HAL
open science

Residual distribution schemes on quadrilateral meshes.

Remi Abgrall, Fabien Marpeau

To cite this version:

Remi Abgrall, Fabien Marpeau. Residual distribution schemes on quadrilateral meshes.. Journal of Scientific Computing, 2007, 30 (1), pp.131-175. inria-00334011

HAL Id: inria-00334011

https://inria.hal.science/inria-00334011

Submitted on 24 Oct 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Residual distribution schemes on quadrilateral meshes

F. Marpeau, R. Abgrall(†)
Mathématiques Appliquées de Bordeaux,
Université Bordeaux 1, 351 Cours de la Libération,
33405 Talence Cedex, France
† Institut Universitaire de France
Fabien.Marpeau@math.u-bordeaux1.fr and abgrall@math.u-bordeaux1.fr

Abstract

We propose an investigation of the residual distribution schemes for the numerical approximation of two dimensional hyperbolic systems of conservation laws on general quadrilateral meshes. In comparison to the use of triangular cells, usual basic features are recovered, an extension of the Upwinding concept is given, and a Lax-Wendroff like theorem is adapted for consistency. We show how to retrieve many variants of standard first and second order accurate schemes. They are proven to satisfy this theorem. An important part of this paper is devoted to the validation of these schemes by various numerical tests for scalar equations and Euler equations system for compressible fluid dynamics on non Cartesian grids.In particular, second order accuracy is reached by an adaptation of the Linear preserving property to quadrangle meshes. We discuss several choices as well as the convergence of iterative method to steady state. We also provide examples of schemes that are not constructed from an upwinding principle.

1 Introduction

The residual distribution schemes, or fluctuation splitting schemes, have emerged these last years for the numerical approximation of hyperbolic conservation laws. They are actually gaining more and more maturity, and are now well documented for steady computations (see for e.g. [1, 2, 3, 4, 5], and the references contained therein), and more recently, for unsteady computations (see [6, 7]). Practical experiments have shown that these schemes are robust, and several works have allowed the construction of schemes along three main principles : *Upwinding property, Monotonicity* and *Linearity Preservation*. However, to our knowledge, researchers often restrict the studies to triangular meshes, and in our opinion, it would be interesting to completely observe their behavior on quadrilateral meshes, for which few analysis have been published until now (see [8]). This would enable the use of hybrid meshes which are often used in industrial applications.
The model consists on a generic hyperbolic system of conservation laws in an open bounded two-dimensional space domain $\Omega \subset \mathbb{R}^2$,

$$\partial_t U(t,X) + \operatorname{div} \mathcal{F}\big(U(t,X)\big) = 0 \ , \ (t,X) \in \mathbb{R}_+ \times \Omega \ , \tag{1}$$

coupled to initial condition at $t = 0$ and boundary conditions on $\partial\Omega$. In this system, t and $X = (x, y)$ denote time and space variables respectively, $U \in \mathbb{R}^m$ is the vector-value unknown, and the nonlinear flux function $\mathcal{F} = (F, G)$ goes from a subset of $\mathbb{R}^m$ to $(\mathbb{R}^m)^2$, and is differentiable with the notations $A := F'$ and $B := G'$. A particular interest is devoted to the underlying steady–state system related to (1),

$$\operatorname{div} \mathcal{F}\big(U(X)\big) = 0 \ , \ X \in \Omega \ , \tag{2}$$

coupled to boundary conditions on $\partial\Omega$.

This paper provides a development of the residual distribution schemes to approximate (1) and mainly (2) on general quadrilateral meshes. In comparison to triangular meshes, all the properties relying on geometrical considerations have to be reconsidered, and the difficulties come from the use of nonlinear finite elements

which complicate the calculations. The Lax-Wendroff type Theorem in [9] is used to recover consistency. However, some assumptions need to be modified. The two-dimensional Upwinding concept was originally based on the fact that the number of vertices per triangle is three, and has consequently to be extended as well. Then, some classical upwind schemes are adapted to quadrangles by using a trick of [10] to recover a conservative formulation of the N scheme. A strategy allows us to engineer many variants of these schemes. All of our results are submitted to numerical tests for scalar conservation laws, and Euler's system for fluid mechanics. In addition to this, an integration lemma is presented to facilitate the implementation of the given schemes for the reader.

Basic features of residual distribution schemes for triangular meshes have become classical. They are briefly recalled in the next section for the sake of clearness; the third section explains how the residual distribution schemes can be naturally adapted to quadrilateral meshes, and examples of usual fluctuation splitting schemes are recovered in section 4. The next two parts of the paper present numerical results for various conservation laws. The last part discuss some difficulties and propose remedies to them. Our technique enable to consider residual distribution schemes that are not constructed thanks to an upwindig principle. Three technical appendices give some interpolation results and prove the consistency of the present schemes along the Lax-Wendroff like theorem.

2 Remarks about the residual distributions schemes on triangular meshes

We consider first a mesh composed of of triangles, $\{T_i\}_{i=1,\ldots,n_e}$, we denote $\{X_i\}_{i=1,\ldots,n_v}$ as the mesh points, and define h as the maximal length of the diameters of the elements. The dual cells, $\{C_i\}_{i=1,\ldots,n_v}$, are constructed for all X_i, $i = 1, \ldots, n_v$, by joining the centroids to the midpoints of the edges of T_j, $j = 1, \ldots, n_e$. When there is no ambiguity, an element, T_j, is denoted by T. The time interval $\mathbb{R}^+$ is usually discretized in sub-intervals (t_n, t_{n+1}), $0 = t_0 < t_1 < \ldots < t_n < t_{n+1} < \ldots$, we set $\Delta t_n = t_{n+1} - t_n$, and $\Delta t = \sup_n(\Delta t_n)$. The vector-value $U_i^n \in \mathbb{R}^m$ denotes the approximation of $U(t_n, X_i)$ and one defines $U^n := \{U_i^n\}_{1\le i\le n_v}$. Throughought the paper, the notation $|\psi|$ represents any norm of ψ if ψ is a vector of any $\mathbb{R}^q$, $q \in \mathbb{N}^\star$, and the measure of ψ, if ψ is a subdomain of any $\mathbb{R}^q$.
System (1) is approximated by the residual distribution scheme

$$U_i^{n+1} = U_i^n - \frac{\Delta t_n}{|C_i|} \sum_{T, X_i \in T} \Phi_i^{T,n}(U^n) \ . \tag{3}$$

In this formula, for any element T, each of the three residuals $\Phi_i^{T,n}(U^n)$, $X_i \in T$, have to be viewed as a fraction distributed to node i of a total fluctuation $\Phi^{T,n}(U^n)$ that is defined by

$$\Phi^{T,n}(U^n) = \int_T \operatorname{div} \mathcal{F}^h(U_h)\, dx \ , \tag{4}$$

where $\mathcal{F}^h$ is an approximation of $\mathcal{F}$ ($\mathcal{F}^h$ is continuous at the interfaces of the elements, $\mathcal{F}^h$ converges towards $\mathcal{F}^h$ in L^1_{loc}, see section 3.1), and U_h is the piecewise constant function having for value U_i^n in each space-time dual cell $[t_n, t_{n+1}[\times C_i$. These residuals have to verify

$$\sum_{i, X_i \in T} \Phi_i^{T,n}(U^n) = \Phi^{T,n}(U^n) \ . \tag{5}$$

Note that the index of the sum symbol in (5) is i such that X_i is a vertex of T, while the index of the sum symbol in (3) is T such that T shares X_i as a vertex. The dependency on n and U^n of $\Phi^{T,n}(U^n)$ and $\Phi_i^{T,n}(U^n)$ are often omitted in the writing, resulting in Φ^T and Φ_i^T. Because the scheme (3) is generally the consequence of the explicit Euler time discretization, its accuracy in time is at most first order. However it is well-known that an iteration process in (3) until convergence in n can lead to a spatial approximation

of (2), for which the accuracy only relies upon a relevant distribution of the Φ_i^T at the vertices of T. In [11], the authors build a method to approximate (2) at each order. To simplify the writing, the vertices $X_{i_1}, X_{i_2}, X_{i_3}$ of a triangle T are replaced by X_1, X_2, X_3, so that the residuals are $\Phi_1^T, \Phi_2^T, \Phi_3^T$, and (5) becomes $\Phi^T = \Phi_1^T + \Phi_2^T + \Phi_3^T$. When there is no ambiguity, since the considered scheme (3) is fully explicit, U_1, U_2, U_3 denote the values of U^n in T instead of U_1^n, U_2^n, U_3^n.

The consistency of fluctuation splitting schemes has the form of a Lax Wendroff like theorem that is proven in [9, 11]. Roughly speaking, under reasonable assumptions of the residuals Φ_i^T, such as existence of the above approximation of the flux $\mathcal{F}^h$ satisfying (4) and (5), if U_h converges to a function U in $\left(L^2_{\text{loc}}(\mathbb{R}_+ \times \mathbb{R}^2)\right)^m$, then U is a weak solution of the Cauchy problem corresponding to (1). In particular, this theorem requires the existence of a constant C independent of h, T and n such that

$$\left|\operatorname{div} \mathcal{F}^h_{|T}(U_h)\right| \leq \frac{C}{h} \sum_{i,j=1}^{3} |U_i - U_j| \ . \tag{6}$$

An important design criterion for standard residual distribution schemes is the Upwinding property, that generalizes the one of finite volume schemes. Consider schemes that are built on the quasilinear form of (2), namely,

$$\big(A(U), B(U)\big) \cdot \nabla U = 0 \ . \tag{7}$$

Let $\bar{U}$ denote an average state of U^n in T, depending on U_1, U_2, U_3, and define $\vec{n}_i$, $i = 1, 2, 3$, as the inward normal to the edge (X_j, X_k), $j, k \neq i$, such that $|\vec{n}_i| = |\overrightarrow{X_j X_k}|$. Define the matrices K_i, $i = 1, 2, 3$, by $K_i = \big(A(\bar{U}), A(\bar{U})\big) \cdot \vec{n}_i$ (if $m = 1$, K_i are scalars), and assume the averaging procedure that gives $\bar{U}$ respects the hyperbolicity of (2), so that K_i have real eigenvalues and is diagonalizable. Setting $K_i = \frac{1}{2} R_i \Lambda_i L_i$ where Λ_i is the diagonal matrix whose values are the eigenvalues of K_i, R_i is the matrix the columns of which are the corresponding eigenvectors, and $L_i = R_i^{-1}$, define

$$K_i^+ = \frac{1}{2} R_i \Lambda_i^+ L_i, \ K_i^- = \frac{1}{2} R_i \Lambda_i^- L_i,$$

where $\Lambda_i^+ = \frac{1}{2}\big(\Lambda_i + |\Lambda_i|\big)$, $\Lambda_i^- = \frac{1}{2}\big(\Lambda_i - |\Lambda_i|\big)$, $|\Lambda_i|$ being the diagonal matrix composed with the modulus of the eigenvalues. A residual distribution scheme is upwind if the residuals are sent only to the upstream nodes, that means: if $K_i^+ = 0$, then $\Phi_i^T = 0$. This property relies strongly upon the choice of the average state $\bar{U}$ in the linearization (7) and the geometry of the triangles for the definition of vectors $\vec{n}_i$. For scalar problems, a triangle for which only one residual is different from 0 is called a one-target element. If two residuals are different from 0, it is called a two-targets element. Since $\vec{n}_1 + \vec{n}_2 + \vec{n}_3 = 0$, it is clear that the three-targets case is impossible.

The monotonicity property stands for the capacity of a scheme to produce l^∞-stable, nonoscilatory, and eventually oscillatory numerical solutions, even in the presence of strong discontinuities. An example of monotone upwind scheme under a CFL-like condition is the N scheme ([4, 5]). Defining the matrices K_i^- as above, if the matrix $\sum_{i=1}^3 K_i^-$ is invertible, its inverse is denoted by N, and the N scheme is defined by

$$\Phi_i^{T,N} = K_i^+ \Big(U_i - N \sum_{j=1}^{3} K_j^- U_j\Big) = \sum_{j=1}^{3} K_i^+ N K_j^- (U_i - U_j) \ . \tag{8}$$

Sometimes the matrix N may not exist, but even in this case, a result in [1] gives a meaning to the system N scheme for any symetrizable system. Although the monotonicity of this scheme is carried out from the nonnegativity of the coefficients $K_i^+ N K_j^-$ for a scalar equation, this property is difficult to generalize for systems. However, the construction of scalar and system N schemes are very similar, numerical experiments prove the system N scheme is monotone shock capturing as its scalar version (see for example [2, 1], and

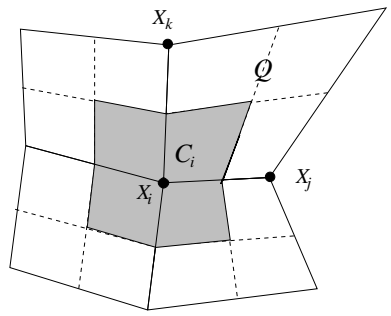


Figure 1: Dual cell C_i.

their references). Note that the conservation requirement (5) is not obvious. It needs the existence of a conservative linearization, that means an average state and an approximation $\mathcal{F}^h$ satisfying

$$\int_T \operatorname{div} \mathcal{F}^h(U_h)\, dx = \sum_{i=1}^{3} K_i U_i \ . \tag{9}$$

When there exists a quadratic variable change $Z \mapsto U$ such that $Z \mapsto \mathcal{F}\big(U(Z)\big)$ is also quadratic, the above conservation holds by setting $\bar{U} = U(\bar{Z})$ and $\mathcal{F}^h(U^h) = \mathcal{F}\big(U(\pi_h^T Z^h)\big)$, where π_h^T denotes the linear P_1–interpolation operator on the triangles. The existence of such a variable change occurs in certain cases, such as some scalar equations or Euler equations (Roe–Struijs–Deconink linearization, see [12]), but may be untrue for general hyperbolic systems of conservation laws. Moreover, (9) relies strongly upon the linearity of the P_1-interpolation with respect to space variable. That is why the N scheme is unacceptable for quadrilateral meshes.

3 Passing to quadrilateral meshes

In this section, we investigate the extension of the residual based schemes to quadrilateral meshes. The mesh elements are now quadrangles, $\{Q_i\}_{i=1,\dots,n_e}$. The nodes, still denoted by $\{X_i\}_{i=1,\dots,n_v}$, are the vertices of these elements, and the dual cells, C_i, $i = 1, \dots, n_v$, are still obtained by joining the centroids to the middles of the edges of the Q_i (see Figure 1). Let the symbol Q refer to a quadrangle. The scheme (3) is naturally extended to

$$U_i^{n+1} = U_i^n - \frac{\Delta t_n}{|C_i|} \sum_{Q, X_i \in Q} \Phi_i^Q \ , \tag{10}$$

and the conservation property (4),(5), becomes

$$\Phi^Q := \int_Q \operatorname{div} \mathcal{F}^h(U_h)\, dx \ ; \ \sum_{i, X_i \in Q} \Phi_i^Q = \Phi^Q \ . \tag{11}$$

With previous notations, we refer to the nodes of a quadrangle Q as X_1, X_2, X_3, X_4, instead of $X_{i_1}, X_{i_2}, X_{i_3}, X_{i_4}$, so that the conservation of the four residuals in Q reads $\Phi_1^Q + \Phi_2^Q + \Phi_3^Q + \Phi_4^Q = \Phi^Q$; when there is no ambiguity, the dependency on n of the four values of the numerical solution in Q is dropped, so that they denoted by U_1, U_2, U_3, U_4. We detail the basic principles of these schemes.

