ASYMPTOTIC HIGH-ORDER SCHEMES FOR 2×2 DISSIPATIVE HYPERBOLIC SYSTEMS

DENISE AREGBA-DRIOLLET[1], MAYA BRIANI[2,3], AND ROBERTO NATALINI[2]

Abstract. We investigate finite difference schemes which approximate 2×2 one dimensional linear dissipative hyperbolic systems. We show that it is possible to introduce some suitable modifications in standard upwinding schemes, which keep into account the long-time behaviour of the solutions, to yield numerical approximations which are increasingly accurate for large times when computing small perturbations of stable asymptotic states, respectively around stationary solutions and in the diffusion (Chapman-Enskog) limit.

1. Introduction

Let us consider a one-dimensional 2×2 linear system with real coefficients

$$\begin{cases} u_t + \alpha u_x + \beta v_x = 0, \\ \\ v_t + \gamma u_x + \delta v_x = \zeta u - \eta v, \end{cases} \tag{1}$$

for $x \in \mathbb{R}$ and $t > 0$, with the initial conditions

$$u(x,0) = u_0(x), \;\; v(x,0) = v_0(x), \quad \text{for } x \in \mathbb{R}. \tag{2}$$

We assume that the system is strictly hyperbolic: the eigenvalues of the matrix $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$,

$$\lambda_\pm = \frac{1}{2}\left[(\alpha+\delta) \pm \sqrt{(\alpha-\delta)^2 + 4\beta\gamma}\right],$$

are real and distinct, which is simply equivalent to the condition

$$(\alpha-\delta)^2 + 4\beta\gamma > 0. \tag{3}$$

Clearly, according to the general theory of linear hyperbolic problems, see for instance [15], for every initial data in some suitable functional class, there exists a corresponding global solution in the same class. However the stability of this solution, i.e. its global boundedness in time, is more involved, and mostly connected to the behavior of the equilibrium problem, when the source term vanishes, namely the problem restricted on the manifold $v = \frac{\zeta}{\eta}u$ with u solving

$$u_t + \alpha^* u_x = 0, \tag{4}$$

for $\alpha^* = \alpha + \frac{\zeta}{\eta}\beta$. It is well-known that a necessary and sufficient stability condition for problem (1) is given by the so-called subcharacteristic condition, see [16, 11]: the coefficient η is positive and the speed of the equilibrium problem has to be bounded by the speeds of problem (1), i.e.:

$$\eta > 0, \quad \lambda_- \le \alpha^* \le \lambda_+. \tag{5}$$

1991 *Mathematics Subject Classification.* Primary: 65M12; Secondary: 35L65.

Key words and phrases. Finite differences methods, dissipative hyperbolic problems, asymptotic behavior, asymptotic high order schemes, Chapman-Enskog expansion.

[1] IMB, Mathématiques Appliquées de Bordeaux, Université Bordeaux 1.

[2] Istituto per le Applicazioni del Calcolo "Mauro Picone", Consiglio Nazionale delle Ricerche.

[3] Università LUISS - Guido Carli, Roma.

This simple example is a useful prototype for more general nonlinear relaxation systems which arise in various contexts, see again [16, 11] for examples and references. Recently, in [8] and [2], a quite complete theory of global existence and of the asymptotic behavior of smooth solutions for this type of systems was developed, actually in a nonlinear and fully multidimensional framework. This theory, which in the case of our special example will be briefly summarized in Section 2, needs for an extra assumption, the so-called Shizuta-Kawashima condition [14, 8], which guarantees for a sufficient coupling between the source and the advection terms. In the present example this condition is just equivalent to impose the strict inequalities in (5). Roughly speaking, under these assumptions, it is possible to prove that for every perturbation of a given stationary solution to problem (1), the corresponding solution decays in the L^p-norm to its unperturbed state as $\mathcal{O}\left(t^{-\frac{1}{2}\left(1-\frac{1}{p}\right)}\right)$, for $p \in [1, \infty]$. Moreover, when the asymptotic state is given by a constant state, it is possible to compare this perturbed solution with the solution $\hat{u}$ of the Chapman-Enskog expansion, which in the present case is given by

$$\hat{u}_t + \alpha^* \hat{u}_x = \nu \hat{u}_{xx}, \tag{6}$$

with $\nu = \frac{\beta}{\eta^2}\left[\gamma\eta - \frac{\beta}{\eta}\zeta^2 - (\alpha - \delta)\zeta\right]$. Notice that, the strict subcharacteristic inequalities are equivalent to the condition $\nu > 0$. So, equation (6) enjoys well-posedeness of the Cauchy problem and, again under the previous assumptions, it is actually possible to prove that the difference between u and the diffusive profile $\hat{u}$ decays in the L^p-norm as $\mathcal{O}\left(t^{-\frac{1}{2}\left(2-\frac{1}{p}\right)}\right)$, for $p \in [1, \infty]$, so a factor $\frac{1}{2}$ better than the simple decay to the steady state. Clearly, if the Shizuta-Kawashima condition is violated, no decay is in general expected, as follows easily for instance in the case $\beta = 0$.

In this paper, numerical approximations related to these asymptotic results will be introduced and discussed. Following the ideas in [5], where the case of a scalar hyperbolic equation was considered, we are looking for *Asymptotic High Order* (AHO) schemes, i.e.: schemes which are high-order accurate, with respect to the local truncation error, when restricted to every element of a given family of stable asymptotic states. Here we show that, for 2×2 dissipative hyperbolic systems, it is possible to introduce AHO schemes which are compatible with the behavior predicted by the qualitative analysis, respectively for the long-time asymptotic and in the Chapman-Enskog regimes. The main idea we are going to apply is to modify standard upwinding schemes to keep into account the long-time behaviour of the solutions. In Section 3, we present some results of convergence and error estimates for general upwinding finite difference schemes for systems of the form (1), under the strict subcharacteristic condition. Section 4 contains the description of our new schemes, which are AHO respectively around the perturbation of general steady states and in the diffusion limit. CFL estimates are presented, to guarantee the monotonicity of these schemes. Some numerical tests are presented and discussed in Section 5, to show the better performance of our schemes with respect to the usual pointwise approximation of the source term, and even with the classical upwinding of the source proposed by Roe in 1986 [13]. In particular, we show by numerical tests that the L^∞ global error of our main diffusive adapted AHO2p-scheme decays as $\mathcal{O}(1/t)$, in agreement with the decay given by inequality (32), for a given fixed space step, against the decay as $\mathcal{O}(1/\sqrt{t})$ of the other schemes.

Finally, let us remark that our approach is somewhat related to the ideas motivating the so-called well-balanced schemes, see for instance [1, 7, 9, 3, 12], and the book [4], which however are more relevant on different time scales and contexts, as for instance for inhomogeneous source terms.

2. Analytical backgrounds

Let us consider system (1), which first will be rewritten in a vector form, by setting $U = (u, v)$ and

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \ B = \begin{pmatrix} 0 & 0 \\ \zeta & -\eta \end{pmatrix}.$$

System (1) reads now

$$U_t + AU_x = BU. \tag{7}$$

The first order system (7) is said to be in Conservative-Dissipative form (CD-form), if the matrix A is symmetric and the coefficient ζ, in the matrix B, vanishes. Our first goal is to reduce every 2×2 system, which verifies some suitable assumptions, to the CD-form, to enter the framework in [2].

Proposition 2.1. *Let $\eta > 0$ and assume the strict subcharacteristic condition*

$$\lambda_- < \alpha^* < \lambda_+. \tag{8}$$

Then there exists a non singular linear change of variable such that system (1) can be set in the CD-form.

Proof. First let us remark that, by an elementary computation, condition (8) is equivalent to the single inequality

$$\beta \left[\gamma\eta - \frac{\beta}{\eta}\zeta^2 - (\alpha - \delta)\zeta \right] > 0. \tag{9}$$

Let us introduce the following change of variable

$$\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ q & r \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

with $q = -\frac{\zeta}{\eta} r$ and r given by

$$r = \left[\beta\eta \left(\gamma\eta - \frac{\beta}{\eta}\zeta^2 - (\alpha - \delta)\zeta \right)^{-1} \right]^{1/2}$$

By condition (9), this quantity is well defined, real and positive. The new unknowns $(\tilde{u}, \tilde{v})$ solve the system

$$\begin{cases} \tilde{u}_t + (\alpha + \frac{\zeta}{\eta}\beta)\tilde{u}_x + \frac{\beta}{r}\tilde{v}_x = 0, \\ \\ \tilde{v}_t + \frac{\beta}{r}\tilde{u}_x + (\delta - \frac{\zeta}{\eta}\beta)\tilde{v}_x = -\eta\tilde{v}, \end{cases} \tag{10}$$

which is in CD-form. □

Therefore, according to the previous result, from now on and without loss of generality, we shall restrict our attention to the following model problem

$$\begin{cases} u_t + au_x + bv_x = 0, \\ \\ v_t + bu_x + cv_x = -dv, \end{cases} \tag{11}$$

where the constants a, b, c, and d are real, $b \neq 0$ and $d > 0$. For system (11), we have some useful estimates. First of all we can prove a contraction principle of the solutions in the L^2-norm.

Proposition 2.2 (L^2-stability)**.** *Under the assumptions given above, the solutions of the Cauchy problem* (11)–(2) *verify*

$$\int_{\mathbb{R}} (|u(x,t)|^2 + |v(x,t)|^2)dx \leq \int_{\mathbb{R}} (|u(x,0)|^2 + |v(x,0)|^2)dx \tag{12}$$

Proof. Starting from system (11) and multiplying the first equation by u and the second one by v respectively, we get,

$$\frac{1}{2}\partial_t(u^2+v^2)+\partial_x(\frac{1}{2}au^2+buv+\frac{1}{2}cv^2)=-dv^2.$$

By integrating we have

$$\int_{\mathbb{R}}(u^2+v^2)(x,t)dx+2\int_0^t\int_{\mathbb{R}}dv^2dxdt=\int_{\mathbb{R}}(u^2+v^2)(x,0)dx,$$

and this ends the proof. □

To obtain the estimates for the L^1, L^∞ and BV norms, we need to diagonalize the system. First let us rewrite system (11) in the vector form (7), with

$$A=\begin{pmatrix} a & b \\ b & c \end{pmatrix},\ B=-\begin{pmatrix} 0 & 0 \\ 0 & d \end{pmatrix}.$$

Now we can diagonalize the matrix A:

$$A=R\Lambda R^{-1}$$

where $\Lambda=\operatorname{diag}(\lambda_1,\lambda_2)$, and $R=(r^{(1)},r^{(2)})$ is the column matrix of right eingenvectors, i.e. $Ar^{(i)}=\lambda_i r^{(i)}$, $i=1,2$. Introducing the notation

$$\mathrm{w}=R^{-1}U, \tag{13}$$

problem (11) becomes,

$$\begin{cases} \mathrm{w}_t+\Lambda\mathrm{w}_t=\tilde{B}\mathrm{w}, \\ \\ \mathrm{w}(x,0)=\mathrm{w}^0(x)=R^{-1}\mathrm{u}^0(x), \end{cases} \tag{14}$$

where $\tilde{B}=R^{-1}BR$.
If we choose the matrix of the diagonalization of the system (7) as

$$R=\begin{pmatrix} 1 & 1 \\ -\dfrac{(a-\lambda_1)}{b} & \dfrac{(\lambda_2-a)}{b} \end{pmatrix}, \tag{15}$$

the corresponding source term is then given by

$$\tilde{B}=\frac{d}{\lambda_2-\lambda_1}\begin{pmatrix} -(a-\lambda_1) & (\lambda_2-a) \\ (a-\lambda_1) & -(\lambda_2-a) \end{pmatrix}. \tag{16}$$

Now we can prove the main result about monotonicity and L^1-contraction properties of system (11). Let us set

$$L=L(\lambda)=\max_{i=1,2}|\lambda_i|.$$

For any given vectors $C,D\in\mathbb{R}^2$, set $C\le D$ if $c_i\le d_i$ $(i=1,2)$; if $C\le D$, let us denote the corresponding interval by $[C,D]$. Also, for each interval $I=(x_1,x_2)\subseteq\mathbb{R}$, we set $\tau=(x_2-x_1)/2L$.

