$\begin{aligned} G^{3,*}_{j+1/2} = {} & \frac{3}{2}G^*(v_j, v_{j+1}) - \frac{3}{10}\left[G^*(v_{j-1}, v_{j+1}) + G^*(v_j, v_{j+2})\right] \\ & + \frac{1}{30}\left[G^*(v_{j-2}, v_{j+1}) + G^*(v_{j-1}, v_{j+2}) + G^*(v_j, v_{j+3})\right]. \end{aligned} \tag{44}$$

Entropy stable numerical fluxes which are $2p$ order accurate are obtained modifying (28) and designing highly accurate diffusion terms $D_{j+1/2}$. It has already been noted that to obtain a q order scheme the diffusion terms must be of order q. Thus the diffusion terms written in the form (29) are modified applying diffusion to reconstructed data. For this purpose, a piecewise polynomial reconstruction operator of degree $q-1$ is applied to the grid values of the entropic variables, obtaining the function $\mathcal{V}(x)$, and the extrapolated values at the cell interfaces are computed as

$$v^+_{j+1/2} = \lim_{x \to x_j + \frac{h}{2}^-} \mathcal{V}(x) \qquad \text{and} \qquad v^-_{j+1/2} = \lim_{x \to x_{j+1} - \frac{h}{2}^-} \mathcal{V}(x).$$

If the function v is smooth, then $v^+_{j+1/2} - v^-_{j+1/2} = O(q+1)$, and thus the new diffusion terms will be given by

$$D_{j+1/2} = \alpha_{j+1/2}\left(v^+_{j+1/2} - v^-_{j+1/2}\right). \tag{45}$$

It is straightforward to prove that the space entropy dissipation (31) for the semidiscrete scheme becomes

$$S_j^{(x)} = -\frac{1}{4h}\left[(v_{j+1}-v_j)^T\alpha_{j+1/2}\left(v^+_{j+1/2}-v^-_{j+1/2}\right)+(v_j-v_{j-1})^T\alpha_{j-1/2}\left(v^+_{j-1/2}-v^-_{j-1/2}\right)\right]. \tag{46}$$

Since the schemes must be entropy stable, the entropy space dissipation must be negative. Therefore, the jump in the reconstructed values must have the same sign as the jump in grid values of v, i.e. $(v_{j+1}-v_j)\cdot(v^+_{j+1/2}-v^-_{j+1/2})\geq 0$. This requirement is called *sign preserving property* in [12], where it is proven that the sign preserving property is enjoyed by the ENO (Essentially Non Oscillatory) reconstruction algorithm of [29], and the standard piecewise linear reconstruction with the MinMod limiter. However, the sign preserving property does not hold for other standard reconstruction algorithms, such as the piecewise linear reconstruction with the Superbee limiter, or the WENO reconstruction [5].

Not much is known on the behavior of the entropy dissipation due to high order time integrators. Usually, high order fully discrete schemes for systems of conservation laws rely on Runge-Kutta schemes for the time discretization. Runge Kutta methods are defined by their Butcher tableax (A,b), where b is a vector of ν coefficients, and A is a $\nu\times\nu$ matrix of coefficients; ν is the number of stages of the Runge-Kutta scheme. For explicit Runge-Kutta methods (ERK) the matrix A is strictly lower triangular. Introducing a Runge-Kutta discretization in time in the conservative semidiscrete scheme

$$\frac{\mathrm{d}}{\mathrm{d}t}u_j+\frac{1}{h}\left(F_{j+1/2}-F_{j-1/2}\right)=0, \tag{47}$$

the numerical solution can be written as

$$U_j^{n+1}=U_j^n-\lambda\sum_{i=1}^{\nu}b_i\left(F^{(i)}_{j+1/2}-F^{(i)}_{j-1/2}\right), \tag{48}$$

where, for an explicit scheme, the numerical fluxes $F^{(i)}_{j+1/2}$ are computed only from the already known stage values, which are given by

$$U_j^{(i)}=U_j^n-\lambda\sum_{k=1}^{i-1}a_{i,k}\left(F^{(k)}_{j+1/2}-F^{(k)}_{j-1/2}\right). \tag{49}$$

Within this framework, it is quite natural to define the cell entropy dissipation due to the fully discrete scheme as:

$$S_j=\frac{1}{k}\left(\eta(U)_j^{n+1}-\eta(U)_j^n\right)+\frac{1}{h}\left(\Psi^n_{j+1/2}-\Psi^n_{j-1/2}\right), \tag{50}$$

where

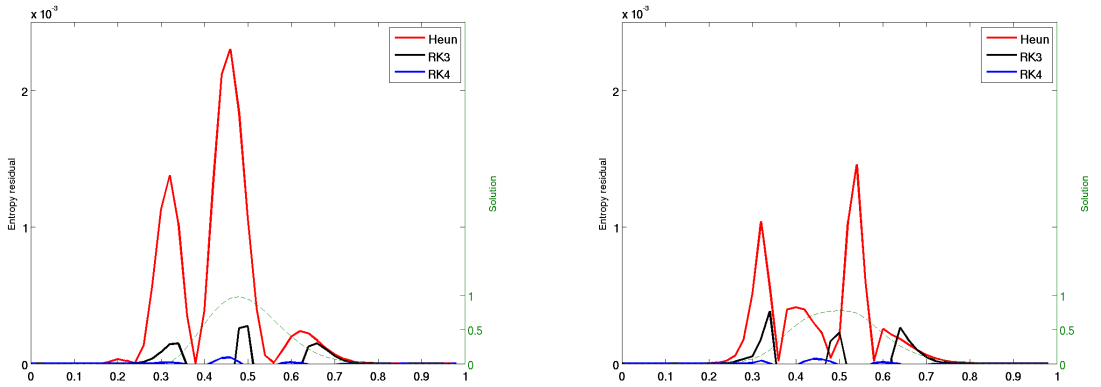
$$\Psi^n_{j+1/2}=\sum_{i=1}^{\nu}b_i\Psi^{(i)}_{j+1/2}$$

$$\Psi^{(i)}_{j+1/2}=\Psi(v^{(i)}_{j-p+1},\dots,v^{(i)}_{j+p})=\sum_{k,r=0}^{p-1}\beta_{r,k}\Psi\left(v^{(i)}_{j-k},v^{(i)}_{j+1+r}\right),$$

and the two-point numerical entropy flux $\Psi(a,b)$ was defined in (23), the notation for the high order flux is established in (42) and clearly $v_j^{(i)}=v(U_j^{(i)})$.