3.1 Consistency

We make assumptions that allow to obtain a Lax-Wendroff like theorem for residual distribution schemes on quadrilateral meshes. These assumptions are nearly the same as those of triangular meshes, but some of

them had to be enlarged. Furthermore these assumptions give a rigorous meaning to the above mentioned approximation $\mathcal{F}^h$. For these reasons we recall them all. Consider a family $\{\mathcal{Q}_h\}_h$, of quadrilateral meshes which make a partition of $\mathbb{R}^2$, where h denotes the maximal length of the diagonals of the cells of the mesh $\mathcal{Q}_h$, and let $n_{v,h}$ and $\{C_{h,i}\}_i$ be the number of mesh points and the set of dual cells of $\mathcal{Q}_h$. The family $\{\mathcal{Q}_h\}_h$ is assumed to be regular in the following sense : there exist $C_1, C_2 > 0$ such that

$$\forall h,\ C_1 \leq \frac{h^2}{\inf_{Q\in\mathcal{Q}_h}\left(|Q|\right)} \leq C_2\ . \tag{12}$$

Define the set

$$E_h = \left\{ V_h \in \left(L^2_{\text{loc}}(\mathbb{R}_+ \times \mathbb{R}^2)\right)^m \text{ such that } V_h \text{ is constant in } [t_n, t_{n+1}[\times C_i \right\}\ .$$

Denote as before $(U_h)_h$ as the family functions of E_h having for value U_i^n in $[t_n, t_{n+1}[\times C_i$, where U_i^n is given by the scheme (10). Assume the scheme (10) satisfies the following hypotheses: there exists $C_3 > 0$ such that

$$\forall n,\ h,\ Q \in \mathcal{Q}_h,\ X_i \in Q,\ |\Phi_i^Q| \leq C_3 h \sum_{j,k,X_j,X_k\in Q} |U_j^n - U_k^n|\ . \tag{13}$$

There exists an approximation $\mathcal{F}^h$ of the flux function $\mathcal{F}$ satisfying: For all $h, Q \in \mathcal{Q}_h$,

$$\sum_{i=1}^4 \Phi_i^Q = \int_Q \operatorname{div} \mathcal{F}^h_{|_Q}(U_h)\, dx\ ; \tag{14}$$

For all function $V_h \in E_h$ converging to a function V in $\left(L^2_{\text{loc}}(\mathbb{R}_+ \times \mathbb{R}^2)\right)^m$ as $h, \Delta t \to 0$, and such that there exists a compact $\mathcal{K} \subset \mathbb{R}^m$ such that $V^h(t,x) \in \mathcal{K}$, then

$$\mathcal{F}^h(V_h) \xrightarrow[h,\Delta t\to 0]{} \mathcal{F}(V) \text{ in } \left(L^1_{\text{loc}}(\mathbb{R}_+ \times \mathbb{R}^2)\right)^m; \tag{15}$$

For all Q_1 and Q_2 neighbors (having a common edge),

$$\mathcal{F}^h(V_h)_{|_{Q_1}} \cdot \vec{n} = \mathcal{F}^h(V_h)_{|_{Q_2}} \cdot \vec{n}, \quad \text{on } \bar{Q}_1 \cap \bar{Q}_2, \tag{16}$$

$\vec{n}$ being an outward unit normal to $\bar{Q}_1 \cap \bar{Q}_2$; There exists $C_4 > 0$ such that

$$\forall n,\ h,\ Q,\ \int_Q \left|\operatorname{div} \mathcal{F}^h_{|_Q}(U_h)\right| dx \leq C_4 h \sum_{i,j=1}^4 |U_i^n - U_j^n|\ . \tag{17}$$

We have the following theorem.

Theorem 1. *Let (12)–(17) hold. Assume $U_0 \in \left(L^\infty(\mathbb{R}^2)\right)^m$ and $\mathcal{F} \in \left(\mathcal{C}^1(\mathcal{O};\mathbb{R}^m)\right)^2$, where $\mathcal{O} \subset \mathbb{R}^m$. Assume there exists a compact $\mathcal{K} \subset \mathcal{O}$ and $U \in \left(L^2_{loc}(\mathbb{R}_+ \times \mathbb{R}^2)\right)^m$ such that :*

$$\forall n, i,\ U_i^n \in \mathcal{K}\ ,$$

$$U_h \xrightarrow[h,\Delta t\to 0]{} U \text{ in } \left(L^2_{loc}(\mathbb{R}_+ \times \mathbb{R}^2)\right)^m\ .$$

Then U is a weak solution of the Cauchy problem corresponding to (1).

The proof follows exactly the same lines as in [9, 11] and relies upon Abel summation arguments, and the use of the first order Q_1 interpolation on quadrilaterals (see appendix A), instead of P_1 interpolation on triangles in the original proof. It shall not be reproduced here. However, since the Q_1 interpolation is not piecewise linear with respect to X, we observed that the proof could not be retrieved with the natural extension of assumption (6), that has been replaced by (17).

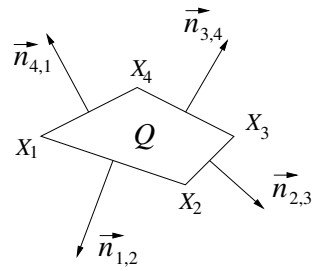


Figure 2: Outward normals to Q.

3.2 Upwinding

The definition of the Upwinding property is directly linked to the geometry of the cells. More precisely, each vertex X_i in a triangle has a unique opposite edge. This allowed to define the vectors $\vec{n}_i$, $i = 1, \dots, 3$, in section 2. In a quadrilateral, each vertex X_i has two opposite edges, thus the Upwinding concept has to be modified. One naturally decides to define $\vec{n}_i$, for $i = 1, \dots, 4$, as the sum of the two inward normals to the two opposite edges of X_i: to be more rigorous, for $i = 1, \dots, 4$, one denotes X_j as the opposite vertex to X_i, and X_k, X_l as the two other vertices. Then, for $i = 1, \dots, 4$, one defines $\vec{n}_i = -\vec{n}_{j,k} - \vec{n}_{j,l}$ where $\vec{n}_{j,k}$ is the outward normal to (X_j, X_k) having for modulus $|\overrightarrow{X_j X_k}|$, and $\vec{n}_{j,l}$ is the outward normal to (X_j, X_l) having for modulus $|\overrightarrow{X_j X_l}|$. In Figure 2, $\vec{n}_1 = -\vec{n}_{2,3} - \vec{n}_{3,4}$, $\vec{n}_2 = -\vec{n}_{3,4} - \vec{n}_{4,1}$, $\vec{n}_3 = -\vec{n}_{4,1} - \vec{n}_{1,2}$, $\vec{n}_4 = -\vec{n}_{1,2} - \vec{n}_{2,3}$. Then denote as before $\bar{U}$ as an average state of U in Q, depending on U_1, U_2, U_3, U_4, and define $K_i = \big(A(\bar{U}), B(\bar{U})\big) \cdot \vec{n}_i$, $i = 1, \dots, 4$, and $K_i^{\pm} = \frac{1}{2} R_i \Lambda_i^{\pm} L_i$. The Upwinding property is naturally generalized as:

$$\text{for } i = 1, \dots, 4, \quad \text{if } K_i^+ = 0, \quad \text{then } \Phi_i^Q = 0 \ .$$

We point out that since $\vec{n}_{1,2} + \vec{n}_{2,3} + \vec{n}_{3,4} + \vec{n}_{4,1} = 0$, one has $\vec{n}_1 = -\vec{n}_3$ and $\vec{n}_2 = -\vec{n}_4$, hence it can only exist one-target and two-targets quadrangles again. The three and four-targets cases are excluded.

3.3 Monotonicity

The monotonicity condition is the same as the one available in the literature, but with another CFL like condition. For a scalar conservation equation, the scheme (10) is said to be monotonic if it can be written on the form

$$\Phi_i^Q = \sum_{j=1}^{4} c_{ij}(U_i - U_j) \ , \quad \text{with} \quad c_{ij} \geq 0, \ i, j = 1, \dots, 4 \ . \tag{18}$$

Since

$$U_i^{n+1} = \sum_{Q, X_i \in Q} \frac{|C_i \cap Q|}{|C_i|} \Big(U_i^n - \frac{\Delta t_n}{|C_i \cap Q|} \Phi_i^Q \Big) \ ,$$

a scheme is l^∞-stable and preserves nonnegative solutions if each term in the previous sum is nonnegative, and direct calculations show that is true for a scheme verifying (18) under the CFL like condition

$$\forall i, \ \forall Q \text{ with } C_i \cap Q \neq \emptyset, \ \Delta t \sum_{j, j \neq i} c_{i,j} \leq |C_i \cap Q| \ .$$

We point out that $|C_i \cap Q| \neq \frac{1}{4} Q$ in general. More generally, for systems, or scalars schemes for which (18) is not true, we still say a scheme is monotonic if it does not produce spurious oscillation and is monotone shock capturing, that means if there is non pre- or post- shock oscillation experimentally.

3.4 Second order accuracy at steady state

Following exactly the same lines as [1], one observes that a scheme satisfying the usual residual property ([13]) produces a second order approximation of the steady state system (2), as soon as $\mathcal{F}^h$ is a second order approximation in space of $\mathcal{F}$, that means, for all smooth function V constant in time, $\mathcal{F}^h\big(V(X)\big) - \mathcal{F}\big(V(X)\big) = \mathcal{O}(h^2)$ (this is generally true when $\mathcal{F}^h$ is based on the Q_1 interpolation in the next section). One recalls a scheme is said to have the residual property if for any sequence $(V_k)_k$, for all Q,

$$\text{if } \Phi^Q(V_k) \xrightarrow[k\to+\infty]{} 0, \text{ then for all } i = 1,\dots,4,\ \Phi_i^Q(V_k) \xrightarrow[k\to+\infty]{} 0 \ ,$$

with the convergences being independent of h. This property occurs in particular for schemes for which the residuals can be written on the form $\Phi_i^Q = \beta_i \Phi^Q$, $i = 1,\dots,4$, where β_i is a uniformly bounded matrix with respect to U^n and to the mesh.

4 Examples of residual distribution schemes for quadrilateral meshes satisfying (13)–(17).

Some usual first and second order fluctuation splitting schemes that exist for triangular meshes are recovered. Excepted in subsection 4.2, our strategy is to construct the upwind shemes from intrinsic functions $\mathcal{F}^h$ that are proven to satisfy assumptions (15)–(17) in appendix B, and then, any approximation $\mathcal{F}^h$ and any averaging procedure produce a variant of these schemes.

4.1 Examples of functions $\mathcal{F}^h$ satisfying (15)–(17).

The main tool in this paragraph is the nonlinear Q_1-interpolation (see appendix A). We denote ϕ_i, $i = 1,\dots,4$, as the basis function of the Q_1 interpolation and $\pi_h f_h$ as the Q_1-interpolation of any function f defined everywhere in $\mathbb{R}_+ \times \Omega$.

4.1.1 The Q_1-interpolation of the flux function

The first approximation of $\mathcal{F}$ is given by

$$\mathcal{F}^h(U_h) = \pi_h \mathcal{F}(U_h) \ , \tag{19}$$

that means $\mathcal{F}^h(U_h) = \sum_{i=1}^4 \mathcal{F}(U_i)\phi_i$. A straightforward application of the integration lemma 4 enables us to calculate the total fluctuation on a quadrangle Q for $\mathcal{F}^h$.

Lemma 1. *The interpolation of the flux yields:* $\displaystyle\int_Q div\,\mathcal{F}^h(U_h)\,dx = \frac{1}{2}\sum_{i=1}^4 \mathcal{F}(U_i)\cdot\vec{n}_i$.

4.1.2 Flux evaluated at an interpolated quantity.

Now define $U \mapsto Z$ as a $\mathcal{C}^1$ change of variable, denote $Z(U)$ as the application $U \mapsto Z$ evaluated in U and $U(Z)$ its inverse mapping evaluated in Z. A second method to approximate $\mathcal{F}$ is:

$$\mathcal{F}^h(U_h) = \mathcal{F}\big(U(\pi_h Z_h)\big), \tag{20}$$

where $Z_h = Z(U_h)$. In practice, we use a variable change such that U and $\mathcal{F}(U)$ are polynomial with respect to Z, and then, since the Q_1 basis functions ϕ_i are linear on the edges of the elements, the total residual $\int_Q \operatorname{div}\mathcal{F}^h(U_h)\,dx$ can be easily exactly computed via quadrature formulas. For example, the using of Green's formula combined with Lemma 4 in appendix C allows to compute Φ^Q up to a fourth degree polynomial dependency in Z for $\mathcal{F}$.

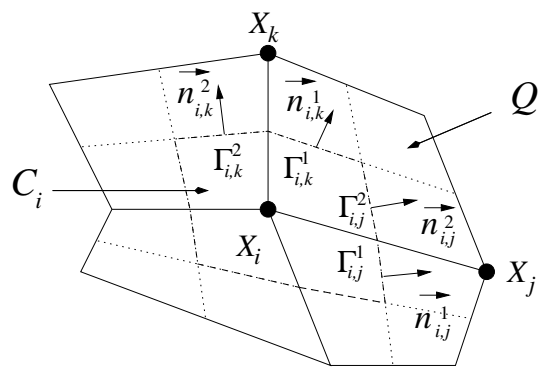


Figure 3: Geometry for the FV scheme.

4.2 The vertex-centered finite volume scheme.

As for the triangular meshes case, the finite volume scheme is a residual distribution scheme. For any neighboring nodes i and j, we set $\Gamma_{i,j} := C_i \cap C_j$ composed with $\Gamma^1_{i,j}, \Gamma^2_{i,j}$, and $\vec{n}^1_{i,j}$ and $\vec{n}^2_{i,j}$ the corresponding outward unit normals to C_i on $\Gamma^1_{i,j}$ and $\Gamma^2_{i,j}$, respectively (see Figure 3). The finite volumes scheme writes :

$$U_i^{n+1} = U_i^n - \frac{\Delta t_n}{|C_i|} \sum_{l,l \text{ neighbor of } i} \left(|\Gamma^1_{i,l}|\, \mathfrak{F}(U_i, U_l, \vec{n}^1_{i,l}) + |\Gamma^2_{i,l}|\, \mathfrak{F}(U_i, U_l, \vec{n}^2_{i,l}) \right)$$

where the numerical flux, $\mathfrak{F}$, is a locally Lipschitz continuous function with respect to U_i, U_l, verifying : for all $\vec{n} \in \mathbb{R}^2$, $V, W \in \mathbb{R}^m$,

$$\text{conservation} : \mathfrak{F}(V, W, \vec{n}) = -\mathfrak{F}(V, W, -\vec{n}) \ , \tag{21}$$

$$\text{consistency} : \mathfrak{F}(V, V, \vec{n}) = \mathcal{F}(V) \cdot \vec{n} \ . \tag{22}$$

Since $\displaystyle\sum_{l,l \text{ neighbor of } i} \sum_{k=1}^{2} |\Gamma^k_{i,l}| \vec{n}^k_{i,l} = 0$ and using the consistency property of $\mathfrak{F}$, we get

$$\sum_{l,l \text{ neighbor of } i} \sum_{k=1}^{2} |\Gamma^k_{i,l}| \mathfrak{F}(U_i, U_i, \vec{n}^k_{i,l}) = 0 \quad ,$$

hence the scheme can be written in term of residual distribution scheme as

$$\begin{aligned} \Phi_i^{Q,\mathrm{FV}} = {} & |\Gamma^1_{i,j}|\, \mathfrak{F}(U_i, U_j, \vec{n}^1_{i,j}) + |\Gamma^2_{i,k}|\, \mathfrak{F}(U_i, U_k, \vec{n}^2_{i,k}) \\ & - |\Gamma^1_{i,j}|\, \mathfrak{F}(U_i, U_i, \vec{n}^1_{ij}) - |\Gamma^2_{ik}|\, \mathfrak{F}(U_i, U_i, \vec{n}^2_{i,k}) \end{aligned} \tag{23}$$

where j and k are the two adjacent vertices to i in the quadrilateral Q, numbered in the same order than in the picture. In appendix B it is proven that the conservation (11) holds with $\mathcal{F}^h$ being the Q_1 interpolation of the flux (19).

4.3 The conservation based N scheme.

For the above mentioned reason, the original form of the N scheme (formula (8), with "4" instead of "3") cannot be used, since the conservation (11) is generally not exactly true for a given approximation $\mathcal{F}^h$. In [10], the authors exhibit that a nonconservative scheme is useless because it can lead to erroneous and misplaced discontinuities. They consequently build and test on several examples a version of the N scheme, the "Contour integral based N scheme", that is always conservative even the used linearization is not. Since

it does not rely upon a linear interpolation, this can fit to our case. We refer to this scheme as "CN", for "conservative N scheme". First, for a quadrangle Q, construct any average state on Q, select any $\mathcal{F}^h$ in subsection (4.1), as (19), or (20), and define Φ^Q from (11). Then, define the upwind residuals

$$\Phi_i^{Q,\mathrm{CN}} = K_i^+(U_i - \tilde{U}), \tag{24}$$

where $\tilde{U}$ must be chosen so that $\sum_{i=1}^4 \Phi_i^{Q,\mathrm{CN}} = \Phi^Q$. Then sum (24) from $i = 1$ to $i = 4$, and rearrange, to obtain

$$\Big(\sum_{i=1}^4 K_i^+\Big)\tilde{U} = \sum_{i=1}^4 K_i^+ U_i - \Phi^Q \ .$$

Since $\sum_{i=1}^4 \vec{n}_i = 0$, one has $\sum_{i=1}^4 K_i^+ = -\sum_{i=1}^4 K_i^-$, thus when the matrix $N := \Big(\sum_{i=1}^4 K_i^-\Big)^{-1}$ exists[1], $\tilde{U}$ is well defined, and the scheme (24) reduces to

$$\Phi_i^{Q,\mathrm{CN}} = K_i^+\Big(U_i + N\Big(\sum_{j=1}^4 K_j^+ U_j - \Phi^Q\Big)\Big) \ . \tag{25}$$

As the original N scheme, this variant does not satisfy the residual property and then stays at most first order accurate. Furthermore, the monotonicity property of the original N scheme is lost, but nevertheless, numerical tests in [10] and sections 5,6 of the present paper show the CN scheme reveals monotonic experimentally.