Proposition 2.3. *Let* w *and* $\tilde{\mathrm{w}}$ *be two weak solutions of the Cauchy problem (14) in* $\mathbb{R}\times(0,T)$ $(T>0)$*, for the initial data* w^0 *and* $\tilde{\mathrm{w}}^0$ *respectively. Let* $[C,D]\subseteq\mathbb{R}^2$ *be an interval with non empty interior such that* $\mathrm{w},\ \tilde{\mathrm{w}}\in[C,D]$*. If the matrix* $\tilde{B}$ *is given by* (16)*, then for each interval* $(x_1,x_2)\subseteq\mathbb{R}$ *and for all* $t\in(0,\min(\tau,T))$*, we have*

$$\sum_{i=1,2}\int_{x_1+tL}^{x_2-tL}[w_i-\tilde{w}_i]_+dx\le\sum_{i=1,2}\int_{x_1}^{x_2}[w_i^0-\tilde{w}_i^0]_+dx. \tag{17}$$

Proof. Define $\mathrm{v} = \mathrm{w} - \tilde{\mathrm{w}}$. Multiplying the i component of the equation (14) by the Heaviside function $H(v_i)$, where $v_i = w_i - \tilde{w}_i$, $i \in \{1, 2\}$ we get

$$\partial_t [v_i]_+ + \lambda_i \partial_x [v_i]_+ = \sum_{j\in\{1,2\}} \tilde{b}_{ij} H(v_i) v_j, \quad i \in \{1, 2\}, \tag{18}$$

where we used the relations $H(\mathrm{v})\mathrm{v} = [\mathrm{v}]_+$ and $H(\mathrm{v})\partial_\cdot \mathrm{v} = \partial_\cdot [\mathrm{v}]_+$.
Let Ω be the domain contained in the plane (x, s), with boundary defined by the union of lines ϕ_1, ϕ_2, ϕ_3, ϕ_4 defined in parametric form as

$$\phi_1 : \begin{cases} x = y \\ s = 0 \end{cases} \quad y \in [x_1, x_2], \qquad \phi_2 : \begin{cases} x = y \\ s = \frac{-y+x_2}{L} \end{cases} \quad y \in [x_2 - Lt, x_2],$$

$$\phi_3 : \begin{cases} x = y \\ s = t \end{cases} \quad y \in [x_1 + Lt, x_2 - Lt], \qquad \phi_2 : \begin{cases} x = y \\ s = \frac{y-x_1}{L} \end{cases} \quad y \in [x_1, x_1 + Lt].$$

Then, by the Green-Gauss formula

$$\iint_\Omega (\partial_s [v_i]_+ + \lambda_i \partial_x [v_i]_+)\, dxds = \int_{\partial\Omega} -[v_i]_+ dx + \lambda_i [v_i]_+ ds$$

$$= -\int_{x_1}^{x_2} [v_i^0]_+ dy + \int_{x_1+Lt}^{x_2-Lt} [v_i(y, t)]_+ dy \tag{19}$$

$$+ \left(1 + \frac{\lambda_i}{L}\right) \int_{x_2-Lt}^{x_2} [v_i(\phi_2)]_+ dy + \left(1 - \frac{\lambda_i}{L}\right) \int_{x_1}^{x_1+Lt} [v_i(\phi_4)]_+ dy.$$

By (18),

$$\begin{aligned} \int_{x_1+Lt}^{x_2-Lt} [v_i(y, t)]_+ dy &= \int_{x_1}^{x_2} [v_i^0]_+ dy + \iint_\Omega \sum_{j\in\{1,2\}} \tilde{b}_{ij} H(v_i) v_j \\ &- \left(1 + \frac{\lambda_i}{L}\right) \int_{x_2-Lt}^{x_2} [v_i(\phi_2)]_+ dy - \left(1 - \frac{\lambda_i}{L}\right) \int_{x_1}^{x_1+Lt} [v_i(\phi_4)]_+ dy. \end{aligned} \tag{20}$$

Since $-L \le \lambda_i \le L$ ($i \in \{1, 2\}$), it is $(1 \pm \lambda_i/L) \ge 0$ and

$$\int_{x_1+Lt}^{x_2-Lt} [v_i(y, t)]_+ dy \le \int_{x_1}^{x_2} [v_i^0]_+ dy + \iint_\Omega \sum_{j\in\{1,2\}} \tilde{b}_{ij} H(v_i) v_j. \tag{21}$$

The thesis is then achieved if

$$\sum_{i,j\in\{1,2\}} \tilde{b}_{ij} H(v_i) v_j \le 0.$$

We have,

$$\sum_{i,j\in\{1,2\}} \tilde{b}_{ij} H(v_i) v_j \le \sum_{i\in\{1,2\}} \tilde{b}_{ii} [v_i]_+ + \sum_{i\in\{1,2\}} \sum_{j\ne i} \tilde{b}_{ij} [v_j]_+ \tag{22}$$

$$= \sum_{i\in\{1,2\}} \Big(\tilde{b}_{ii} + \sum_{j\ne i} \tilde{b}_{ji}\Big) [v_i]_+. \tag{23}$$

The proof follows from (16). □

Clearly, from Proposition 2.3 we can easily deduce, by standard arguments, L^1 and L^∞-estimates, as well as obvious comparison properties of solutions, and, as the system (11) is linear, the same estimates hold for the derivatives of solutions. More precisely we use the following result in the sequel:

Corollary 2.4. *Let* $\mathrm{w}^0 \in C^m(\mathbb{R})$ *such that*

$$\frac{d^k \mathrm{w}^0}{dx^k} \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}) \quad \text{for } \ 0 \leq k \leq m. \tag{24}$$

Then the Cauchy problem (14) has a unique solution $\mathrm{w} \in C^0([0,+\infty[, L^1(\mathbb{R}))$ *and there exists a constant* $K > 0$*, not depending on* w^0*, such that for* $0 \leq k \leq m$*,* $p = 1$ *and* $p = \infty$*:*

$$\|\partial_x^k \mathrm{w}(.,t)\|_p \leq K \|\frac{d^k \mathrm{w}^0}{dx^k}\|_p\,, \quad \|\partial_t \mathrm{w}(.,t)\|_p \leq K(\|\frac{d\mathrm{w}^0}{dx}\|_p + \|\mathrm{w}^0\|_p). \tag{25}$$

Now, we are in position to investigate the asymptotic behavior of solutions to problem (11). Here we refer to the results in [8, 2] about the Cauchy problem for a general $n \times n$ hyperbolic symmetrizable one dimensional system of balance laws

$$U_t + f(U)_x = g(U), \tag{26}$$

with the initial condition

$$U(x,0) = U_0(x), \tag{27}$$

where $U = (u_1, u_2) \in \Omega \subseteq \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, with $n_1 + n_2 = n$. It is also assumed that there are n_1 conservation laws in the system, namely that we can take

$$g(U) = \begin{pmatrix} 0 \\ q(U) \end{pmatrix}, \quad \text{with } q(U) \in \mathbb{R}^{n_2}. \tag{28}$$

In [8], it was proposed a general framework of sufficient conditions which guarantee the global existence in time of smooth solutions, by showing that dissipation, at least for small data, can prevent shock formation. Actually, for the systems which are endowed with a strictly convex entropy function $\mathcal{E} = \mathcal{E}(U)$, a first natural assumption is the *entropy dissipation condition*, see [6], namely for every U, $\overline{U} \in \Omega$, with $g(\overline{U}) = 0$,

$$\big(\mathcal{E}'(U) - \mathcal{E}'(\overline{U})\big) \cdot g(U) \leq 0.$$

Unfortunately, it is easy to see that this condition is too weak to prevent the formation of singularities. A quite natural supplementary condition can be imposed to entropy dissipative systems, following the approach first proposed by Shizuta and Kawashima [10, 14], which for system (26) reads

$$Ker\ Dg(\overline{U}) \cap \{\text{eigenspaces of } Df(\overline{U})\} = \{0\}, \tag{29}$$

for every $\overline{U} \in \Omega$, with $g(\overline{U}) = 0$. It is possible to prove that this condition, which is satisfied in many interesting examples, is also sufficient to establish a result of global existence for small perturbations of equilibrium constant states, see [8]. Clearly also system (11) fits this framework. In fact the system is entropy dissipative for the entropy function $\mathcal{E}(u,v) = \frac{1}{2}(u^2 + v^2)$, if and only if $d > 0$, and the condition (29) is satisfied if $b \neq 0$. Therefore we can apply all the results proved in [2] about the asymptotic behavior of dissipative hyperbolic systems to system (11). Here we just state two main results of time decay. More details and proofs are contained in [2]. For $s > 0$, set $E_s = \max\Big\{\|U(0)\|_{L^1}, \|U(0)\|_{H^s}\Big\}$.

Theorem 2.5. *Let* $(u(t), v(t))$ *be a smooth global solution to problem* (11)–(2)*, with* $b \neq 0$ *and* $d > 0$ *. Let* $p \in [1, \infty]$*. The following decay estimate holds, for some positive constant* $C > 0$*:*

$$\|D^\beta(u(t), v(t))\|_{L^p} \leq C \min\Big\{1, t^{-\frac{1}{2}(1-\frac{1}{p}) - |\beta|/2}\Big\} E_{|\beta|+2}. \tag{30}$$

For the dissipative part $v(t)$, we have a more precise estimate, for some positive constant $C > 0$:

$$\|D^\beta v(t)\|_{L^p} \le C \min\Big\{1, t^{-\frac{1}{2}(1-\frac{1}{p})-1/2-|\beta|/2}\Big\} E_{|\beta|+2}, \tag{31}$$

Another interesting result concerns the convergence to the Chapman-Enskog expansion of problem (11), which is given by the parabolic equation (6).

Theorem 2.6. *Let $\hat{u}$ be the solution of problem* (6), *with $\alpha^* = a$ and $\nu = \frac{b^2}{d}$, and the initial condition given by* (2). *Under the assumptions of Theorem 2.5, we have the following more accurate decay estimate, for some positive constant $C > 0$:*

$$\|D^\beta(u(t) - \hat{u}(t))\|_{L^p} \le C \min\Big\{1, t^{-\frac{1}{2}(1-\frac{1}{p})-|\beta|/2-1/2}\Big\} E_{|\beta|+2}. \tag{32}$$

3. Numerical approximation

Here we introduce and discuss general finite difference approximations for system (11) and we prove some convergence properties and error estimates.

3.1. **Construction of the schemes.** We approximate the differential part following the direction of the characteristic velocities, so first we study the methods for the system in diagonal form (14) obtained by the choice (15). We denote by $\mathrm{w} = (w^1, w^2)$ the exact solution.

We denote by h the uniform mesh-length and by $x_l = l\,h$ the spatial grid points for all $l \in \mathbb{Z}$. The time levels t_n, with $t_0 = 0$, are also spaced uniformly with mesh-length $\Delta t = t_{n+1} - t_n$ for $n \in \mathbb{N}$. We denote by δ the CFL ratio $\delta = \Delta t/h$, which is taken constant through the paper.

We consider the Cauchy problem (14). The initial data w^0 is supposed to be smooth and is approximated by its node values. The approximate solution $W_l^n = (w^1_{l,n}, w^2_{l,n})^T$ is given by

$$\begin{aligned} &\frac{W_l^{n+1} - W_l^n}{\Delta t} + \frac{\Lambda}{2h}\left(W_{l+1}^n - W_{l-1}^n\right) - \frac{\tilde{Q}}{2h}\left(W_{l+1}^n - 2W_l^n + W_{l-1}^n\right) \\ &= \tilde{\mathcal{B}}_{-1} W_{l-1}^n + \tilde{\mathcal{B}}_0 W_l^n + \tilde{\mathcal{B}}_1 W_{l+1}^n, \quad l \in \mathbb{Z},\ n \in \mathbb{N} \\ &W_l^0 = \mathrm{w}^0(x_l), \quad l \in \mathbb{Z}, \end{aligned} \tag{33}$$

where $\tilde{Q} = \mathrm{diag}(\tilde{q}_1, \tilde{q}_2)$ is the diagonal matrix of the artificial diffusion terms $\tilde{q}_i \ge 0$ ($i \in \{1,2\}$), and $\tilde{\mathcal{B}}_{-1} = (\tilde{\beta}_{ij}^{-1})_{i,j=1,2}$, $\tilde{\mathcal{B}}_0 = (\tilde{\beta}_{ij}^{0})_{i,j=1,2}$ and $\tilde{\mathcal{B}}_1 = (\tilde{\beta}_{ij}^{1})_{i,j=1,2}$ are 2×2 constant matrices that define the source approximation. Those matrices may depend on h.

