Finite volume schemes for nonhomogeneous scalar conservation laws : error estimate

Claire Chainais-Hillairet [1] Sylvie Champier [2]

Abstract

In this paper, we study finite volume schemes for the nonhomogeneous scalar conservation law $u_t + \mathrm{div} F(x,t,u) = q(x,t,u)$ with initial condition $u(x,0) = u_0(x)$. The source term may be either stiff or nonstiff. In both cases, we prove error estimates between the approximate solution given by a finite volume scheme (the scheme is totally explicit in the nonstiff case, semi-implicit in the stiff case) and the entropy solution. The order of these estimates is $h^{\frac{1}{4}}$ in space-time L^1-norm (h denotes the size of the mesh). Furthermore, the error estimate does not depend on the stiffness of the source term in the stiff case.

1 Introduction

1.1 Presentation of the problem

We consider the initial value problem for the scalar nonhomogeneous conservation law in N space dimensions ($N \geq 1$) :

$$(1) \qquad \left\{ \begin{array}{rcll} u_t(x,t) + \mathrm{div}(F(x,t,u(x,t))) &=& q(x,t,u(x,t)), & \forall x \in \mathbb{R}^N, \forall t \in \mathbb{R}_+, \\ u(x,0) &=& u_0(x), & \forall x \in \mathbb{R}^N \end{array} \right.$$

where

$$F \; : \; \begin{array}{rcl} \mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R} & \to & \mathbb{R}^N, \\ (x,t,s) & \mapsto & F(x,t,s) \end{array}$$

and

$$q \; : \; \begin{array}{rcl} \mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R} & \to & \mathbb{R}^N, \\ (x,t,s) & \mapsto & q(x,t,s). \end{array}$$

The purpose of this paper is to study finite volume schemes for this problem and to prove error estimates between the approximate solutions given by different numerical schemes and the entropy solution of (1).

Many problems of physics are governed by such equations, in many different topics. The source terms can appear due to physical or geometrical effects. For instance, the models

[1]UMPA-ENSLyon, 46, allée d'Italie, 69364 Lyon Cedex 7, FRANCE. e-mail : chillair@umpa.ens-lyon.fr

[2]Université de St-Etienne, 23 rue Dr P. Michelon, 42023 St-Etienne, FRANCE. e-mail : champier@anumsun1.univ-st-etienne.fr

of combustion, of gas dynamics with heat transfer, of water waves with friction force on the bottom lead to a hyperbolic system of conservation laws with source terms. T.P. Liu has been the main contributor to the theoretical study of such problems [16], especially for the case of gas dynamics with source terms providing from a varying section of a duct [17]. Here, we restrict our study to the scalar case. Such an equation can be found for instance in porous media modelling.

There is already a wide literature about numerical schemes for nonhomogeneous hyperbolic equations. Following M. Crandall and A. Majda [8], numerous authors as T. Tang and Z.-H. Teng [23, 22], A. Chalabi [2, 3], H. J. Schroll and R. Winther [20] have developed different kinds of finite difference schemes for the problem (1) with $F(x,t,s) = F(s)$ and $q(x,t,s) = q(s)$. These authors have proved convergence results or error estimates for their schemes in one space dimension.

It appears that two kinds of hypotheses on the source terms arise. Either the source terms are stiff or they are nonstiff. As it is explained by R. Pember in [18, 19] and by R. J. LeVeque and H. C. Yee in [15], the source terms are stiff when they induce a time scale that is much faster than the scale on which the solution is evolving (this slower scale is related to the characteristic speed $\frac{\partial F}{\partial s}$). For instance, the generic form of stiff source terms is described by $q(x,t,s) = \frac{1}{\delta} g(s)$ with $\delta \ll 1$. Then, it is crucial that the very small parameter δ does not affect the error estimate. Indeed, if it appeared in the error estimate, a strong limitation on the time step would follow. Therefore, in the case of stiff source terms, it is important to discretize them in an implicit way.