4.4 The LDA scheme.

The well-known upwind LDA scheme ([14]) is also retrieved. Taking $\mathcal{F}^h$ as above and defining an average state in Q,

$$\Phi_i^{Q,\mathrm{LDA}} = -K_i^+ N \Phi^Q \ . \tag{26}$$

The conservation property (11) is obviously satisfied. One recall moreover the uniform boundedness of the distribution coefficients $-K_i^+ N$ makes this scheme a second order approximation of (2).

4.5 Blended CN-LDA schemes.

A blending between the previous two upwind schemes is

$$\Phi_i^{Q,\mathrm{Bl}} = \Theta\Phi_i^{Q,\mathrm{CN}} + (\mathrm{Id} - \Theta)\Phi_i^{Q,\mathrm{LDA}}, \tag{27}$$

where Id stands for the identity matrix in $\mathbb{R}^m$, and Θ is a bounded matrix so that the scheme $\Phi_i^{Q,\mathrm{Bl}}$ satisfies both the residual property and the monotonicity condition. Here, the use of the non monotonic CN scheme may make the monotonicity of the resulting blended scheme not hold, even for the examples in [1]. Nevertheless section 5 shows that such a blended scheme reveals quite robust in practice. As it is a convex combination of CN and LDA schemes, the exact conservation of the total fluctuation is still satisfied. In [2, 10], the authors propose the so-called B scheme by introducing the diagonal blending matrix

$$\Theta_{kk}^{\mathrm{B}} = \frac{|(\Phi^Q)_k|}{\sum_{j=1}^4 |(\Phi_j^{\mathrm{CN}})_k|} \text{ if } (\Phi^Q)_k \neq 0 \ , \quad \text{and} \quad 0 \quad \text{otherwise} \ ,$$

$(\Phi^Q)_k$ and $(\Phi_j^{\mathrm{CN}})_k$ denoting the kth coordinate of Φ^Q and $\Phi_i^{Q,\mathrm{CN}}$, respectively. This scheme is not monotonic, even when written in terms of N scheme instead of CN, but this bending procuces good numerical results in general.

[1] or more precisely, $K_j^+ N$ can always been given a meaning if the system (1) is symetrizable.

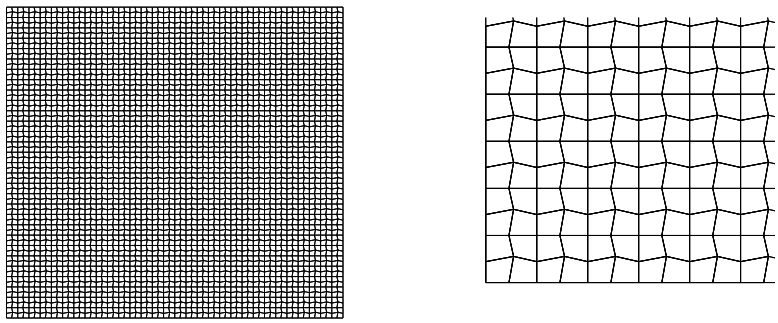


Figure 4: Mesh (3600 elements) and zoom of the mesh at the bottom left corner of Ω for scalar test cases.

4.6 Limited schemes.

A high compressive limitation formula in [15, 16] provides monotone schemes endowed with the residual property and does not rely upon the number of cell nodes. This naturally lends itself to quadrilateral meshes. Consider a scalar monotonic scheme Φ_i^Q, define as in [5, 16] the limiting weights

$$\beta_i^\star = \frac{\left(\Phi_i^Q/\Phi^Q\right)^+}{\sum_{j=1}^4 \left(\Phi_j^Q/\Phi^Q\right)^+} \text{ if } \Phi^Q \neq 0, \quad \text{and 0 otherwise .} \tag{28}$$

The resulting scheme, $\Phi^{Q,\star} = \beta_i^\star \Phi^Q$, is still conservative and monotonic since $\sum_{i=1}^4 \beta_i^\star = 1$ and $\beta_i^\star/\beta_i \geq 0$, with $\beta_i = \Phi_i^Q/\Phi^Q$. It is second order accurate at steady state since for all i, $0 \leq \beta_i^\star \leq 1$. Since $\Phi_i^{Q,\star} = 0$ if $\Phi_i^Q = 0$, then $\Phi^{Q,\star}$ is upwind if Φ_i^Q is upwind. References [15, 16] generalize this scheme for systems as follows: select a unit basis $\{e_j\}_{j=1,\ldots,m}$ of $\mathbb{R}^m$, and express a monotonic scheme Φ_i^Q and its corresponding total fluctuation as

$$\Phi_i^Q = \sum_{l=1}^m \varphi_{i,l} e_l \ , \ \Phi^Q = \sum_{l=1}^m \varphi_l e_l \ , \tag{29}$$

so that the component-wise conservation $\sum_{i=1}^4 \varphi_{i,l} = \varphi_l$ holds for all l. Then define $\Phi_i^{Q,\star}$ by using the previous scalar limitation formula as

$$\Phi_i^{Q,\star} = \sum_{l=1}^m \varphi_{i,l}^\star e_l \ , \ \varphi_{i,l}^\star = \frac{\left(\varphi_{i,l}/\varphi_l\right)^+}{\sum_{j=1}^4 \left(\varphi_{j,l}/\varphi_l\right)^+} \ \varphi_l \ . \tag{30}$$

As before, the limitation of the CN scheme may not yield a theoretical monotonic scheme, but is robust in practice. We refer to this scheme as "limited-CN" scheme.

5 Numerical results for scalar conservation laws.

First, the previous upwind schemes are tested for steady scalar equations, in the domain $\Omega = (0,1) \times (0,1)$, with a structured, but non-uniform mesh (Figure 4). In addition, since each quadrangle is at most a two-target element, we also coded a blended scheme of [1] which relies upon Roe's and Sidilkover's analysis (see [17]): if $\Phi_i^{Q,\mathrm{CN}} = 0$ for all $i = 1,\ldots,4$, we keep $\Phi_i^{Q,\mathrm{Bl}} = 0$, $i = 1,\ldots,4$. If Q is a one-target element, it follows from the conservation of both CN and LDA schemes that any value of Θ in (27) works. If two elements $\Phi_1^{Q,\mathrm{CN}}$ and $\Phi_2^{Q,\mathrm{CN}}$ are different from 0, setting $r_i = \Phi_i^{Q,\mathrm{CN}}/\Phi_i^{Q,\mathrm{LDA}}$, $i = 1,2$, we define the blending coefficient in (27) as

$$\Theta^{\mathrm{Bl}} = \min\Big(1, \max\left(\psi(r_1), \psi(r_2)\right)\Big) \ , \ \psi(r_i) = \frac{|r_i|}{1+|r_i|} \ .$$

From now, we will refer to this scheme as "Bl scheme". The numerical solution of (2) is computed by iterating the scheme (10) until convergence of the sequence U^n. For these next two test cases, the convergence criterion with the discrete l^∞ norm reached the computer zero for all the schemes, as the same manner as triangular meshes. The CFL condition was as follows: for all $i = 1, \ldots, nv$, for all Q such that $C_i \cap Q \neq \emptyset$, $\Delta t \sum_{j=1}^{4} |K_j| \leq |C_i \cap Q|$.

5.1 Constant convection.

Consider the simplest case

$$\operatorname{div} \lambda U = 0,\ U \in \mathbb{R},\ \lambda \in \mathbb{R}^2\ ,$$

with $\lambda = (1, 0.7)$, coupled to the boundary conditions $U(0, y) = 0$, $U(x, 0) = 1$, for $x, y \in (0, 1)$, which is equivalent to a space-time one-dimensional linear transport equation. The solution is obvious, and features a discontinuity along the line $y = 0.7\,x$, which separates the two constant values 0 and 1. We select $\bar{U} = \frac{1}{4}(U_1 + U_2 + U_3 + U_4)$ as average state, and the flux evaluated at the Q_1 interpolation of U_h, $\lambda \pi_h U_h$, as consistent approximation of $\mathcal{F}$; since λ is constant, this function $\mathcal{F}^h$ is equivalent to the Q_1 interpolation of the flux. With previous notations, one has $\Phi^Q = \sum_{i=1}^{4} K_i U_i$, and therefore, in this simple test, the conservative CN scheme (25) reduces to the original N scheme (8): it is monotone, positive, and the Bl scheme verifies upwind, monotonicity and residual properties, as well as the limited N scheme : this is the well-known PSI scheme [12].

The initial value for the iterative procedure is $U(0, ., .) = 0$ in Ω. The results are plotted in Figures 5 and 6. The first remark is that residual distribution schemes work even the cells are quadrilateral and have the same behaviours as triangular ones: monotonicity and diffusivity of the N scheme across the shock line, second order accuracy but oscillations for the LDA scheme, which captures the shock better than the N scheme, and both monotonicity and second order accuracy for both Bl and PSI schemes, that are very close together and capture the shock much better than the first order N scheme. Note that the B scheme is quite oscillatory in this case, but it solves the shock accurately.

5.2 Burger's equation.

The second test consists of the nonlinear equation

$$\operatorname{div}\left(U, \frac{U^2}{2}\right) = 0,$$

with $U(x, 0) = 1.5 - 2x$, $U(0, y) = 1.5$, $U(1, y) = -0.5$. The solution of this equation has a shock along the line connecting $(\frac{3}{2}, \frac{1}{2})$ and $(1, 1)$, and a fan at the bottom part of the domain. The computations have been performed with the same arithmetic average state as above, $\mathcal{F}^h(U_h) = \left(\pi_h U_h, \frac{1}{2}(\pi_h U_h)^2\right)$, and $U(0, x, y) = 1.5 - 2x$ as starting value. The CN scheme is used here since it is now a different formula from the original N scheme, which is no longer conservative. Figures 7 and 8 show the experimental monotonicity of the CN scheme, that is oscillation free, although we are not able to write it as (18). Consequently, the shock capturing limited-CN and Bl schemes are also monotone. Here, we point out that the B scheme gives results very close to both Bl and limited-CN schemes, although there is small overshoots for its extrema (Table 1). Small wiggles are observed in the fan region for the three second order schemes. This phenomenon has been related by other authors (see for example [15]). We come back to this in section 7. The LDA scheme is unstable near the shock. To investigate the dependency on the choice of $\mathcal{F}^h$ in the construction of the schemes, this test was also performed using the interpolation of the flux, namely, $\mathcal{F}^h(U_h) = \pi_h(U_h, U_h^2)$. Isovalues were exactly superposed to those of Figure 7. However, slight differences between the extrema were observed (see Table 1).

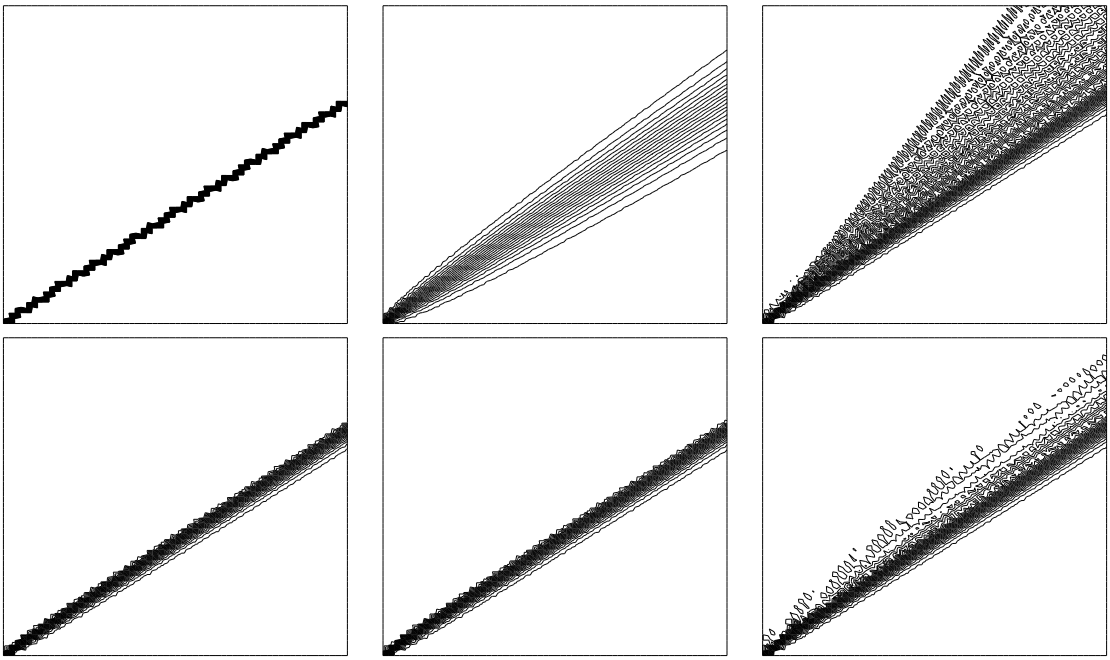


Figure 5: Constant convection contours: exact solution (top-left), numerical solution computed with the N scheme (top-center), LDA scheme (top-right), PSI scheme (bottom-left), Bl scheme (bottom-center) and B scheme (bottom-right), using twenty isovalues between the extrema.

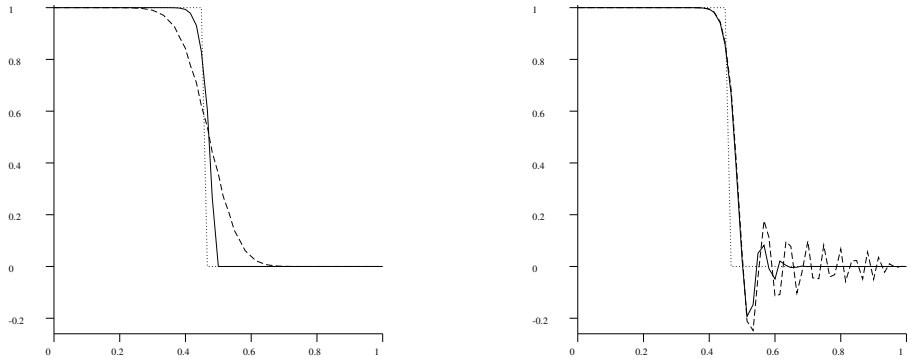


Figure 6: Constant convection: cross section along the line $x = 0.65$ of exact solution (dotted line), numerical solution computed with the N scheme (left,dashed line), PSI scheme (left, solid line), LDA scheme (right, dashed line), B scheme (right, solid line).

Table 1: Extrema for the two approximations $\mathcal{F}^h$.

	$\mathcal{F}^h(U_h) = \left(\pi_h U_h, \frac{1}{2}(\pi_h U_h)^2\right)$	$\mathcal{F}^h(U_h) = \pi_h(U_h, U_h^2)$
CN, limited-CN, Bl	$-0.5 \leq U \leq 1.5$	$-0.5 \leq U \leq 1.5$
LDA	$-1.02 \leq U \leq 1.74$	$-0.89 \leq U \leq 1.77$
B	$-0.5 \leq U \leq 1.68$	$-0.5 \leq U \leq 1.61$

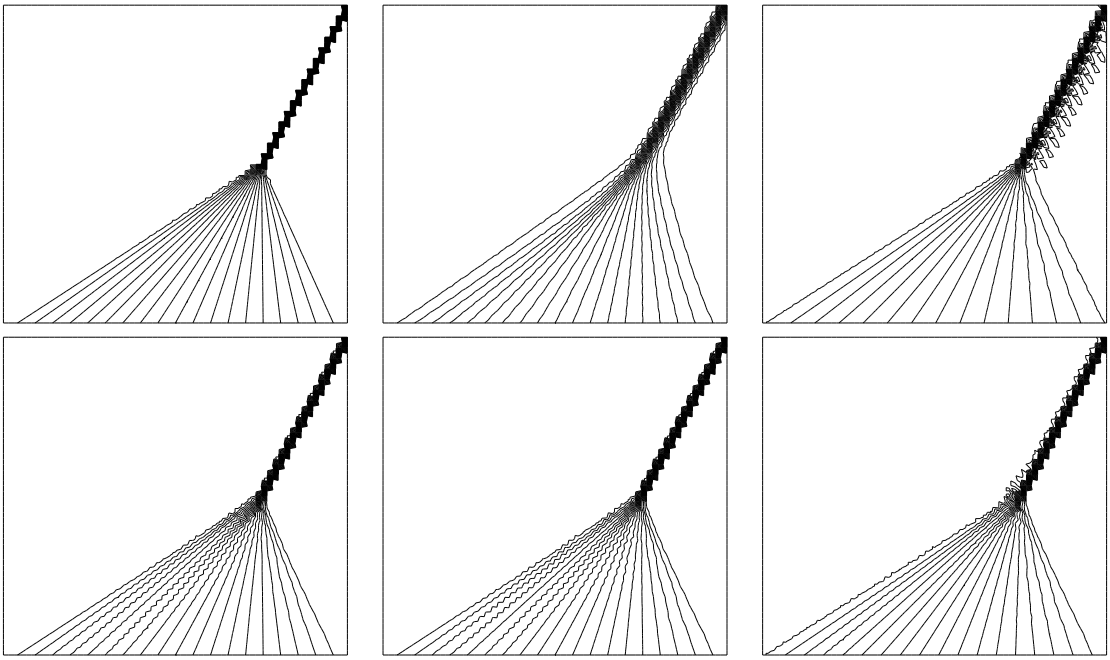


Figure 7: Burgers' equation contours: exact solution (top-left), numerical solutions computed with the CN, LDA, limited-CN, Bl, and B schemes, from left to right and top to bottom respectively, using twenty isovalues between the extrema.