We set

$$W_\Delta^n(x) = W_l^n \quad \text{if } x \in [x_l, x_{l+1}[$$

and

$$W_\Delta(x,t) = W_\Delta^n(x) \quad \text{if } t \in [t_n, t_{n+1}[.$$

The scheme (33) can be seen as a linear function

$$W_\Delta^{n+1} = \mathbf{S}_{\Delta t}(W_\Delta^n). \tag{34}$$

More precisely we have

$$w^i_{l,n+1} = S^i(W_{l-1}^n, W_l^n, W_{l+1}^n), \quad i \in \{1,2\}, \quad l \in \mathbb{Z} \tag{35}$$

with

$$\begin{aligned} S^i(W^n_{l-1},W^n_l,W^n_{l+1}) &= \alpha^i_{-1}w^i_{l-1,n}+\alpha^i_0 w^i_{l,n}+\alpha^i_1 w^i_{l+1,n} \\ &+\Delta t\sum_{j\in\{1,2\}}\left(\tilde\beta^{-1}_{ij}w^j_{l-1,n}+\tilde\beta^0_{ij}w^j_{l,n}+\tilde\beta^1_{ij}w^j_{l+1,n}\right), \end{aligned} \tag{36}$$

where, for all $i\in\{1,2\}$,

$$\alpha^i_{-1}=\frac{\delta}{2}\lambda_i+\frac{\delta}{2}\tilde q_i,\quad \alpha^i_0=1-\delta\tilde q_i,\quad \alpha^i_1=-\frac{\delta}{2}\lambda_i+\frac{\delta}{2}\tilde q_i\ . \tag{37}$$

For any function ψ defined on $\mathbb{R}$ with values in $\mathbb{R}^2$, we set

$$\overline{\psi}(x)=(\psi(x-h),\psi(x),\psi(x+h)). \tag{38}$$

The function $\mathbf{S}_{\Delta t}$ is then defined by

$$\mathbf{S}_{\Delta t}(\psi)(x)=\left(S^1(\overline{\psi}(x)),S^2(\overline{\psi}(x))\right). \tag{39}$$

3.2. **Convergence for smooth data.** We assume that the scheme satisfies the two following properties:

(1) **Consistency**
The scheme (33) is consistent with problem (14), i.e

$$\tilde{\mathcal{B}}_{-1}+\tilde{\mathcal{B}}_0+\tilde{\mathcal{B}}_1=\tilde B+h\tilde C, \tag{40}$$

where $\tilde C=(\tilde c_{ij})_{i,j=1,2}$ is a 2×2 constant matrix not depending on h and Δt.

(2) **Monotonicity**
For all $i\in\{1,2\}$, the operators S^i, defined in (35), are monotone nondecreasing in all their components, i.e.

$$\tilde\beta^{-1}_{ij},\tilde\beta^0_{ij},\tilde\beta^1_{ij}\ge 0\quad \forall\, i\ne j, \tag{41}$$

$$A_{i,0}=1-\delta\tilde q_i+\Delta t\tilde\beta^0_{ii}\ge 0,\quad \forall\ i\in\{1,2\}, \tag{42}$$

$$\left\{\begin{aligned} A_{i,-1}&=\frac{\delta\lambda_i}{2}+\frac{\delta\tilde q_i}{2}+\Delta t\tilde\beta^{-1}_{ii}\ge 0,\\ A_{i,1}&=-\frac{\delta\lambda_i}{2}+\frac{\delta\tilde q_i}{2}+\Delta t\tilde\beta^{1}_{ii}\ge 0,\end{aligned}\right.\qquad \forall\ i\in\{1,2\}. \tag{43}$$

For $t\ge 0$ and $x\in\mathbb{R}$, by using the notation (38) for w, the local truncation error $T(x,t)$ is defined by

$$T(x,t)=\frac{\mathrm{w}(x,t+\Delta t)-\mathbf{S}_{\Delta t}\mathrm{w}(.,t)(x)}{\Delta t}$$

which can also be written

$$T^i(x,t)=\frac{w^i(x,t+\Delta t)-S^i(\overline{\mathrm{w}}(x,t))}{\Delta t},\quad i\in\{1,2\}.$$

It is obtained straightforwardly by Taylor expansions:

$$\begin{aligned} T^i(x,t) &= \frac{\Delta t}{2}\partial_{tt}w^i(x,t+\theta_i\Delta t)-h\frac{\tilde q_i}{2}\partial_{xx}w^i(x,t)-h\sum_{j\in\{1,2\}}\tilde c_{ij}w^j(x,t) \\ & -h\sum_{j\in\{1,2\}}(\tilde\beta^1_{ij}-\tilde\beta^{-1}_{ij})\partial_x w^j(x,t)+\mathcal{O}(h^2). \end{aligned} \tag{44}$$

In this formula, $\theta_i\in[0,1]$ and the last term depends on the spatial derivatives of w, up to third order.

In view of proving convergence, we need for the discrete truncation error, which is defined as

$$T_\Delta(x,t) = T(x_l, t) \quad for \quad x \in [x_l, x_{l+1}[.$$

The scheme is first order in the following sense:

Proposition 3.1 (Accuracy for smooth data)**.** *Consider a smooth initial data* w^0 *such that*

$$\frac{d^k \mathrm{w}^0}{dx^k} \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}) \quad \text{for } \ 0 \le k \le 4. \tag{45}$$

Then the discrete local truncation error $T_\Delta(x,t)$ *for the scheme (33) satisfies*

$$\sum_{i\in\{1,2\}} \left[\|T^i_\Delta(.,t)\|_1 + \|T^i_\Delta(.,t)\|_\infty \right] \le C(\Delta t + h). \tag{46}$$

Here

$$C \le C_1 \sup_{0\le k\le 4} \left(\left\| \frac{d^k \mathrm{w}^0}{dx^k} \right\|_1 + \left\| \frac{d^k \mathrm{w}^0}{dx^k} \right\|_\infty \right) \tag{47}$$

and C_1 *is a uniform constant.*

Proof. We can express $\partial_{tt}\mathrm{w}$ as a linear function of w, $\partial_x \mathrm{w}$, $\partial_{xx}\mathrm{w}$. As a consequence, there exists a constant C_0,

$$C_0 \le C_0' \sup_{k\le 3} \sup_{t\in[0,+\infty[} (\|\partial^k_x \mathrm{w}(.,t)\|_1 + \|\partial^k_x \mathrm{w}(.,t)\|_\infty)$$

such that

$$\sum_{i\in\{1,2\}} \left[\|T^i(.,t)\|_1 + \|T^i(.,t)\|_\infty \right] \le C_0(\Delta t + h). \tag{48}$$

Moreover, it is easy to see that for $p = 1$ or $p = \infty$:

$$\|T^i(.,t) - T^i_\Delta(.,t)\|_p \le h \|\partial_x T^i(.,t)\|_p, \quad i \in \{1,2\}$$

which, with corollary 2.4, gives the estimate (46). □

We prove now that the scheme is stable.

Proposition 3.2 (L^1 stability)**.** *For the matrix* $\tilde{C}$ *defined in (40), we introduce the quantity*

$$\tilde{\gamma} = \left[\sup_{i\in\{1,2\}} \{ \tilde{c}_{ii} + \sum_{j\ne i} \tilde{c}_{ji} \} \right]_+ . \tag{49}$$

Under the monotonicity assumptions (41-43), the scheme (33) is L^1*-stable: for all* $W^0_\Delta \in L^1(\mathbb{R})$ *it holds*

$$\|W^n_\Delta\|_1 \le e^{\tilde{\gamma} h t_n} \|W^0_\Delta\|_1, \quad \forall n \ge 1. \tag{50}$$

Proof. The norm of W^{n+1}_Δ can be written as

$$\begin{aligned} \|W^{n+1}_\Delta\|_1 &= h \sum_{i=1,2} \sum_{l\in\mathbb{Z}} \Bigg| w^i_{l,n} - \delta \frac{\lambda_i}{2}(w^i_{l+1,n} - w^i_{l-1,n}) + \delta \frac{\tilde{q}_i}{2}(w^i_{l+1,n} - 2w^i_{l,n} + w^i_{l-1,n}) \\ &\qquad + \Delta t \sum_{j=1,2} \left(\tilde{\beta}^{-1}_{ij} w^j_{l-1,n} + \tilde{\beta}^0_{ij} w^j_{l,n} + \tilde{\beta}^1_{ij} w^j_{l+1,n} \right) \Bigg| \\ &\le h \sum_{i=1,2} \sum_{l\in\mathbb{Z}} |w^i_{l,n}| A_i. \end{aligned} \tag{51}$$

Here A_i is defined by

$$A_i = |A_{i,0}| + |A_{i,-1}| + |A_{i,1}| + \Delta t \sum_{j\neq i}(|\tilde{\beta}_{ji}^{-1}| + |\tilde{\beta}_{ji}^{0}| + |\tilde{\beta}_{ji}^{1}|)$$

with quantities $A_{i,k}$ defined in (41-43). By those assumptions and by the definition of $\tilde{C}$ we obtain that

$$A_i \leq 1 + \tilde{\gamma} h\, \Delta t.$$

Hence

$$\|W_\Delta^{n+1}\|_1 \leq (1+\tilde{\gamma} h\, \Delta t)\|W_\Delta^n\|_1$$

and this gives (50). □

We have also a uniform bound for the scheme:

Proposition 3.3 (L^∞ stability)**.** *Under the monotonicity assumptions (41-43), the scheme (33) is L^∞-stable: for all $W_\Delta^0 \in L^\infty(\mathbb{R})$ it holds*

$$\|W_\Delta^n\|_\infty \leq (1+\tilde{\gamma} h\, \Delta t)\|W_\Delta^{n-1}\|_\infty \leq e^{\tilde{\gamma} h t_n}\|W_\Delta^0\|_\infty, \quad \forall n \geq 1. \tag{52}$$

The proof is similar as the one of proposition (3.2) so we omit it.

We have proved consistency and stability, so we have convergence.

Theorem 3.4. *Let* w^0 *be a smooth initial data satisfying (45). We suppose that the monotonicity assumptions (41-43) are verified and that $\delta = \Delta t/h$ is fixed.*

For all $T > 0$ the scheme (33) converges in $L^\infty([0,T], L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}))$ towards the solution of the Cauchy problem (14): there exists $h_0(T) > 0$ such that for all $h < h_0(T)$:

$$\sup_{t\in[0,T]} (\|W_\Delta(.,t) - \mathrm{w}(.,t)\|_1 + \|W_\Delta(.,t) - \mathrm{w}(.,t)\|_\infty) \leq (2T+1)\, C\, (\Delta t + h) \tag{53}$$

where C is defined in (47).

