INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS, VOL. 25, 825–845 (1997)

ON THE RELATION BETWEEN THE ENTROPY BALANCE AND THE NUMERICAL SOLUTIONS OF SYSTEMS OF CONSERVATION LAWS

F. GRASSO,[1] M. MANNA,[2] C. MEOLA[3] AND A. PASCARELLI[3]*

[1] *Dipartimento di Meccanica e Aeronautica, Università 'La Sapienza', Roma, Italy*
[2] *Dipartimento di Ingegneria Meccanica per l'Energetica, Università degli Studi di Napoli 'Federico II', Napoli, Italy*
[3] *Dipartimento di Energetica, Termofluidodinamica Applicata e Condizionamenti Ambientali, Università degli Studi di Napoli 'Federico II', Napoli, Italy*

SUMMARY

This work deals with the relation between the numerical solutions of hyperbolic systems of conservation laws and the associated entropy evolution. An analysis of the continuum problem by means of variational calculus clearly emphasizes the consequences of the adopted reconstruction procedure on the induced entropy balance. A methodology is proposed that allows for *a posteriori* local and global spurious entropy production estimates on the basis of an additional equation representing a discrete approximation to the entropy inequality. The problem of defining a consistent approximation of the numerical entropy flux is also addressed in detail. Properly designed numerical experiments support the analysis and contribute to providing a more comprehensive evaluation of the numerical entropy dynamics. © 1997 John Wiley & Sons, Ltd.

Int. J. Numer. Meth. Fluids, **25**: 825–845 (1997).

No. of Figures: 10. No. of Tables: 0. No. of References: 21.

KEY WORDS: conservation laws; entropy inequality; Euler equations

1. INTRODUCTION

In this work we consider some relevant aspects of the numerical solution of the one-dimensional Euler equations of gas dynamics

$$\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = \mathbf{0} \quad \forall (x, t) \in \mathbb{R} \times (0, T), \tag{1a}$$

$$\mathbf{u}(x, 0) = \mathbf{u}_0(x), \quad \forall x \in \mathbb{R}, \tag{1b}$$

where the k-column vectors $\mathbf{u}(x, t)$ (conservative variables) and $\mathbf{f}(\mathbf{u})$ (fluxes) are given by

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho q \\ \rho E \end{pmatrix}, \qquad \mathbf{f}(\mathbf{u}) = \begin{pmatrix} \rho q \\ \rho q^2 + p \\ q(\rho E + p) \end{pmatrix} \tag{2}$$

and $\rho, q, p, E = \varepsilon + q^2/2$ and ε stand for density, velocity, pressure and total and internal energy per unit mass respectively. Moreover, an ideal gas with constant specific heats (whose ratio is γ) is

* Correspondence to: A. Pascarelli, DETEC, University of Naples, Piazzale Tecchio 80, I-80125 Naples, Italy.

CCC 0271–2091/97/070825–21 $17.50

Received March 1996
Revised November 1996

considered so that pressure is related to ε by the *equation of state* $p = (\gamma - 1)\rho\varepsilon$.

The system of conservation laws (1a) can also be cast in the non-conservative form

$$\mathbf{u}_t + \mathbf{A}(\mathbf{u})\mathbf{u}_x = \mathbf{0}, \qquad \mathbf{A}(\mathbf{u}) = \mathbf{f}_\mathbf{u}, \tag{3}$$

where $\mathbf{A}$ is the Jacobian matrix of $\mathbf{f}$ and its eigenvalues are real and distinct; hence equation (3) form a strictly hyperbolic system. Owing to the non-linearity of the flux function $\mathbf{f}(\mathbf{u})$, the solutions of (1) can develop discontinuities at a finite time $t = t^*$ in the form of shock waves and/or contact discontinuities, even when the initial data are very smooth. It follows that for $t > t^*$, $\mathbf{u}(x, t)$ is not a *classic* solution of (1a) any longer and it is therefore necessary to introduce the concept of *generalized* or *weak* solutions which satisfy (1) in the sense of distribution theory, i.e. the integral conservative relations[1]

$$\int_a^b [\mathbf{u}(x, t_2) - \mathbf{u}(x, t_1)]\mathrm{d}x + \int_{t_1}^{t_2} [\mathbf{f}(\mathbf{u}(b, t)) - \mathbf{f}(\mathbf{u}(a, t))]\mathrm{d}t = \mathbf{0}, \quad \forall (a, b) \in \mathbb{R}, \quad \forall (t_1, t_2) \in (0, T). \tag{4}$$

As a consequence, discontinuous weak solutions of (1) satisfy the Rankine–Hugoniot (R–H) condition

$$\mathbf{f}(\mathbf{u}_\mathrm{R}) - \mathbf{f}(\mathbf{u}_\mathrm{L}) = C(\mathbf{u}_\mathrm{R} - \mathbf{u}_\mathrm{L}), \tag{5}$$

where C is the speed of propagation of the discontinuity and $\mathbf{u}_\mathrm{L}$ and $\mathbf{u}_\mathrm{R}$ are the left and right states of the discontinuity.[2]

Such weak solutions are not uniquely determined by their initial data, so additional criteria are required in order to select the unique physically relevant solution. One such selection principle is to look for those solutions which satisfy a *regularized* problem, i.e. weak solutions $\mathbf{u}(x, t)$ are admissible when

$$\mathbf{u}(x, t) = \lim_{\varepsilon \to 0} \mathbf{u}^\varepsilon(x, t), \tag{6}$$

where $\mathbf{u}^\varepsilon(x, t)$ is the solution of

$$\mathbf{u}^\varepsilon_t + \mathbf{f}(\mathbf{u}^\varepsilon)_x = \varepsilon \mathbf{u}^\varepsilon_{xx}, \tag{7a}$$

$$\mathbf{u}^\varepsilon(x, 0) = u_0(x) \tag{7b}$$

and ε is a positive parameter (i.e. an artificial diffusivity coefficient). Such limit solutions, which should be stable under small perturbations of the initial data, can be characterized with the aid of a *mathematical entropy function* $U(\mathbf{u})$ as proposed by Lax.[2] Here we consider systems of conservation laws that are equipped with a non-empty set of *entropy functions*, i.e. functions which satisfy the conditions

$$U_{\mathbf{uu}} > 0, \tag{8a}$$

$$(U_\mathbf{u})^\mathrm{T} \mathbf{f}_\mathbf{u} = (G_\mathbf{u})^\mathrm{T}, \tag{8b}$$

where $G = G(\mathbf{u})$ is the so-called *entropy flux* (the superscript $^\mathrm{T}$ stands for transpose). Multiplying (1a) from the left by $U^\mathrm{T}_\mathbf{u}$ and accounting for (8b), it is easy to see that each smooth solution of (1a,b) satisfies

$$U(\mathbf{u})_t + G(\mathbf{u})_x = 0. \tag{9a}$$

Likewise, the limit solution (6) satisfies in the weak sense the *entropy condition* (or *entropy inequality*)

$$U(\mathbf{u})_t + G(\mathbf{u})_x \leqslant 0. \tag{9b}$$

Integrating (9b) over all rectangles $(a, b) \times (t_1, t_2)$ yields

$$\int_a^b [U(\mathbf{u}(x, t_2)) - U(\mathbf{u}(x, t_1))]\mathrm{d}x + \int_{t_1}^{t_2} [G(\mathbf{u}(b, t)) - G(\mathbf{u}(a, t))]\mathrm{d}t \leqslant 0. \tag{9c}$$

In fact, multiplying (7a) from the left by $U_{\mathbf{u}}^{\mathrm{T}}$ and again assuming (8b), we obtain

$$U(\mathbf{u}^\varepsilon)_t + [G(\mathbf{u}^\varepsilon) - \varepsilon(U_{\mathbf{u}}^{\mathrm{T}}\mathbf{u}_x^\varepsilon)]_x = -\varepsilon(\mathbf{u}_x^\varepsilon)^{\mathrm{T}} U_{\mathbf{uu}}\mathbf{u}_x^\varepsilon, \tag{10}$$

which shows that the mathematical entropy associated with the regularized solutions verifies a balance equation that presents both a diffusive flux and a negative (owing to (8a)) production. It can then be shown that the entropy inequality (9b) (or (9c)) follows from (10) in the limit of $\varepsilon \to 0$. Therefore admissible discontinuities satisfy the entropy jump inequality

$$G(\mathbf{u}_\mathrm{R}) - G(\mathbf{u}_\mathrm{L}) \leqslant C[U(\mathbf{u}_\mathrm{R}) - U(\mathbf{u}_\mathrm{L})]. \tag{11}$$

Inequalities (9b,c) and (11) can be used to rule out unphysical discontinuities. In the scalar case ($k = 1$), Lax[2] has shown that they also ensure the uniqueness of the weak solutions of (1) in the range of the admissible ones. These results are weaker for $k > 1$, even though they are valid in the small.