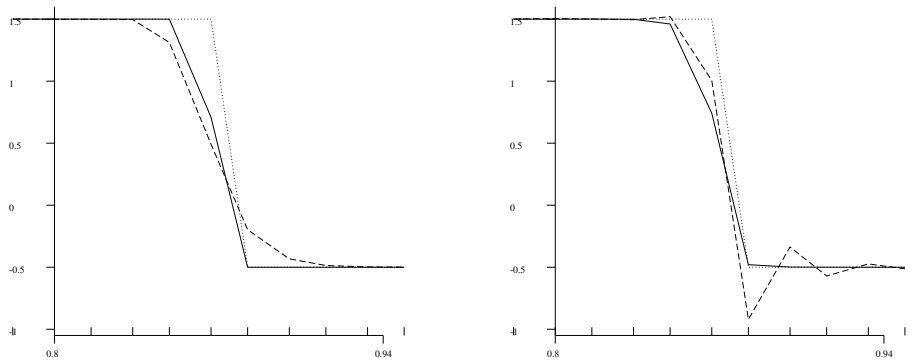


Figure 8: Burger's equation: cross section of the shock along the line $y = 0.736$: exact solution (dotted line), CN scheme (left, dashed line), limited-CN and Bl scheme (left, solid line), LDA scheme (right, dased line) and B scheme (right, solid line). The abscissa tics are the intersection between the axis $y = 0.736$ and the mesh.

6 Numerical results for the Euler equation system.

Let us consider the steady Euler equations system for compressible fluid dynamics in conserved variables,

$$\operatorname{div} \mathcal{F}(U) = 0,\ U \in \mathbb{R}^4, \tag{31}$$

with $\mathcal{F} = (F, G)$,

$$U = (\rho, \rho u, \rho v, E)^T,\ F(U) = (\rho u, \rho u^2 + p, \rho uv, u(E+p))^T,$$
$$G(U) = (\rho v, \rho uv, \rho v^2 + p, v(E+p))^T,$$

where ρ is the gas density, u, v stand for the two components of the velocity, E is the total energy and p the gas pressure. The total enthalpy, H, and the speed of sound, c, are defined by $H = \dfrac{E+p}{\rho}$ and $c = \sqrt{\dfrac{\gamma p}{\rho}}$. These systems are closed with the equation of state

$$p = (\gamma - 1)\Big(E - \frac{1}{2}\rho u^2 + v^2\Big)\ ,\ \gamma = 1.4 \ \text{ for perfect air.}$$

For all $V \in \mathbb{R}^4$, one denotes V^i as the ith component of V. For physical reasons, we want $\rho > 0$, $p > 0$ (no vacuum), hence the natural definition domain for $\mathcal{F}$ is

$$\mathcal{O} = \Big\{(U^1, U^2, U^3, U^4) \in \mathbb{R}^4,\ U^1 > 0,\ U^4 > \frac{1}{2}\Big(\frac{(U^2)^2}{U^1} + \frac{(U^3)^2}{U^1}\Big)\Big\}\ ,$$

The function $\mathcal{F}$ is a $\mathcal{C}^1$ mapping, the system is hyperbolic in $\mathcal{O}$, symetrizable, and the eigenvalues and eigenvectors of A, B with respect to U are known (we refer the reader to [18], for example).

The primitive variables, P, and the Roe-Struijs-Deconink (RSD) variables, Z, are defined by

$$P = (\rho, u, v, p),\ Z = \sqrt{\rho}(1, u, v, H),$$

so that U can be expressed as $U(P)$ and $U(Z)$, with

$$U(P) = \Big(P^1, P^1P^2, P^1P^3, \frac{1}{\gamma-1}P^4 + \frac{1}{2}P^1(P^2P^2 + P^3P^3)\Big),$$
$$U(Z) = \Big(Z^1Z^1, Z^1Z^2, Z^1Z^3, \frac{1}{\gamma}\big(Z^1Z^4 + \frac{\gamma-1}{2}Z^2Z^3\big)\Big)\ .$$

The choice of the average state $\bar{U} = (\bar{\rho}, \bar{\rho}\bar{u}, \bar{\rho}\bar{v}, \bar{E})^T$ that defines matrices K_i is free, provided that they are diagonalizable with real eigenvalues. Three examples are used in this paper :

- The linearization on the conserved variables, $\bar{U} = \frac{1}{4}(U_1 + U_2 + U_3 + U_4)$, preserves the hyperbolicity since $\mathcal{O}$ is a convex subset of $\mathbb{R}^4$. For the computation of the Jacobians,

$$\bar{u} = \frac{\rho_1 u_1 + \rho_2 u_2 + \rho_3 u_3 + \rho_4 u_4}{\rho_1 + \rho_2 + \rho_3 + \rho_4},\ \bar{v} = \frac{\rho_1 v_1 + \rho_2 v_2 + \rho_3 v_3 + \rho_4 v_4}{\rho_1 + \rho_2 + \rho_3 + \rho_4},$$
$$\bar{p} = (\gamma - 1)\big(\bar{E} - \frac{1}{2}\bar{\rho}(\bar{u}^2 + \bar{v}^2)\big),\ \bar{H} = (\bar{E} + \bar{p})/\bar{\rho},\ \bar{c}^2 = \gamma\bar{p}/\bar{\rho}\ .$$

- The linearization of the primitive variables reads $\bar{U} = U(\bar{P})$, with $\bar{P} = \frac{1}{4}(P_1 + P_2 + P_3 + P_4)$. The average density, velocities and pressure, $\bar{\rho}$, $\bar{u}$, $\bar{v}$, $\bar{p}$, are obvious, and

$$\bar{H} = \frac{\gamma}{\bar{\rho}(\gamma-1)}\bar{p} + \frac{1}{2}(\bar{u}^2 + \bar{v}^2)\ ,\ \bar{c}^2 = \gamma\frac{\bar{p}}{\bar{\rho}}\ .$$

- The Roe-Struijs-Deconink (RSD) linearization is defined by $\bar{U} = U(\bar{Z})$, with $\bar{Z} = \frac{1}{4}(Z_1+Z_2+Z_3+Z_4)$, and then

$$\bar{\rho} = \left(\sqrt{\rho_1}+\sqrt{\rho_2}+\sqrt{\rho_3}+\sqrt{\rho_4}\right)^2,\ \bar{u} = \frac{\sqrt{\rho_1}u_1+\sqrt{\rho_2}u_2+\sqrt{\rho_3}u_3+\sqrt{\rho_4}u_4}{\sqrt{\rho_1}+\sqrt{\rho_2}+\sqrt{\rho_3}+\sqrt{\rho_4}},$$
$$\bar{v} = \frac{\sqrt{\rho_1}v_1+\sqrt{\rho_2}v_2+\sqrt{\rho_3}v_3+\sqrt{\rho_4}v_4}{\sqrt{\rho_1}+\sqrt{\rho_2}+\sqrt{\rho_3}+\sqrt{\rho_4}},\ \bar{H} = \frac{\sqrt{\rho_1}H_1+\sqrt{\rho_2}H_2+\sqrt{\rho_3}H_3+\sqrt{\rho_4}H_4}{\sqrt{\rho_1}+\sqrt{\rho_2}+\sqrt{\rho_3}+\sqrt{\rho_4}},$$
$$\bar{c}^2 = (\gamma-1)\left(\bar{H}-\frac{1}{2}(\bar{u}^2+\bar{v}^2)\right),\ \bar{p} = \frac{1}{\gamma}\bar{\rho}\,\bar{c}^2\ .$$

  The fact that the hyperbolicity of the system is preserved with this average state is a classical result (see for example [12, 1, 18]).

The choice of an approximation $\mathcal{F}^h$ is also free; we use three of them: the Q_1 interpolation of the flux, and the flux applied to the Q_1 interpolation of the primitive variables, and the RSD variables, namely,

$$\mathcal{F}^h(U_h) = \pi_h\mathcal{F}(U_h)\ ,\quad \mathcal{F}^h(U_h) = \mathcal{F}\big(U(\pi_h P_h)\big)\ ,\ \mathcal{F}^h(U_h) = \mathcal{F}\big(U(\pi_h Z_h)\big)\ ,$$

respectively. Here, P_h and Z_h denote the piece-wise constant functions of E_h constructed with the set of values $P_i^n := P(U_i^n)$, $Z_i^n := Z(U_i^n)$. The flux $\mathcal{F}$ is clearly a fourth degree polynomial with respect to P, and a second degree polynomial with respect to Z, which permits the calculation of the total fluctuation $\int_Q \operatorname{div}\mathcal{F}^h\,dx\,dy$ with Lemma 4.
Although the use of the RSD variables for both linearization and interpolation provides the conservative linearization (9) for triangle meshes endowed with the linear P_1 interpolation, the choice of this variable change is a priori not more advantageous than others, when using quadrilateral cells.
As mentioned above, for the construction of CN and LDA schemes, the matrix $N = \left(\sum_{i=1}^4 K_i^-\right)^{-1}$ may not exist in some cases. In [1] is shown that it can only occur when $\bar{u} = \bar{v} = 0$, and that if $\bar{u}^k, \bar{v}^k$, are two sequences converging to 0, then $K_i^+(\bar{U}^k)N(\bar{U}^k)$ converge, with $\bar{U}^k = (\bar{\rho}, \bar{\rho}\bar{u}^k, \bar{\rho}\bar{v}^k, \bar{E})^T$. This gives a meaning to CN and LDA schemes even when $\bar{u} = \bar{v} = 0$. From [1], a subtle way of computing K_j^+N can be put forward. Here, however, we have chosen a crude method : when $\bar{u} = \bar{v} = 0$, we decide to replace $\bar{U}$ by

$$\widetilde{\bar{U}} = (\bar{\rho}, \bar{\rho}\varepsilon, \bar{\rho}\varepsilon, \bar{E})^T,$$

where $\varepsilon > 0$ is chosen small enough so that $\widetilde{\bar{U}} \in \mathcal{O}$, that means $\bar{E} - \frac{1}{2}\bar{\rho}(\varepsilon^2+\varepsilon^2) > 0$. This allows to define N, and by construction, the CN and LDA schemes still satisfy the exact conservation of the total fluctuation Φ^Q on any quadrangle and their consistency is preserved.

The CN, LDA, B, and limited-CN schemes are tested in the three examples below. There is no clear rule to choose a basis to construct the limited-CN scheme in (29) and (30). We select in the first set of examples the set of eigenvectors of the matrix $\cos(\theta)A(\bar{U}) + \sin(\theta)B(\bar{U})$ with $\theta = \pi/2$. Another choice is the direction of the local velocity, we use this value in section 7 in order to point out some problems and to propose solutions. However, it can be (numericaly) checked that the non–oscillatory nature of the results is independant of the choice of θ. Our results confirm this claim.

Some other basis (for example the canonic basis of $\mathbb{R}^4$) can lead to code blow up for negative pressure. Although the convergence criterion for steady flow computations reached the computer zero for the CN scheme, this is not the case for both B and limited-CN schemes, for which the convergence is moreover slow and not monotonic. This lack of convergence for nonlinear system-schemes is common and was expected from previous studies ([10, 1, 15]). We also come back to this point in section 7. As for CFL condition, we use for all X_i, for all Q such that $C_i \cap Q \neq \emptyset$, $\Delta t \sum_{j=1}^4 \lambda_j \leq |C_i \cap Q|$, where λ_j is the maximum of the modulus of the eigenvalues of the matrix K_j defined for Q.

6.1 Steady shock tube.

The first computation for Euler's system is a supersonic steady flow in a shock tube. This test has been performed for example in [1]. The domain Ω is still the unit square, and the unsteady system in (31) is solved

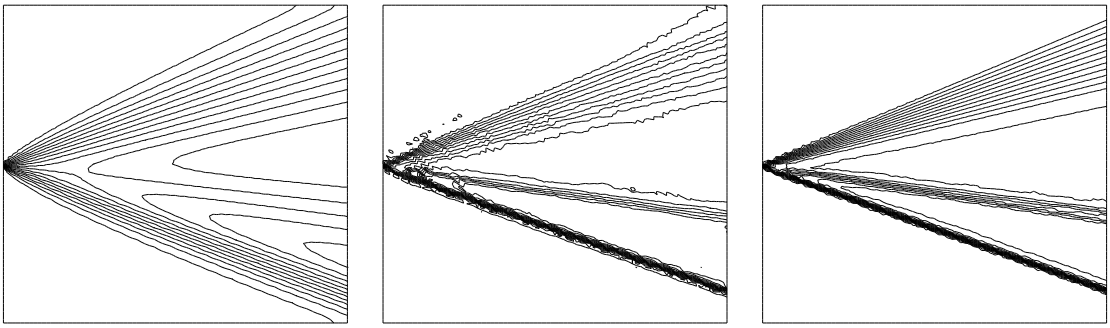


Figure 9: Steady shock tube: density isolines computed with linearization and interpolation of the primitive variables, for the CN (left, $0.7 \leq \rho \leq 1.4$), B (center, $0.6 \leq \rho \leq 1.47$) and limited-CN (right, $0.69 \leq \rho \leq 1.4$) schemes, using twenty contours between the extrema.

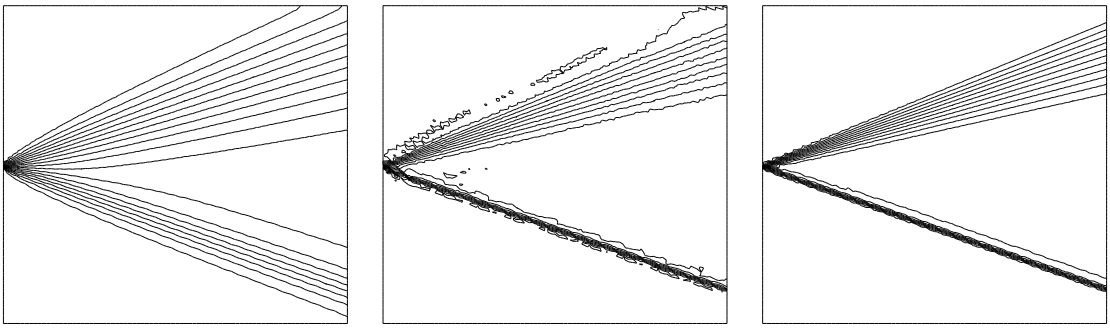


Figure 10: Steady shock tube: pressure isolines computed with linearization and interpolation of the RSD variables, for the CN (left, $0.25 \leq p \leq 1$), B (center, $0.148 \leq p \leq 1.03$) and limited-CN (right, $0.25 \leq p \leq 1$) schemes, using twenty contours between the extrema.

until convergence, from initial conditions $\rho(x, y) = \gamma/2$, $p(x, y) = 0.25$, Mach number$(x, y) = 4$, $v = 0$, for $0 \leq y \leq \frac{1}{2}$, and $\rho = \gamma$, $p = 1$, Mach number$= 2.4$, $v = 0$, for $\frac{1}{2} \leq y \leq 1$, for $0 \leq x \leq 1$. Boundary conditions are the same as initial conditions at $x = 0$, and nothing is imposed elsewhere. The solution of this problem contains a shock wave, a contact discontinuity and a fan, thus constituting a relevant test for the schemes. The mesh has the same appearance than the one in Figure 4, but with 6400 elements, instead of 3600.

This test has been performed with all possible combinations of the above average states and approximations of the flux. The various numerical results obtained were very close together. Sometimes the convergence histories, or the necessary CFL constraint were a bit different from one to another, oscillations and extrema of LDA and B schemes were not exactly the same, but yet we were not able to define a clear rule to choose the best average state and best interpolation of $\mathcal{F}^h$.

Numerical densities computed with the linearization and interpolation of the primitive variables are plotted in Figure 9, and pressure contours computed with linearization and interpolation of the RSD variables are drawn in Figure 10. Figure 11 presents a cross section of these two quantities.

The numerical results favour the limited-CN scheme that resolves the three waves accurately, while the CN scheme is numerically monotonic. The B scheme does not look monotonic in this test. The CN scheme is very robust contrary to the B scheme that was performed after using one thousand iterations of the CN scheme to obtain the above results. The LDA scheme was too oscillatory and consequently is not represented. The convergence with the l^∞ norm stagnated at the order of 10^{-6} for the limited-CN scheme, 10^{-5} (and 10^{-6} with L^1 norm) for the LDA scheme, and 10^{-4} (and 10^{-6} with the L^1 norm) for the B scheme.