In the previous papers [2, 3], [20], [23, 22], the authors study finite difference schemes for the nonhomogeneous problem. But finite difference schemes are necessarily constructed on cartesian grids. The proofs of convergence and of error estimates rest on strong BV estimates verified by the approximate solutions. Because of the source terms, the BV norm of the approximate solution is increasing and the schemes are not TVD but all the same it is possible to get a locally-in-time BV estimate. This technique is really a 1-D technique and cannot be extended to any mesh in multi-D.

An other approach that is well-adapted to the multidimensional case is the discretization by finite volume schemes. It has already been done for the homogeneous conservation law ($q = 0$) by R. Eymard, T. Gallouët, R. Herbin [10, 11], [9] (with M. Ghilani) in the case where $F(x,t,s) = v(x,t)f(s)$ and by B. Cockburn, F. Coquel, P. LeFloch [6, 7], D. Kröner, M. Rokyta [12] and J.-P. Vila [24] in the case where $F(x,t,s) = F(s)$. In a former paper [1], C. Chainais considered the homogeneous problem

$$(2) \qquad \left\{ \begin{array}{rcll} u_t(x,t) + \mathrm{div}(F(x,t,u(x,t))) & = & 0, & x \in \mathbb{R}^N, t \in \mathbb{R}_+, \\ u(x,0) & = & u_0(x), & x \in \mathbb{R}^N \end{array} \right.$$

under the hypothesis $\mathrm{div}_x F = 0$. She studied finite volume schemes and proved their convergence towards the entropy solution with an error estimate of order $h^{1/4}$ in a space-time L^1-norm (h is the size of the mesh). S. Champier in [4] considered the hyperbolic equation with nonstif source terms in the case where $F(x,t,s) = v(x,t)f(s)$ (with $\mathrm{div}_x v = 0$). She also proved error estimate between the entropy solution and the approximate solution given by an explicit finite volume scheme.

The aim of the present paper is to extend these two results to (1) without the hypothesis $\mathrm{div}_x F = 0$ and with stiff or nonstiff source terms. If the source term is stiff, it is important to note that the error estimate does not depend on its stiffness.

1.2 Preliminaries

1.2.1 Hypotheses and definition

We first give the hypotheses on the data and then we define the notion of solution.

The hypotheses on u_0 and F are the following :

$$
(3) \qquad \left\{
\begin{array}{ll}
(i) & u_0 \in L^\infty(\mathbb{R}^N), \\[2mm]
(ii) & F \in C^1(\mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}) \text{ and } \dfrac{\partial F}{\partial s} \text{ is locally Lipschitzian,} \\[2mm]
(iii) & \text{for every compact } \mathcal{K} \subset \mathbb{R}, \text{ there exists } V_{\mathcal{K}} < +\infty \text{ such that} \\
& \left|\dfrac{\partial F}{\partial s}(x,t,s)\right| \leq V_{\mathcal{K}} \text{ for a.e. } (x,t,s) \in \mathbb{R}^N \times \mathbb{R}_+ \times \mathcal{K}, \\[2mm]
(iv) & \mathrm{div}_x F \text{ is locally Lipschitzian and there exist } D^0 \text{ and } D^1 \text{ such that :} \\
& \sup\limits_{\mathbb{R}^N \times \mathbb{R}_+} |\mathrm{div}_x F(x,t,0)| \leq D^0 \text{ and } \sup\limits_{\mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}} |\dfrac{\partial \mathrm{div}_x F}{\partial s}(x,t,s)| \leq D^1.
\end{array}
\right.
$$

The hypotheses on the source term q are different whether we assume the source term is nonstiff or stiff. In the nonstiff case, we only assume that :

$$
(4) \qquad \left\{
\begin{array}{ll}
(i) & q \in C^1(\mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}), \\[2mm]
(ii) & q \text{ is locally Lipschitzian and there exist } Q^0 \text{ and } Q^1 \text{ such that :} \\
& \sup\limits_{\mathbb{R}^N \times \mathbb{R}_+} |q(x,t,0)| \leq Q^0 \text{ and } \sup\limits_{\mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}} \left|\dfrac{\partial q}{\partial s}(x,t,s)\right| \leq Q^1.
\end{array}
\right.
$$

In this case, we allow a Q^1-depending error estimate. But, for stiff source terms, the quantity Q^1 may be related to a very small parameter δ with, for instance, the relationship $Q^1 = O(\delta^{-1})$. Under these circumstances, an error estimate proportionate to Q^1 is unsatisfactory.