For a generic system of conservation laws there is no *a priori* guarantee of the existence of a mathematical entropy function satisfying relations (8a,b). However, in the case of the smooth solutions of the Euler equations it is well known that the *specific thermodynamic entropy* $s = s(\varepsilon, v)$, where $v = 1/\rho$ is the specific volume, verifies

$$s_t + qs_x = 0 \tag{12}$$

and use of the continuity equation

$$(\rho s)_t + (\rho q s)_x = 0, \tag{13}$$

which shows that the quantity ρs (entropy per unit volume) satisfies the condition (8b). In the presence of discontinuities the second law of thermodynamics implies

$$s_t + qs_x \geqslant 0. \tag{14}$$

Hence it is plausible to assume as a mathematical entropy function

$$U(\mathbf{u}) = -\rho(\eta - \eta_0), \tag{15a}$$

where $\eta = s/c_\mathrm{v}$ is the non-dimensional thermodynamic entropy, with c_v the specific heat at constant volume, and η_0 is a constant reference value. For an ideal gas, $\eta = \log(p/\rho^\gamma)$, thus yielding

$$U(\mathbf{u}) = -u_1 \log\left(\frac{(\gamma - 1)(u_3 - \frac{1}{2}u_2^2/u_1)}{u_1^\gamma}\right) + u_1\eta_0, \tag{15b}$$

where u_i are the components of the unknown vector $\mathbf{u}$. This $U(\mathbf{u})$ is a convex function of $\mathbf{u}$ and satisfies (5).[2–11] A more general entropy function can be defined by considering an arbitrary smooth function $\sigma = \sigma(s)$. Left-multiplying (12) by $\mathrm{d}\sigma/\mathrm{d}s$ yields

$$\sigma_t + q\sigma_x = 0, \tag{16}$$

whose conservative form is

$$(\rho\sigma)_t + (\rho q\sigma)_x = 0.$$

Therefore for the Euler equations of gas dynamics it is possible to define a family of entropy functions $U(\mathbf{u}) = -\rho\sigma(\eta)$, under some constraints on σ and its derivatives, in order to ensure the convexity of $U(\mathbf{u})$.[12] Thus expressions other than (15a) may be used for U.

Despite the mathematical and physical evidence of the entropy inequalities that must be satisfied by the admissible weak solutions, it is not easy to verify such inequalities at a discrete level. In this study we deal with the problem of evaluating the consistency of numerical solutions of the Euler equations with the entropy inequalities. More precisely, we aim to devise a methodology which, on the basis of an additional induced equation representing a discrete approximation of the entropy balance, would allow for an *a posteriori* local and global spurious entropy production estimate. The analysis is supported by analytical and numerical considerations and contributes to providing a more comprehensive evaluation of the numerical entropy dynamics. The motivations for this study stem from analogous research carried out by Tadmor[11] and, in a computational framework, by Cox and Argrow,[13] with which it shares some similarities.

The outline of the paper is as follows. In the first part an analysis of the relation between the entropy balance and the reconstruction of a finite-dimensional state vector is carried out by means of variational calculus. A discrete level analysis follows and the results for systems of conservation laws are given. Next the problem of defining a suitable numerical entropy flux is addressed and local and global entropy balances consistent with the underlying numerical schemes are introduced. Some test problems are devised in order to assess the effects of the order of accuracy of finite volume time-marching algorithms upon the entropy production.

2. NUMERICAL PRELIMINARIES

The computational approach to the continuum problem requires some preliminary choices for the finite-dimensional representation of the unknown functions and the discretization of the equations.

The aim of the discretization is to replace the functional requirements expressed by equations (1) and (2) (sometimes also accounting for (9c)) by a set of algebraic relations for the fully discretized approach or by a system of ordinary differential equations for the semidiscretized approach. In both cases the numerical solution is represented by a *finite* set of real parameters, the so-called *state vector* $\{\mathbf{v}_j^n\}$ (here the indices j and n refer to the spatial and temporal discretizations respectively). The meaning of the scalar components of $\{\mathbf{v}_j^n\}$ depends on the adopted representation and also constitutes a distinguishing feature of the various computational methods. These components can be defined in general by means of the theory of projections onto suitable basis functions. More commonly, however, the unknowns $\mathbf{v}_j^n$ are defined as either pointwise values or cell averages or even as real coefficients of the chosen shape functions. It is worth noticing that a functional space and/or space–time description of the quantities associated with the state vectors may be useful even in cases where the description itself is not strictly linked to the representation, as in finite difference or finite volume methods. For these methods it may be necessary to associate some interpolation or reconstruction (either local or global) of the state vector with the representation and discretization procedures.

The design of a time-marching finite volume discretization consistent with the conservation law (1a) essentially requires

(i) the selection of a finite-dimensional state, e.g. state vector of nodal values, state vector of values averaged over computational cells, etc.

(ii) the selection of finite-dimensional flux vectors, i.e. approximation of the temporal integration of $\mathbf{f}(\mathbf{u})$ over the time step at the cell boundaries
(iii) the analysis of the influence of the spatial and temporal discretizations on stability, accuracy and physical relevance of the numerical solutions (i.e. consistency with the entropy inequality (9c)).

Items (i) and (ii) have been widely addressed in the literature (nevertheless, some issues related to (ii) will be briefly recalled in a later section). With regard to (iii) we recall that the most successful methods for computing discontinuous solutions are those derived from a conservative and consistent finite volume formulation. This is a consequence of the Lax–Wendroff theorem,[14] which shows that, by using conservative difference schemes and consistent numerical fluxes, the numerical solution converges boundedly almost everywhere. Furthermore, its limit is a weak solution of (1) and satisfies automatically the jump conditions across discontinuities. However, the Lax–Wendroff theorem does not guarantee that the obtained numerical approximation will always converge to the physically relevant weak solution, i.e. it may not satisfy the entropy condition. Equations such as (9c) are not usually included in the discretization, although the numerical solution should eventually satisfy the entropy condition.

Over the last two decades many efforts have been devoted to the construction of numerical schemes which possess some special stability properties such as *monotonicity*, *l_1-contractivity*, *total variation diminishing* (TVD)[15] and *total variation bounded* (TVB)[8] requirements. Although of extreme importance with respect to the convergence towards a weak solution of a conservative scheme, these properties do not always ensure consistency with the entropy inequality (9c). We remark that many popular schemes do not satisfy the entropy condition unless an *entropy correction* is introduced. Among these is the popular approximate Riemann solver (ARS) of Roe,[6] which admits expansion shocks as steady solutions.

Furthermore, it is difficult to simultaneously provide accurate and physically relevant solutions owing to the limitations dictated by the Godunov theorem (and its Harten–Keyfitz extension for non-linear equations).[16] The only attempts (of which the authors are aware) at designing entropy-satisfying schemes are due to Osher[5] and Tadmor[10] and these (implicitly) satisfy the entropy inequality only in an approximate sense. In particular, for the scalar case, Osher has introduced the *E-schemes* which are shown to be TVD and entropy-satisfying, thus converging to a physically relevant solution. Unfortunately, such an *E*-consistency requirement restricts the approximation to first-order accuracy. Tadmor has generalized this idea by introducing the *entropy conservative scheme* which satisfy a discrete cell entropy equality. He has shown that those conservative schemes that have more viscosity than an entropy conservative scheme are also entropy stable.[10] It is obvious that the *E*-schemes introduced by Osher fall within this category by comparison with the Godunov scheme which is known to satisfy the entropy inequality.

We therefore conclude that the design of a numerical scheme capable of simultaneously providing accurate and physically relevant solutions is not an easy task. Hence the increasing demand for a wider and more detailed analysis of the relation between the state vector $\{\mathbf{v}_j^n\}$ obtained from any finite volume discretization procedure and the entropy inequality (9c) appears justified.

3. THE ENTROPY RECONSTRUCTION RELATION

In the design of modern time-marching finite volume algorithms it is possible to distinguish a *projection* or *reconstruction stage* and an *evolution stage*. The first step consists of a local or global continuous space description of the approximate solution at a given time by means of a finite-dimensional state vector. In the evolution stage the reconstruction allows the evaluation of the

numerical fluxes required to upgrade the state vector at the next time level. Obviously the global reconstruction of the unknown quantities also allows for the reconstruction of other dependent variables such as the entropy and therefore provides the possibility of estimating the entropy evolution at various time levels. We remark that in principle the reconstruction process contains an unlimited number of degrees of freedom. In the following we intend to study the relation between the choice of reconstruction stage and the related entropy production.

For the sake of simplicity we refer to the scalar case and let $D = (a, b)$ be a finite interval of $\mathbb{R}$ and U: $\mathbb{R} \to \mathbb{R}$ a twice continuously differentiable function, i.e. $U \in C^2(\mathbb{R})$, and convex with respect to w, i.e. $U_{ww} > 0$. Consider the differentiable functional

$$\Phi: w \in C^0(\bar{D}) \to \int_D U(w(x))\mathrm{d}x.$$

The following proposition holds.