To illustrate the independence of the choice of linearization and interpolation, Figure 12 draws Mach number contours obtained for three variant of the limited-CN scheme.

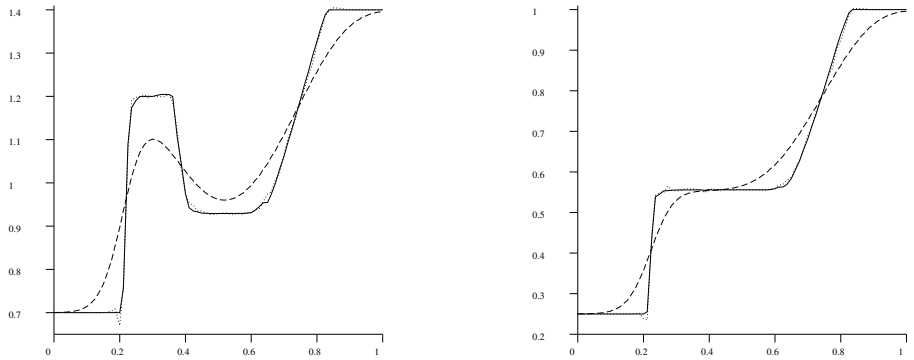


Figure 11: Steady shock tube: cross section along the line $x = 0.7$ for the density (left), and the pressure (right), using the CN scheme (dashed line), the B scheme (dotted line), and limited-CN scheme (solid line).

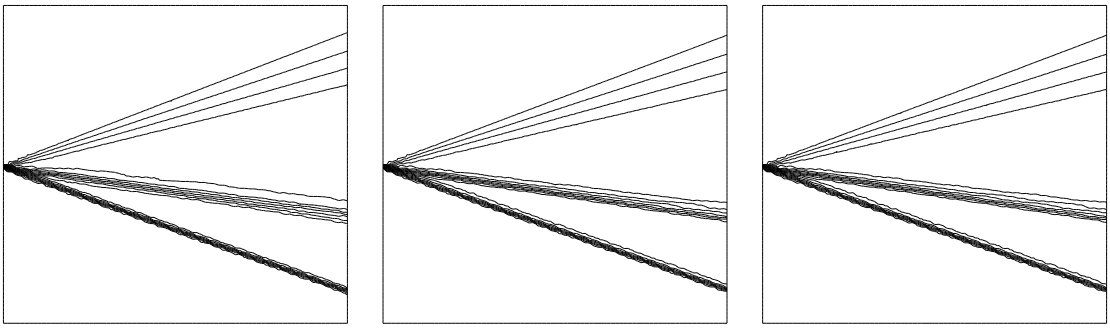


Figure 12: Steady shock tube: Mach number ($2.839 \leq \text{Mach} \leq 4.73$) computed with the limited-CN scheme and linearization and interpolation of the primitive variables (left), linearization and interpolation of the RSD variables (center), linearization of the primitive variables and interpolation of the flux (right). Twenty contours between the extrema are plotted.

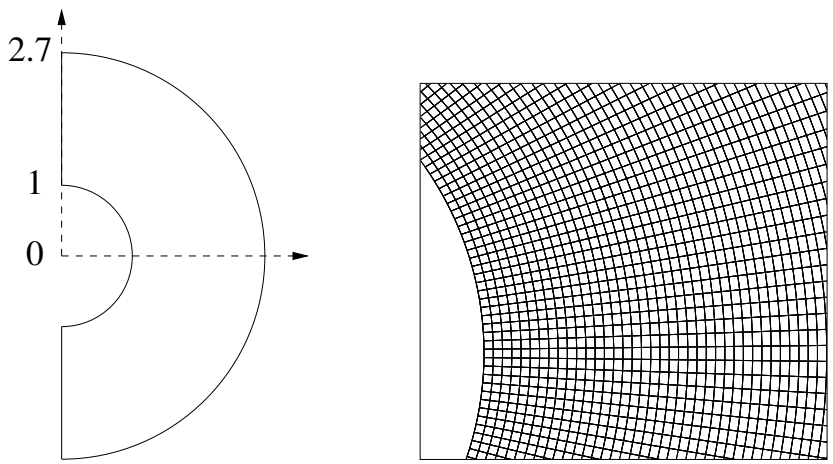


Figure 13: Domain and zoom of the mesh (8840 elements) for the Mach 15 motion of a cylinder problem.

6.2 Hypersonic flow over a cylinder (Mach$_\infty$=15).

Consider a cylinder of radius unity that is moving with a Mach number equal to 15. The domain and regular mesh are drawn in Figure 13, and the following boundary conditions are imposed : inflow condition at infinity, with $M_\infty = 15$, $v_\infty = 0$, $\rho_\infty = 1.22499903071317$, $p_\infty = 1.1325 \times 10^5$, wall conditions on the cylinder, and nothing is imposed at the vertical left boundary of the domain.

This test could only be handled using the CN and limited-CN schemes. The other schemes gave negative pressure that blew up the code. Furthermore, the choice of the average state and interpolation procedure was crucial. With this particular form of mesh, computations could only be performed with the average of the conservative variables and the Q_1 interpolation of the flux. Others also blew up. The bow shock around the cylinder is observed in Figure 14, for the pressure field, p, and the tangential velocity, u. The convergence criterion for the limited-CN scheme could not be decreased under the order of magnitude 10^{-5} with the l^∞ norm. Isolines for the limited-CN scheme show small wiggles. In [15], the authors believe this comes from the fact that the limiter (28) (30) is too compressive. The advantage of the limited scheme is that the shock is better captured than the first order accurate scheme, particularly at the top of the cylinder. These remarks are illuminated in Figure 15.

7 Additional remarks on the stabilisation of the RD schemes.

We come back to the small wiggles that can be observed in Figure 7 but also on Figures 9 and 10 if one looks very carefully in the fan. A carefull observation of this figure shows that the shocks, and more generally the discontinuities, are perfectly resolved while the smooth part of the solution exibits sometimes wiggles. This is a typical behavior of RD schemes, which is analysed in [19] for triangular meshes.

In this paper, it is shown that if one starts from a monotone first order schemes (for example the Residual based scheme constructed from the Lax Friedrichs schemes using the construction of section 4.2), which accuracy is upgraded following the procedure of section 4.6, one obtains a scheme that is l^∞ stable (by construction), which resolves perfectly (in one cell) discontinuities and does a very disapointing job on the smooth part of the solution. This is exactly the opposite situation as for standard high order scheme for which the smooth part of the solution is easier to compute than its non–smooth parts. This has nothing to do with the non–linear nature of the problem, since for the very simple example of a solid advection, this phenomena already occurs. An example is provided in Figure 16 where we show a cross–section of the

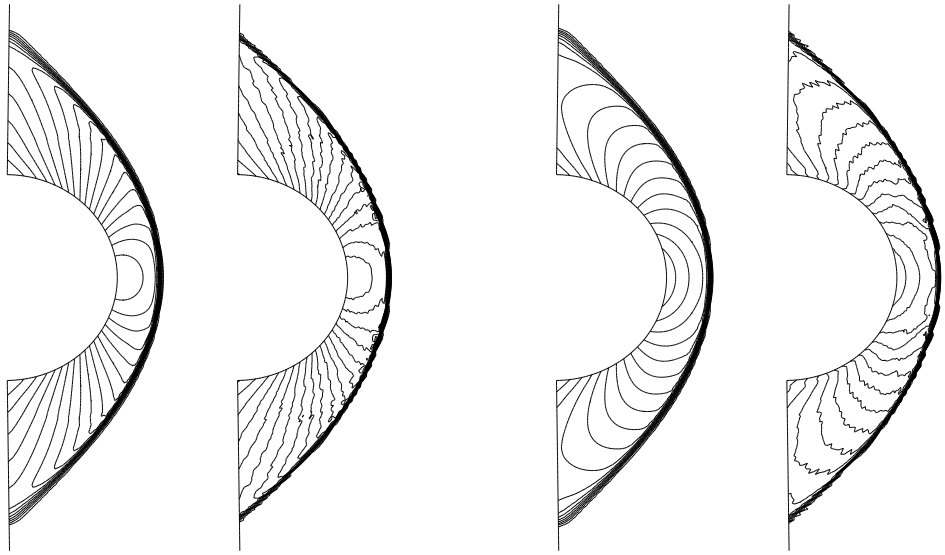


Figure 14: Flow over a cylinder : from left to right, pressure contour, using twenty isovalues between $1013.25.10^5$ and $2.9.10^7$, for the CN scheme, for the limited-CN scheme, and tangential velocity contour, using twenty isovalues between -5148 and -34.9, for the CN scheme, for the limited-CN scheme.

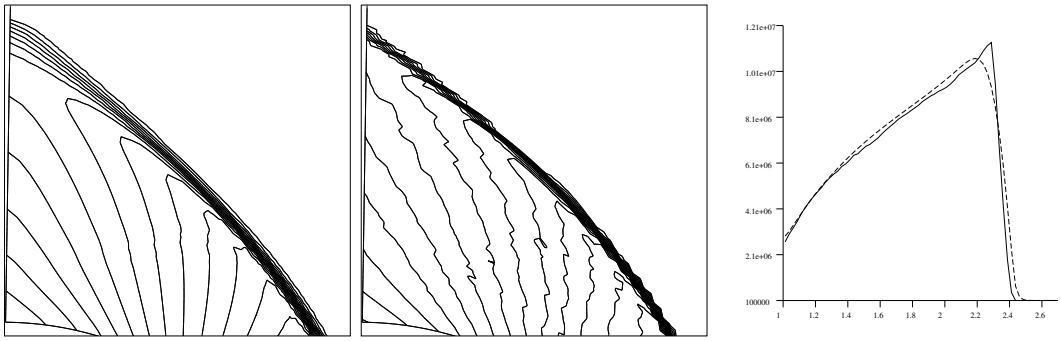
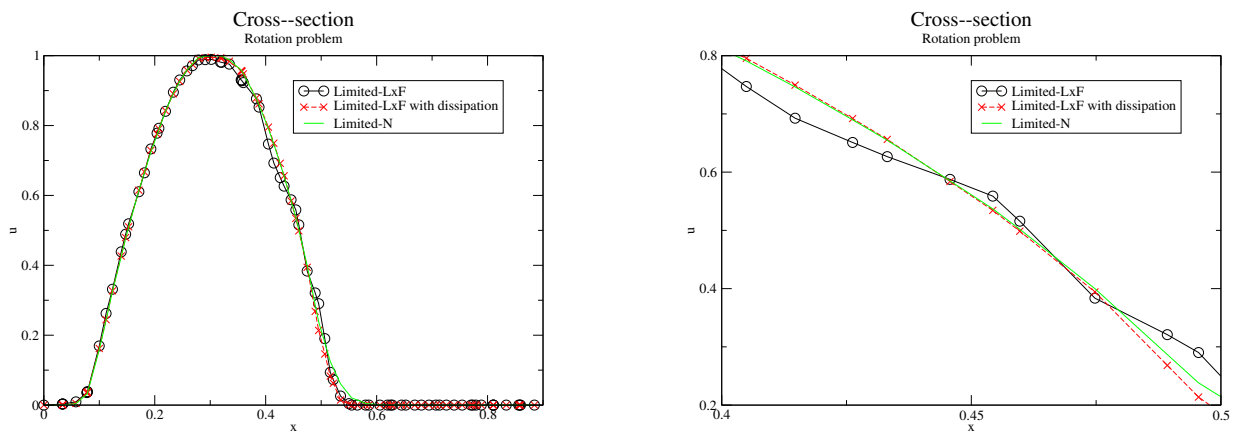


Figure 15: Flow over a cylinder: zoom of pressure contours in Figure 14 for the CN scheme (left), and the limited-CN scheme (center); cross section of the shock at the top of the cylinder (right) for the CN scheme (dashed line), and the limited-CN scheme (solid line).

solution after half a rotation. The local velocity at point (x, y) is $(-y, x)$, and the exact solution is

$$U(0,x) = \begin{cases} -\sin\left(\pi \dfrac{x-0.7}{0.6}\right) & \text{if } x \in [0.1, 0.7] \\ 0 & \text{else.} \end{cases}$$

One can see that the solution constructed from the N scheme is good, while the one constructed from the



zoom in $[0.4, 0.5]$

Figure 16: Cross–section for a rotation problem for a triangular mesh. The second order solution constructed on the N scheme is plotted with plain lines, the one constructed from the Lax–Freidrichs scheme with circles. The Limited LxF scheme with the additional dissipation (33) is shown with crosses.

Lax–Friedrichs scheme has a stair–case like behavior.

7.1 Modification of the limited scheme

In [19], a solution to this problem is given. Consider the scalar problem

$$\operatorname{div} \mathcal{F}(U) = 0 \ , \tag{32}$$

supplemented by inflow conditions. Let T denote a generic triangle and for simplicity, we denote its vertices by their indices. Starting from a monotone scheme which residual in T are Φ_1^T, Φ_2^T and Φ_3^T, we construct a second order scheme by setting

$$\Phi_j^{T,\star} = \beta_j^{T,\star}\, \Phi^T$$

where Φ^T is the total residual in T and (for example)

$$\beta_j^{T,\star} = \frac{\beta_j^+}{\displaystyle\sum_{k=1}^{3} \beta_k^+}$$

where

$$\beta_j = \frac{\Phi_j^T}{\Phi^T}.$$

This scheme is L^∞ stable by construction but exhibits small spurious oscillations. For a steady problem, the iterative convergence of the method may be very poor, as we have seen here in section 5. However, it can also be observed that when the solution presents a discontinuity, the scheme (32) construct very good solution. Last, the iterative convergence problem occurs only because of the smooth regions. We refer to [19] for detailled calculations.

For the problem (32), these oscillations can be removed, and the iterative convergence problem be solved, if we modify the limited residual by a least square modification, namely

$$\Phi_j^{T,\star\star} = \Phi_j^{T,\star} + \Theta(U_h) h_T \int_T \Big(\nabla\mathcal{F}(\pi_h^T U_h)\cdot\nabla\varphi_j\Big)\Big(\nabla\mathcal{F}(\pi_h^T U_h)\cdot\nabla\pi_h^T U_h\Big)dx. \tag{33}$$

In (33), φ_j is the P_1 shape function at M_j. The parameter h_T is set to $\sqrt{|T|}$ and $\Theta(U_h)$ is a function such that $\Theta(U_h)\simeq 1$ when the solution is smooth and $\Theta(U_h)\simeq 0$ else. In practice, we choose

$$\Theta(U_h) = 1 - \frac{\Big|\max_{j\in T} U_j - \min_{j\in T} U_j\Big|}{\Big|\max_{j\in T} U_j\Big| + \Big|\min_{j\in T} U_j\Big| + \varepsilon} \tag{34}$$

where ε is a small number (of the order of 10^{-10}). Once more, we refer to [19] for more detail.

In the case of quadrangles, it is simple to adapt this idea. In equations (33) and (34), we just change the triangle T by the quadrangle Q, while the residual $\Phi_j^{Q,\star}$ is the limited residual described in section 4.6. In (33), the interpolant $\pi_h^T U_h$ is replaced by the Q_1 interpolant, so that the integral maybe complex to evaluate. A variant, which is the one we have implemented in the numerical tests, is

1. to cut the quadrangle Q into two triangles T_Q^1 and T_Q^2 (there are two possible choices, and our numerical results are independant of that choice),
2. replace the integral by

$$\int_{T_Q^1}\Big(\nabla\mathcal{F}\big(\pi_h^{T_Q^1}U_h\big)\cdot\nabla\varphi_j\Big)\Big(\nabla\mathcal{F}\big(\pi_h^{T_Q^1}U_h\big)\cdot\nabla\pi_h^{T_Q^1}U_h\Big)dx + \int_{T_Q^2}\Big(\nabla\mathcal{F}\big(\pi_h^{T_Q^2}U_h\big)\cdot\nabla\varphi_j\Big)\Big(\nabla\mathcal{F}\big(\pi_h^{T_Q^2}U_h\big)\cdot\nabla\pi_h^{T_Q^2}U_h\Big)dx. \tag{35}$$

When dealing with systems, the limited scheme is constructed as above. The only modification is about the function Θ. In that case, we have chosen to define Θ as the diagonal matrix which i–th entrie is defined as in (34) when U_h is replaced but the i–th component of the conserved variable. Other choices are possible, we have experimented very few sensitivity to the actual choice.

7.2 Numerical experiments

We re–do the experiment on the Burger equation and the Euler ones (jet case). In each case, the results obtained with the modified limited scheme are obtained by the same iterative scheme, this time the results are all converged. We have considered two versions of the first order scheme. The first one is the CN scheme above, the second one is a Lax–Freidrichs (LxF(type of scheme,

$$\Phi_i^Q = \frac{1}{4}\Big(\Phi^Q + \alpha\sum_{j=1}^{4}(U_i - U_j)\Big). \tag{36}$$

The parameter α is chosen as

$$\alpha \geq \max_{j=1,4}|K_j|$$

in the scalar case and

$$\alpha \geq \max_{j=1,4} \rho(K_j)$$

in the system case[2]. The rest follows as above.