Then, as in many theoretical papers (see [3], [5], [16], [20], [21]) , we assume that

$$(5)\qquad \begin{cases} (i)\ q \text{ only depends on } s \text{ and } q \in C^1(\mathbb{R}), \\ (ii)\ q(0) = 0 \text{ and } q' \leq 0. \end{cases}$$

It means that $u = 0$ is an equilibrium state and that the source term is dissipative in the sense of Chen, Levermore and Liu [5]. These hypotheses seem to be relevant because of the fact that they are satisfied in many practical problems. As a matter of fact, we could also allow the dependence of q upon x and t ; in this case, we would have to distinguish the influence of the Lipschitz constants of q with respect to x and t from the one with respect to s. But, as in the generic case $q(x,t,s) = \frac{1}{\delta} g(s)$, it is easier to cut out the dependence of q upon x and t.

It is now well known that the good notion of solution for a hyperbolic equation is the notion of entropy solution. Indeed, the entropy solution is the only weak solution which is physically relevant.

Definition 1 *We say that u is an entropy solution to (1) if $u \in L^\infty(\mathbb{R}^N \times]0,T[)$ for all $T > 0$ and if, for all $\kappa \in \mathbb{R}$, for all $\varphi \in C_c^1(\mathbb{R}^N \times \mathbb{R}_+, \mathbb{R}_+)$:*

$$\begin{aligned} \int_{\mathbb{R}^N \times \mathbb{R}_+} \Big[|u(x,t) - \kappa| \varphi_t(x,t) + (F(x,t,u(x,t) \top \kappa) - F(x,t,u(x,t) \bot \kappa)).\nabla \varphi(x,t) \Big] dxdt \\ - \int_{\mathbb{R}^N \times \mathbb{R}_+} sign(u(x,t) - \kappa) \Big(div_x F(x,t,\kappa) - q(x,t,u(x,t)) \Big) \varphi(x,t) dxdt \\ + \int_{\mathbb{R}^N} |u_0(x) - \kappa| \varphi(x,0) dx \geq 0, \end{aligned} \qquad (6)$$

where $a \top b$ denotes $\max\{a,b\}$ and $a \bot b$ $\min\{a,b\}$.

We recall that the $sign$ function is defined by $sign(x) = \dfrac{x}{|x|}$ for $x \neq 0$ and $sign(0) = 0$.

1.2.2 Numerical schemes

Let $\mathcal{T}$ be a mesh of $\mathbb{R}^N$ such that the common interface between two cells of $\mathcal{T}$ is included in a hyperplane of $\mathbb{R}^N$. We assume that there exist $h > 0$ and $\alpha > 0$ such that, for any $p \in \mathcal{T}$:

$$(7)\qquad \begin{aligned} \alpha h^N &\leq m(p), \\ m(\partial p) &\leq \frac{1}{\alpha} h^{N-1}, \\ \delta(p) &\leq h. \end{aligned}$$

where $m(p)$ denotes the N-dimensional Lebesgue measure of the cell p, $m(\partial p)$ denotes the $(N-1)$-dimensional Lebesgue measure of its boundary and $\delta(p)$ denotes its diameter.

With these notations, the parameter h defines the size of the mesh and α its regularity.

Let $k > 0$ be the time step and $t^n = nk$ for all $n \in \mathbb{N}$.