Proposition 1

The functional $\Phi(w)$ admits an extremum if and only if $U(w)$ has an extremum.

Proof. Suppose that*

$$\exists w_0 \in \mathbb{R} : \mathrm{d}U(w_0)/\mathrm{d}w = 0.$$

Then $\Phi(w)$ has a relative extremum at the extremal function** $w(x) = w_e(x) = w_0 = \text{constant}$. In addition, owing to its convexity and to the consequent strictly monotone behaviour of its first derivative, U has a unique minimum value at the point w_0. As a consequence of this, Φ has the unique minimum value at the point $w_e(x)$. On the other hand, if U does not have an extremal value w_0, then Φ does not have an extremum either. In conclusion, if U has (resp. does not have) lower or upper bounds, then Φ has (resp. does not have) lower or upper bounds as well. □

The above reasoning also applies when U is a convex function of a vector-valued function $\mathbf{w}(x)$. In this case it is again easy to prove that $\Phi(\mathbf{w})$ has a minimum at $\mathbf{w}_0$ if and only if $\mathbf{w}_0$ is an extremal point for $U(\mathbf{w})$. Likewise, the boundedness (or unboundedness) of $U(\mathbf{w})$ corresponds to the boundedness (or unboundedness) of Φ.

Let the set $\{I_i\}_{i\in\{1,\ldots,N\}}$ constitute a partition of D and let S_D be the space of *piecewise continuous* functions*** over D. Here the subintervals $I_i = (x_{i-1/2}, x_{i+1/2})$ have size h_i, with $x_{1/2} = a$ and $x_{N+1/2} = b$. In order to understand the features of some numerical schemes for conservation laws, we

* If Φ has a relative extremum at the point w, then

$$\int_D \frac{\mathrm{d}U(w)}{\mathrm{d}w} v \,\mathrm{d}x = 0, \quad \forall v \in C^0(\bar{D}).$$

Hence

$$\mathrm{d}U(w)/\mathrm{d}w = 0,$$

which is the *Euler equation of the functional* Φ.

** An extremal function (or extremal point) is a function which makes a given functional stationary.

*** We say that a function $w(x)$ is *piecewise continuous* over a given interval if the interval can be divided into a finite number of subintervals, within which w is continuous, with finite limits at the left and right endpoints, and

$$S_D = \{w: w(x) = w_i(x), \forall x \in I_i, \text{and } w_i \in C^0([x_{i-1/2}, x_{i+1/2}]), \forall I_i\}.$$

If, in addition, $w(x)$ is continuously differentiable over each closed subinterval, then we say that it is *piecewise smooth*.

attempt to analyse the consequence of adopting either an interpolation or a reconstruction procedure of the unknown by extending the continuum level analysis. We first assume that w is a scalar function which obeys some conditions, i.e. w belongs either to the set of piecewise continuous interpolations

$$T_D = \{w \in S_D : w(x_i) = w_i, \forall I_i\}$$

or to the set of piecewise continuous reconstructions

$$V_D = \left\{ w \in S_D : \frac{1}{h_i} \int_{-h_i/2}^{h_i/2} w(x_i + \eta) \mathrm{d}\eta = \bar{w}_i, \forall I_i \right\},$$

where x_i is the centre of I_i. Note that the sets T_D and V_D are typical of finite difference and finite volume schemes respectively.

Let $U = U(w) \in C^2$ be a convex function with an extremum. From Proposition 1 the following proposition immediately ensues.

Proposition 2

The functional

$$\Phi : w \in T_D \to \sum_{i=1}^{N} \int_{I_i} U(w(x)) \mathrm{d}x$$

does not admit any extrema, with the trivial exception of the case $w_i = w_0$.

The situation is different in the case of reconstruction, as illustrated by the following proposition.

Proposition 3

Let $w = w(x) \in V_D$ and $p = p(x)$ be the primitive function of $w(x)$:

$$p(x) = \int_a^x w(\xi) \mathrm{d}\xi, \quad x \in (a, b). \tag{17}$$

Then the functional

$$\Phi : p \to \sum_{i=1}^{N} \int_{I_i} U(p'(x)) \mathrm{d}x$$

has a minimum.

Proof. To prove the proposition, we again use the variational calculus. We start by noticing that from (17) it immediately follows that

$$p(x_{i+1/2}) - p(x_{i-1/2}) = h_i \bar{w}_i, \tag{18}$$

$$p(x_{i+1/2}) = \sum_{k=1}^{i} h_k \bar{w}_k. \tag{19}$$

As a consequence, the values of the primitive function $p(x)$ are constrained at all cell interfaces $x_{i\pm1/2}$ by (19). The necessary condition for a constrained relative extremum of Φ then reduces to (*the Euler equation*)

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}U}{\mathrm{d}p'}\right) = 0 \quad \Leftrightarrow \quad \frac{\mathrm{d}^2 U}{\mathrm{d}p'^2} p'' = 0, \quad \forall I_i.$$

Let p_0 be an extremal function for Φ. It then follows that

$$p_0'' = 0 \quad \Rightarrow \quad p_0 = \alpha_i x + \beta_i, \tag{20}$$

$U_{p'p'}$ being a positive quantity. The coefficients α_i and β_i are easily obtained through (18), thus giving the piecewise linear function

$$p_0(x) = \bar{w}_i(x - x_{i-1/2}) + p_0(x_{i-1/2}), \quad \forall x \in [x_{i-1/2}, x_{i+1/2}], \tag{21}$$

which implies

$$w_e(x) = \bar{w}_i, \quad \forall x \in [x_{i-1/2}, x_{i+1/2}].$$

Moreover, owing to the convexity of U, Φ has a minimum at p_0. □

The previous analysis can be extended to the case in which U: $\mathbb{R}^k \to \mathbb{R}$. Again a vector function $\mathbf{p}(x)$, whose components are the primitives of the corresponding components of the vector $\mathbf{w}(x)$, can be introduced. Similarly to the scalar case, the following Euler equation holds:

$$(\mathbf{p}'')^{\mathrm{T}} \cdot U_{\mathbf{p}'\mathbf{p}'} = \mathbf{0}, \quad \forall I_i. \tag{22}$$

Let $\mathbf{p}_0$ be an extremal point for $\Phi(\mathbf{p})$. Then, being the Hessian $U_{\mathbf{p}'\mathbf{p}'}$ a symmetric positive definite matrix, equation (22) implies $\mathbf{p}_0'' = \mathbf{0}$ on each subinterval I_i. Hence each component of $\mathbf{p}_0$ is a piecewise linear function as given by (20) and (21) and also in this case Φ has a minimum at $\mathbf{p}_0$.

It is useful to recognize here an immediate consequence for the gas dynamics equations; namely, in the case of a single interval $I_1 = D$ the distribution of the variables $\mathbf{u}$ which ensures the minimum of $\Phi(\mathbf{u})$ (maximum of the global thermodynamic entropy), under the condition of conservation of the integral of mass, momentum and energy, is the constant one.

The significance of the above considerations towards the understanding of the finite volume reconstruction procedure as a source of spurious entropy production will be further clarified in Section 5.

4. THE FULLY DISCRETIZED FINITE VOLUME FORMULATION

Hereafter we concentrate on the numerical solution of (1) obtained by two-level fully explicit schemes in conservative form. For the sake of simplicity we consider a uniform partition of D and denote by h and τ the mesh size and the time step respectively. Let $x_{i+1/2} = (i + \frac{1}{2})h$ and $t^n = n\tau$, $\forall n \geqslant 0$, and $i \in \{1, \ldots, N\}$. The *fully discretized finite volume formulation of* (1) is

$$\mathbf{v}_i^{n+1} - \mathbf{v}_i^n = -\frac{\tau}{h}(\hat{\mathbf{f}}_{i+1/2} - \hat{\mathbf{f}}_{i-1/2}). \tag{23}$$

Here $\hat{\mathbf{f}}_{i+1/2}$ is the approximation of the true flux at cell face $x_{i+1/2}$, based on a support of $2m$ values, averaged between t^n and t^{n+1},

$$\hat{\mathbf{f}}_{i+1/2} = \hat{\mathbf{f}}(\mathbf{v}_{i-m+1}^n, \ldots, \mathbf{v}_{i+m}^n) \approx \hat{\mathbf{f}}(x_{i+1/2}, t, \mathbf{u}) = \frac{1}{\tau}\int_0^\tau \mathbf{f}(\mathbf{u}(x_{i+1/2}, t + \xi))\mathrm{d}\xi, \tag{24}$$

which satisfies the consistency requirement

$$\hat{\mathbf{f}}(\mathbf{u}, \ldots, \mathbf{u}) = \mathbf{f}(\mathbf{u}), \tag{25}$$

and $\mathbf{v}_i^n$ is an approximation to the average of the exact solution $\mathbf{u}(x, t)$ in cell I_i,

$$\mathbf{v}_i^n \approx \bar{\mathbf{u}}_i^n = \frac{1}{h}\int_{-h/2}^{h/2} \mathbf{u}(x_i + \eta, t^n)\mathrm{d}\eta. \tag{26}$$