We first present the isolines of the solution, Figure 17. We can see that the wiggles have totaly disapeared, and more astonishing maybe, the results from the LxF scheme and the CN scheme, in their second order version, are very similar.

This is confirmed by the cross–sections along the discontinuity ($y = 0.75$, Figure 18) and in the fan ($y = 0.25$, Figure 19).

Then, we move to the jet problem for the Euler equations. We first present the density and the pressure isolines, figures 20 and 21. We also provide the cross–sections of the density (Figure 22) and the pressure (Figure 23). These Figures shows that the choice of the angle θ in the limitation (here $\cos\theta, \sin\theta) = \frac{(u,v)}{\sqrt{u^2+v^2}}$ instead of $\theta = \frac{\pi}{2}$ in section 6) does not play a role in the flow discontinuity, but has a clear influence on the smooth part (compare Figures 20 and 9 for example). These figures show the very clear improvement of adding a dissipation to remove the spurious oscillations that are consequence of the over–compressive nature of the limitation technique. This is important because since we do not know the best choice of θ, we need to have a robust method for increasing the order of accuracy.

Last, we note a very small undershoot/overshoot of the density and pressure in the new scheme. These are very small, as it can be seen from Table 2. In fact standard schemes also present such "oscillations".

Density

Scheme	LxF	Limited–LxF	Limited–LxF+dissipation	CN	Limited–CN	Limited–CN+disspiation
Min	0.7	0.7	0.688	0.7	0.7	0.683
Max	1.4	1.4	1.406	1.4	1.4	1.409

Pressure

Scheme	LxF	Limited–LxF	Limited–LxF+dissipation	CN	Limited–CN	Limited–CN+disspiation
Min	0.25	0.25	0.243	0.25	0.25	0.24
Max	1	1.	1.006	1.	1.	1.008

Table 2: Minimum and maximum of the density and the pressure for the different schemes.

8 Conclusion

We have constructed stable and accurate residual distribution schemes on quadrilateral meshes for steady problems. The numerical tests of the previous sections demonstrate, in our opinion, the correctness of our approach. Our construction starts from a quasi–monotone first order scheme. We show how to upgrade the accuracy, and propose a solution to some of the numerical problems we have to face, in particular to construct stable *and* accurate schemes. Two classes of schemes are considered in this paper. One relies on an extension of the N scheme following the ideas of [10], the other one starts from a Lax Friedrichs like scheme. We show numericaly that the second order schemes give very similar results for boths classes.

This paper illustrates above all the underlying flexibility of the construction of residual distribution schemes, based on the mesh points, and not on the geometry of cells, contrary to finite volume.

Unsteady problems can be handled with the same techniques combined with the ideas of [6, 19].

[2] $\rho(A)$ denotes the spectral radius of the matrix A.

Acknowledgements

The second author would like to thanks Pr. Th. Sonar for the kind hospitality he offered at the Institut for Computational mathematics in TU Braunschweig, Germany. This has enabled to finish the paper.

A Review of the Q_1 interpolation

Polynomial interpolation is the topic of many books. The reader is referred to [20], for example. We recall the basic statements of the first order interpolation for general quadrilaterals. Let $\hat{Q}$ be the quadrangle of reference $\big(\hat{X}_1(0,0)$,$\hat{X}_2(1,0)$,$\hat{X}_3(1,1)$,$hat X_4(0,1)\big)$. The bilinear Q_1 interpolation functions on $\hat{Q}$ are defined by

$$\hat{\phi}_1(\hat{x},\hat{y}) = (1-\hat{x})(1-\hat{y}),\ \hat{\phi}_2(\hat{x},\hat{y}) = \hat{x}(1-\hat{y}),\ \hat{\phi}_3(\hat{x},\hat{y}) = \hat{x}\hat{y},\ \hat{\phi}_4(\hat{x},\hat{y}) = (1-\hat{x})\hat{y}\ .$$

Next, let Q be any quadrilateral the vertices of which are $X_1(x_1,y_1)$, $X_2(x_2,y_2)$, $X_3(x_3,y_3)$, $X_4(x_4,y_4)$ where the points X_i are ordered as follows:

$$y_1 \le y_4,\ y_2 \le y_4,\ y_1 \le y_3,\ x_4 \le x_3,\ y_2 \le y_3,\ x_1 \le x_2\ .$$

It is well known there exists an invertible bilinear polynomial σ_Q that maps $\hat{Q}$ into Q such that $\sigma_Q(\hat{X}_i) = X_i$,$i = 1,\dots,4$.This mapping, σ_Q, is defined by

$$\sigma_Q\begin{pmatrix}\hat{x}\\ \hat{y}\end{pmatrix} = \begin{pmatrix} a_1\hat{x}\hat{y} + b_1\hat{x} + c_1\hat{y} + d_1 \\ a_2\hat{x}\hat{y} + b_2\hat{x} + c_2\hat{y} + d_2\end{pmatrix}\ ,$$

and has to verify $\sigma_Q(\hat{X}_i) = X_i$, thus

$$\begin{array}{c} a_1 = x_3 + x_1 - x_2 - x_4,\ a_2 = y_3 + y_1 - y_2 - y_4,\ b_1 = x_2 - x_1, \\ b_2 = y_2 - y_1,\ c_1 = x_4 - x_1,\ c_2 = y_4 - y_1,\ d_1 = x_1,\ d_2 = y_1\ . \end{array} \tag{37}$$

The Q_1 interpolation basis functions on Q are $\phi_i^Q = \hat{\phi}_i \circ \sigma_Q^{-1}$ for $i = 1,\dots,4$. Then, letting $\mathcal{Q}$ denote a mesh, the Q_1 interpolation of any function f in $\mathcal{Q}$ is the continuous mapping in $\mathcal{Q}$ defined for all $Q \in \mathcal{Q}$ by

$$\pi_h f(x,y) = \sum_{i=1}^{4} f_i \phi_i^Q(x,y),\ (x,y) \in Q, \quad \text{with } f_i = f(X_i)\ .$$

Note that the superscript Q in ϕ_i^Q is often omitted. If the function f also depends on time, we let $\pi_h f$ also denote the piecewice constant function in time having the value $\pi_h f(t_n,.,.)$ in each sub-interval $[t_n, t_{n+1}[$.

B The schemes in section 4 verify (13)–(17)

The aim of this appendix is to show the consistency given by Theorem 1 holds with the schemes in paragraph 4. One shall begin our proof by showing the approximations $\mathcal{F}^h$ in subsection 4.1 satisfy (15)–(17). Two technical lemmas are needed.

Lemma 2. *Let Q be a quadrilateral the diagonals of which are bounded by h. There exists a constant $C > 0$ independent of h such that for all function W defined from Q to $\mathbb{R}$,*

$$\int_Q |\nabla \pi_h W|\, dx\, dy \le Ch \sum_{i,j=1}^{4} |W_i - W_j|\ .$$

Proof. Define the variable change $(x,y)=\sigma_Q(\hat{x},\hat{y})$, and denote $\hat{\nabla}$ as the gradient operator with respect to $(\hat{x},\hat{y})$ in $\hat{Q}$. For all $(x,y)\in Q$, $i=1,\dots,4$,

$$\nabla\phi_i(x,y)=\hat{\nabla}\big(\phi_i\circ\sigma_Q\big)(\hat{\nabla}\sigma_Q)^{-1}(\hat{x},\hat{y})=\frac{1}{\det\big(\hat{\nabla}\sigma_Q(\hat{x},\hat{y})\big)}\hat{\nabla}\hat{\phi}_i\ J(\hat{x},\hat{y}),$$

where according to (37),

$$J(\hat{x},\hat{y})=\begin{pmatrix} a_2\hat{x}+c_2 & -a_1\hat{x}-c_1 \\ -a_2\hat{y}-b_2 & a_1\hat{y}+b_1 \end{pmatrix}\ .$$

Then integrate $|\nabla\pi_h W|$ over Q to get

$$\int_Q|\nabla\pi_h W|\,dx\,dy=\int_{\hat{Q}}\Big|\Big(\sum_{i=1}^4 W_j\nabla\hat{\phi}_i\Big)\ J(\hat{x},\hat{y})\Big|\,d\hat{x}\,d\hat{y}\ . \tag{38}$$

Next, since $\hat{\nabla}\hat{\phi}_1=(\hat{y}-1,\hat{x}-1)$, $\hat{\nabla}\hat{\phi}_2=(1-\hat{y},-\hat{x})$, $\hat{\nabla}\hat{\phi}_3=(\hat{y},\hat{x})$, $\hat{\nabla}\hat{\phi}_4=(-\hat{y},1-\hat{x})$, there holds

$$\sum_{i=1}^4 W_i\hat{\nabla}\hat{\phi}_i(\hat{x},\hat{y})=(W_1-W_2)\big(\hat{y}-1,\hat{x}\big)+(W_3-W_4)\big(\hat{y},\hat{x}\big)+(W_4-W_1)\big(0,1\big)\ ,$$

so that

$$\Big|\sum_{i=1}^4 W_j\nabla\hat{\phi}_i(\hat{x},\hat{y})\Big|\le 2\sum_{i,j=1}^4|W_i-W_j|\ . \tag{39}$$

Returning to the definition of a_1, a_2, b_1, b_2, c_1, c_2 in (37), the fact that the diagonals of Q are bounded by h yields $\big\|J(\hat{x},\hat{y})\big\|_{\mathcal{M}_2(\mathbb{R})}\le C'h$ in $\hat{Q}$ where C' is independent of h, for any norm in $\mathcal{M}_2(\mathbb{R})$. This, together with (38) and (39), concludes on the existence of our desired constant C. □

Lemma 3. *Let $(V_h)_h$ be a family of functions in E_h. We assume that there exists a compact $\mathcal{K}\subset\mathbb{R}^m$ and $V\in\big(L^2_{loc}(\mathbb{R}^+\times\mathbb{R}^2)\big)^m$ such that :*

$$V_h(t,x,y)\in\mathcal{K},\ \textit{for}\ (t,x,y)\in\mathbb{R}_+\times\mathbb{R}^2,\quad \textit{and}\quad \lim_{h,\Delta t\to 0}\|V_h-V\|_{\left(L^2_{loc}(\mathbb{R}^+\times\mathbb{R}^2)\right)^m}=0\ .$$

Let be $p\in\mathbb{N}$, $q\in\mathbb{N}$, S a $\mathcal{C}^1$ mapping from $\mathbb{R}^m$ to $\mathbb{R}^p$, and R a $\mathcal{C}^1$ mapping from $\mathbb{R}^p$ to $\mathbb{R}^q$. Then

$$R\big(\pi_h S(V_h)\big)-R\big(S(V_h)\big)\xrightarrow[h,\Delta t\to 0]{}0\ \textit{in}\ \big(L^1_{loc}(\mathbb{R}^+\times\mathbb{R}^2)\big)^q\ .$$

Proof. Let $K=[0,T]\times\omega$ be a compact subset in $\mathbb{R}^+\times\mathbb{R}^2$. Let $\bar{h}>0$, $\bar{\omega}$ a compact in $\mathbb{R}^2$ be such that $\omega\subset\bar{\omega}$, and the distance between $\partial\omega$ and $\partial\bar{\omega}$ is lower than $\bar{h}$. Let be $\bar{\Delta t}>0$, $\bar{T}=T+\bar{\Delta t}$. From now, we assume $h<\bar{h}$, $\Delta t<\bar{\Delta t}$. From the regularity of R and the use of the finite growth theorem,

$$\Big\|R\big(\pi_h S(V_h)\big)-R\big(S(V_h)\big)\Big\|_{L^1(K)}\le C\sum_{n=0}^N\Delta t_n\sum_{Q\subset\bar{\omega}}\int_Q\big|\pi_h S(V_h)-S(V_h)\big|\,dx\,dy\ , \tag{40}$$

where $C>0$ is independent of h, Δt, and $N=\max\limits_{t_n\in[0,\bar{T}]}(n)$. Because for all Q, for all $(x,y)\in Q$, $\sum_{i=1}^4\phi_i(x,y)=1$ and $|\phi_i(x,y)|\le 1$,

$$\big|\pi_h S(V_h)-S(V_h)\big|=\Big|\sum_{i=1}^4\big(S_i-S(V_h)\big)\phi_i\Big|\le C\sum_{i=1}^4|V_i-V_h|\ \text{in}\ Q\ ,$$

where C is another constant coming from the finite growth theorem. Integrating this result over a quadrangle Q,

$$\int_Q \big|\pi_h S(V_h) - S(V_h)\big|\, dx\, dy \le C \sum_{j=1}^{4} \int_{C_j \cap Q} \sum_{i=1}^{4} |V_i - V_j|\, dx\, dy \le C' |Q| \sum_{i=1,j}^{4} |V_i - V_j|,$$

and therefore, returning to (40), one has

$$\Big\| R\big(\pi_h S(V_h)\big) - R\big(S(V_h)\big) \Big\|_{L^1(K)} \le C' \sum_{n=0}^{N} \Delta t_n \sum_{Q \subset \bar{\omega}} |Q| \sum_{i,j=1}^{4} |V_i - V_j| \ .$$

¿From the convergence of V_h towards V, we may adapt arguments in [9, 11] to allow the right-hand side of the above inequality tend to 0 and conclude the proof. □

Let $(V_h)_h$, $V_h \in E_h$, be a family of functions converging in $L^2_{\rm loc}(\mathbb{R}^+ \times \mathbb{R}^2)^m$ and such that $V_h(t,x,y) \in \mathcal{K} \subset \mathbb{R}^m$, $\mathcal{K}$ being compact.
Assume the hypotheses of Theorem 1 are satisfied. Towards the end, one denotes W^k as the $k^{\rm th}$ component of any function W, while W_i is still denoting the value of W at location X_i. Let $\mathcal{F}^h$ be defined as the interpolation of the flux function (19). According to Lemma 2, one has

$$\begin{aligned} \int_Q \big|\operatorname{div} \mathcal{F}^h(V_h)\big|\, dx\, dy &\le \sum_{k=1}^{m} \int_Q |\nabla F^{h,k}| + |\nabla G^{h,k}|\, dx\, dy \\ &\le Ch \sum_{i,j=1}^{4} \Big(\big|F(V_i) - F(V_j)\big| + \big|G(V_i) - G(V_j)\big| \Big) \end{aligned}$$

and one concludes from the finite growth theorem that $\mathcal{F}^h$ satisfies (17). The convergence (15) is a direct consequence of Lemma 3 using $S = \mathcal{F}$ and $R = I_d(\mathbb{R}^{2\times m})$.
Next, define $\mathcal{F}^h = \mathcal{F}\big(U(\pi_h Z_h)\big)$, where $U \mapsto Z$ is a $\mathcal{C}^1$ variable change (subsection 4.1.2). Using the fact that V_h is uniformly bounded and Lemma 2 successively,

$$\int_Q \Big|\operatorname{div} \mathcal{F}\big(U(\pi_h Z_h)\big)\Big|\, dx\, dy \le C \sum_{k=1}^{m} \int_Q \big|\nabla \pi_h Z^{h,k}\big|\, dx\, dy \le C' h \sum_{i,j=1}^{4} |Z_i - Z_j| \ ,$$

and (17) follows from the finite growth theorem. Another use of Lemma 3, with $S = Z$ and $R = \mathcal{F}$ leads to (15).
Finally, the fact that these two different approximations $\mathcal{F}^h$ verify (16) is a consequence of the continuity of any interpolated function.