Proof. Let us fix $t \in [0,T]$. There exists an integer n such that $t_n \leq t < t_{n+1}$. We set

$$\mathrm{w}_\Delta(x,t) = \mathrm{w}(x_l,t) \quad \mathit{for} \quad x \in [x_l, x_{l+1}[.$$

For $p=1$ and $p=\infty$, we have

$$\begin{aligned} \|W_\Delta(.,t) - \mathrm{w}(.,t)\|_p \quad &\leq \quad \|W_\Delta^n - \mathrm{w}_\Delta(.,t_n)\|_p \\ &+ \quad \|\mathrm{w}_\Delta(.,t_n) - \mathrm{w}(.,t_n)\|_p \\ &+ \quad \|\mathrm{w}(.,t_n) - \mathrm{w}(.,t)\|_p\ . \end{aligned} \tag{54}$$

The two last terms of this inequality are easily estimated:

$$\|\mathrm{w}_\Delta(.,t_n) - \mathrm{w}(.,t_n)\|_p + \|\mathrm{w}(.,t_n) - \mathrm{w}(.,t)\|_p \leq C(h + \Delta t).$$

In order to study the other term, we introduce the following notation:

$$V_\Delta^m = \mathrm{w}_\Delta(.,t_m), \quad m = 0,\dots,n.$$

We have

$$\begin{aligned} \|W_\Delta^n - \mathrm{w}_\Delta(.,t_n)\|_p \quad &\leq \quad \|\mathbf{S}_{\Delta t}^{n-1}\mathbf{S}_{\Delta t}(V_\Delta^0) - \mathbf{S}_{\Delta t}^{n-1}(V_\Delta^1)\|_p \\ &+ \quad \|\mathbf{S}_{\Delta t}^{n-2}\mathbf{S}_{\Delta t}(V_\Delta^1) - \mathbf{S}_{\Delta t}^{n-2}(V_\Delta^2)\|_p \\ &+ \quad \dots \\ &+ \quad \|\mathbf{S}_{\Delta t}(V_\Delta^{n-1}) - V_\Delta^n\|_p\ . \end{aligned}$$

By L^p stability:

$$\begin{aligned} \|W^n_\Delta - \mathrm{w}_\Delta(.,t_n)\|_p &\leq \|\mathbf{S}_{\Delta t}(V^0_\Delta) - V^1_\Delta\|_p e^{\tilde{\gamma} h t_{n-1}} \\ &+ \|\mathbf{S}_{\Delta t}(V^1_\Delta) - V^2_\Delta\|_p e^{\tilde{\gamma} h t_{n-2}} \\ &+ \dots \\ &+ \|\mathbf{S}_{\Delta t}(V^{n-1}_\Delta) - V^n_\Delta\|_p \; . \end{aligned} \tag{55}$$

We then sum over $p \in \{1, \infty\}$. We may apply the consistency result (46) to each term in the right-hand side:

$$V^{j+1}_\Delta(x) - \mathbf{S}_{\Delta t}(V^j_\Delta)(x) = \Delta t T_\Delta(x, t_j), \quad j = 0, \dots, n-1.$$

Summing up all the terms gives

$$\|W^n_\Delta - \mathrm{w}_\Delta(.,t_n)\|_1 + \|W^n_\Delta - \mathrm{w}(.,t_n)\|_\infty \leq \Delta t C(\Delta t + h) \frac{e^{\tilde{\gamma} h T} - 1}{e^{\tilde{\gamma} h \Delta t} - 1} \, .$$

As δ is constant we have

$$\lim_{h \to 0} \Delta t \frac{e^{\tilde{\gamma} h T} - 1}{e^{\tilde{\gamma} h \Delta t} - 1} = T$$

which ends the proof. □

Remark 3.5. If $\tilde{\gamma} = 0$, the scheme is L^1 contracting and we have for all h and Δt:

$$\sup_{t \in [0,T]} \left(\|W_\Delta(.,t) - \mathrm{w}(.,t)\|_1 + \|W_\Delta(.,t) - \mathrm{w}(.,t)\|_\infty \right) \leq (T+1) \; C \; (\Delta t + h).$$

4. Asymptotic High-Order Schemes

In this section, we study the discretization of the source term, defined by coefficients $\tilde{\mathcal{B}}_{-1,0,1}$ introduced in (14), to present some schemes which are increasingly accurate for large times, with respect to the asymptotic behavior of solutions. This property of accuracy is required in order to get better results for large time simulations when computing perturbations of non constant stable states. Actually, given a stable solution z (or a family of stable solutions), we say that a scheme is *Asymptotic High-Order* of order q (AHOq) with respect to this stable state if the scheme is accurate of order q when restricted to this solution, see [5].

4.1. The construction of an AHO4 scheme for stationary solutions.

Here, we consider the case of a perturbation of a generic stationary solution. Let us first develop the formula (44) for a smooth solution w:

$$\begin{aligned} T(x,t) = \; & \frac{\Lambda}{2h} \left(\frac{h^3}{3} \partial_x^3 \mathrm{w}(x,t) + \mathcal{O}(h^5) \right) \\ & - \frac{\tilde{Q}}{2h} \left(h^2 \partial_x^2 \mathrm{w}(x,t) + \frac{h^4}{12} \partial_x^4 \mathrm{w}(x,t) + \mathcal{O}(h^6) \right) \\ & - h \tilde{C} \mathrm{w}(x,t) - h \left(\tilde{\mathcal{B}}_1 - \tilde{\mathcal{B}}_{-1} \right) \partial_x \mathrm{w}(x,t) - \frac{h^2}{2} \left(\tilde{\mathcal{B}}_1 + \tilde{\mathcal{B}}_{-1} \right) \partial_x^2 \mathrm{w}(x,t) \\ & - \frac{h^3}{6} \left(\tilde{\mathcal{B}}_1 - \tilde{\mathcal{B}}_{-1} \right) \partial_x^3 \mathrm{w}(x,t) + \mathcal{O}(h^4) + \Delta t \int_0^1 (1-\theta) \partial_t^2 \mathrm{w}(x, t + \theta \Delta t) d\theta. \end{aligned} \tag{56}$$

The terms $\mathcal{O}(h^k)$ involve only spatial derivatives of w. Let us now consider the stationary solution z, namely the solution of problem

$$\Lambda \partial_x \mathrm{z} = \tilde{B} \mathrm{z}. \tag{57}$$

First, we remark that if $\lambda_1\lambda_2 = 0$, *i.e.* $ac - b^2 = 0$, then all the stationary solutions are constant with respect to x. Hence for such solutions:

$$T(x,t) = -h\tilde{C}\mathrm{z}.$$

Any scheme with $\tilde{C} = 0$ is an AHO-∞ scheme. In the following we suppose that

$$ac - b^2 \neq 0. \tag{58}$$

For smooth data we get

$$\partial_x^k \mathrm{z} = (\Lambda^{-1}\tilde{B})^k \mathrm{z}. \tag{59}$$

We compute the local truncation error (56) on this particular solution. It does not depend on t anymore and we denote it by $T_{stat}(x)$:

$$\begin{aligned} T_{stat}(x) = \; & -\frac{h}{2}\Big(\tilde{Q}(\Lambda^{-1}\tilde{B})^2 + 2\tilde{C} + 2\big(\tilde{\mathcal{B}}_1 - \tilde{\mathcal{B}}_{-1}\big)(\Lambda^{-1}\tilde{B})\Big)\,\mathrm{z}(x) \\ & -\frac{h^2}{6}\Big(-\tilde{B} + 3\big(\tilde{\mathcal{B}}_1 + \tilde{\mathcal{B}}_{-1}\big)\Big)\,(\Lambda^{-1}\tilde{B})^2\mathrm{z}(x) \\ & -\frac{h^3}{24}\Big(\tilde{Q}(\Lambda^{-1}\tilde{B}) + 4\big(\tilde{\mathcal{B}}_1 - \tilde{\mathcal{B}}_{-1}\big)\Big)\,(\Lambda^{-1}\tilde{B})^3\mathrm{z}(x) + \mathcal{O}(h^4). \end{aligned} \tag{60}$$

Example 4.1 (pointwise approximation of the source term (AHO1-UP))**.** Fixing, for the differential terms, the upwind approximation $\tilde{Q} = \mathrm{diag}(|\lambda_1|, |\lambda_2|)$, the basic scheme gives a first-order approximation even on the stationary solution:

$$\tilde{B}_{up}^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \tilde{B}_{up}^{0} = \begin{bmatrix} \tilde{b}_{11} & \tilde{b}_{12} \\ \tilde{b}_{21} & \tilde{b}_{22} \end{bmatrix}, \quad \tilde{B}_{up}^{1} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{61}$$

Example 4.2 (upwinding of the source term (AHO2-ROE), [13])**.** An improvement on the previous example is given by a second order upwinding scheme:

$$\begin{gathered} \tilde{B}_{roe}^{-1} = \frac{1}{2}\begin{bmatrix} H(\lambda_1)\tilde{b}_{11} & H(\lambda_1)\tilde{b}_{12} \\ H(\lambda_2)\tilde{b}_{21} & H(\lambda_2)\tilde{b}_{22} \end{bmatrix}, \quad \tilde{B}_{roe}^{0} = \frac{1}{2}\begin{bmatrix} \tilde{b}_{11} & \tilde{b}_{12} \\ \tilde{b}_{21} & \tilde{b}_{22} \end{bmatrix}, \\ \tilde{B}_{roe}^{1} = \frac{1}{2}\begin{bmatrix} (1-H(\lambda_1))\tilde{b}_{11} & (1-H(\lambda_1))\tilde{b}_{12} \\ (1-H(\lambda_2))\tilde{b}_{21} & (1-H(\lambda_2))\tilde{b}_{22} \end{bmatrix}, \end{gathered} \tag{62}$$

where $H(\cdot)$ is the Heaviside function and where, as above, $\tilde{Q} = \mathrm{diag}(|\lambda_1|, |\lambda_2|)$.

A more accurate scheme is given by our AHO4 scheme, which is fourth order on the stationary solution, as soon as all terms between round brackets in (60) vanish. The AHO4-coefficients $\tilde{\mathcal{B}}_{-1}$, $\tilde{\mathcal{B}}_0$, $\tilde{\mathcal{B}}_1$ are then uniquely defined by

$$\left\{ \begin{array}{l} \tilde{\mathcal{B}}_{-1} = \tilde{B}/6 + \tilde{Q}\Lambda^{-1}\tilde{B}/8, \\ \tilde{\mathcal{B}}_{0} = 2\tilde{B}/3 - h\tilde{Q}(\Lambda^{-1}\tilde{B})^2/4, \\ \tilde{\mathcal{B}}_{1} = \tilde{B}/6 - \tilde{Q}\Lambda^{-1}\tilde{B}/8, \end{array} \right. \tag{63}$$

or equivalently

$$\begin{aligned} \tilde{\mathcal{B}}_{aho}^{-1} &= \frac{1}{8}\begin{bmatrix} \tilde{b}_{11}\big(4/3 + \tilde{q}_1/\lambda_1\big) & \tilde{b}_{12}\big(4/3 + \tilde{q}_1/\lambda_1\big) \\ \tilde{b}_{21}\big(4/3 + \tilde{q}_2/\lambda_2\big) & \tilde{b}_{22}\big(4/3 + \tilde{q}_2/\lambda_2\big) \end{bmatrix}, \\ \tilde{\mathcal{B}}_{aho}^{0} &= \frac{2}{3}\begin{bmatrix} \tilde{b}_{11} & \tilde{b}_{12} \\ \tilde{b}_{21} & \tilde{b}_{22} \end{bmatrix} + \frac{adh}{4\lambda_1^2\lambda_2^2}\begin{bmatrix} \lambda_2\tilde{q}_1\tilde{b}_{11} & \lambda_2\tilde{q}_1\tilde{b}_{12} \\ \lambda_1\tilde{q}_2\tilde{b}_{21} & \lambda_1\tilde{q}_2\tilde{b}_{22} \end{bmatrix}, \\ \tilde{\mathcal{B}}_{aho}^{1} &= \frac{1}{8}\begin{bmatrix} \tilde{b}_{11}\big(4/3 - \tilde{q}_1/\lambda_1\big) & \tilde{b}_{12}\big(4/3 - \tilde{q}_1/\lambda_1\big) \\ \tilde{b}_{21}\big(4/3 - \tilde{q}_2/\lambda_2\big) & \tilde{b}_{22}\big(4/3 - \tilde{q}_2/\lambda_2\big) \end{bmatrix}. \end{aligned}$$

The scheme is of fourth order for long times, as expressed in the following proposition:

Proposition 4.3 (Asymptotic high order)**.** *Let* z *be a smooth stationary solution of system (14). Let* w_0 *be a small and smooth perturbation of* z*.*

Then, there exists a constant C_4 *such that the local truncation error* $T(x,t)$ *of the scheme (14)-(63) satisfies:*

$$\|T(.,t)\|_\infty \leq C_4 \min\left(\Delta t + h, h^4 + \frac{1}{\Delta t\sqrt{t}}\right). \tag{64}$$

Proof. This result is an easy consequence of the uniform boundedness property (52) and of Theorem 2.5: there exists a constant $K > 0$ such that for all $t > 0$:

$$\|\mathrm{w}(.,t) - \mathrm{z}\|_\infty \leq \frac{K}{\sqrt{t}}.$$

Hence, using the notation (38) for w and z, we can write:

$$\|T(.,t)\|_\infty \leq \frac{\|\mathrm{w}(.,t+\Delta t) - \mathrm{z}\|_\infty}{\Delta t} + \|T_{stat}\|_\infty + \frac{\|S_{\Delta t}(\overline{Z} - \overline{W})\|_\infty}{\Delta t}.$$

So, we find

$$\|T(.,t)\|_\infty \leq \frac{K}{\Delta t\sqrt{t+\Delta t}} + C' h^4 + \frac{K}{\Delta t\sqrt{t}}.$$

To conclude, it is enough to use the consistency result (48) in the proof of proposition 3.1. □

We address now the problem of the monotonicity of our schemes. Clearly the schemes AHO1-UP, given by (61), and AHO2-ROE, given by (62), are easily seen to be monotone under the standard upwind CFL condition. For the scheme AHO4 we have the following result.