Since the expression for Ψ is known, the quantity S_j given by (50) is computable, once the numerical solution has been updated. Due to the high non-linearity of the Runge-Kutta scheme, it is quite hard to separate the space from the time contributions to S_j, as was done for the simpler schemes before. However, in analogy with the case of the Backward and the Forward Euler scheme, and keeping into account the structure of the Runge Kutta scheme, the space contribution to the entropy dissipation can be *defined* as

$$S_j^{(x)}=-\frac{1}{4h}\sum_{i=1}^{\nu}b_i\left[(v^{(i)}_{j+1}-v^{(i)}_j)^T\alpha_{j+1/2}\left(v^{+,(i)}_{j+1/2}-v^{-,(i)}_{j+1/2}\right)+(v^{(i)}_j-v^{(i)}_{j-1})^T\alpha_{j-1/2}\left(v^{+,(i)}_{j-1/2}-v^{-,(i)}_{j-1/2}\right)\right]. \tag{51}$$



(a) (b)

Figure 6: Linear advection of a smooth function. Numerical solution (dashed line). Time entropy residual for schemes of order 2, 3, and 4 (from top to bottom). (a) Entropy conservative, (b) Entropy stable.

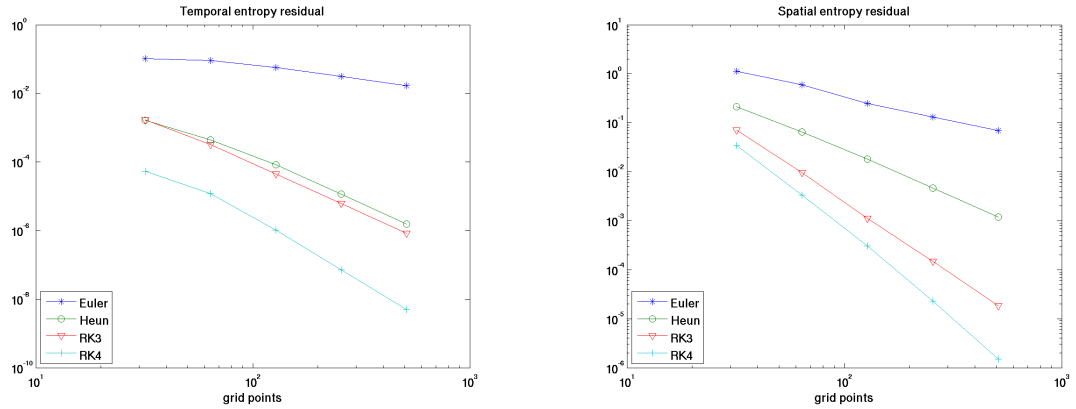
In this case, the time contribution to the cell entropy dissipation can be estimated only a-posteriori as

$$S_j^{(t)} = S_j - S_j^{(x)}. \tag{52}$$

The definition above of the contribution to the entropy dissipation given by the time discretization is not exact, because space and time are non linearly coupled in a high order scheme with Runge-Kutta time advancement, and at present it is not known how to separate the two contributions. However, we can have an idea of how reliable the construction proposed really is, comparing the entropy dissipation due to an entropy conservative (in space) semidicrete scheme integrated in time with a Runge-Kutta method, and the time entropy dissipation $S_j^{(t)}$ of an entropy stable scheme, with $S_j^{(t)}$ defined in (52). In the first case, we find the *exact* entropy behavior of the time integrator, because the space entropy dissipation is zero by construction. This procedure can be applied only for smooth flows, because the numerical solution would start oscillating in the presence of discontinuities and/or high gradients. In the second case, we are only estimating the time contribution to the entropy production, but we are working on a scheme that can be applied also to non linear problems developing shocks. The differences in the two approaches are shown in Fig. 6. Both plots show the time entropy production due to several Runge-Kutta schemes, for a fixed number of grid points in space. On the left, we show the total entropy production due to a fully discrete scheme built with entropy conservative fluxes; note that, by construction, the entropy production of this scheme is due only to the time discretization. On the right, we plot $S_j^{(t)}$ for a fully discrete scheme using entropy stable fluxes. As the figure shows, the two results do not coincide, but they have the same behavior. In particular, in both cases, the explicit Runge-Kutta schemes give a *positive contribution* to the entropy production, which therefore must be counterbalanced by entropy dissipation from the space discretization. However, the spurious entropy production due to the time discretization decreases very fast increasing the order of the time integrator. Note that the two numerical solutions (dashed line in the background) are quite different: these results were obtained with a small number of grid points ($N = 50$), and the second order scheme. The solution on the left does not suffer from numerical diffusion, and is very close to the exact solution, on the right the effect of the diffusive fluxes, needed to achieve entropy stability, smears the profile.

In the following part of this section, we want to show a few results concerning the entropy behavior of high order Runge Kutta schemes for the time integration of scalar conservation laws, coupled with high order entropy stable fluxes. In each cell entropy dissipation is computed with (50), the space contribution is given by (51), and the time contribution is found by subtraction as in (52). We underline that the splitting we propose between the space and the time dissipation is only approximate, since both terms are intertwined in the Runge-Kutta time advancement. Still, the data we obtain permit to compute the total entropy dissipation, and to give ideas of what is the order of magnitude of the terms involved.

Now, we consider a non linear test. The equation is Burgers' equation with initial data $u_0(x) = \sin(x)$, and periodic boundary conditions. In this problem, a shock develops at $T = 1/(2\pi)$ starting from zero strength, reaching a



(a) (b)

Figure 7: Burgers' solution before shock formation. Rate of convergence of the entropy residuals for fully discrete Runge-Kutta schemes with entropy stable fluxes. (a) $\max |S_j^{(t)}|$, (b) $\max |S_j^{(x)}|$.

maximum strength which later decreases with time. Fig. 7 shows the rate of convergence of the entropy residuals for fully discrete entropy stable schemes, from order 1 to order 4, before shock formation. Both residuals converge to zero with the expected rate, but it is clear that the entropy dissipation due to the space discretization is orders of magnitude larger than the entropy production due to the time integrators. Next, Fig. 8 shows the behavior of the entropy with time as the solution develops a shock, for several schemes. The solution is obtained with $N = 50$ grid points. On the left, we show the difference $\boldsymbol{\eta}(t)$ between the total entropy present in the computational domain as a function of time and the initial entropy. Since the scheme dissipates entropy, the quantity $\boldsymbol{\eta}(t)$ is negative, and it further decreases with time. Here the absolute value of this quantity is shown. The vertical line indicates the time of shock formation. On the right, the absolute value of the total entropy dissipated in each time step appears. The different curves correspond to the different schemes. As long as the flow remains smooth, the entropy dissipation decreases in absolute value as the order is increased. After shock formation however, all curves converge towards the exact entropy dissipation. In particular, the exact entropy dissipation per time step for a single shock is given by, see (8):

$$S_{\text{exa}}(t) = \frac{1}{h}\,[\![\eta]\!]\left(-\frac{[\![f]\!]}{[\![u]\!]} + \frac{[\![\psi]\!]}{[\![\eta]\!]}\right),$$