Now let us check whether both (13) and (14) hold for all the schemes in section 4. Consider first the finite volume scheme (23) with the geometry of Figure 3, and define $\mathcal{F}^h$ as the Q_1 interpolation of the flux function , (19). One has

$$|\Phi_i^{Q,\rm FV}| \le |\Gamma^1_{ij}| \, \big|\mathcal{F}(U_i, U_j, \vec{n}^1_{ij}) - \mathcal{F}(U_i, U_i, \vec{n}^1_{ij})\big| + |\Gamma^2_{ik}| \, \big|\mathcal{F}(U_i, U_k, \vec{n}^2_{ik}) + \mathcal{F}(U_i, U_i, \vec{n}^2_{ik})\big|,$$

and (13) follows since $|\Gamma^1_{ij}| \le h$, $|\Gamma^2_{ik}| \le h$, the numerical flux is locally Lipschitz continuous, and U_h is

uniformly bounded. Denoting p, q, l, m as the vertices of a quadrangle Q numbered in the direct order,

$$\begin{aligned}\sum_{i, X_i \in Q} \Phi_i^{Q,\mathrm{FV}} \quad = \quad & |\Gamma^1_{p,m}|\, \mathcal{F}(U_p, U_m, \vec{n}^1_{p,m}) + |\Gamma^2_{p,q}|\, \mathcal{F}(U_p, U_q, \vec{n}^2_{p,q}) + |\Gamma^1_{q,p}|\, \mathcal{F}(U_q, U_p, \vec{n}^1_{q,p}) \\ & + |\Gamma^2_{q,l}|\, \mathcal{F}(U_q, U_l, \vec{n}^2_{q,l}) + |\Gamma^1_{l,q}|\, \mathcal{F}(U_l, U_q, \vec{n}^1_{l,q}) + |\Gamma^2_{l,m}|\, \mathcal{F}(U_l, U_m, \vec{n}^2_{l,m}) \\ & + |\Gamma^1_{m,l}|\, \mathcal{F}(U_m, U_l, \vec{n}^1_{m,l}) + |\Gamma^2_{m,p}|\, \mathcal{F}(U_m, U_p, \vec{n}^2_{m,p}) - |\Gamma^1_{p,m}|\, \mathcal{F}(U_p, U_p, \vec{n}^1_{p,m}) \\ & - |\Gamma^2_{p,q}|\, \mathcal{F}(U_p, U_p, \vec{n}^2_{p,q}) - |\Gamma^1_{q,p}|\, \mathcal{F}(U_q, U_q, \vec{n}^1_{q,p}) - |\Gamma^2_{q,l}|\, \mathcal{F}(U_q, U_q, \vec{n}^2_{q,l}) \\ & - |\Gamma^1_{l,q}|\, \mathcal{F}(U_l, U_l, \vec{n}^1_{l,q}) - |\Gamma^2_{l,m}|\, \mathcal{F}(U_l, U_l, \vec{n}^2_{l,m}) - |\Gamma^1_{m,l}|\, \mathcal{F}(U_m, U_m, \vec{n}^1_{m,l}) \\ & - |\Gamma^2_{mp}|\, \mathcal{F}(U_m, U_m, \vec{n}^2_{mp}) \ .\end{aligned}$$

Since for all a, b, $\vec{n}^1_{a,b} = -\vec{n}^2_{b,a}$, and $|\Gamma^1_{a,b}| = |\Gamma^2_{b,a}|$ the eight first terms of the right-hand side of the above inequality cancel by using the conservation (21). Then one may use the consistency (22) on the other terms to reduce the script as

$$\begin{aligned}\sum_{i=1}^{4} \Phi_i^{Q,\mathrm{FV}} = & -\mathcal{F}(U_p) \cdot \big(|\Gamma^1_{p,m}|\vec{n}^1_{p,m} + |\Gamma^2_{p,q}|\vec{n}^2_{p,q}\big) - \mathcal{F}(U_q) \cdot \big(|\Gamma^1_{q,p}|\vec{n}^1_{q,p} + |\Gamma^2_{q,l}|\vec{n}^2_{q,l}\big) \\ & - \mathcal{F}(U_l) \cdot \big(|\Gamma^1_{l,q}|\vec{n}^1_{l,q} + |\Gamma^2_{l,m}|\vec{n}^2_{l,m}\big) - \mathcal{F}(U_m) \cdot \big(|\Gamma^1_{m,l}|\vec{n}^1_{m,l} + |\Gamma^2_{m,p}|\vec{n}^2_{m,p}\big) \ .\end{aligned}$$

Finally, remark that $|\Gamma^1_{p,m}|\vec{n}^1_{p,m} + |\Gamma^2_{p,q}| \cdot \vec{n}^2_{p,q} = \frac{1}{2}\vec{n}_p$, $|\Gamma^1_{q,p}|\vec{n}^1_{q,p} + |\Gamma^2_{q,l}|\vec{n}^2_{q,l} = \frac{1}{2}\vec{n}_q$, $|\Gamma^1_{l,q}|\vec{n}^1_{l,q} + |\Gamma^2_{l,m}|\vec{n}^2_{l,m} = \frac{1}{2}\vec{n}_l$, $|\Gamma^1_{m,l}|\vec{n}^1_{m,l} + |\Gamma^2_{m,p}|\vec{n}^2_{m,p} = \frac{1}{2}\vec{n}_m$, and then transform the right-hand side of the above equation in $\frac{1}{2}\sum_{i=1}^4 \mathcal{F}(U_i) \cdot \vec{n}_i$, that is $\int_Q \operatorname{div} \mathcal{F}^h(U_h)\, dx\, dy$ (Lemma 1).

For the other schemes, we recall $\mathcal{F}^h$ has to be predefined, so that the conservation of the total fluctuation is obviously satisfied, and it only remains to check (13). Moreover, since $U_h(t, x, y)$ is supposed to belong to a compact, there exists a constant $C > 0$ independent of U_h and h such that

$$\big\|A(\bar{U})\big\|_{\mathcal{M}_m(\mathbb{R})} + \big\|B(\bar{U})\big\|_{\mathcal{M}_m(\mathbb{R})} + \|K_i^+ N\|_{\mathcal{M}_m(\mathbb{R})} \le C \ , \ \|K_i^- \|_{\mathcal{M}_m(\mathbb{R})} \le Ch \ . \tag{41}$$

Consider now the conservation based N scheme (25), and write it as a perturbation of the original N scheme (8), as

$$\Phi_i^{Q,\mathrm{CN}} = \sum_{j=1}^{4} K_i^+ N K_j^- (U_i - U_j) - K_i^+ N \Phi^Q + K_i^+ N \Phi'^Q, \tag{42}$$

where $\Phi'^Q = \sum_{i=1}^4 K_i U_i$. From (41), the first term of the above right-hand side is bounded by a term of the form $Ch \sum_{j=1}^4 |U_i - U_j|$. Next, since $\mathcal{F}^h$ verifies (17) , one also has

$$|\Phi^Q| \le \int_Q \big|\operatorname{div} \mathcal{F}^h(U_h)\big|\, dx\, dy \le Ch \sum_{j,k=1}^{4} |U_j - U_k| \ , \tag{43}$$

and therefore, from (41), the second term of the right-hand side of (42) is bounded by $Ch \sum_{j,k=1}^4 |U_j - U_k|$. To estimate the last term in (42), one remarks $\Phi'^Q = \big(A(\bar{U}), B(\bar{U})\big) \cdot (\vec{n}_1 U_1 + \vec{n}_2 U_2 + \vec{n}_3 U_3 + \vec{n}_4 U_4)$. Using "$\vec{n}_1 = -\vec{n}_3$, $\vec{n}_2 = -\vec{n}_4$" and (41) successively,

$$|\Phi'^Q| = \Big|\big(A(\bar{U}), B(\bar{U})\big) \cdot \big(\vec{n}_1(U_1 - U_3) + \vec{n}_2(U_2 - U_4)\big)\Big| \le C'h\big(|U_1 - U_3| + |U_2 - U_4|\big),$$

which finally proves (13).

The LDA (26) and limited schemes (28),(29), can be written as $\Phi_i^Q = \beta_i \Phi$, where β_i is a uniformly bounded matrix with respect to h and U_h. Therefore (13) follows from (43). Both CN and LDA schemes satisfy (13), consequently, so does the Blended scheme (27) with a bounded matrix Θ.

C Total fluctuation calculations

The following lemma enables the calculation of $\int_Q \operatorname{div} \mathcal{F}^h \, dx\, dy$ by using the Stokes formula, for a given $\mathcal{F}^h$ which is polynomial with respect to the basis functions ϕ_i. The proof is straightforward and omitted. It is carried out by the fact that functions ϕ_i are linear on the edges of the quadrangles.

Lemma 4. *Let Q be a quadrilateral $\big(M_1, M_2, M_3, M_4\big)$ with $M_i = (x_i, y_i)$, and ϕ_i are the Q_1 interpolation functions on Q, for $i = 1, \ldots, 4$. Then*

$$\int_{\partial Q} \phi_i \vec{n}\, d\Gamma = \frac{1}{2}(\vec{n}_{i,p} + \vec{n}_{i,q});$$

$$\int_{\partial Q} \phi_i \phi_j \vec{n}\, d\Gamma = \begin{cases} \frac{1}{3}(\vec{n}_{i,p} + \vec{n}_{i,q}) & \text{if } i = j, \\ \frac{1}{6}\vec{n}_{i,j} & \text{if } i \neq j \text{ and } i \text{ and } j \text{ adjacent}, \\ 0 & \text{otherwise;} \end{cases}$$

$$\int_{\partial Q} \phi_i \phi_j \phi_k \vec{n}\, d\Gamma = \begin{cases} \frac{1}{4}(\vec{n}_{i,p} + \vec{n}_{i,q}) & \text{if } i = j = k, \\ \frac{1}{12}\vec{n}_{i,k} & \text{if } i = j \neq k, \text{ and } i \text{ and } k \text{ adjacent}, \\ 0 & \text{if } i \text{ and } k \text{ opposite}; \end{cases}$$

$$\int_{\partial Q} \phi_i \phi_j \phi_k \phi_l \vec{n}\, d\Gamma = \begin{cases} \frac{1}{5}(\vec{n}_{i,p} + \vec{n}_{i,q}) & \text{if } i = j = k = l, \\ \frac{1}{20}\vec{n}_{i,j} & \text{if } i = j = k \neq l \text{ and } i \text{ and } l \text{ adjacent}, \\ \frac{1}{30}\vec{n}_{i,j} & \text{if } i = j \neq k = l \text{ and } i \text{ and } l \text{ adjacent}, \\ 0 & \text{if } i \text{ and } j \text{ are opposite.} \end{cases}$$

where M_p and M_q are the two adjacent vertices to M_i, $\vec{n}$ and the $\vec{n}_{a,b}$, for all a, b, are the unit outward normal to ∂Q and the outward normal to $[M_a M_b]$ having for modulus $|\overrightarrow{M_a M_b}|$, respectively.

References

[1] R. Abgrall. Toward the ultimate conservative scheme : following the quest. *J. Comp. Phys.*, 167:277–315, 2001.

[2] H. Deconinck, K. Sermeus, and R. Abgrall. Status of multidimensional upwind residual distribution schemes and applications in aeronautics. AIAA Paper 2000-2328J, 2000.

[3] H. Deconinck, R. Struijs, G. Bourgeois, and P.L. Roe. Compact advection schemes on unstructured meshes. Computational Fluid Dynamics, 1993. VKI Lecture series 1993-04.

[4] H. Paillère and H. Deconinck. *Euler and Navier-Stokes solvers using multi-dimensional upwind schemes and multigrid acceleration*, chapter Compact cell vertex convection schemes on unstructured meshes, pages 1–49. Vieweg, Braunschweig, 1997.

[5] E. van der Weide and H. Deconinck. Positive matrix distribution schemes for hyperbolic systems, with application to the euler equations. In *Computational Fluid Dynamics*, New York, 1996. 3rd ECCOMAS CFD Conference, Wiley.

[6] R. Abgrall and M. Mezine. Construction of second order accurate monotone and stable residual distribution schemes for unsteady problems. *J. Comp. Phys.*, 188(1):16–55, 2003.

[7] A. Csík and H. Deconinck. Space time residual distribution schemes for hyperbolic conservation laws on unstructured linear finite elements. In M J Baines, editor, *Numerical Methods for Fluid Dynamics VII*, pages 557–564, Oxford, 2002.

[8] E. van der Weide and H. Deconinck. Fluctuation splitting schemes for the euler equations on quadrilateral grids. In *Numerical methods for fluid dynamics V*, Oxford, UK, 1995.

[9] R. Abgrall, K. Mer, and B. Nkonga. A Lax–Wendroff type theorem for residual schemes. In M. Hafez and J.J Chattot, editors, *Innovative methods for numerical solutions of partial differential equations*, pages 243–266. World Scientific, 2002.

[10] Á Csík, M. Ricchiuto, and H. Deconinck. A conservative formulation of the multidimensional upwind residual distribution schemes for general conservation laws. *J. Comp. Phys.*, 179(1):286–312, 2003.

[11] R. Abgrall and P.L. Roe. Construction of very high order fluctuation scheme. *J. Scientific Computing*, 19(1–3):3–36, Dec 2003.

[12] H. Deconinck, P.L. Roe, and R. Struijs. A multidimensional generalisation of Roe's difference splitter for the Euler equations. *Computer and Fluids*, 22:215–222, 1993.

[13] H. Paillère. *Multi-dimensional Upwind Residual Distribution Schemes for the Euler and Navier-Stokes Equations on Unstructures Grids.* PhD thesis, Université Libre de Bruxelles, 1995.

[14] R. Struijs, H. Deconinck, and P.L. Roe. Fluctuation splitting schemes for the 2D Euler equations. Computational Fluid Dynamics, 1991. VKI Lecture series 1991-01.

[15] R. Abgrall and M. Mezine. Construction of second order accurate monotone and stable residual distribution schemes for steady problems. *J. Comp. Phys.*, 195(2):474–507, 2004.

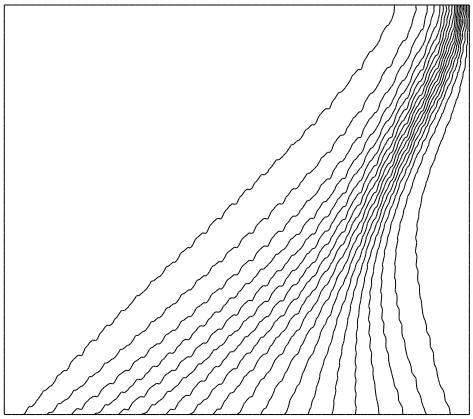
[16] M. Mezine. *Conception de schémas distributifs pour l'aérodynamique stationnaire et instationnaire.* PhD thesis, Université Bordeaux 1, 2002.

[17] P.L. Roe and D. Sidilkover. Optimum positive linear schemes for advection in two and three dimensions. *SIAM J. Numer. Anal.*, 29(6):15–42, 1992.

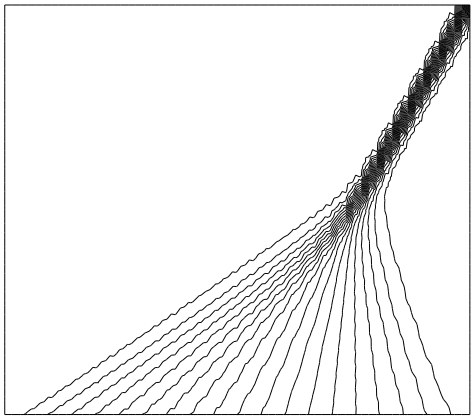
[18] E. Godlewski and P.A. Raviart. *Numerical Approximation of Hyperbolic Systems of Conservation Laws.*, volume 118 of *Applied Mathematical Sciences.* Springer, 1995.

[19] R. Abgrall. Essentially non oscillatory residual distribution schemes for hyperbolic problems. *J. Comp. Phys.*, 2005. submitted.

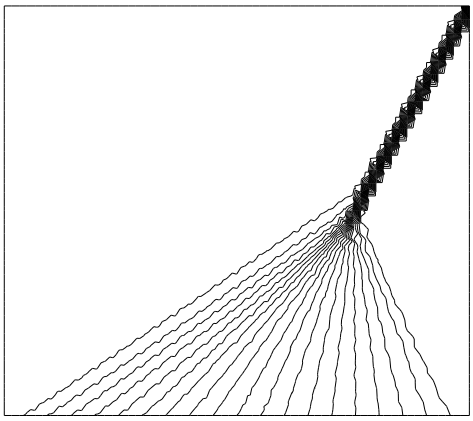
[20] P.G. Ciarlet. *The finite element method for elliptic problems.* North Holland Publishing Company, 1978.