One of the earliest approaches to evaluating the numerical flux function is due to Godunov.[17] We recall that in the Godunov method the fundamental properties of the local exact solution of the conservation laws are accounted for in the discretization procedure. More precisely, the cell averages are advanced in time by solving at each time step a Riemann problem at each cell interface, assuming a piecewise constant reconstruction. For sufficiently small CFL number there are no interactions between the different Riemann problems and a solution $\mathbf{w}(x, t)$ is obtained, $\forall x \in I_i, 0 \leqslant t \leqslant \tau$. The approximation at time level $t^{n+1} = t_n + \tau$ can then be written in the conservation form (23) with the numerical flux given by

$$\hat{\mathbf{f}}^{\mathrm{God}}_{i+1/2} = \hat{\mathbf{f}}(\mathbf{v}^n_i, \mathbf{v}^n_{i+1}) = \mathbf{f}(\mathbf{w}(0^+; \mathbf{v}^n_i, \mathbf{v}^n_{i+1})), \tag{27}$$

where $\mathbf{w}(x, t) = \mathbf{w}((x - x_{i+1/2})/(t - t^n); \mathbf{v}^n_i, \mathbf{v}^n_{i+1})$ is the exact self-similar solution of the Riemann problem at cell interface $x_{i+1/2}$ at time t^n. We observe that the only source of error in the Godunov scheme arises from the averaging of the unknowns over a computational cell when assuming a piecewise constant state representation.

In recent years a class of schemes which share some similarities with the method of Godunov has been developed. These schemes, usually referred to as 'upwind' difference schemes, take into account the essential physical properties of the hyperbolic conservation laws. Upwind schemes have been cast in a general form after Harten *et al.*,[1] according to the conservative discretization (23). We briefly recall the class of upwind schemes based on (conservative) flux vector splitting, which consists of decomposing the flux $\mathbf{f}$ in the form

$$\mathbf{f}(\mathbf{u}) = \mathbf{f}^+(\mathbf{u}) + \mathbf{f}^-(\mathbf{u}),$$

where $\mathbf{f}^{\pm} = \mathbf{A}^{\pm}(\mathbf{u})\mathbf{u}$ and $\mathbf{A}^{\pm}(\mathbf{u})$ are such that $\mathbf{A} = \mathbf{A}^+ + \mathbf{A}^-$ and $\pm\mathbf{A}^{\pm}(\mathbf{u})$ have real and non-negative eigenvalues. The resulting numerical flux function is

$$\hat{\mathbf{f}}_{i+1/2} = \hat{\mathbf{f}}(\mathbf{v}_i, \mathbf{v}_{i+1}) = \mathbf{f}^+(\mathbf{v}_i) + \mathbf{f}^-(\mathbf{v}_{i+1}).$$

In a general finite volume scheme the numerical flux function $\hat{\mathbf{f}}_{i+1/2}$, evaluated by introducing a local (not necessarily constant) reconstruction of the averaged cell values, is followed by a time integration. Consequently, the numerical flux can be recast as

$$\hat{\mathbf{f}}_{i+1/2} = \hat{\mathbf{f}}(\mathbf{R}^n_i(x_{i+1/2}), \mathbf{R}^n_{i+1}(x_{i+1/2})), \tag{28}$$

where $\mathbf{R}^n_i(x)$ and $\mathbf{R}^n_{i+1}(x)$ are the vector-valued reconstructions of the state vector within cells I_i and I_{i+1} respectively. The components of $\mathbf{R}_i(x)$ are obtained by using a polynomial function of degree $r - 1$ (yielding a conservative rth-order-accurate approximation in the sense of Harten[15]) and the global approximation of $\mathbf{u}(x)$ is obviously a piecewise polynomial function. In most instances the flux function is approximated by

$$\mathbf{f}^{\mathrm{R}}(\mathbf{R}^n_i(x_{i+1/2}), \mathbf{R}^n_{i+1}(x_{i+1/2})), \tag{29}$$

where $\mathbf{f}^{\mathrm{R}}(\mathbf{a}, \mathbf{b})$ denotes the flux at $x_{i+1/2}$ associated with an approximate solution of the Riemann problem whose initial states are $(\mathbf{a}, \mathbf{b})$. Here it must be emphasized that irrespective of the type of numerical flux function employed, the accuracy and main properties of (23) are dictated by the reconstruction procedure.

In the present work we analyse schemes whose design relies on either a second- or a third-order reconstruction procedure, of which we give a brief description hereafter. The conservative reconstruction is obtained by means of Legendre polynomials and we apply the reconstruction in a component-wise fashion, i.e. we assign to each component of the vector of unknowns an independent stencil according to an essentially non-oscillatory (ENO) approach.[18,19] From the given local cell

average vector of the conserved variables $\mathbf{v}_i = (\bar{\rho}, \overline{\rho q}, \overline{\rho E})_i^{\mathrm{T}}$ we approximate the cell averages of the primitive variables $(\bar{\rho}, \bar{q}, \bar{p})_i^{\mathrm{T}}$ through

$$\bar{q}_i = \frac{(\overline{\rho q})_i}{\bar{\rho}_i}, \qquad \bar{p}_i = (\gamma - 1)\left((\overline{\rho E})_i - \frac{\bar{\rho}_i \bar{q}_i^2}{2}\right).$$

The generic primitive variable f is then reconstructed either by

$$f_i(x) = \bar{f}_i + f_i^a(x - x_i), \quad \forall x \in I_i,$$

or by

$$f_i(x) = \bar{f}_i + f_i^a(x - x_i) + \tfrac{1}{2} f_i^b[(x - x_i)^2 - h^2/12], \quad \forall x \in I_i,$$

in the case of piecewise linear ($r = 2$) and quadratic ($r = 3$) reconstructions respectively. The quantities f_i^a in the linear reconstruction are obtained by imposing

$$\frac{1}{h}\int_{I_j} f_i(x)\mathrm{d}x = \bar{f}_j,$$

while in the quadratic case f_i^a and f_i^b follow by requiring

$$\frac{1}{h}\int_{I_j} f_i(x)\mathrm{d}x = \bar{f}_j \quad \text{and} \quad \frac{1}{h}\int_{I_k} f_i(x)\mathrm{d}x = \bar{f}_k,$$

where the indices j and k are selected via an ENO stencil selection procedure.[18,19]

5. GLOBAL AND LOCAL ENTROPY VERIFICATION

Before turning to the issue of defining the discrete counterpart of an entropy inequality, we wish to recall some definitions and results related to the locally induced entropy balance. Following Reference 16, the admissibility criterion for the exact weak solution of (1) is expressed in terms of an inequality such as (9c). The scheme (23) is said to be *entropy stable* if the following *cell entropy inequality* holds:

$$U_i^{n+1} - U_i^n + \frac{\tau}{h}(\hat{G}_{i+1/2} - \hat{G}_{i-1/2}) \leqslant 0, \tag{30}$$

where U_i^k and $\hat{G}_{i+1/2}$ are the approximations

$$U_i^k \approx \frac{1}{h}\int_{-h/2}^{h/2} U(\mathbf{u}(x_i + \xi, t^k))\mathrm{d}\xi, \tag{31}$$

$$\hat{G}_{i+1/2} \approx \hat{G}(x_{i+1/2}, t, \mathbf{u}) = \frac{1}{\tau}\int_0^{\tau} G(\mathbf{u}(x_{i+1/2}, t + \xi))\mathrm{d}\xi. \tag{32}$$

Note that although in some instances (e.g. the Euler equations of gas dynamics) an entropy flux function $G(\mathbf{u})$ can be identified at least for the analytical problem (1), at the discrete level (24) and (25) do not provide any guideline to define $\hat{G}_{i+1/2}$. In practice, for each entropy function U the existence of a *numerical entropy flux* (function of $2m$ arguments)

$$\hat{G}_{i+1/2} = \hat{G}(\mathbf{v}_{i-m+1}^n, \ldots, \mathbf{v}_{i+m}^n) \tag{33}$$

which is consistent with the differential one

$$\hat{G}(\mathbf{u}, \ldots, \mathbf{u}) = G(\mathbf{u}) \tag{34}$$

does not follow in a natural way from the chosen $\hat{\mathbf{f}}$ but must in some sense be postulated. This is a crucial point; indeed, when the schemes are not designed such that they naturally obey a discrete cell entropy inequality, the only alternative route is to carry out an *a posteriori* analysis of the numerical solutions.