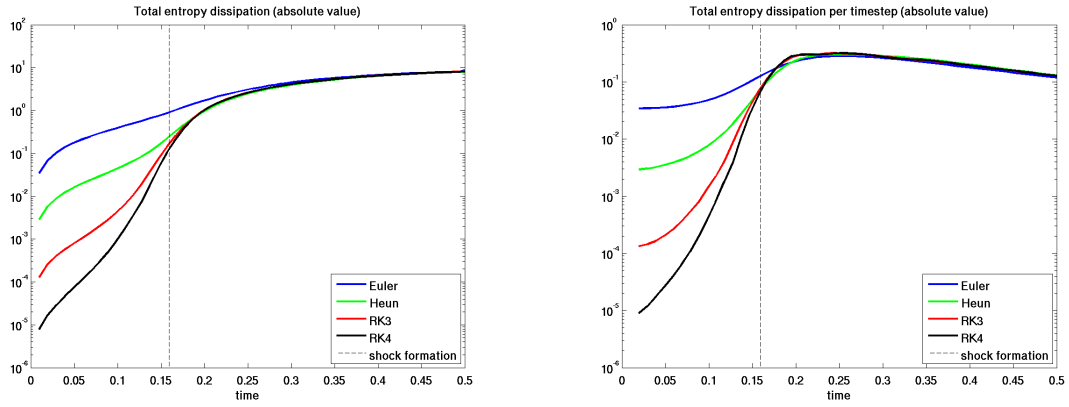
where $[\![u]\!]$ denotes the jump in the variable u across the shock, since only the shocked cell gives a contribution. This is the value the curves in the plots converge to. Note that the entropy on the shock per time step changes with time. Initially, in this test problem, $S_{\text{exa}}(t)$ increases as the shock develops starting from zero strength, and later it decreases as the shock strength slowly decays.

4. Entropy residual as an a posteriori error indicator

In this section we show how to compute an error indicator using the entropy production due to a finite volume scheme. The indicator can then be used as a tool to drive grid adaptivity. This technique was presented in [13] for central schemes based on staggered grids, and later extended to standard finite volume schemes in [14]. A more rigorous approach which however applies only to scalar conservation laws can be found in [15].

The idea is to use the entropy production in each cell, induced by a standard finite volume scheme, to estimate the local residual. This information can be exploited to build an *a posteriori error indicator* to drive the construction of an adaptive mesh. For a standard finite volume scheme, the numerical solution satisfies equation (10) in all cells. Using the information needed to compute the numerical flux F, a consistent entropy flux Ψ can be computed, and the entropy production in each cell corresponding to the scheme above can be defined as

$$S_j^n = \frac{\overline{\eta(U)}_j^{n+1} - \overline{\eta(U)}_j^n}{k} + \frac{1}{h}\left(\Psi_{j+1/2} - \Psi_{j-1/2}\right). \tag{53}$$



(a) (b)

Figure 8: Burgers' solution and shock formation. Total entropy as a function of time (left) minus initial entropy, and total entropy dissipation per time step (right) for the fully discrete Runge-Kutta schemes with entropy stable fluxes, for order from 1 to 4. (a) Total entropy , (b) Entropy production per time step.

However, even on smooth flows $S_j^n \neq 0$. It will be shown that on smooth flows

$$S_j^n = O(h)^p,$$

where p is the accuracy of the underlying numerical scheme. On shocks,

$$S_j^n \simeq \frac{1}{h}.$$

Thus S_j^n has the same size of the local error on smooth flows, while it is large on discontinuities.

These ideas can be applied on a standard high order finite volume method, based on Runge-Kutta time adavancement and piecewise polynomial reconstructions. Suppose the RK method has ν stage values, and it is based on the Runge-Kutta tableaux defined by the coefficient matrices (A, b). Then, equation (48) is applied to the cell averages of the solution, obtaining the updated cell averages as

$$\overline{U}_j^{n+1} = \overline{U}_j^n - \lambda \sum_{i=0}^{\nu} b_i \left(F_{j+1/2}^{(i)} - F_{j-1/2}^{(i)} \right), \tag{54}$$

where the numerical fluxes are computed with a standard first order numerical flux $F(a, b)$, such as Lax-Friedrichs or Godunov, to the boundary extrapolated data of the solution at the same stage:

$$F_{j+1/2}^{(i)} = F(U_{j+1/2}^{(i),-}, U_{j+1/2}^{(i),+}).$$

The boundary extrapolated data are obtained applying a non-oscillatory reconstruction procedure $\mathcal{R}(x)$ to the cell averages, obtaining the values at the cell edges:

$$\mathcal{R} : \{\overline{U}\} \to \mathcal{R}(x), \qquad U_{j+1/2}^{+} = \lim_{x \to x_j + \frac{h}{2}^-} \mathcal{R}(x), \qquad U_{j+1/2}^{-} = \lim_{x \to x_{j+1} - \frac{h}{2}^-} \mathcal{R}(x).$$

Thus the boundary extrapolated data $U_{j+1/2}^{(i),\pm}$ are computed from the cell averages of the stage values at level (i), which are given by:

$$\overline{U}_j^{(i)} = \overline{U}_j^n - \lambda \sum_{k=1}^{i-1} a_{ik} \left(F_{j+1/2}^{(k)} - F_{j-1/2}^{(k)} \right).$$

Therefore a single Runge-Kutta time step requires ν evaluations of the numerical fluxes and of the reconstruction $\mathcal{R}$ at each cell interface. For high order schemes, it is known that the numerical flux chosen influences the solution less

and less as the order of the scheme is increased, [5]. So one chooses simple and fast numerical fluxes, such as the local Lax Friedrichs flux. On the other hand the reconstruction becomes more complex as the order of accuracy is increased. As a consequence, most of the computational effort is concentrated on the reconstruction of the boundary data.

Once the solution U has been updated, the entropy production can be computed. Choose a numerical entropy flux Ψ, consistent with $\psi(u)$. At each Runge-Kutta stage, evaluate

$$\Psi^{(i)}_{j+1/2} = \Psi(U^{(i),-}_{j+1/2}, U^{(i),+}_{j+1/2}). \tag{55}$$

Note that here the numerical entropy flux is applied to the boundary extrapolated values which have already been reconstructed for computing $F^{(i)}_{j+1/2}$. Then the numerical entropy production in the j^{th} cell is defined as

$$S^n_j = \frac{1}{k}\left\{\overline{\eta\left(U^{n+1}\right)}_j - \overline{\eta\left(U^n\right)}_j + \lambda \sum_{i=1}^{\nu} b_i \left(\Psi^{(i)}_{j+1/2} - \Psi^{(i)}_{j-1/2}\right)\right\}, \tag{56}$$

where

$$\overline{\eta\left(U^n\right)}_j = \frac{1}{h}\int_{x_j-h/2}^{x_j+h/2} \eta\left(\mathcal{R}(x)\right)\, \mathrm{d}x.$$

This integral will be estimated with quadrature, using the same reconstruction operator used to compute the boundary extrapolated data. For order up to 2, $\overline{\eta(U^n)}_j$ can be evaluated with the midpoint quadrature rule: $\overline{\eta\left(U^n\right)}_j = \eta\left(\overline{U}^n_j\right) + O(h)^2$. For higher accuracy, an improved quadrature rule must be chosen to evaluate $\overline{\eta(U^n)}_j$. It is important to note that the evaluation of the entropy indicator (56) costs *less* than the actual update of the solution (54), because no extra reconstructions are needed.