Proposition 4.4 (Monotonicity)**.** *The scheme (14), coupled with (63) satisfies the monotonicity requirements (41 - 43) under the following assumptions:*

$$h < \min_{1\leq i\leq 2} \frac{3|\lambda_i|}{|\tilde{b}_{ii}|}, \tag{65}$$

$$h \leq \left|\frac{2\lambda_1\lambda_2}{ad}\right| \quad \text{if} \quad \min(a\lambda_1, a\lambda_2) < 0, \tag{66}$$

$$1 + h\frac{|\tilde{b}_{ii}|}{3(4|\lambda_i| + h|\tilde{b}_{ii}|)} \leq \frac{\tilde{q}_i}{|\lambda_i|} \leq \frac{4}{3}, \quad i = 1, 2, \tag{67}$$

and

$$\frac{\Delta t}{h} \leq \min_{1\leq i\leq 2}\left(\tilde{q}_i + h|\tilde{b}_{ii}|\left(\frac{adh\tilde{q}_i}{4\lambda_1\lambda_2\lambda_i} + \frac{2}{3}\right)\right)^{-1}. \tag{68}$$

Moreover these conditions are not empty.

Proof. We recall that for system (14) we have $\tilde{b}_{11}, \tilde{b}_{22} < 0$, $\tilde{b}_{12}, \tilde{b}_{21} > 0$ and that $\lambda_1\lambda_2 \neq 0$.

We begin with condition (43). It can be written as

$$\begin{cases} \dfrac{\lambda_i}{2h} + \dfrac{\tilde{b}_{ii}}{6} + \dfrac{\tilde{q}_i}{\lambda_i}\left(\dfrac{\lambda_i}{2h} + \dfrac{\tilde{b}_{ii}}{8}\right) \geq 0, \\ -\dfrac{\lambda_i}{2h} + \dfrac{\tilde{b}_{ii}}{6} + \dfrac{\tilde{q}_i}{\lambda_i}\left(\dfrac{\lambda_i}{2h} - \dfrac{\tilde{b}_{ii}}{8}\right) \geq 0 \end{cases} \quad i = 1, 2. \tag{69}$$

Adding those inequalities shows that necessarily $\tilde{q}_i > 0$. Let us denote

$$r_i = \frac{|\lambda_i|}{2h}, \ \ s_i = -\frac{\tilde{b}_{ii}}{2}, \ \ \alpha_i = \frac{\tilde{q}_i}{|\lambda_i|}.$$

Conditions (69) read as

$$\begin{cases} r_i - \dfrac{s_i}{3} + \alpha_i \left(r_i - \dfrac{s_i}{4} \right) \geq 0, \\ -r_i - \dfrac{s_i}{3} + \alpha_i \left(r_i + \dfrac{s_i}{4} \right) \geq 0 \end{cases} \quad i = 1, 2.$$

Condition (65) insures that $r_i > \frac{s_i}{3} > \frac{s_i}{4}$, so that we have to verify that

$$\alpha_i \geq \frac{r_i + s_i/3}{r_i + s_i/4}, \quad i = 1, 2$$

which also writes

$$1 + h \frac{|\tilde{b}_{ii}|}{3(4|\lambda_i| + h|\tilde{b}_{ii}|)} \leq \frac{q_i}{|\lambda_i|}, \quad i = 1, 2. \tag{70}$$

Consequently, under the assumptions of the proposition, condition (43) is satisfied. Condition (41) is satisfied if and only if

$$\begin{cases} \dfrac{4}{3} \pm \dfrac{\tilde{q}_i}{\lambda_i} \geq 0, \\ \dfrac{2}{3} + \dfrac{adh}{4|\lambda_i|\lambda_j} \dfrac{\tilde{q}_i}{|\lambda_i|} \geq 0, \end{cases} \quad i = 1, 2, \quad j \neq i. \tag{71}$$

The first inequality is ensured by assumption (67). If $a\lambda_j$ is non negative, the second inequality is satisfied. Else, we have to insure that

$$h \leq \frac{8|\lambda_i \lambda_j|}{3|a|d} \frac{|\lambda_i|}{\tilde{q}_i}.$$

As $\frac{\tilde{q}_i}{|\lambda_i|} \leq \frac{4}{3}$, assumption (66) implies that this inequality is satisfied. It remains to verify that (42) holds. Let us denote

$$\mathcal{A}_i = \frac{2}{3} + \frac{adh}{4\lambda_1\lambda_2} \frac{\tilde{q}_i}{\lambda_i}.$$

This quantity is non negative and (42) can be written as

$$1 - \frac{\Delta t}{h} \left(\tilde{q}_i - h\tilde{b}_{ii}\mathcal{A}_i \right) \geq 0, \quad i = 1, 2.$$

Taking into account the facts that $\tilde{q}_i > 0$ and $\tilde{b}_{ii} < 0$, this is equivalent to the CFL restriction (68). To end the proof we remark that the sign properties insure that the conditions (65-68) are not empty. □

4.2. **A AHO scheme for the diffusion limit.** In this section we construct a numerical scheme for system (11) which is high-order accurate with respect to the parabolic asymptotic problem (6). The numerical error decay in time faster then the error of other schemes.

We consider for simplicity only the case $c = -a$, so looking at the adimensional problem satisfied by $U(x, b^{-1}t)$:

$$\begin{cases} u_t + \tilde{a}u_x + v_x = 0, \\ \\ v_t + u_x - \tilde{a}v_x = -\tilde{d}v, \end{cases} \tag{72}$$

where $\tilde{a} = a/b$ and $\tilde{d} = b^{-1}d$. The Chapman-Enskog limit of system (72) is given by

$$\begin{cases} \hat{u}_t + \tilde{a}\hat{u}_x = \frac{1}{\tilde{d}}\hat{u}_{xx}, \\ \\ \hat{v} = -\frac{1}{\tilde{d}}\hat{u}_x. \end{cases} \tag{73}$$

Let $U_j^n = (u_j^n, v_j^n)$ be the numerical approximation of the analytical solution $U = (u, v)$ to problem (72). Let us consider the following finite difference approximation

$$\begin{aligned}\frac{U_j^{n+1}-U_j^n}{\Delta t} + \frac{A}{2h}\left(U_{j+1}^n - U_{j-1}^n\right) - \frac{R\tilde{Q}R^{-1}}{2h}\left(U_{j+1}^n - 2U_j^n + U_{j-1}^n\right) \\ = R\tilde{\mathcal{B}}_{-1}R^{-1}U_{j-1}^n + R\tilde{\mathcal{B}}_0R^{-1}U_j^n + R\tilde{\mathcal{B}}_1R^{-1}U_{j+1}^n.\end{aligned} \tag{74}$$

Let $\pm\lambda = \pm\sqrt{1+\tilde{a}^2}$ be the eigenvalues of matrix A and

$$R = \begin{pmatrix} 1 & 1 \\ -(\tilde{a}+\lambda) & \lambda - \tilde{a} \end{pmatrix}.$$

Taking

$$\tilde{Q} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix},$$

we have that

$$Q = R\tilde{Q}R^{-1} = \tilde{Q}.$$

We set also

$$C = R\tilde{C}R^{-1}, \;\; \mathcal{B}_k = R\tilde{\mathcal{B}}_kR^{-1}, \;\; k = -1, 0, 1.$$

By Taylor expansion, the numerical approximation (74) is consistent with system

$$U_t + AU_x - BU = h\left[\frac{Q}{2}U_{xx} + \left(\mathcal{B}_1 - \mathcal{B}_{-1}\right)U_x + CU\right] + \mathcal{O}(h^2 + \Delta t). \tag{75}$$

Going into details, for different schemes, the Chapman-Enskog limit of system (75) gives:

- for the source term pointwise approximation (AHO1-UP), given by (61), we have

$$\mathcal{B}_1 - \mathcal{B}_{-1} = 0, \quad C = 0,$$

$$\begin{cases} \hat{u}_t + \tilde{a}\hat{u}_x = \left(\frac{1}{\tilde{d}} + \frac{h\lambda}{2}\right)\hat{u}_{xx}, \\ \\ \hat{v} = -\frac{1}{\tilde{d}}\hat{u}_x, \end{cases} \tag{76}$$

- for the upwinding of the source term (AHO2-ROE), given by (62), we have

$$\mathcal{B}_1 - \mathcal{B}_{-1} = \frac{\tilde{d}}{2\lambda}\begin{pmatrix} 0 & 1 \\ 0 & -\tilde{a} \end{pmatrix}, \quad C = 0,$$

$$\begin{cases} \hat{u}_t + \tilde{a}\hat{u}_x = \left(\frac{1}{\tilde{d}} + h\frac{\lambda}{2} - h\frac{1}{2\lambda}\right)\hat{u}_{xx}, \\ \\ \hat{v} = -\frac{1}{\tilde{d}}\hat{u}_x, \end{cases} \tag{77}$$

- for the Asymptotic High Order (AHO4) given by (63), we have

$$\mathcal{B}_1 - \mathcal{B}_{-1} = \frac{\tilde{d}}{4\lambda}\begin{pmatrix} 0 & 1 \\ 0 & -\tilde{a} \end{pmatrix}, \quad C = \frac{\tilde{a}\tilde{d}^2}{4\lambda^3}\begin{pmatrix} 0 & 1 \\ 0 & -\tilde{a} \end{pmatrix},$$

$$\begin{cases} \hat{u}_t + \tilde{a}\left(1 + h\frac{\tilde{a}\tilde{d}}{\left(4\lambda^3 + h\tilde{a}^2\tilde{d}\right)}\right)\hat{u}_x = \left[\frac{4\lambda^3}{\tilde{d}\left(4\lambda^3 + h\tilde{a}^2\tilde{d}\right)} + h\frac{\lambda}{2} - h\frac{\lambda^2}{\left(4\lambda^3 + h\tilde{a}^2\tilde{d}\right)}\right]\hat{u}_{xx}, \\ \\ \hat{v} = -\frac{4\lambda^3}{\tilde{d}\left(4\lambda^3 + h\tilde{a}^2\tilde{d}\right)}\hat{u}_x. \end{cases} \tag{78}$$

Notice that, for all schemes the consistency with the parabolic system (73) is of order $\mathcal{O}(h)$. In all cases, the diffusive term has a perturbation of order $\mathcal{O}(h)$. In particular, the pointwise approximation is the most diffusive one. Moreover, we observe that the AHO4 approximation has a perturbation of order $\mathcal{O}(h)$ also on the drift term.