The discretization of the flux term does not differ whether the flux F is divergence free or not. Then, as in [1], we consider some numerical fluxes $F^n_{p,q} \in C(\mathbb{R}^2, \mathbb{R}) : (u,v) \to F^n_{p,q}(u,v)$ that verify the following assumptions of monotony, conservativity, consistency and regularity :

$$(8) \quad \begin{cases} (i) \ F^n_{p,q}(u,v) \text{ is nondecreasing with respect to } u \text{ and nonincreasing with respect to } v, \text{ for } (u,v) \in \mathbb{R}^2, \\ (ii) \ F^n_{p,q}(u,v) = -F^n_{q,p}(v,u) \text{ for all } (u,v) \in \mathbb{R}^2, \\ (iii) \ F^n_{p,q}(u,v) \text{ is Lipschitz continuous over } \mathbb{R}^2 \text{ with the same Lipschitz constant w.r.t. } u \text{ and } v : m(\sigma_{pq})M, \text{ where } M \text{ only depends on } F \text{ and } u_0, \\ (iv) \ F^n_{p,q}(s,s) = \dfrac{1}{k}\displaystyle\int_{t^n}^{t^{n+1}} \int_{\sigma_{pq}} F(\gamma,t,s).n_{p,q} d\gamma dt \text{ for all } s \in \mathbb{R}. \end{cases}$$

We note here that (8 iv) implies :

$$\sum_{q \in N(p)} F^n_{p,q}(s,s) = \frac{1}{k}\int_{t^n}^{t^{n+1}} \int_p \mathrm{div}_x F(x,t,s) dx dt, \ \forall s \in \mathbb{R}.$$

The discrete unknowns are denoted by $(u^n_p)_{p \in \mathcal{T}, n \in \mathbb{N}}$. We set :

$$(9) \qquad u^0_p = \frac{1}{m(p)}\int_p u_0(x)dx, \ p \in \mathcal{T}.$$

If the source term is nonstiff, we discretize it in an explicit way : the scheme is defined by (9) and

$$(10) \quad u^{n+1}_p - u^n_p + \frac{k}{m(p)}\sum_{q \in N(p)} F^n_{p,q}(u^n_p, u^n_q) = \frac{1}{m(p)}\int_{t^n}^{t^{n+1}}\int_p q(x,t,u^n_p)dxdt, \ p \in \mathcal{T}, \ n \in \mathbb{N}.$$

From now on, we will use the following notation :

$$q^n_p = \frac{1}{km(p)}\int_{t^n}^{t^{n+1}} \int_p q(x,t,u^n_p)dxdt.$$

If the source term is stiff, we discretize it implicitly and the scheme is then defined by (9) and

$$(11) \qquad u^{n+1}_p - u^n_p + \frac{k}{m(p)}\sum_{q \in N(p)} F^n_{p,q}(u^n_p, u^n_q) = kq(u^{n+1}_p), \ p \in \mathcal{T}, \ n \in \mathbb{N}.$$

Both schemes are studied under the same Courant-Friedrichs-Lewy condition :

$$\text{there exists } \xi \in]0,1[\text{ such that } k \leq \frac{(1-\xi)\alpha^2 h}{4M}. \tag{12}$$

The approximate solution is denoted by $u_{\mathcal{T},k}$ and is defined by :

$$u_{\mathcal{T},k}(x,t) = u_p^{n+1} \text{ for } x \in p \text{ and } t \in [t^n, t^{n+1}[. \tag{13}$$

We also set :

$$u_{\mathcal{T},0}(x) = u_p^0 \text{ for } x \in p. \tag{14}$$

In [1], the approximate solution is defined by $u_{\mathcal{T},k}(x,t) = u_p^n$ in $p \times [t^n, t^{n+1}[$. But here, because of the source term and of the fact that $\text{div}_x F \neq 0$, the function *sign* plays an important role in the definition of the entropy solution (6). This function is not smooth in 0 and we may note later that the definition of $u_{\mathcal{T},k}$ by (13) is crucial for the proof of the entropy inequalities (45) and (54).

1.3 Main results

We present here the theorems that are proved in Section 5. In Theorem 1, we give the error estimate between the approximate solution given by the explicit scheme and the entropy solution for nonstiff source terms.