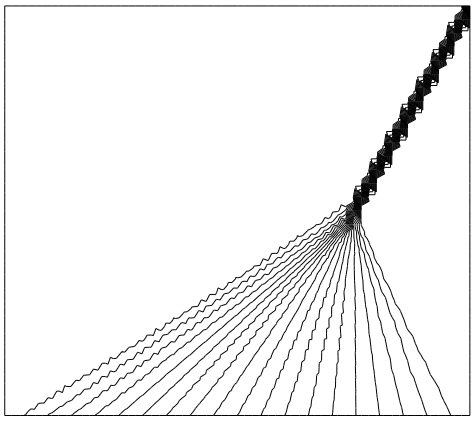
LxF scheme



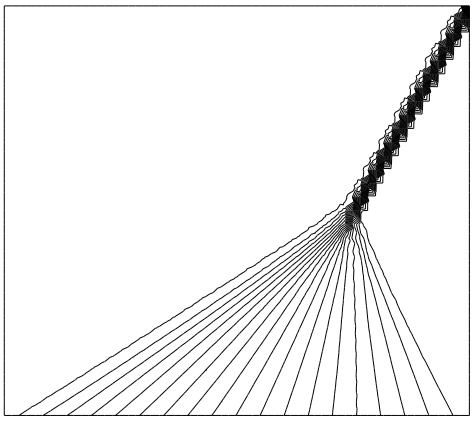
CN scheme



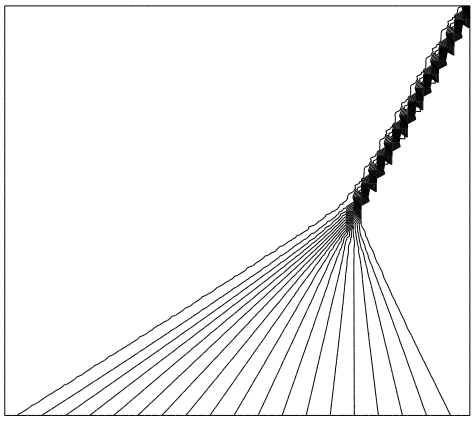
Limited–LxF scheme



Limited–CN scheme

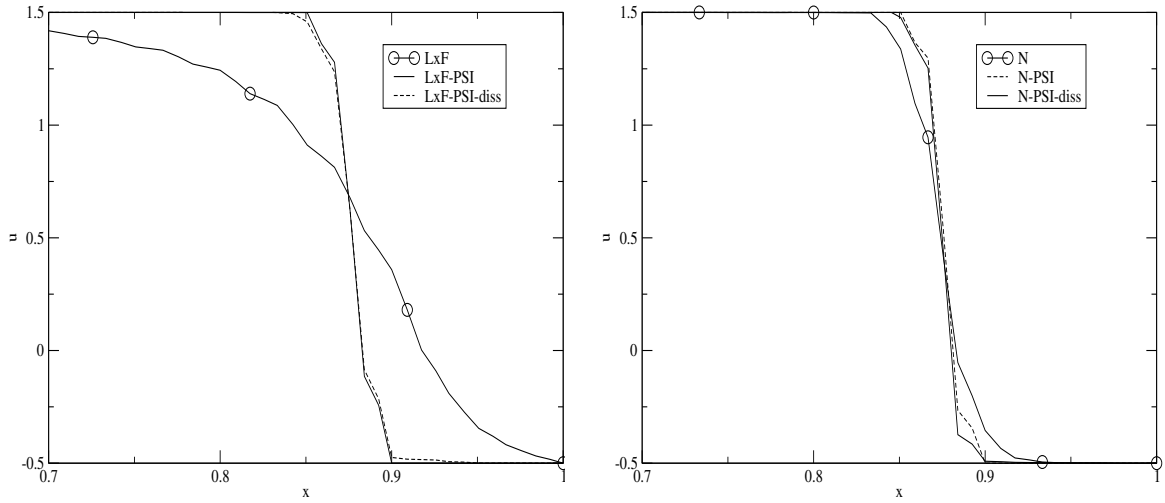


Limited–LxF scheme with dissipation (33)



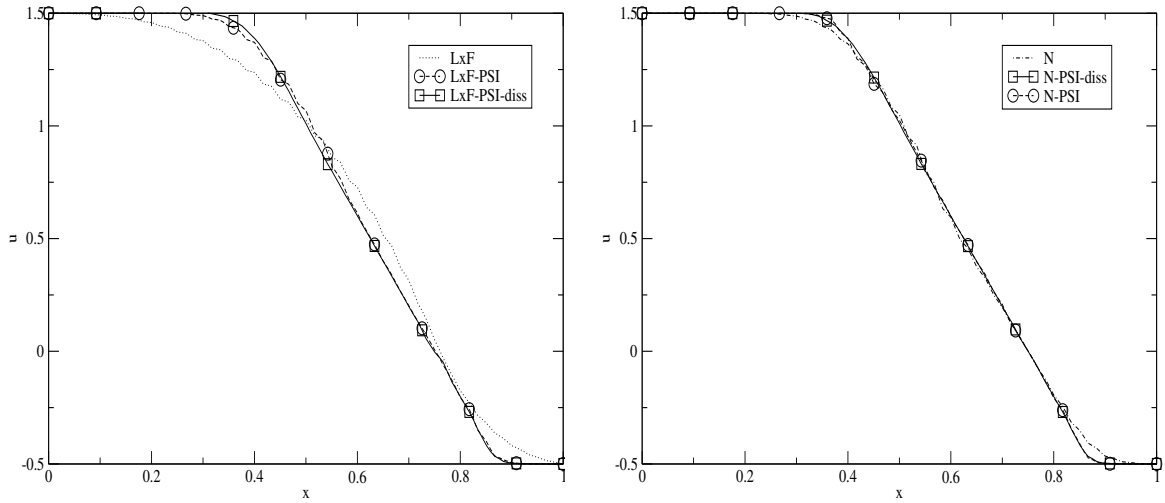
Limited–CN scheme with dissipation (33)

Figure 17: The Burger problem with the first order, second order and second order with additional dissipation. The baseline first order scheme are either the LxF scheme or the N scheme



LxF CN

Figure 18: Cross sections at $y = 0.75$, Burger equation.



LxF CN

Figure 19: Cross sections at $y = 0.25$, Burger equation.

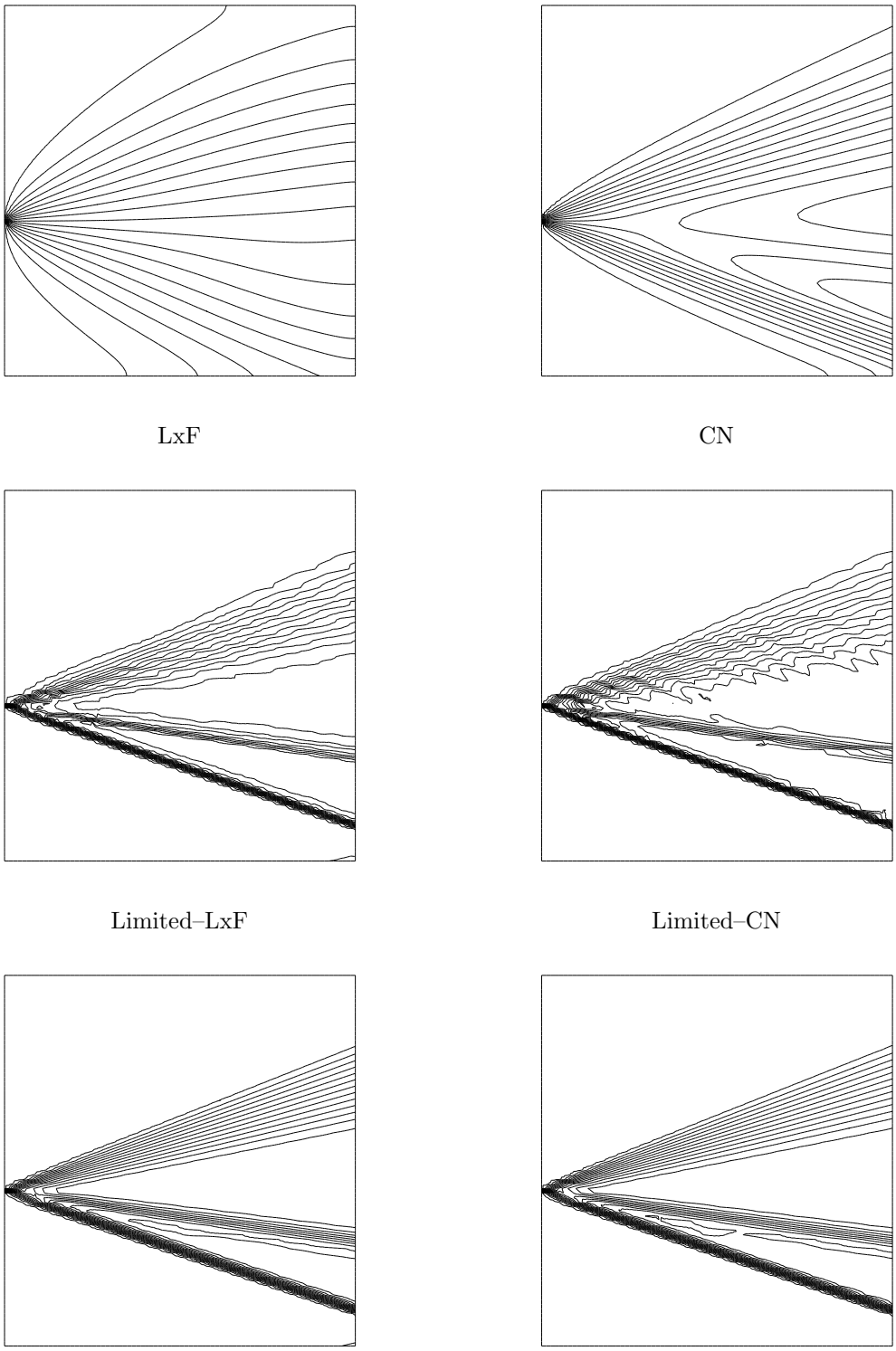


Figure 20: Isolines of the density for the schemes constructed from the LxF scheme and the CN scheme.

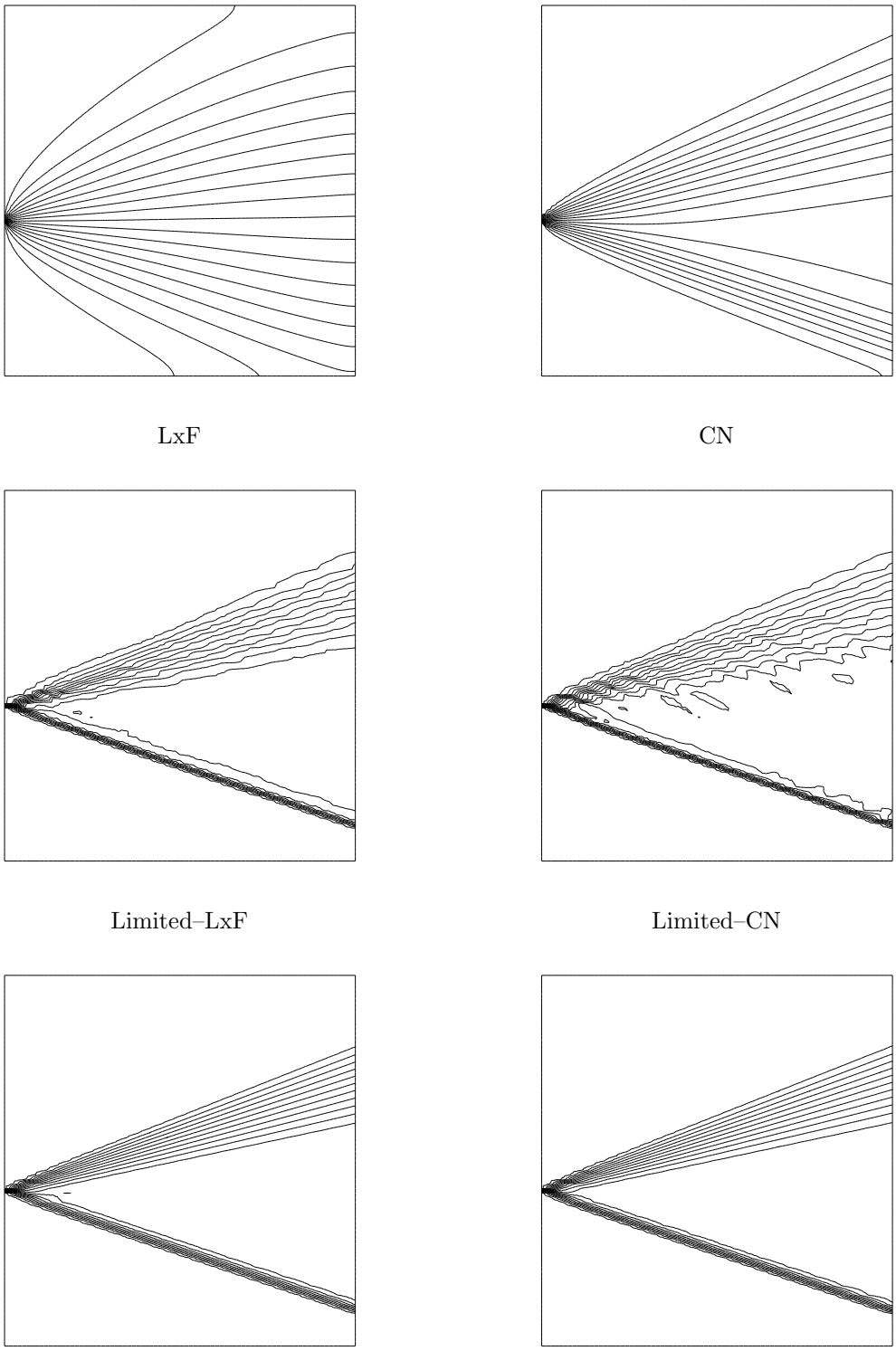


Figure 21: Isolines of the pressure for the schemes constructed from the LxF scheme and the CN scheme.

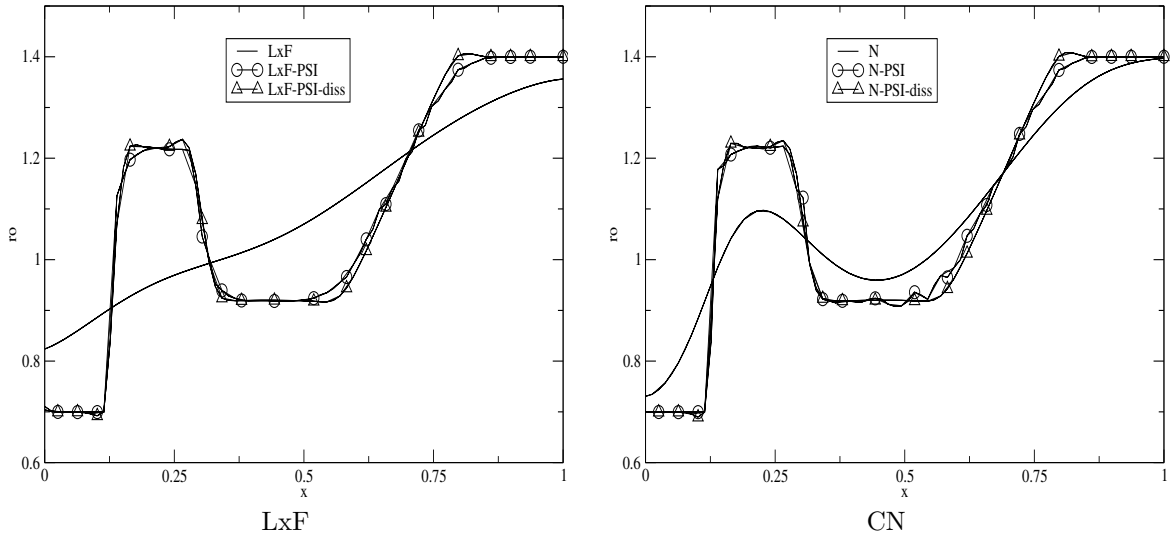


Figure 22: Cross–section of the density for the schemes based on the LxF scheme and the CN scheme

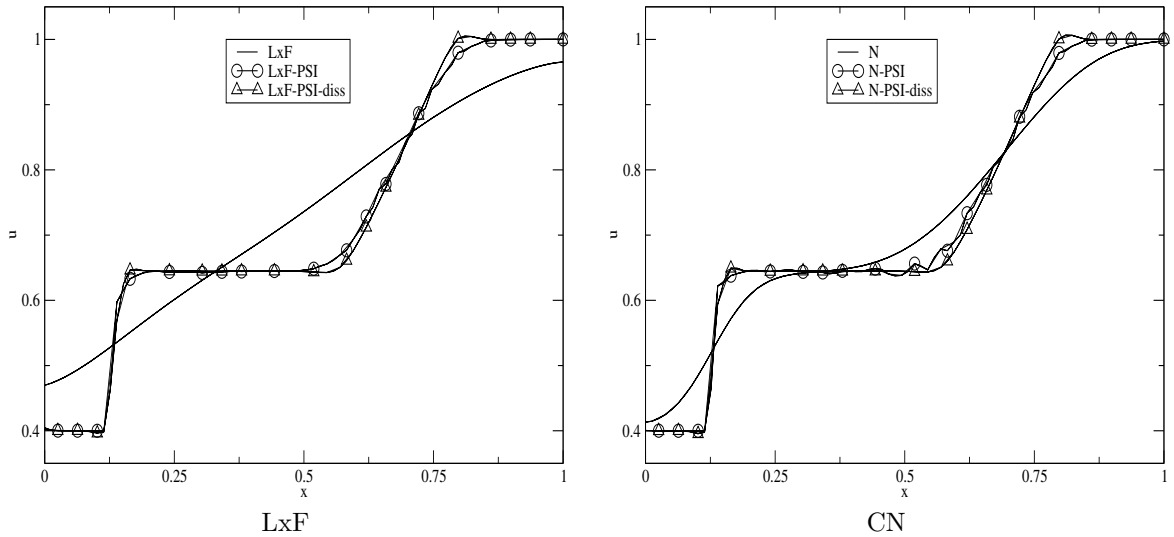


Figure 23: Cross–section of the pressure for the schemes based on the LxF scheme and the CN scheme