According to this discussion, it is worth looking for a scheme which is at least of order $\mathcal{O}(h^2)$ with respect to the parabolic problem (73). Consider system (75) with $Q = \tilde{Q} = diag(q, q)$,

$$\begin{cases} u_t + \tilde{a}u_x + v_x = & h\left[\frac{q}{2}u_{xx} + (\beta^1_{11} - \beta^{-1}_{11})u_x + (\beta^1_{12} - \beta^{-1}_{12})v_x + c_{11}u + c_{12}v\right] \\ & +\mathcal{O}(h^2 + \Delta t), \\ v_t + u_x - \tilde{a}v_x + \tilde{d}v = & h\left[\frac{q}{2}v_{xx} + (\beta^1_{21} - \beta^{-1}_{21})u_x + (\beta^1_{22} - \beta^{-1}_{22})v_x + c_{21}u + c_{22}v\right] \\ & +\mathcal{O}(h^2 + \Delta t). \end{cases} \tag{79}$$

Assuming enough regularity on the data, the decay properties of the solution, stated in Theorem 2.6 allow us to consider, when t goes to $+\infty$, the following reduced system,

$$\begin{cases} u_t + \tilde{a}u_x + v_x = h\left[\frac{q}{2}u_{xx} + (\beta^1_{11} - \beta^{-1}_{11})u_x + (\beta^1_{12} - \beta^{-1}_{12})v_x + c_{11}u + c_{12}v\right], \\ u_x + \tilde{d}v = h\left[(\beta^1_{21} - \beta^{-1}_{21})u_x + c_{21}u + c_{22}v\right]. \end{cases} \tag{80}$$

Then, choosing $c_{11} = c_{21} = 0$ and setting $\gamma_{ij} = \beta^1_{ij} - \beta^{-1}_{ij}$, $i, j = 1, 2$, from the second equation we have

$$v_x = \frac{1 - h\gamma_{21}}{-\tilde{d} + hc_{22}} u_{xx}.$$

Replacing this expression in the first equation and neglecting second order terms, we get

$$u_t + \left[\tilde{a} - h\left(\gamma_{11} - c_{12}\frac{1 - h\gamma_{21}}{\tilde{d} - hc_{22}}\right)\right]u_x = \left[\frac{1 - h\gamma_{21}}{\tilde{d} - hc_{22}}(1 - h\gamma_{12}) + h\frac{q}{2}\right]u_{xx} + \mathcal{O}(h^2 + \Delta t),$$

We now look for coefficients $c_{\cdot}$, and $\gamma_{\cdot}$ such that the drift and the diffusion terms are equal to $\tilde{a}$ and $1/\tilde{d}$ respectively.

- studying the drift term, we have

$$\tilde{a}' = \tilde{a} + h\left(c_{12}\frac{1 - h\gamma_{21}}{\tilde{d} - hc_{22}} - \gamma_{11}\right) = \tilde{a} + h\frac{c_{12}(1 - h\gamma_{21}) - \gamma_{11}(\tilde{d} - hc_{22})}{\tilde{d} - hc_{22}}. \tag{81}$$

  Then, to have $c_{\cdot}$, and $\gamma_{\cdot}$ independent of h, we choose

$$c_{12} = \gamma_{11}\tilde{d}, \quad c_{22} = \gamma_{21}\tilde{d}.$$

- studying the diffusion term, we have

$$\tilde{d}' = \frac{1 - h\gamma_{21}}{\tilde{d} - hc_{22}}(1 - h\gamma_{12}) + h\frac{q}{2} = \frac{2(1 - h\gamma_{21})(1 - h\gamma_{12}) + hq(\tilde{d} - hc_{22})}{2(\tilde{d} - hc_{22})}. \tag{82}$$

  We then write,

$$\tilde{d}' = \frac{1}{\tilde{d}} + \tilde{d}'',$$

  therefore,

$$\tilde{d}'' = \frac{h(-2\tilde{d}(\gamma_{12} + \gamma_{21}) + 2c_{22} + \tilde{d}^2q) + h^2(2\tilde{d}\gamma_{21}\gamma_{12} - \tilde{d}qc_{22})}{2\tilde{d}(\tilde{d} - hc_{22})}.$$

For $c_{22} = \gamma_{21}\tilde{d}$, we obtain that $\tilde{d}'' = 0$ if

$$\gamma_{12} = \frac{\tilde{d}q}{2}.$$

To summarize, a class of finite different schemes of the form (74), consistent of order $\mathcal{O}(h^2 + \Delta t)$ with respect the asymptotic problem (73) is defined by selecting

$$c_{11} = c_{21} = 0, \quad c_{12} = \gamma_{11}\tilde{d}, \quad c_{22} = \gamma_{21}\tilde{d}, \quad \gamma_{12} = \frac{\tilde{d}q}{2}. \tag{83}$$

There are still some parameters to define, which will be chosen to guarantee the monotonicity properties of the scheme. Let us focus on the particular case $q = \lambda$ (upwind approximation for the first order derivatives), which arises if we look for the minimal artificial diffusion in the scheme.

Proposition 4.5 (Monotonicity for the AHO2p scheme)**.** *Let $q = \lambda$ and set $a_\pm = \lambda \pm \tilde{a}$. The scheme* (33) *with coefficients $\tilde{\mathcal{B}}_{-1}$, $\tilde{\mathcal{B}}_0$, $\tilde{\mathcal{B}}_1$ defined by*

$$\tilde{\mathcal{B}}_0 = \frac{\tilde{d}}{4\lambda}\begin{pmatrix} -2a_+(\lambda^2+1) & a_- \\ a_+ & -2a_-(\lambda^2+1) \end{pmatrix},$$

$$\tilde{\mathcal{B}}_{\pm 1} = \tfrac{1}{2}\left(\tilde{B} - \tilde{\mathcal{B}}_0 \pm \tilde{\Gamma} + h\tilde{C}\right), \tag{84}$$

$$\tilde{\Gamma} = \tfrac{\tilde{d}q}{2}\begin{pmatrix} -a_+ & 0 \\ 0 & a_- \end{pmatrix}, \quad \tilde{C} = \tfrac{\tilde{d}^2 q}{4\lambda}\begin{pmatrix} a_+ & -a_- \\ -a_+ & a_-, \end{pmatrix},$$

satisfies the monotonicity assumptions (41 - 43) under the conditions:

$$h < \frac{4}{\tilde{d}}, \quad \Delta t \le \frac{2\lambda h}{2\lambda^2 + h\tilde{d}(\lambda + \tilde{a})(\lambda^2+1)}. \tag{85}$$

Proof. Let us consider system (72) in diagonal form

$$\begin{cases} w_t - \lambda w_x = \frac{\tilde{d}}{2\lambda}\left[-(\tilde{a}+\lambda)w + (\lambda - \tilde{a})z\right], \\ \\ z_t + \lambda z_x = \frac{\tilde{d}}{2\lambda}\left[(\tilde{a}+\lambda)w - (\lambda - \tilde{a})z\right], \end{cases} \tag{86}$$

and the general numerical finite difference approximation (33). Starting from relations (83), we then construct the AHO2p-coefficients $\tilde{\mathcal{B}}_{-1}$, $\tilde{\mathcal{B}}_0$, $\tilde{\mathcal{B}}_1$ (84) to satisfy conditions (41 - 43).

First of all, we look at condition (41). We have to prove that

$$\tilde{\beta}^k_{ij} \ge 0, \text{ for } i \neq j.$$

Clearly, from consistency (40) and setting $\tilde{\Gamma} = \tilde{\mathcal{B}}_1 - \tilde{\mathcal{B}}_{-1}$, we have

$$\tilde{\mathcal{B}}_{\pm 1} = \frac{1}{2}\left(\tilde{B} - \tilde{\mathcal{B}}_0 \pm \tilde{\Gamma} + h\tilde{C}\right). \tag{87}$$

Since h goes to 0, in practice we have to impose that

$$\tilde{\beta}_{ij} - |\tilde{\gamma}_{ij}| > \tilde{\beta}^0_{ij} \ge 0, \text{ for } i \neq j. \tag{88}$$

Setting $a_\pm = \lambda \pm \tilde{a}$, we can write

$$\tilde{B} = -\frac{\tilde{d}}{2\lambda}\begin{pmatrix} a_+ & -a_- \\ -a_+ & a_- \end{pmatrix}.$$

So, to verify the inequalities (88), it is enough to show

$$\frac{\tilde{d}}{2\lambda}a_- > |\tilde{\gamma}_{12}|, \ \frac{\tilde{d}}{2\lambda}a_+ > |\tilde{\gamma}_{21}|, \tag{89}$$

taking then the off-diagonal values of $\tilde{\mathcal{B}}_0$ small enough, for instance

$$\tilde{\beta}^0_{ij} = \frac{1}{2}\left(\tilde{\beta}_{ij} - |\tilde{\gamma}_{ij}|\right), \text{ for } i \neq j. \tag{90}$$

Now we can compute $\tilde{\Gamma}$. We have

$$\begin{aligned}\tilde{\Gamma} &= R^{-1}\Gamma R = \frac{1}{2\lambda}\begin{pmatrix} a_- & -1 \\ a_+ & 1 \end{pmatrix}\begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -a_+ & a_- \end{pmatrix} \\ &= \frac{1}{2\lambda}\begin{pmatrix} a_-(\gamma_{11} - a_+\gamma_{12}) - (\gamma_{21} - a_+\gamma_{22}) & a_-(\gamma_{11} + a_-\gamma_{12}) - (\gamma_{21} + a_-\gamma_{22}) \\ a_+(\gamma_{11} - a_+\gamma_{12}) + (\gamma_{21} - a_+\gamma_{22}) & a_+(\gamma_{11} + a_-\gamma_{12}) + (\gamma_{21} + a_-\gamma_{22}) \end{pmatrix}.\end{aligned}$$

So inequalities (89) read now

$$\begin{aligned}\tilde{d}a_- &> |a_-(\gamma_{11} + a_-\gamma_{12}) - (\gamma_{21} + a_-\gamma_{22})|, \\ \tilde{d}a_+ &> |a_+(\gamma_{11} - a_+\gamma_{12}) + (\gamma_{21} - a_+\gamma_{22})|.\end{aligned} \tag{91}$$

The best we can do is to put to 0 the right hand side in both the inequalities, which gives, recalling that $\gamma_{12} = \frac{\tilde{d}q}{2}$:

$$-a_-(\gamma_{11} - \gamma_{22}) + \gamma_{21} = a_-^2\frac{\tilde{d}q}{2},$$

$$a_+(\gamma_{11} - \gamma_{22}) + \gamma_{21} = a_+^2\frac{\tilde{d}q}{2}.$$

Solving the linear system, we find

$$\gamma_{11} - \gamma_{22} = \tilde{d}q\tilde{a}, \ \gamma_{21} = \frac{\tilde{d}q}{2} = \gamma_{12}. \tag{92}$$

We can quantify the influence of the matrix $\tilde{C}$ in the monotonicity. Let us just consider the parabolic case. It is clear, that since the limit $h \to 0$ is relevant, we can start from conditions (92). Now we have that

$$\begin{aligned}\tilde{C} &= R^{-1}CR = \frac{\tilde{d}}{2\lambda}\begin{pmatrix} a_- & -1 \\ a_+ & 1 \end{pmatrix}\begin{pmatrix} 0 & \gamma_{11} \\ 0 & \gamma_{21} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -a_+ & a_- \end{pmatrix} \\ &= \frac{\tilde{d}}{2\lambda}\begin{pmatrix} -a_+a_-\gamma_{11} + a_+\gamma_{21} & a_-^2\gamma_{11} - a_-\gamma_{21} \\ -a_+^2\gamma_{11} - a_+\gamma_{21} & a_+a_-\gamma_{11} + a_-\gamma_{21} \end{pmatrix}.\end{aligned}$$

Then inequalities (88), can be refined to

$$\tilde{\beta}_{ij} - |\tilde{\gamma}_{ij}| + h\tilde{c}_{ij} > \tilde{\beta}^0_{ij} \geq 0, \text{ for } i \neq j. \tag{93}$$

Using conditions (92), we need to verify both

$$1 + \frac{h}{2\lambda}(a_-\gamma_{11} - \gamma_{21}) > 0, \ 1 - \frac{h}{2\lambda}(a_+\gamma_{11} + \gamma_{21}) > 0, \tag{94}$$

which are equivalent to

$$1 - \frac{h}{4\lambda}\tilde{d}q \pm \frac{h}{2\lambda}a_\mp\gamma_{11} > 0. \tag{95}$$

So, the optimal choice is given by

$$\gamma_{11} = 0, \ \gamma_{22} = -\tilde{d}q\tilde{a}, \ \gamma_{21} = \gamma_{12} = \frac{\tilde{d}q}{2} \tag{96}$$

with the restriction on the space step

$$h < \frac{4\lambda}{\tilde{d}q}. \tag{97}$$

To summarize, up to now we have defined

$$\Gamma = \frac{\tilde{d}q}{2}\begin{pmatrix} 0 & 1 \\ 1 & -2\tilde{a} \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 \\ 0 & \gamma_{21}\tilde{d}, \end{pmatrix}.$$

From these two relations we obtain in diagonal variables, $\tilde{\Gamma}$ and $\tilde{C}$ defined by (84).