A few properties of the entropy indicator defined in (56) now follow.

Theorem 4.1. Rate of convergence. *Consider a conservative scheme of order p, of the form* (54)*. Choose a numerical entropy flux* Ψ *consistent with the entropy flux* ψ*. Evaluate the entropy production* S^n_j *as in* (56)*. Then, if the solution is smooth in the* $j-$*th cell:*

$$S^n_j = O(h^p) \ \textit{for}\ \ h \to 0,$$

If the solution is not smooth, in the $j-$*th cell,*

$$S^n_j = O(1/h)$$

Thus the theorem ensures that the entropy production is of the same order of the local truncation error of the scheme, where the solution is smooth, and the grid is fine enough to resolve the accuracy of the scheme. On the other hand, $S^n_j \sim \frac{C}{h}$ on shocks. The proof is based on the fact that S^n_j is the residual of a finite volume scheme for $\partial_t\eta(u) + \partial_x\psi(u) = 0$.

We give the proof using the midpoint rule as a quadrature rule for the evaluation of $\overline{\eta(U^n)}_j$. The general case is similar and it can be found in [14].

Proof. Using the midpoint rule for evaluating $\overline{\eta(u)}$, the time part of the entropy residual can be written as

$$\overline{\eta(U^n)}_j = \eta(\overline{U}^n_j) + \frac{h^2}{24}\,\eta''\big|_{U(x_j)}\, U_x|_{x_j} + O(h^3).$$

Let $u(t,x)$ be the exact solution and take $\overline{U}^n_j = \overline{u}^n_j$, so that the local error is $\overline{U}^{n+1} = \overline{u}^{n+1} + O(h^{p+1})$ for a scheme of order p. Then

$$\begin{aligned}
\overline{\eta\left(U^{n+1}\right)}_j - \overline{\eta\left(U^n\right)}_j &= \eta(\overline{U}^{n+1}_j) + \cancel{\frac{h^2}{24}\eta'' U_x} - \eta(\overline{U}^n_j) - \cancel{\frac{h^2}{24}\eta'' U_x} + O(h^3)\\
&= \eta(\overline{u}^{n+1}_j) + O(h^{p+1}) - \eta(\overline{u}^n_j) + O(h^3)\\
&= \overline{\eta(u)}^{n+1}_j - \overline{\eta(u)}^n_j + O(h^{\min(3,p+1)}) = \frac{1}{h}\int_{I_j}\Big[\eta(u(t^{n+1},x)) - \eta(u(t^n,x))\Big]\mathrm{d}x + O(h^{\min(3,p+1)})\\
&= \frac{1}{h}\int_{I_j}\int_{t^n}^{t^{n+1}} \frac{\partial\eta(u(t,x))}{\partial t}\mathrm{d}x\,\mathrm{d}t + O(h^{\min(3,p+1)}).
\end{aligned}$$

Now the space part of S_j^n is estimated. Since a Runge-Kutta scheme can be written as a time quadrature rule, and the numerical flux $\Psi_{j+1/2}$ is consistent with the exact entropy flux ψ,

$$\begin{aligned}
\lambda \sum_{i=1}^{\nu} b_i \left(\Psi_{j+1/2}^{(i)} - \Psi_{j-1/2}^{(i)}\right) &= \frac{1}{h}\int_{t^n}^{t^{n+1}} \left(\Psi(U_{j+1/2}^+(t), U_{j+1/2}^-(t)) - \Psi(U_{j-1/2}^+(t), U_{j-1/2}^-(t))\right) \mathrm{d}t + O(k^{p+1}) \\
&= \frac{1}{h}\int_{t^n}^{t^{n+1}} \left[\Psi(u(x_{j+1/2},t), u(x_{j+1/2},t)) - \Psi(u(x_{j-1/2},t), u(x_{j-1/2},t))\right] \mathrm{d}t + O(k^{p+1}) + O(h^{p+1}) \\
&= \frac{1}{h}\int_{t^n}^{t^{n+1}} \left[\psi(u(x_{j+1/2},t)) - \psi(u(x_{j-1/2},t))\right] + O(k^{p+1}) + O(h^{p+1}) \\
&= \frac{1}{h}\int_{t^n}^{t^{n+1}} \int_{I_j} \frac{\partial \psi(u)}{\partial x} \mathrm{d}x\mathrm{d}t + O\left(h^{p+1}, k^{p+1}\right),
\end{aligned}$$

where $k = \lambda h$ was used together with the assumption that the scheme is $p-$th order accurate on smooth solutions. Now, putting together the two contributions

$$S_j^n = \frac{1}{k}\left\{\frac{1}{h}\int_{t^n}^{t^{n+1}} \int_{I_j} \left(\cancel{\frac{\partial \eta(v)}{\partial t}} + \cancel{\frac{\partial \psi(v)}{\partial x}}\right) \mathrm{d}x\mathrm{d}t + O\left(h^{\min(3,p+1)}, k^{p+1}\right)\right\} = O(h^{\min 2,p}) \tag{57}$$

The proof can be extended to higher order schemes, improving the quadrature for $\overline{(\eta)}$ as in [14], thus removing the $O(h^2)$ term.

If the cell j contains a discontinuity, each term contributing to the definition of S_j^n is bounded. Thus $S_j^n \sim \frac{C}{h}$ on shocks. ■

The entropy indicator preserves the correct sign of the entropy dissipation, provided the numerical entropy flux is not only consistent, but accurately tailored to the particular numerical flux used to update the solution.

Theorem 4.2. First order scheme with Lax-Friedrichs flux. *Suppose a scalar conservation law is integrated with the Forward Euler scheme and the Lax Friedrichs flux, with* $\alpha = \max|f'(u)|$. *Choose*

$$\begin{aligned}
F_{j+1/2} &= \frac{1}{2}\left[f(U_{j+1}) + f(U_j) - \alpha(U_{j+1} - U_j)\right], \\
\Psi_{j+1/2} &= \frac{1}{2}\left[\psi(U_{j+1}) + \psi(U_j) - \alpha(\eta(U_{j+1}) - \eta(U_j))\right].
\end{aligned}$$

Then

$$S_j^n \le 0$$

away from local extrema, while small positive overshoots may occour in those cells containing a local extremum, with $|S_j^n| = O(h^4)$

The proof is quite technical and appears in [14]. However, the line of the proof can be appreciated from the proof of the following result.