Therefore, in order to verify the consistency of a numerical solution with (30), it is necessary to evaluate both U_i^k and $\hat{G}_{i\pm1/2}$. While the former can be readily obtained via reconstruction by means of (31), the latter opens the way to ambiguities, i.e. it is unclear how to define a numerical entropy flux which is consistent with the numerical flux functions $\hat{\mathbf{f}}_{i+1/2}$. To show how one can overcome such a difficulty, we consider an initial value problem for (1a) in a finite interval $D=[0,1]$ subject to periodic boundary conditions

$$\mathbf{u}(0,t)=\mathbf{u}(1,t).$$

This is representative of a non-linear convection problem in an annular ring of unit length where the curvature effects are neglected. Then a global analysis may be carried out by looking at the time evolution of the quantity

$$U_D(t)=\int_D U(x,t)\mathrm{d}x.$$

The periodicity of the boundary conditions ensures that the instantaneous net flux of the conserved variables in D is exactly zero. Integrating (10) over the region $A=\{x\in D, 0\leqslant t\leqslant T\}$, we find

$$\int_D U(\mathbf{u}^\varepsilon(x,t))\mathrm{d}x-\int_D U(\mathbf{u}^\varepsilon(x,0))\mathrm{d}x=-\varepsilon\int\!\!\int_A (\mathbf{u}_x^\varepsilon)^{\mathrm{T}}U_{\mathbf{uu}}\mathbf{u}_x^\varepsilon\ \mathrm{d}x\ \mathrm{d}t$$

and from the convexity of U we obtain the estimate

$$\int_D U(\mathbf{u}(x,t))\mathrm{d}x\leqslant\int_D U(\mathbf{u}(x,0))\mathrm{d}x.$$

For the discrete level analysis let us introduce the cell (non-positive) space–time average (of the rate) of entropy production π_i and recast (30) in the form

$$U_i^{n+1}-U_i^n+\frac{\tau}{h}(\hat{G}_{i+1/2}-\hat{G}_{i-1/2})=\tau\pi_i. \tag{35}$$

If we then sum (35) over the N control volumes, we get a global discrete entropy inequality that does not require any evaluation of the numerical entropy flux, i.e.

$$\sum_{i=1}^N (U_i^{n+1}-U_i^n)\equiv\tau\sum_{i=1}^N \pi_i\leqslant 0. \tag{36}$$

The left-hand side of (36) can be used to evaluate the global entropy production associated with any finite volume scheme. One might conjecture that the verification of (36) for any initial condition would also allow for the determination of the (assumed non-positive) local cell entropy productions π_i. The non-linear character of the equations makes it difficult to prove such a conjecture and, moreover, makes it impossible to consider a realistic exhaustive set of different initial conditions. As a consequence, one has also to face the problem of verifying the local entropy condition (35).

For a local analysis we first observe that if the time rate of change of the entropy function is known in each cell (as in the periodic test problem discussed above), then (35) reduces to a system of N equations in $2N$ unknowns: the entropy productions π_i and the numerical entropy fluxes $\hat{G}_{i+1/2}$.

Therefore, if one could estimate the local entropy productions, one might attempt to determine the numerical entropy fluxes from

$$\hat{G}_{i+1/2} - \hat{G}_{i-1/2} = h\left(\pi_i - \frac{U_i^{n+1} - U_i^n}{\tau}\right), \quad i = 1, \ldots, N, \tag{37}$$

which is a linear system of N equations in the N unknowns $\hat{G}_{i+1/2}$ that does not admit a unique solution, since each row is a linear combination of all the others. This is not a surprising result owing to the periodic character of the problem ($\mathbf{v}_1 = \mathbf{v}_{N+1}$ and $\hat{G}_{1/2} = \hat{G}_{N+1/2}$). Nevertheless, setting for instance $\hat{G}_{N+1/2} = 0$ and eliminating the Nth equation, it is easy to see that (37) has the solution

$$\hat{G}_{i+1/2} = h \sum_{k=1}^{i} \left(\pi_k - \frac{U_k^{n+1} - U_k^n}{\tau}\right), \quad i = 1, \ldots, N-1, \tag{38}$$

which shows the cumulative effect of the errors on the local estimates of $\hat{G}_{i+1/2}$. Hence a reliable evaluation of the numerical entropy fluxes would be possible if an accurate approximation of the numerical entropy productions were available. The complexity of the numerical entropy production mechanism (due to the presence of discontinuities, the reconstruction procedure, etc.) makes it difficult to find a suitable distribution of the entropy production.

On the contrary, it seems more reasonable to find a consistent approximation to $\hat{G}_{i+1/2}$ and then determine the local entropy productions π_i. We recall that for the Euler equations of gas dynamics a unique solution of the Riemann problem exists as long as a vacuum does not form. Such an exact solution $\mathbf{w}(x, t)$ satisfies the entropy condition (9b):

$$\int_{-h/2}^{h/2} U(\mathbf{w}(x_i + \eta, t^{n+1}))\mathrm{d}\eta \leqslant hU(\mathbf{v}_i^n) - \tau[G(\mathbf{w}(0^+; \mathbf{v}_i^n, \mathbf{v}_{i+1}^n)) - G(\mathbf{w}(0^+; \mathbf{v}_{i-1}^n; \mathbf{v}t^n))].$$

We then note that when the Godunov scheme is used, the entropy flux

$$\hat{G}_{i+1/2}^{\mathrm{God}} \equiv G(\mathbf{w}(0^+; \mathbf{v}_i^n, \mathbf{v}_{i+1}^n)) \tag{39}$$

is evaluated exactly. Furthermore, as previously shown, the piecewise constant representation of the data implies that the following inequality is satisfied:

$$hU(\bar{\mathbf{w}}_i) \leqslant \int_{-h/2}^{h/2} U(\mathbf{w}(x_i + \eta))\mathrm{d}\eta,$$

which yields the obvious consequence that the Godunov scheme verifies the entropy inequality (30).

From the foregoing we can conclude that the Godunov scheme allows us to safely estimate the local entropy productions π_i; however, a more general procedure that does not rely on the use of an exact Riemann solver is desirable. In the following we therefore address the issue of defining a suitable approximation of the entropy flux function. A possibility might consist of assuming

$$\hat{G}_{i+1/2} = G(\mathbf{u}^*), \tag{40}$$

with $\mathbf{u}^*$ the state vector such that

$$\mathbf{f}(\mathbf{u}^*) = \hat{\mathbf{f}}_{i+1/2}. \tag{41}$$

However, (41) does not admit in general a unique solution: the Rankine–Hugoniot relations for a steady shock discontinuity provide an example of two different states (and two values of G)

corresponding to a unique flux vector **f**. To find an expression for the numerical entropy flux function, we reason that one could exploit the homogeneity property of Euler fluxes. Indeed, we have

$$G(\mathbf{u}) = -\eta(\mathbf{u})f^1(\mathbf{u}) = f^1(-\eta(\mathbf{u})\mathbf{u}), \tag{42}$$

where f^1 is the mass flux (i.e. the first component of the flux function **f**). A numerical entropy flux function, hereafter referred to as the mass consistent entropy flux (MCEF), can then be devised according to

$$\hat{G}_{i+1/2} = \hat{G}(\mathbf{v}_i, \mathbf{v}_{i+1}) = \hat{f}^1(-\eta(\mathbf{v}_i)\mathbf{v}_i, -\eta(\mathbf{v}_{i+1})\mathbf{v}_{i+1}) \tag{43a}$$

for a low-order scheme and

$$\hat{G}_{i+1/2} = \hat{f}^1(-\eta(\mathbf{R}_i(x_{i+1/2}))\mathbf{R}_i(x_{i+1/2}), -\eta(\mathbf{R}_{i+1}(x_{i+1/2}))\mathbf{R}_{i+1}(x_{i+1/2})) \tag{43b}$$

for a high-order method. Equations (43a,b) provide consistent numerical approximations to the entropy flux; besides, they do not require any linearization procedure while employing the same numerical mass flux as the baseline scheme.

Finally, we can also introduce the quantity

$$\delta_i^{\text{const}} = U(\bar{\mathbf{w}}_i^{n+1}) - \frac{1}{h}\int_{-h/2}^{h/2} U(\mathbf{w}(x_i + \eta, t^{n+1}))\mathrm{d}\eta \leqslant 0,$$

which represents a spurious cell entropy production due to a piecewise constant representation of the data (note that all first-order finite volume schemes suffer from such a contribution). According to Proposition 3 discussed in Section 3, we claim that any other functional representation of the unknowns within the cell will lead to a $\delta_i \geqslant \delta_i^{\text{const}}$, which of course implies that the reconstruction contribution to the term $\tau\pi_i$ can be either negative (source of entropy) or positive (sink of entropy).

6. NUMERICAL EXPERIMENTS

In the previous sections we have shown that a thorough analysis of a finite volume scheme would also require the study of the numerical entropy dynamics (and the relation between the entropy balance and the reconstruction process) as well as the quantification of the local entropy production. For that purpose three initial value problems (IVPs) for the Euler equations of gas dynamics have been selected.