To complete the construction of the scheme and the study of monotonicity, we need to verify condition (42) and (43). In this case, they are given by

$$1 - \frac{\Delta t}{h} q + \Delta t \tilde{\beta}^0_{ii} \geq 0, \quad i = 1, 2, \tag{98}$$

$$\frac{\lambda}{2h} + \frac{q}{2h} + \tilde{\beta}^{-1}_{ii} \geq 0, \ -\frac{\lambda}{2h} + \frac{q}{2h} + \tilde{\beta}^{1}_{ii} \geq 0, \quad i = 1, 2. \tag{99}$$

From (87), we have

$$\begin{aligned}
\tilde{\beta}^{-1}_{11} &= \frac{1}{2}\left(\tilde{\beta}_{11} - \tilde{\beta}^0_{11} - \tilde{\gamma}_{11} + h\tilde{c}_{11}\right) = \frac{\tilde{d}}{4}a_+\left(q - \frac{1}{\lambda}\right) - \frac{\tilde{\beta}^0_{11}}{2}, \\
\tilde{\beta}^{-1}_{22} &= \frac{1}{2}\left(\tilde{\beta}_{22} - \tilde{\beta}^0_{22} - \tilde{\gamma}_{22} + h\tilde{c}_{22}\right) = -\frac{\tilde{d}}{4}a_-\left(q + \frac{1}{\lambda}\right) - \frac{\tilde{\beta}^0_{22}}{2}, \\
\tilde{\beta}^{1}_{11} &= \frac{1}{2}\left(\tilde{\beta}_{11} - \tilde{\beta}^0_{11} + \tilde{\gamma}_{11} + h\tilde{c}_{11}\right) = -\frac{\tilde{d}}{4}a_+\left(q + \frac{1}{\lambda}\right) - \frac{\tilde{\beta}^0_{11}}{2}, \\
\tilde{\beta}^{1}_{22} &= \frac{1}{2}\left(\tilde{\beta}_{22} - \tilde{\beta}^0_{22} + \tilde{\gamma}_{22} + h\tilde{c}_{22}\right) = \frac{\tilde{d}}{4}a_-\left(q - \frac{1}{\lambda}\right) - \frac{\tilde{\beta}^0_{22}}{2}.
\end{aligned}$$

From (99), the coefficient $\tilde{\beta}^0_{11}$ has to verify the relations,

$$\begin{aligned}
&\frac{1}{h}(q+\lambda) + \frac{\tilde{d}}{2}a_+(q - \frac{1}{\lambda}) - \tilde{\beta}^0_{11} \geq 0, \\
&\frac{1}{h}(q-\lambda) - \frac{\tilde{d}}{2}a_+(q + \frac{1}{\lambda}) - \tilde{\beta}^0_{11} \geq 0,
\end{aligned} \tag{100}$$

and $\tilde{\beta}^0_{22}$,

$$\begin{aligned}
&\frac{1}{h}(q+\lambda) - \frac{\tilde{d}}{2}a_-(q + \frac{1}{\lambda}) - \tilde{\beta}^0_{22} \geq 0, \\
&\frac{1}{h}(q-\lambda) + \frac{\tilde{d}}{2}a_-(q - \frac{1}{\lambda}) - \tilde{\beta}^0_{22} \geq 0.
\end{aligned} \tag{101}$$

For $q = \lambda$, a good choice for $\tilde{\beta}^0_{11}$ and $\tilde{\beta}^0_{22}$ should be given by

$$\tilde{\beta}^0_{11} = -\frac{\tilde{d}}{2}a_+(\frac{\lambda^2+1}{\lambda}), \quad \tilde{\beta}^0_{22} = -\frac{\tilde{d}}{2}a_-(\frac{\lambda^2+1}{\lambda}). \tag{102}$$

Substituting these two expression in condition (98), we get the following restriction on the time step Δt,

$$\begin{aligned}
&1 - \frac{\Delta t}{h}\lambda - \frac{\Delta t}{2}\tilde{d}a_+(\frac{\lambda^2+1}{\lambda}) \geq 0, \\
&1 - \frac{\Delta t}{h}\lambda - \frac{\Delta t}{2}\tilde{d}a_-(\frac{\lambda^2+1}{\lambda}) \geq 0.
\end{aligned} \tag{103}$$

Therefore,

$$\Delta t \leq \frac{2\lambda h}{2\lambda^2 + h\tilde{d}a_+(\lambda^2+1)}. \tag{104}$$

To conclude, from relations (90) and (102) we get the AHO2p-coefficients $\tilde{\mathcal{B}}_{-1}$, $\tilde{\mathcal{B}}_0$, $\tilde{\mathcal{B}}_1$ defined by (84), under the monotonicity assumptions (97) with $q = \lambda$ and (104). □

5. Numerical tests

In this Section our aim is to show how, for large time simulations, AHO schemes give better numerical results than standard approximations. Therefore, we focus our attention on the numerical error as a function of time: for all tests, we fix the grid steps $h = \mathcal{O}(10^{-2})$ and Δt satisfying the CFL conditions, and we plot the error as the time $t = n\Delta t$ increases.

For all schemes under consideration we set $\tilde{Q} = \mathrm{diag}(|\lambda_1|, |\lambda_2|)$. The reference solution is obtained by the Roe-type scheme AHO2-ROE (62), with $h = \mathcal{O}(10^{-3})$.

5.1. **Test n.1.** Let us consider system (11) around the stationary and time-asymptotic stable solutions given by

$$u = u_0 - bv_0 e^{px}/a, \quad v = v_0 e^{px}, \tag{105}$$

for some constants u_0, v_0 and for $p = -ad/(ac - b^2)$.

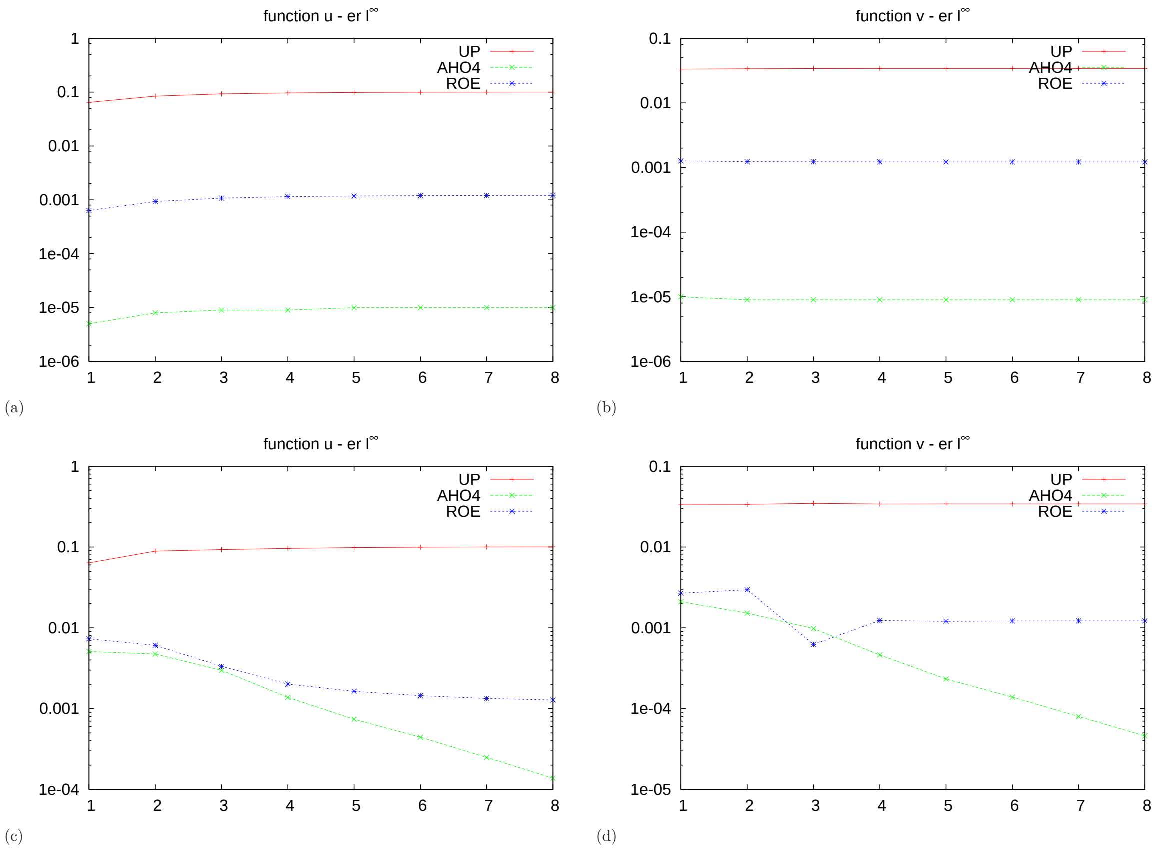
First, we compare AHO4 scheme, with the standard first-order pointwise upwind scheme (61), that is actually just an AHO1-UP scheme, and with the AHO2-ROE scheme (62). The parameters in the test are $a = 1$, $c = -a$, $b = 1$ and $d = 5$. In Figure 1, we plot the l^∞-error for problem as a function of time. Remark that, even if the total time of simulation is $T = 5$, this interval is divided in 8 sub-intervals to plot the error curve.

In figures (a) and (b) we have the evolution of the error, when the initial data are taken to be the stationary solution. In this case all the schemes do not exactly preserve this stable state, but anyway AHO schemes show a better accuracy, which is proportional to their formal asymptotic order. In figures (c) and (d) we show the evolution of the error when the initial data are given by small compactly supported perturbations of the stationary solutions. In that case, the errors evolve in time, and the AHO2-ROE and the AHO4 improve clearly their accuracy for large times. Moreover, we show that the AHO4 scheme gives a good approximation also at intermediate times, when the perturbation is still present in the solution. In figure 2, we plot the solution U as obtained by the AHO4 scheme at time $T = 2.5$ (see the bottom figure 2). To see the properties of the AHO4 scheme, we plot at the top of these plots, the pointwise difference between the approximates solutions, given by schemes AHO1-UP, AHO2-ROE and AHO4, and the reference solution, which is obtained by the Roe-type scheme AHO2-ROE (62), with $h = \mathcal{O}(10^{-3})$. Again, the AHO4 scheme is better performing.

5.2. **Test n.2.** Let us consider for system (11), the constant equilibrium state $u = 1$ and $v = 0$. The parameters in the test are again $a = 1$, $c = -a$, $b = 1$ and $d = 5$. We consider a small compactly supported perturbation of this constant solution as initial data, and we expect a diffusive behavior of the solution, near the corresponding solution of problem (73). Actually what we look for is always a better approximation of the hyperbolic problem, but to obtain that, we need to force the high-order consistency with the main terms in the asymptotic behavior, which are given by the diffusive expansion.

More specifically, here we compare the AHO2p scheme (84), with the standard first-order pointwise upwind scheme (61), and with the AHO2-ROE scheme (62). The reference solution is always obtained by the AHO2-ROE scheme with $h = \mathcal{O}(10^{-3})$.

In Figures 3-(a) and 3-(b), we plot the different approximations of the function u at times $T = 43$ and $T = 350$. In Figures 3-(c) and 3-(d), we plot respectively the l^1 and l^∞ errors as a function of time. Again, even if the total time of simulation is $T = 350$, this interval is divided in 8 sub-intervals to plot the error curve.



(a) (b)

(c) (d)

FIGURE 1. Test n. 1. (a)-(b) evolution of the l^∞-error, when the initial data are taken to be the stationary solutions (105). (c)-(d) evolution of the error, when the initial data are given by a small compactly supported perturbation of a stationary solution.

The numerical results show a clear better performance of the AHO2p-scheme in particular for large times, as expected by our asymptotic analysis. To end, in Figure 4 we compare all the considered schemes with the solution of the diffusion limit problem (73).

Actually this better performance is not surprising, and can be explained in terms of a better time decay of the error for the AHO2p-scheme. We denote respectively by $\hat{u}^P$ for the analytical solution of the parabolic problem (73) and by u^H the solution of (72). From Theorem (2.6), we know that

$$\|(\hat{u}^P - u^H)(t)\|_{L^\infty} \leq Ct^{-1}, \tag{106}$$

for a positive constant C.
By construction, for large time, the scheme AHO2p closely follows the parabolic asymptotic problem (73), and so we expect that the same decay property (106) should then hold for the error function, namely

$$e(t) = \|(U^h_{AHO2p} - u^H)(t)\|_{L^\infty} \leq Ct^{-1}, \tag{107}$$

where U^h_{AHO2p} is the numerical solution of (72) obtained with the AHO2p scheme (84).