Theorem 4.3. First order scheme with Upwind flux. *Suppose a scalar conservation law with* $f'(u) > 0$ *is integrated with the Forward Euler scheme and the Lax Friedrichs flux. Choose*

$$\begin{aligned}
F_{j+1/2} &= f(U_j), \\
\Psi_{j+1/2} &= \psi(U_j).
\end{aligned}$$

Then $S_j^n \le 0$.

Proof. Thanks to the particular choice of the numerical fluxes, the space part of the entropy production is

$$\Psi_{j+1/2} - \Psi_{j-1/2} = \psi(U_j^n) - \psi(U_{j-1}^n) = \int_{U_{j-1}^n}^{U_j^n} \psi'(u)\mathrm{d}u = \int_{U_{j-1}^n}^{U_j^n} \eta'(u) f'(u)\mathrm{d}u.$$

while, for the time part

$$\eta\left(U_j^{n+1}\right) - \eta\left(U_j^n\right) = \int_{U_j^n}^{U_j^{n+1}} \eta'(v)\mathrm{d}v.$$

Introducing the change of variables $v(u) = U_j^n - \lambda[f(U_j^n) - f(u)]$, the time part becomes

$$\eta\left(U_j^{n+1}\right) - \eta\left(U_j^n\right) = \int_{U_j^n}^{U_{j-1}^n} \eta'(v(u))\lambda f'(u)\mathrm{d}u.$$

Adding the two terms, dropping the index n, and applying twice the mean value theorem one gets:

$$kS_j^n = -\int_{U_{j-1}}^{U_j} \eta''(\xi)\left[1 - \lambda f'(\eta)\right]\lambda f'(u)(U_j - u)\,\mathrm{d}u,$$

where ξ and η are contained in the interval formed by U_{j-1} and U_j. Finally, recalling that $\eta'' \geq 0$, that the CFL condition ensures that the first parenthesis is positive, and that $f'\lambda \geq 1$ by hypothesis, the sign of S_j^n coincides with the sign of $-\int_{U_{j-1}}^{U_j}(U_j - u)\,\mathrm{d}u$ which is negative, as required. ■

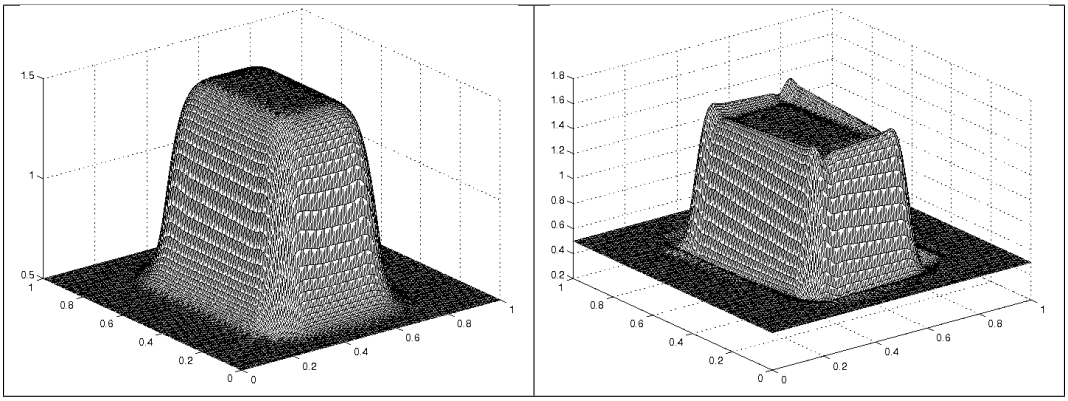


Figure 9: Linear rotation of an initial rectangular patch. Solution with limiters (left), and without limiters (right)

We illustrate this result showing the entropy production obtained integrating on a fixed grid, with a second order scheme based on a piecewise linear reconstruction, the linear equation

$$\partial_t u + \partial_x\left(-(y - \tfrac{1}{2})\tfrac{\pi}{2}u\right) + \partial_y\left((x - \tfrac{1}{2})\tfrac{\pi}{2}u\right) = 0,$$

with initial data consisting of a rectangular patch. Fig 9 shows the solution with and without limiters. As expected, the vertical sides of the patch are smoothed by numerical diffusion, and the solution without limiters presents spurious oscillations. Fig. 10 shows the entropy production corresponding to the numerical solutions shown in Fig. 9. The indicator signals clearly the presence of the discontinuities, but it also detects the presence of spurious oscillations, which correspond to the positive overshoots in the entropy production. The non-oscillatory solution results in entropy dissipation with the correct sign. On the other hand, Fig. 11 shows the error indicator by Karni-Kurganov-Petrova (KKP) presented in [25], again for the numerical solutions shown in Fig. 9. In this case the indicator detects the presence of singularities, but it does not give indications to distinguish the entropic solution from the solution with spurious oscillations.

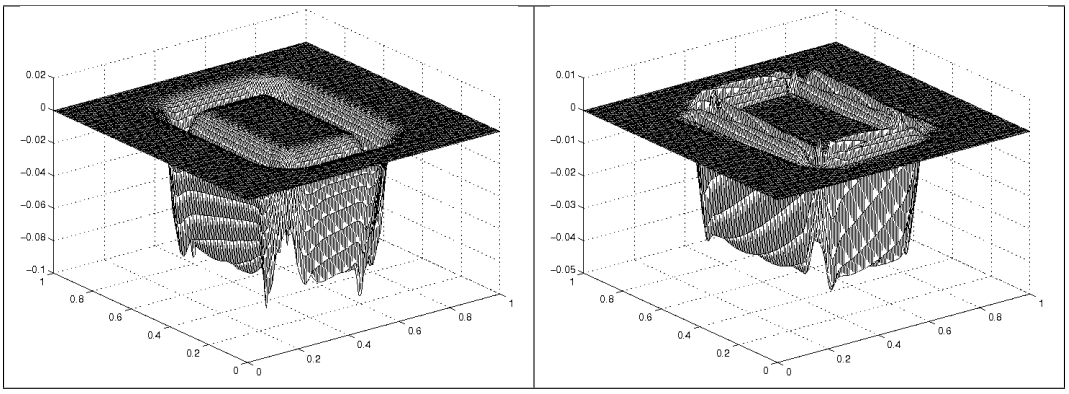


Figure 10: Linear rotation of an initial rectangular patch. Entropy production for the solution with limiters (left), and without limiters (right)