Problem (a)

This corresponds to an IVP with discontinuous initial conditions in the density field,

$$\text{Problem (a)} \quad \begin{cases} (\rho_i = 0{\cdot}125, q_i = 1, p_i = 1), & 1 \leqslant i \leqslant N/2, \\ (\rho_i = 1, q_i = 1, p_i = 1), & N/2 < i \leqslant N, \end{cases}$$

and periodic boundary conditions. The unsteady exact solution of this problem consists of two rotating constant states separated by two contact discontinuities. Problem (a) has first been solved using some well-known numerical flux functions and adopting a piecewise constant reconstruction of the data. In particular, we have analysed the exact Riemann solver of Godunov,[17] the approximate Riemann solvers of Osher[4] and Roe,[6] the flux-vector-splitting methods of Steger and Warming[9] and van Leer,[20] the Lax–Friedrichs method[2] and the advection upstream splitting method[3] (AUSM). All the computations have been carried out using 100 cells ($h = 0{\cdot}01$), $\tau = 10^{-3}$ ($CFL \approx 0{\cdot}4$) and the forward Euler explicit time-differencing procedure.

In Figure 1 we evaluate the performance of the different numerical schemes by reporting the ratio ϕ of the global thermodynamic entropy over its initial value versus time in the interval $[0, T]$ (where $T=1$ is the time necessary for a fluid particle to travel over the entire ring). The discrete global entropy at time t^n is defined as

$$U_{\mathrm{D}}^n = \sum_{j=1}^{N} \int_{I_j} \rho_j^n(x)\eta_j^n(x)\mathrm{d}x. \tag{44}$$

In the figure we also report the exact value represented by the $\phi = 1$ line. As expected, the Godunov method provides the best performance, i.e. is the time rate of change of ϕ is the slowest. We also note that both the Roe and Osher methods yield solutions that are virtually indistinguishable from the Godunov one, while all other schemes prove to be more dissipative. The unavoidable smearing of the two contact discontinuities leads to an increasing time-dependent spurious entropy production. Indeed, while in a Lagrangian formulation a contact discontinuity can be crisply captured, this is difficult to achieve with a Eulerian formulation without resorting to some special treatments such as the subcell resolution proposed in Reference 21.

Next we discuss the effects of the *order* of the reconstruction for a given flux function. In particular, we consider the constant, linear and quadratic ENO reconstruction procedures described in Section 4, adopting the exact Riemann solver of Godunov. All computations have been carried out using the same time step, number of cells and time integration procedure as before. In Figure 2 the numerical and exact solutions are compared after 100 time steps. As expected, a substantial improvement is observed with the quadratic reconstruction. In Figure 3 we again report the variation in ϕ with respect to time. For the higher-order reconstruction the integral in (44) is carried out relating each interval I_j with the reference interval $C = [-1, 1]$ through the transformation

$$x = \tfrac{1}{2}[(x_{j+1/2} + x_{j-1/2}) + \alpha(x_{j+1/2} - x_{j-1/2})], \quad \alpha \in [-1, 1].$$

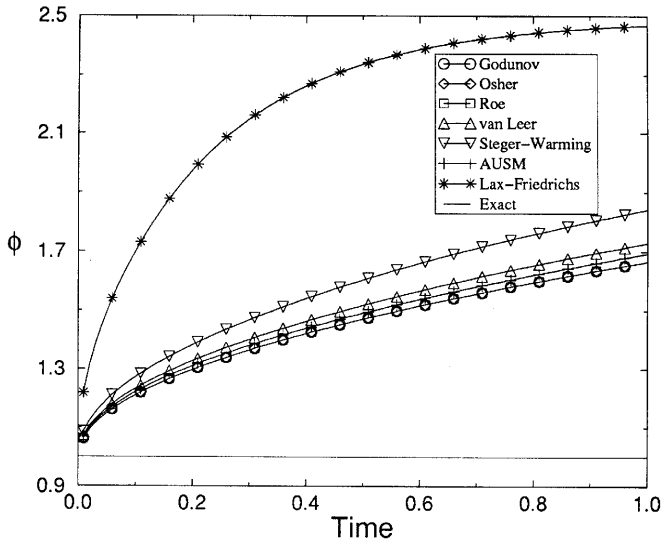

Figure 1. Time history of global entropy variable for various flux functions; test problem (a)

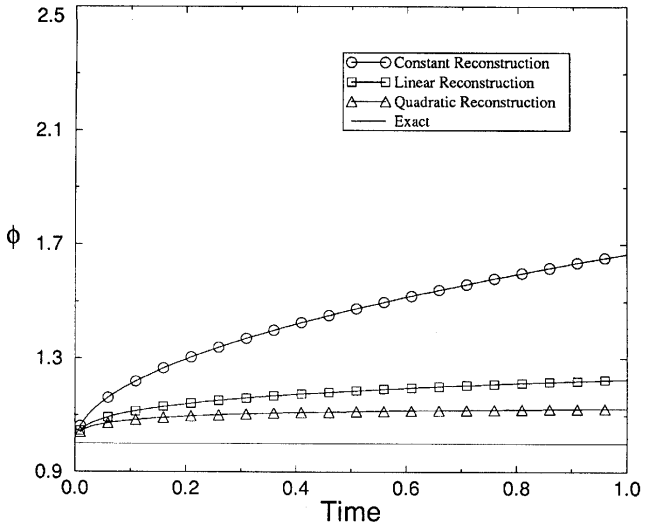

Figure 2. Time history of global entropy variable for various reconstruction procedures; test problem (a)

Thus the integral of a generic function $f(x)$ on I_j is computed as

$$\int_{I_j} f(x)\mathrm{d}x = \frac{h_j}{2}\int_C f(x(\alpha))\mathrm{d}\alpha.$$

In the above expression the integration over C is done numerically using a two-point Gauss–Legendre quadrature formula*

$$\int_{I_j} f(x)\mathrm{d}x \approx \frac{h_j}{2}[f(x(\alpha_1)) + f(x(\alpha_2))], \quad \alpha_1 = -\alpha_2 = -\frac{1}{\sqrt{3}}. \tag{45}$$

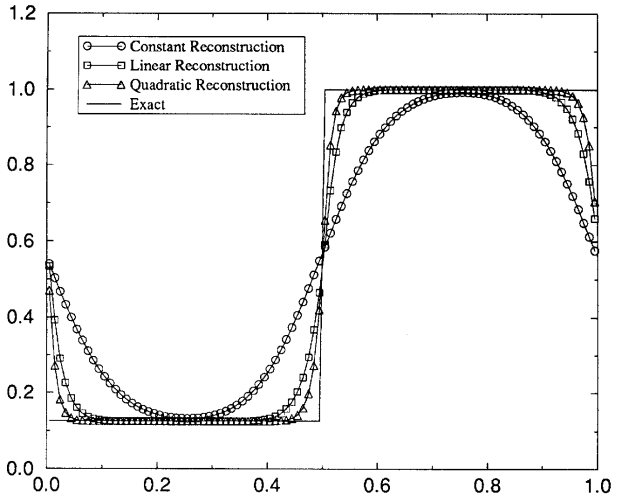

Figure 3. Exact and computed density distributions at time $T = 1$: test problem (a)

* The quadrature (45) is exact when the function f is a polynomial of degree less than or equal to three.

Recalling that the exact entropy production is zero for this test problem, it appears that the time rate of change of the spurious entropy production is drastically reduced with the order of reconstruction.

Problem (b)

To get some indications on the local entropy balance and its relation with the numerical solutions, we have considered two well-known test cases for the Euler equations of gas dynamics, i.e. the Sod[7] and Lax[15] problems, hereafter referred to as problems (b1) and (b2), which are defined by the following initial conditions:

$$\text{problem (b1)} \quad \begin{cases} (\rho_i = 1, q_i = 0, p_i = 1), & 1 \leqslant i \leqslant N/2, \\ (\rho_i = 0{\cdot}125, q_i = 0, p_i = 0{\cdot}1), & N/2 < i \leqslant N, \end{cases}$$

$$\text{problem (b2)} \quad \begin{cases} (\rho_i = 0{\cdot}445, q_i = 0{\cdot}698, p_i = 3{\cdot}528), & 1 \leqslant i \leqslant N/2, \\ (\rho_i = 0{\cdot}5, q_i = 0, p_i = 0{\cdot}571), & N/2 < i \leqslant N. \end{cases}$$

The exact solution of problem (b1) consists of a discontinuity which breaks into a weak shock wave followed by a contact discontinuity and a rarefaction wave. In problem (b2) the initial discontinuity breaks into a moderately strong shock followed by a density level far above its initial state. The contact discontinuity then lowers the density, which undergoes a transition to the left initial state through a rarefaction wave. Because of the monotone time-decreasing density profile, the numerical solution of problem (b1) does not exhibit the difficulties associated with the creation of an intermediate state (i.e. outside the bounds of the initial values), typical of problem (b2).