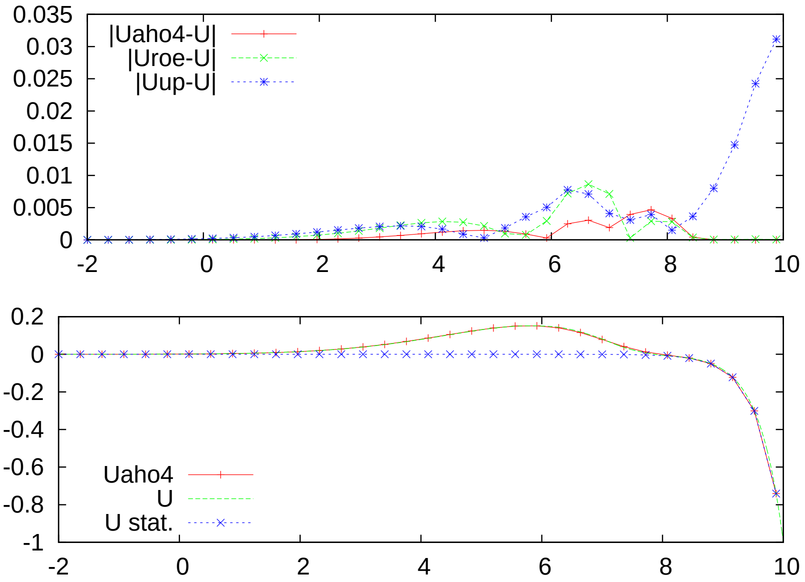


Figure 2. Test n. 1. Bottom: plot of the solution U as obtained by the AHO4 scheme at time $T = 2.5$, starting from a small perturbation of the stationary solution. Top: pointwise difference between the approximates solutions, given respectively by schemes AHO1-UP, AHO2-ROE and AHO4, and the reference solution

Then we look, from numerical tests, for a constant C and an exponent of decay γ which best fit the equality

$$e(t) = \|(u^H - U^h)(t)\|_{L^\infty} = Ct^{-\gamma}, \tag{108}$$

for different numerical approximations U^h.
Using two instants of time t_1 and t_2, to evaluate γ, we have the following estimate

$$\gamma \approx \frac{\ln\left[e(t_1)/e(t_2)\right]}{\ln[t_2/t_1]},$$

and C is estimated by

$$C \approx t_1^\gamma e(t_1) \approx t_2^\gamma e(t_2).$$

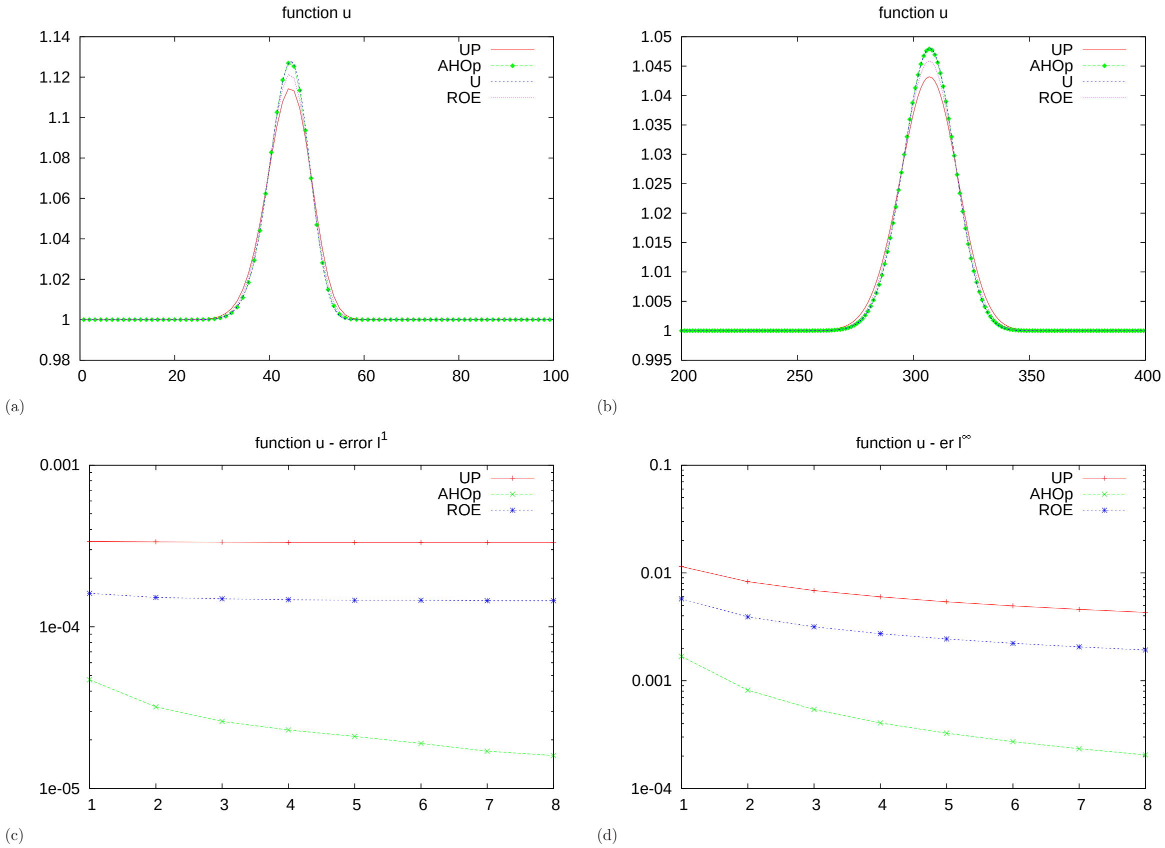
In Table 1, we show the values obtained for different schemes, where u^H is our reference solution of (72), always obtained by the AHO2-ROE scheme with a fine grid, to say $h = \mathcal{O}(10^{-3})$, and, where the solution $\hat{u}^P$ of (73), has been computed by Crank-Nicolson scheme with the same fine grid. The table shows that the AHO2p-scheme, unlike all the other schemes, closely follows the theoretical time decay behavior of the analytical solution.

Acknowledgments

We want to thanks warmly Bernard Hanouzet, in particular for some useful discussions and his sometimes irritating questions, which as usual have been decisive to develop the scheme AHO2p. The research activity reported in this paper has been partially conducted with the support of the project "Problemi iperbolici non lineari" for the year 2006 of INDAM-GNAMPA.

	$\Vert(u-u^H)(t)\Vert_{L^\infty}=Ct^{-\gamma}$	
u	γ	C
$\hat{u}^P$	0.942868	0.385312
AHO2p	0.934315	0.050421
UP	0.480781	0.071943
ROE	0.497644	0.035540

TABLE 1. Evaluation of constants γ and C for relation (108), where the solution u is given, first by the approximation of $\hat{u}^P$ computed by Crank-Nicolson scheme with a fine grid and then respectively by the schemes AHO2p (84), AHO1-UP (61), and AHO2-ROE 62.



(a) (b)

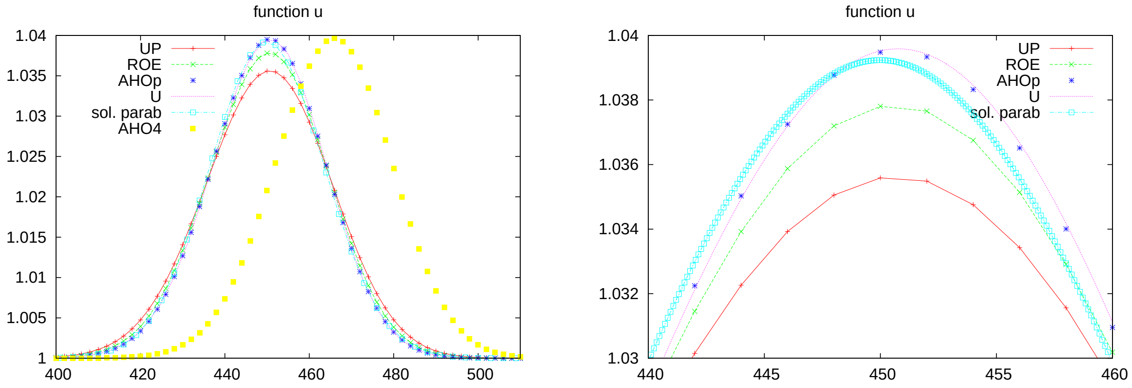
(c) (d)

FIGURE 3. Test n. 2. (a)-(b) plot of different approximations of the function u at times $T = 43$ and $T = 350$. (c)-(d) time evolution of the l^1 and l^∞ errors. for the different approximations.

REFERENCES

[1] A. Bermudez and M.E. Vasquez. Upwind methods for hyperbolic conservation laws with source terms. *Comput. and Fluids*, 23(8):1049–1071, 1994.

[2] S. Bianchini, B. Hanouzet, R. Natalini, Asymptotic behavior of smooth solutions for partially dissipative hyperbolic systems with a convex entropy, IAC Report 79 (11/2005); to appear in *Comm. Pure and Appl. Math.*; PDF file available at `http://www.iac.rm.cnr.it/~natalini/ps/BRSprint.pdf`



(a)

(b)

Figure 4. Test n. 2. (a) Different approximations of the u component, against the reference solution of the hyperbolic problem (11) and of the parabolic problem (73). (b) zoom on the top of the solution, to magnify the different performances.

[3] R. Botchorishvili, B. Perthame, and A. Vasseur. Equilibrium schemes for scalar conservation laws with stiff sources. *Math. Comput.*, 72(241):131–157, 2003.

[4] F. Bouchut. *Nonlinear stability of finite volume methods for hyperbolic conservation laws*, and well-balanced schemes for sources. Birkhäuser, Frontiers in Mathematics series Birkhäuser, ISBN 3-7643-6665-6, 2004.

[5] M. Briani, R. Natalini, Asymptotic high-order schemes for integro-differential problems arising in markets with jumps, *Commun. Math. Sci.* **4** (2006), 81-96.

[6] G.-Q. Chen, C.D. Levermore, and T.-P. Liu, Hyperbolic conservation laws with stiff relaxation terms and entropy, *Comm. Pure and Appl. Math.* **47** (1994), 787–830.

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

[8] B. Hanouzet and R. Natalini, Global existence of smooth solutions for partially dissipative hyperbolic systems with a convex entropy, *Arch. Ration. Mech. Anal.* **169** (2003), no. 2, 89–117.

[9] S. Jin. A steady-state capturing method for hyperbolic systems with geometrical source terms. *Math. Model. Num. Anal.*, 35:631–646, 2001.

[10] S. Kawashima, Large-time behaviour of solutions to hyperbolic-parabolic systems of conservation laws and applications, *Proc. Roy. Soc. Edinburgh* Sect. A **106** (1987), no. 1-2, 169–194.

[11] R. Natalini, *Recent results on hyperbolic relaxation problems*, Analysis of systems of conservation laws (Aachen, 1997), Chapman & Hall/CRC, Boca Raton, FL, 1999, pp. 128–198.

[12] B. Perthame and C. Simeoni. A kinetic scheme for the Saint-Venant system with a source term. *Calcolo*, 38(4):201–231, 2001.

[13] P.L. Roe. Upwind differencing schemes for hyperbolic conservation laws with source term. *Nonlinear Hyperbolic Problems, C.Carasso, P.A.Raviart and D.Serre editors, Lecture Notes in Math.*, 1270:41–51, 1987.

[14] Y. Shizuta and S. Kawashima, Systems of equations of hyperbolic-parabolic type with applications to the discrete Boltzmann equation, *Hokkaido Math. J.* **14** (1985), no. 2, 249–275.

[15] F. Trèves, *Basic linear partial differential equations*. Pure and Applied Mathematics, Vol. 62. Academic Press, New York-London, 1975.

[16] G.B. Whitham, *Linear and nonlinear waves*, A Wiley & Sons, New York, 1974.