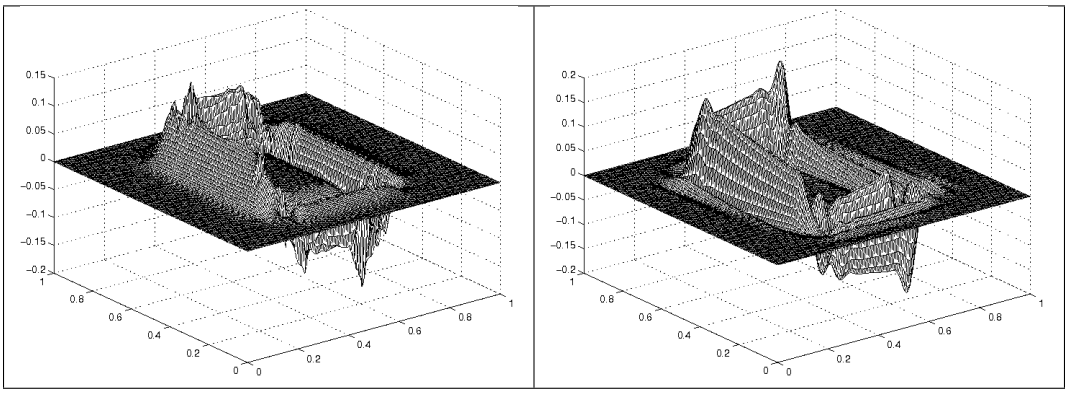


Figure 11: Linear rotation of an initial rectangular patch. Karni-Kurganov-Petrova error indicator for the solution with limiters (left), and without limiters (right)

The rate of convergence of the entropy indicator is shown in Fig. 12, for a second and a for fourth order scheme, applied to a scalar conservation law. The figure exhibits the rate of convergence on a shock, on a contact discontinuity and on a smooth transition, together with the reference slopes. It is clear that $|S_j| \simeq O(\frac{1}{h})$ on a shock, $|S_j| \simeq O(h^0)$ on a contact discontinuity, while $|S_j| \simeq O(h^p)$ on smooth solutions. Moreover, the figure illustrates clearly that $|S_j|$ varies by several orders of magnitude on the different solutions. This characteristic is important for adaptive strategies, because it means that an adaptive grid driven by the entropy indicator will not be too sensitive to the threshold chosen.

4.1. Adaptive algorithm

In this section, we show results obtained applying the entropy indicator as an a posteriori error indicator to select an adaptive grid that changes in time according to the behavior of the solution. The grid used here is based on mesh widths that change as powers of 2 from a maximum length h_0, so that a cell at level l will be characterized by a mesh width $h_l = h_0 2^{-l}$. The non uniform grid on which the numerical scheme is applied is stored in a tree structure, in order to have easy access to each given cell and its neighbors. The memory structure with which the grid is stored is shown in Fig. 13. Cartesian grids are used in these examples, as in the Conservation Laws Pack (CLAWPACK) [30], but, unlike CLAWPACK, in this case only the data for active cells is kept in memory, i.e. only one single grid with cells of different sizes is stored, as opposed to patches of overlayed uniform grids, as in [30]. The error indicator determines the local mesh size. The time step can be chosen following two strategies.

- dt-mode: use the same time step everywhere. With this choice, the CFL condition imposes that the time step

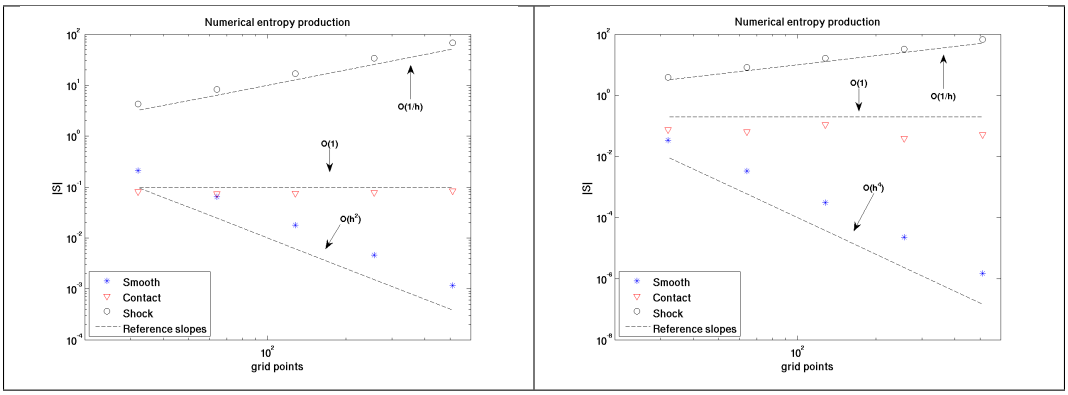


Figure 12: Rate of decay of the entropy indicator versus grid size for second order (left), and fourth order (right) schemes. The dashed lines indicate the different rates: from top to bottom h^{-1} , h^0, h^p, with $p = 2$ on the left and $p = 4$ on the right. The markers give the entropy productions due to shocks (top, circles), a contact discontinuity (triangles), and a smooth transition (stars). Note the difference in the scale of the vertical axis between the two plots.

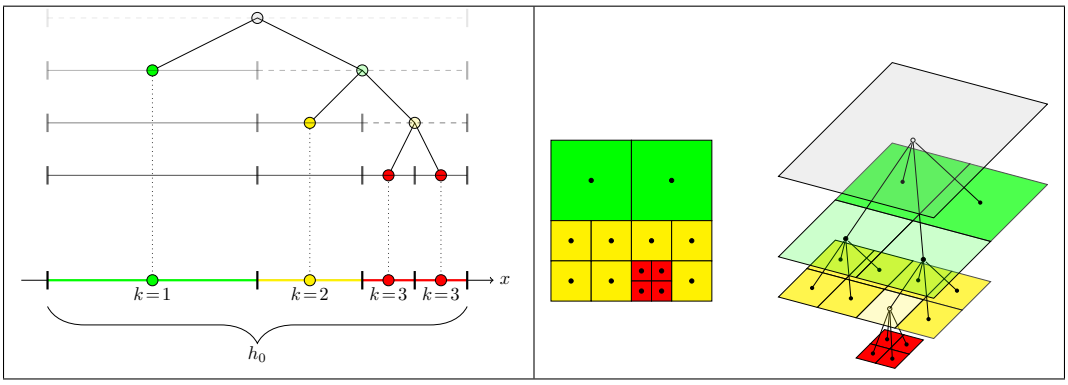


Figure 13: Dyadic one-dimensional grid (left) and Quad-tree grid for two-dimensional problems (right)

is dictated by the *smallest* cells. Thus $k \le \lambda h_l = \lambda h_0 2^{-l} \Longrightarrow k = \lambda h_0 2^{-L_{\max}}$, where $L_{\max}$ denotes the maximum level of refinement. This choice is followed for instance by [16]. This strategy is simple to implement, but can result in very small time steps on coarse cells, if the grid is highly non uniform.