Theorem 1 *We assume the data F, u_0 and q satisfy (3), (4), the mesh (7), the numerical fluxes (8) and the time step the C.F.L. condition (12). Let u be the entropy solution to :*

$$\left\{ \begin{array}{rcl} u_t(x,t) + div(F(x,t,u(x,t))) & = & q(x,t,u(x,t)), \\ u(x,0) & = & u_0(x), \end{array} \right.$$

defined by (6) and let $u_{\mathcal{T},k}$ be the approximate solution defined by the explicit scheme (9), (10), (13).
If $u_0 \in L^\infty \cap BV_{loc}(\mathbb{R}^N)$, we get the following error estimate : for every compact set $E \subset \mathbb{R}^N \times \mathbb{R}_+$, there exists K depending only on E, F, u_0, q, M, α, ξ such that :

$$\int_E |u_{\mathcal{T},k}(x,t) - u(x,t)| dxdt \leq K h^{\frac{1}{4}}. \tag{15}$$

In Theorem 1, we do not explicit the dependence of K upon q. Mainly, it may depend on the Lipschitz constant of q with respect to s. For stiff source terms, we cannot accept this dependence. Thus, Theorem 2 deals with the error estimate between the approximate solution given by the semi-implicit scheme and the entropy solution for stiff source terms. In Theorem 2, it is crucial to see that the quantity K depends on q only through $\|q(u_0)\|_{L^1(E_0)}$ and $\|q(u_{\mathcal{T},0})\|_{L^1(E_0)}$, where E_0 is a ball containing $E \cap \{t=0\}$. Then, if $\|q(u_0)\|_{L^1(E_0)}$ and $\|q(u_{\mathcal{T},0})\|_{L^1(E_0)}$ do not depend on the Lipschitz constant of q with respect to s, we get a good error estimate. Some additional assumptions may be required on u_0 ; more remarks are reported in the last section.

Theorem 2 *Assume that the data F, u_0 and q satisfy (3), (5), the mesh (7), the numerical fluxes (8) and the time step (12). Let u be the entropy solution to*

$$\left\{\begin{array}{rcl} u_t(x,t) + div(F(x,t,u(x,t))) &=& q(x,t,u(x,t)), \\ u(x,0) &=& u_0(x), \end{array}\right.$$

defined by (6) and let $u_{\mathcal{T},k}$ be the approximate solution defined by the semi-implicit scheme (9), (11), (13).
If $u_0 \in L^\infty \cap BV_{loc}(\mathbb{R}^N)$, we get the following error estimate : for every compact set $E \subset \mathbb{R}^N \times \mathbb{R}_+$, there exists E_0 depending only on E, F and u_0 and K depending only on E, F, u_0, $\|q(u_0)\|_{L^1(E_0)}$, $\|q(u_{\mathcal{T},0})\|_{L^1(E_0)}$, M, α, ξ such that :

$$\int_E |u_{\mathcal{T},k}(x,t) - u(x,t)|dxdt \leq Kh^{\frac{1}{4}}. \tag{16}$$

The outline of the paper is the following. In Section 2 and Section 3, we prove some stability properties $-L^\infty$ stability and BV-weak stability– verified by the schemes. In both sections, we have to distinguish between the explicit scheme for nonstiff source terms and the semi-implicit scheme for stiff source terms.

The aim of Section 4 is to prove some inequalities, (45), (54), satisfied by the approximate solutions. These "entropy inequalities" are close to the definition of the entropy solution, (6). They permit to compare approximate and entropy solutions : in Section 5, we show how these entropy inequalities lead to the error estimates (15) and (16). It uses a regularization technique introduced by S.N. Kruskov [13] and developed for numerical schemes by N.N. Kuznetsov [14]. In this last section, we give the proofs of Theorem 1 and Theorem 2 and conclude with some remarks about these results.

2 L^∞-stability of the schemes

In this section, we prove the L^∞ stability of the schemes. For the semi-implicit scheme, we simultaneously prove existence and uniqueness of the approximate solution $u_{\mathcal{T},k}$.