Following the methodology described in Section 4, we analyse the local entropy production and the effects of the reconstruction. The computations have been carried out using 100 equally spaced cells ($h = 0{\cdot}01$), $CFL = 0{\cdot}8$ and the forward Euler explicit time-differencing procedure. In Figures 4 and 5 the numerical results obtained with the solver of Godunov are compared with the exact

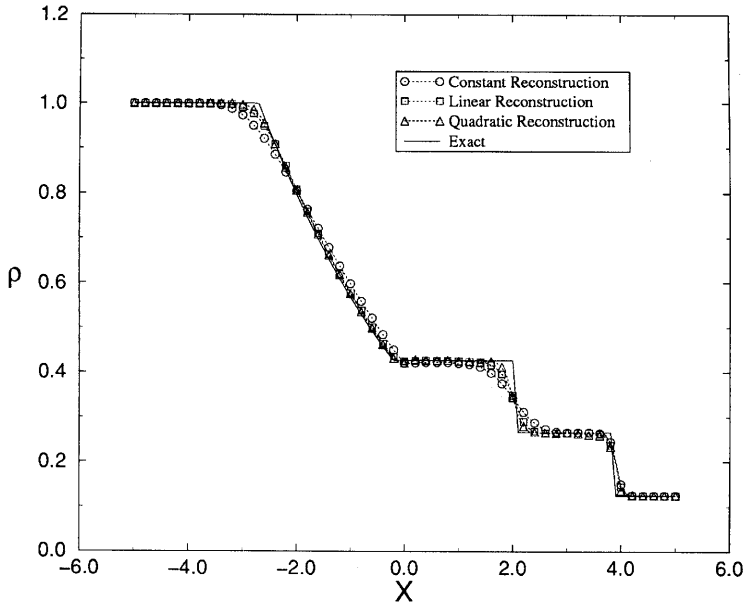

Figure 4. Exact and computed density distributions: test problem (b1)

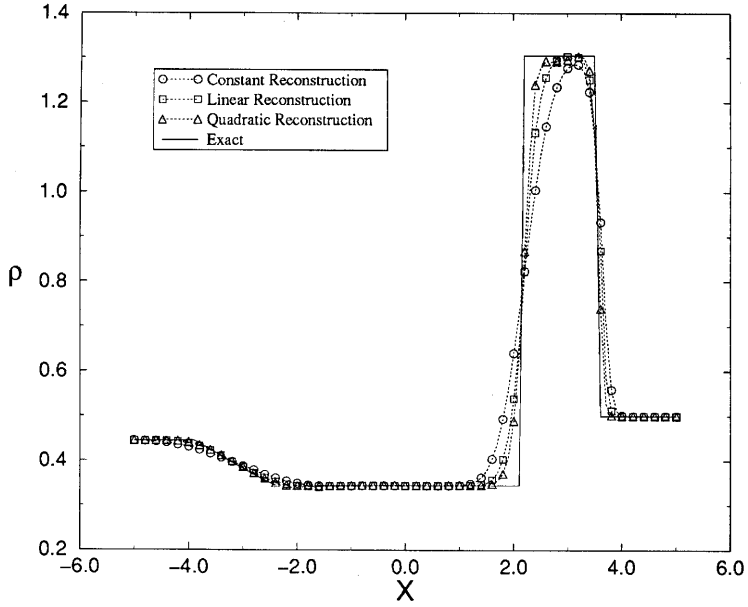

Figure 5. Exact and computed density distributions: test problem (b2)

solutions after 60 and 85 time steps for problems (b1) and (b2) respectively. The figures show the improvement in the resolution as the order of reconstruction increases. The local numerical entropy productions $\{\pi_j\}$ are reported in Figures 6 and 7 respectively for problems (b1) and (b2) after 580 (resp. 380) time steps with a fixed τ equal to 0·0025 (resp. 0·006) (in both cases $CFL \approx 0{\cdot}1$); the exact shock location is also illustrated in the figures. Note that the numerical entropy flux in (35) is exactly computed by means of the solver of Godunov (39). The figures show that spurious entropy

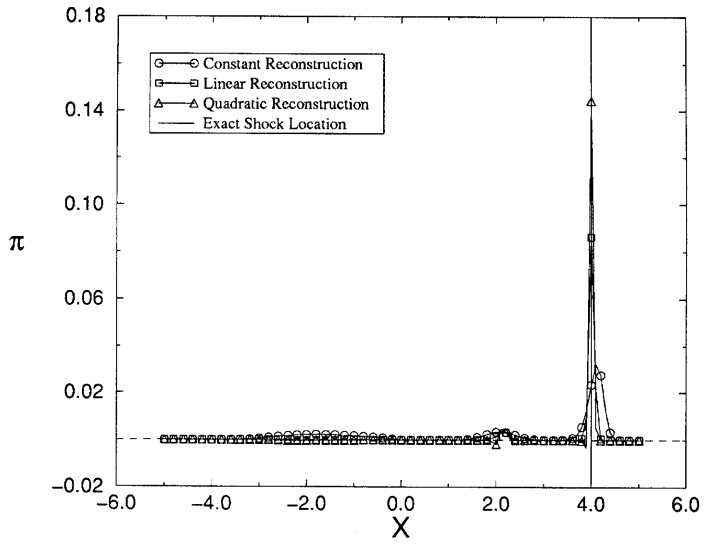

Figure 6. Local entropy production distributions using Godunov flux function: test problem (b1)

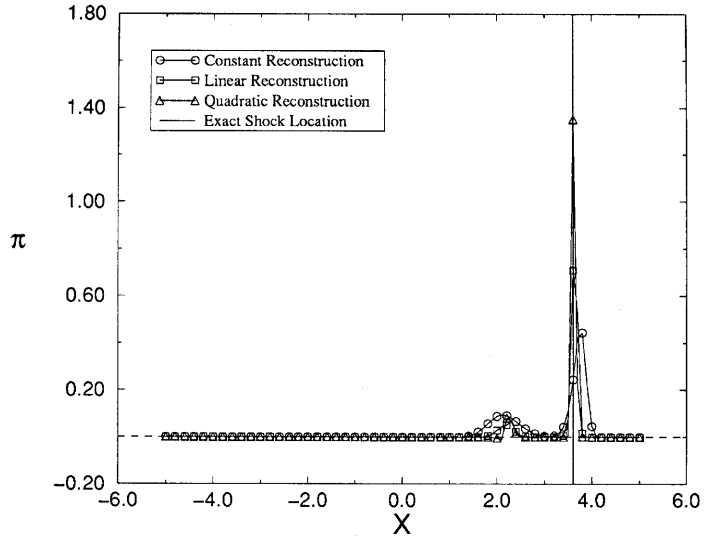

Figure 7. Local entropy production distributions using Godunov flux function: test problem (b2)

productions reduce substantially with increasing order of reconstruction. In agreement with Proposition 3 discussed in Section 3, we further observe that increasing the order of reconstruction may occasionally lead to negative values of π_i.

The local entropy productions corresponding to problem (b2) with the numerical entropy flux computed with the mass consistent entropy flux (MCEF) formulae (43a) and (43b), for constant and quadratic reconstructions respectively, are reported in Figures 8 and 9 and compared with those

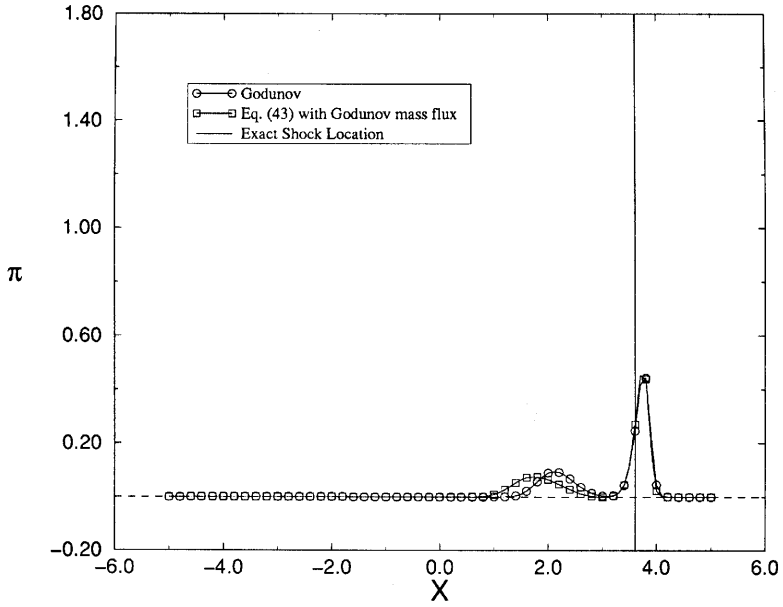

Figure 8. Local entropy production distributions using piecewise constant reconstruction: test problem (b2)