- CFL-mode: local timestepping. In this case, the mesh ratio λ is fixed on all cells. Thus a cell of width h_l has a time step of length $k_l = \lambda h_l$. This choice is followed by, among others, [30], [31], [32]. In [14], we proposed time advancement based on CFL mode with a novel approach, based on updating the solution starting from the smallest cells, up to the coarsest cells. Our approach has the advantage of being exactly conservative, and does not require to embed refined cells on a hierarchy of layers containing less and less refined cells, up to the coarsest level of the grid.

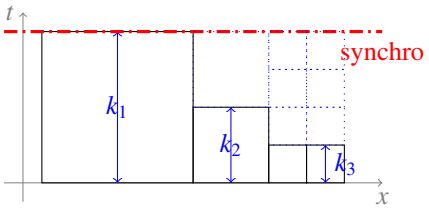
These ideas were implemented on a second order scheme, based on a piecewise linear reconstruction, with Local Lax-Fredrichs flux, and Heun (second order RK scheme) time advancement. A threshold C_{ref} is fixed. Suppose that at the time t^n, at the end of a syncronization step, the solution $\overline{U}^n$ is known on the whole grid. A time step of the adaptive scheme is:

1. Compute $\overline{U}_j^{n+1}$ and S_j^n on all grid points.
2. Recursively refine those cells where

$$S_j^n > C_{\text{ref}} \quad \text{and} \quad h_j > h_{MIN}$$

$k_l = \lambda h_l = \lambda h_0 2^{-l} = k_0 2^{-l}$:

a cell of level l makes 2^l steps, before reaching the syncronization time K_0, where the solution is computed over the whole mesh.



The new initial values for $\overline{U}_j^n$ after refinement are computed evaluating the new cell averages on the refined cells with the aid of the reconstruction. Then $\overline{U}^{n+1}$ are recomputed with the new initial data on each refined cell.

3. Coarsen any previously refined cell if

$$(S_j^n + S_{j+1}^n) < C_{\text{coa}} = \frac{C_{\text{ref}}}{4}$$

The values $\overline{U}^{n+1}$ after coarsening are computed averaging the values of the cell averages in those cells which have been merged together.

We illustrate the performance of the adaptive scheme using the classical Sod' shock tube problem, from gas dynamics. In this problem, the initial data are discontinuous, and three waves arise from the initial discontinuity: a rarefaction wave moving towards the left, a contact discontinuity and a shock wave ahead, moving towards the right. A few

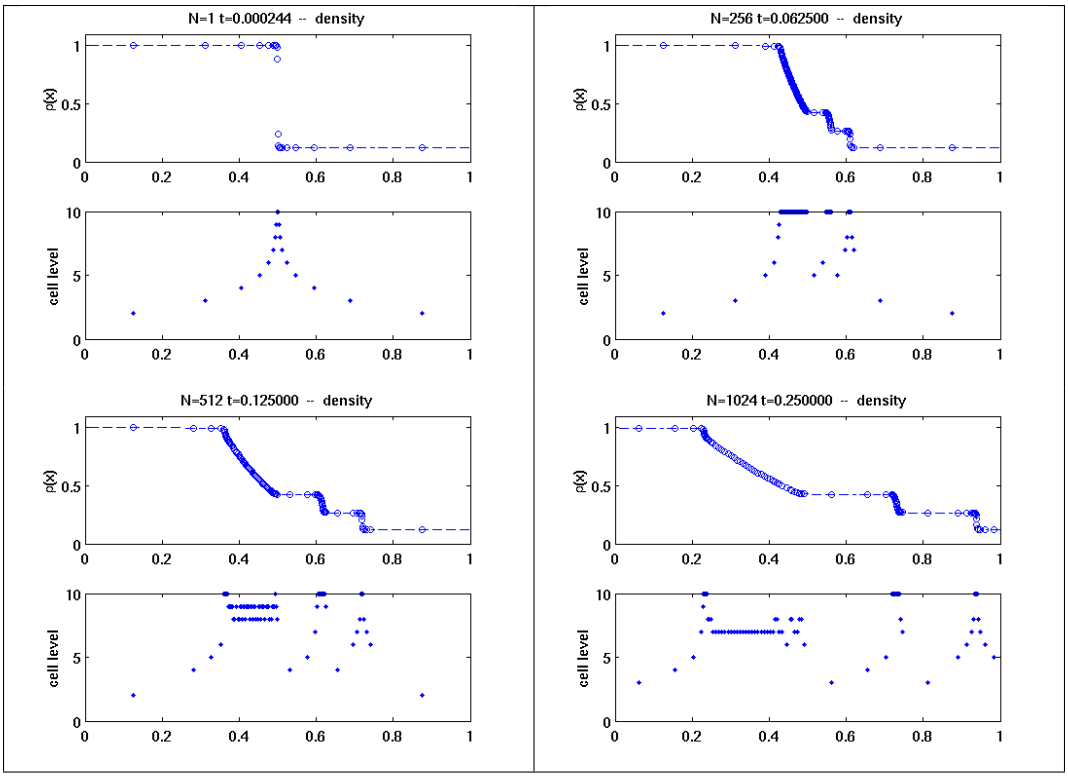


Figure 14: Sod' shock tube problem, evolution of the density profile, from the initial time (top left) to the final time (bottom right). Below each profile the corresponding grid used by the adaptive algorithm is shown

snapshots of the solution and of the corresponding grid chosen by the algorithm appear in Fig. 14. The figure shows

the density profile obtained at several times, and, immediately below, the corresponding grid chosen by the algorithm. At the initial time, the computation is initialized with only two cells, and during the first time step the algorithm recursively refines the grid around the initial step, obtaining the solution and the grid shown on the top left. As the waves separate, the patches containing the grid at the maximum level of refinement move with the waves, while in the intermediate constant states the grid is coarsened. As the two bottom figures show, the grid is coarsened also within the rarefaction wave, which is detected as smooth, except at the head and tail of the rarefaction, where discontinuities in the first derivatives occur.