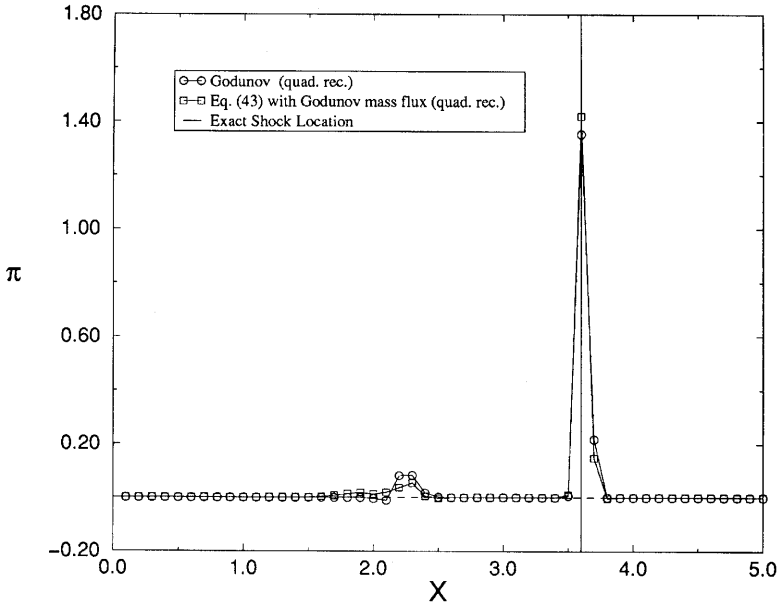

Figure 9. Local entropy production distributions using piecewise quadratic reconstruction: test problem (b2)

obtained with (39). The comparison demonstrates that the proposals (43a) and (43b) agree both qualitatively and quantitatively with (39) irrespective of the order of accuracy of the scheme. Finally, in Figure 10 we analyse the effects of different numerical mass flux functions for determining the local numerical entropy production. The results compare well with those arising from the exact

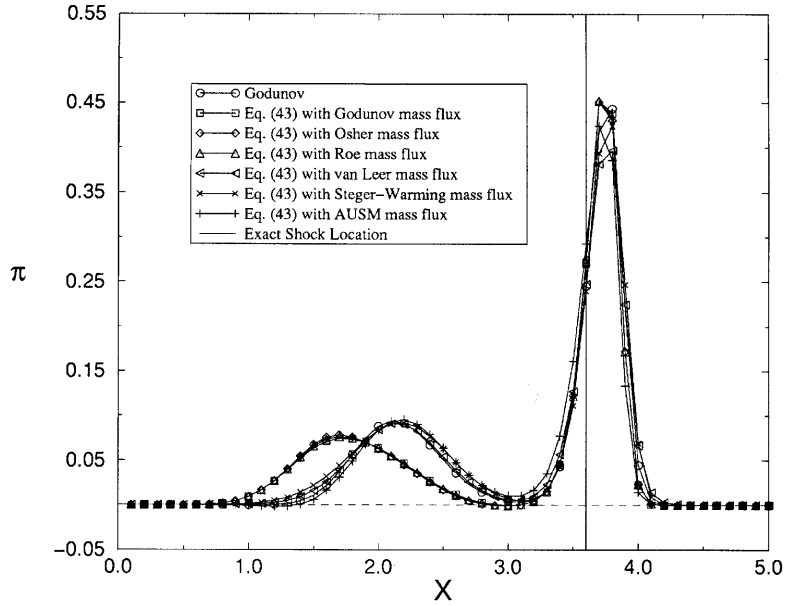

Figure 10. Local entropy production distributions computed by MCEF formulae using various flux functions: test problem (b2)

procedure described previously, thus confirming that one can exploit the homogeneity property of the Euler equations to determine a consistent numerical entropy flux.

7. CONCLUSIONS

This work has attempted to investigate in detail the relation between the numerical solution of systems of conservation laws and the associated entropy evolution. Such evolution may provide information on the physical relevance as well as on the quality of the numerical solution of conservation laws, much in the same way as the so-called *modified* (or *model*) equation furnishes a useful technique for studying the behaviour of the discrete solutions.

While for the Euler equations of gas dynamics an entropy balance can be obtained, its discrete counterpart cannot be identified, mainly owing to the difficulties in defining a numerical entropy flux.

A class of periodic test problems which admit an analytical solution with zero local and global entropy productions has been devised in order to correlate, at least at the global level, the entropy time evolution with the entropy production, while avoiding the necessity of defining a numerical entropy flux. The availability of the exact entropy production rates serve the purpose of quantitative assessment of the performances of different numerical methods.

The discussed numerical experiments suggest that in order to properly estimate the local entropy evolution, it is mandatory to distinguish between the physical entropy production and the numerical one and to locate the different contributions of the latter.

An original analysis of the effect of the reconstruction procedure of a finite-dimensional state vector on the induced entropy balance has been carried out by means of variational calculus. The analysis has led to the interesting conclusion that only the piecewise constant distribution of the unknown variables ensures the maximum of the global thermodynamic entropy under the conditions of mass, momentum and energy conservation.

Finally, the problem of defining a suitable numerical entropy flux has been addressed in detail. A consistent approximate expression for the entropy flux, which does not require any linearization procedure, while employing the same numerical mass flux as the baseline scheme, has been devised. The proposed flux formulation is found to compare well against reference data.

REFERENCES

1. A. Harten, P. D. Lax and B. van Leer, 'On upstream differencing and Godunov-type schemes for hyperbolic conservation laws', *SIAM Rev.*, **25**, 35–61 (1983).
2. P. D. Lax, *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*, SIAM Regional Series on Applied Mathematics, No. 11, SIAM, Philadelphia, PA, 1973.
3. M. S. Liou and C. J. Steffen, 'A new Flux Splitting Scheme' *J. Comput. Phys.*, **107**, 23–39 (1993).
4. S. Osher and F. Solomon, 'Upwind difference schemes for hyperbolic systems of conservation laws', *Math. Comput.*, **39**, 339–374 (1982).
5. S. Osher, 'Riemann solvers, the entropy condition, and difference approximations', *SIAM J. Numer. Anal.*, **21**, 217–235 (1984).
6. P. L. Roe, 'Approximate Riemann solvers, parameter vectors, and difference schemes', *J. Comput. Phys.*, **43**, 357–372 (1981).
7. G. A. Sod, 'A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws', *J. Comput. Phys.*, **27**, 1–31 (1978).
8. C. Shu, 'TVB uniformly high-order schemes for conservation laws', *Math. Comput.*, **49**, 105–121 (1987).
9. J. L. Steger and R. F. Warming, 'Flux vector splitting of the inviscid gasdynamic equation with application to finite difference methods', *J. Comput. Phys.*, **40**, 263–293 (1981).
10. E. Tadmor, 'The numerical viscosity of entropy stable schemes for systems of conservation laws. I', *Math. Comput.*, **49**, 91–103 (1987).
11. E. Tadmor, 'A minimum entropy principle in the gas dynamics equations', *ICASE Rep. N. 86/18*, 1986.
12. A. Harten, 'On the symmetric form of systems of conservation laws with entropy', *J. Comput. Phys.*, **49**, 151–164 (1983).

13. R. A. Cox and B. M. Argrow, 'Entropy production in finite-difference schemes', *AIAA J.*. **31**, 100–118 (1993).
14. P. D. Lax and B. Wendroff, 'Systems of conservation laws', *Commun. Pure Appl. Math.*, **13**, 217–237 (1960).
15. A. Harten, 'High resolution schemes for hypberbolic conservation laws', *J. Comput. Phys.*, **49**, 357–393 (1983).
16. A. Harten, J. M. Hyman and P. D. Lax, 'On finite-difference approximations and entropy conditions for shocks', *Commun. Pure Appl. Math.*, **29**, 297–322 (1976).
17. S. K. Godunov, 'On finite-difference approximations and entropy conditions for shocks', *Mat. Sb.*, **47**, 271–290 (1959).
18. A. Harten and S. Osher, 'Uniformly high-order accurate nonoscillatory schemes I', *SIAM J. Numer. Anal.*, **2**, 279–309 (1987).
19. A. Harten, B. Engquist, S. Osher and S. R. Chakravarthy, 'Uniformly high order accurate essentially non-oscillatory schemes, III', *J. Comput. Phys.*, **71**, 231–303 (1987).
20. B. van Leer, 'Flux vector splitting for the Euler equations', *Lecture Notes in Physics*, Vol. 170, 1982, pp. 507–512, Springer Verlag, Berlin, 1982.
21. A. Harten, 'ENO schemes with subcell resolutions', *J. Comput. Phys.*, **83**, 148–184 (1987).