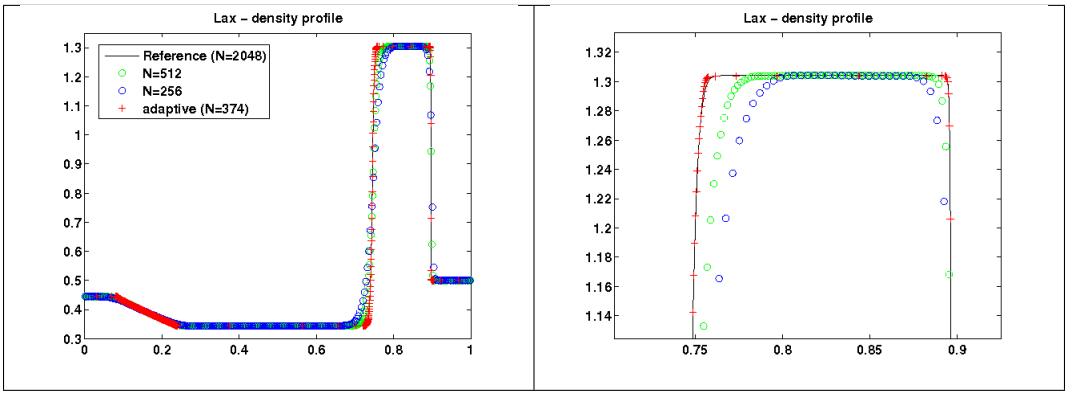


Figure 15: Lax' shock tube problem, density profile at the final time and zoom on the contact discontinuity and the shock wave (right). The red markers denote the adaptive scheme solution, which, at the final time, required 374 cells. The green and the blue markers are obtained with a fixed grid, using, respectively 256 and 512 points. The black continuous line is a reference solution, computed on a fixed grid with 2048 cells.

Finally, we show the density profile for Lax' shock tube problem. Fig. 15 compares the solution obtained by the adaptive algorithm using at the end 374 cells, (red markers) with two solutions computed on a fixed grid using slightly less ($N = 256$, green markers), and slightly more ($N = 512$, blue markers) grid points. It is clear that the adaptive solution has a better resolution than both of them. Actually, the adaptive solution has an improved resolution even with respect to the reference solution (black) which was computed with $N = 2048$ grid points, as the zoom of the density peak on the right of the figure shows.

[1] C. M. Dafermos, Hyperbolic conservation laws in continuum physics, 3rd Edition, Springer, Heidelberg, 2010.
[2] J. Smoller, Shock waves and reaction diffusion equations, Springer, 1983.
[3] E. Godlewski, P.-A. Raviart, Numerical approximation of hyperbolic systems of conservation laws, Vol. 118 of Applied Mathematical Sciences, Springer-Verlag, New York, 1996.
[4] R. J. LeVeque, Numerical methods for conservation laws, 2nd Edition, Vol. 11 of Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1992.
[5] C.-W. Shu, High order ENO and WENO schemes for computational fluid dynamics, in: High-order methods for computational physics, Vol. 9 of Lect. Notes Comput. Sci. Eng., Springer, Berlin, 1999, pp. 439–582.
[6] A. Harten, On the symmetric form of systems of conservation laws with entropy, J. Comput. Phys. 49 (1983) 151–164.
[7] S. Osher, Riemann solvers, the entropy condition and difference approximations, SINUM 22 (1984) 947–961.
[8] E. Tadmor, Numerical viscosity and the entropy condition for conservative difference schemes, Math. Comp. 43 (1984) 369–381.
[9] E. Tadmor, The numerical viscosity of entropy stable schemes for systems of conservation laws. I, Math. Comp. 49 (179) (1987) 91–103.
[10] E. Tadmor, Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems, Acta Numer. 12 (2003) 451–512.
[11] P. LeFloch, J. M. Mercier, C. Rohde, Fully discrete entropy conservative schemes of arbitrary order, SIAM J. Num. Anal.. 40 (5) (2002) 1968–1992.
[12] U. Fjordholm, S. Mishra, E. Tadmor, Arbitrarily high-order accurate entropy stable essentially non-oscillatory schemes for systems of conservation laws, 2011, Preprint.
[13] G. Puppo, Numerical entropy production for central schemes, SIAM J. Sci. Comput. 25 (4) (2003/04) 1382–1415.
[14] G. Puppo, M. Semplice, Numerical entropy and adaptivity for finite volume schemes, Comm. in Comp. Phys.
[15] D. Kröner, M. Ohlberger, A posteriori error estimates for upwind finite volume schemes for nonlinear conservation laws in multidimensions, Math. Comp. 69 (229) (2000) 25–39.
[16] A. Dedner, C. Makridakis, M. Ohlberger, Error control for a class of Runge-Kutta Discontinuos Galerkin methods for non linear conservation laws, SIAM J. Numer. Anal. (2007) 514–538.

[17] M. Ohlberger, A review of a posteriori error control and adaptivity for approximations of non-linear conservation laws, Internat. J. Numer. Methods Fluids 59 (3) (2009) 333–354.
[18] F. Bouchut, Construction of BGK models with a family of kinetic entropies for a given system of conservation laws, J. Stat. Phys. 95 (1/2) (1999) 113–170.
[19] D. I. Pullin, Direct simulation methods for compressible inviscid ideal-gas flow, J. Comput.Phys. 34 (1980) 231–244.
[20] K. Xu, A gas-kinetic BGK scheme for the Navier-Stokes equations and its connection with artificial dissipation and Godunov method, Journal of Computational Physics 171 (2001) 289–335.
[21] J.-L. Guermond, R. Pasquetti, B. Popov, Entropy viscosity method for nonlinear conservation laws, J. Comput.Phys. 230 (11) (2011) 4248–4267.
[22] P. D. Lax, Hyperbolic systems of conservation laws and the mathematical theory of shock waves, Regional Conference Series in Applied Mathematics, SIAM, 1972.
[23] P. D. Lax, Shock waves and entropy, in: N. Y. Academic Press (Ed.), Contributions to nonlinear functional analysis, 1971, pp. 603–634.
[24] M. B. Giles, E. Süli, Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality, Acta Numer. 11 (2002) 145–236.
[25] S. Karni, A. Kurganov, G. Petrova, A smoothness indicator for adaptive algorithms for hyperbolic systems, J. Comput. Phys. 178 (2) (2002) 323–341.
[26] M. G. Crandall, A. Majda, Monotone difference approximations for scalar conservation laws, Math. Comp. 34 (1980) 1–21.
[27] A. Szepessy, An existence result for scalar conservation laws using measure valued solutions, Comm. PDEs 14 (1989) 1329–1350.
[28] P. LeFloch, C. Rohde, High order schemes, entropy inequalities and non classical shocks, SIAM J. Num. Anal.. 37 (6) (2000) 2023–2060.
[29] A. Harten, B. Engquist, S. Osher, S.and Chakravarty, Uniformly high order accurate essentially non-oscillatory schemes, J. Comput. Phys. (1987) 231–303.
[30] R. J. LeVeque, CLAWPACK Version 4.3 Users Guide, http://www.amath.washington.edu/~claw/ (2006).
[31] S. Müller, Y. Stiriba, Fully adaptive multiscale schemes for conservation laws employing locally varying time stepping, J. Sci. Comput. 30 (3) (2007) 493–531.
[32] C. Dawson, R. Kirby, High resolution schemes for conservation laws with locally varying time steps, SIAM J. Sci. Comput. 22 (6) (2001) 2256–2281.