We give some notations used in all the sequel :

$$\begin{array}{rcl} N_T &=& \max\{n \in \mathbb{N},\ n \leq \frac{T}{k}\}, \\ \mathcal{T}_R &=& \{p \in \mathcal{T},\ p \subset B(0,R)\}, \\ \mathcal{E}^n &=& \{(p,q) \in \mathcal{T}^2,\ q \in N(p),\ u_p^n > u_q^n\}, \\ \mathcal{E}_R^n &=& \{(p,q) \in \mathcal{T}^2,\ p \text{ or } q \in \mathcal{T}_R,\ q \in N(p),\ \sigma_{pq} \subset B(0,R) \text{ and } u_p^n > u_q^n\}. \end{array}$$

2.1 The explicit scheme in the nonstiff case

Proposition 1 *Under the hypotheses (3), (4) on the data, (7) on the mesh, (8) on the numerical fluxes and under the C.F.L. condition (12), the approximate solution $u_{\mathcal{T},k}$ defined* *by (9),(10) and (13) verifies :*

$$(17) \quad |u_p^n| \leq (1+k(D^1+Q^1))^n(\|u_0\|_\infty + \frac{D^0+Q^0}{D^1+Q^1}) - \frac{D^0+Q^0}{D^1+Q^1}, \ \forall n \in \mathbb{N}, \ \forall p \in \mathcal{T}$$

and, for every $T>0$, there exists $U(T) < +\infty$, depending on T, $\|u0\|_\infty$, D^0, D^1, Q^0, Q^1, such that

$$(18) \qquad \|u_{\mathcal{T},k}\|_{L^\infty(\mathbb{R}^N\times[0,T[)} \leq U(T).$$

Proof

We first remark that (18) is a straightforward consequence of (17). Indeed,

$$\|u_{\mathcal{T},k}\|_{L^\infty(\mathbb{R}^N\times[0,T[)} \leq \max\{u_p^{n+1}, \ 0 \leq n \leq N_T, \ p \in \mathcal{T}\}$$

and thus

$$\|u_{\mathcal{T},k}\|_{L^\infty(\mathbb{R}^N\times[0,T[)} \leq (1+k(D^1+Q^1))^{N_T+1}(\|u_0\|_\infty + \frac{D^0+Q^0}{D^1+Q^1}) - \frac{D^0+Q^0}{D^1+Q^1}.$$

As $(1+k(D^1+Q^1))^{N_T+1} \leq \exp(2(D^1+Q^1)T)$, we get (18).
We prove (17) by induction. The definition of $(u_p^0)_{p\in\mathcal{T}}$, (9), implies that $|u_p^0| \leq \|u_0\|_\infty$ for all $p \in \mathcal{T}$. We set $U_0 = \|u_0\|_\infty$. Let assume $|u_p^n| \leq U_n$ for all $p \in \mathcal{T}$. The scheme (10) rewrites :

$$u_p^{n+1} = G(u_p^n, u_q^n \ q \in N(p)) - \frac{1}{m(p)} \int_{t^n}^{t^{n+1}} \int_p (\mathrm{div}_x F - q)(x,t,u_p^n) dx dt,$$

where G is a convex combination of u_p^n and the u_q^n, $q \in N(p)$, under the CFL condition (12). Therefore, $|G(u_p^n, u_q^n \ q \in N(p))| \leq U_n$. Moreover, thanks to $(3iv)$ and $(4ii)$, we have :

$$|\frac{1}{m(p)} \int_{t^n}^{t^{n+1}} \int_p (\mathrm{div}_x F - q)(x,t,u_p^n) dx dt| \leq k(D^0+Q^0+(D^1+Q^1)U_n).$$

Then, we get $|u_p^{n+1}| \leq U_{n+1}$ for all $p \in \mathcal{T}$, with :

$$\left\{ \begin{array}{rcl} U_{n+1} & = & (1+k(D^1+Q^1)U_n) + k(D^0+Q^0) \\ U_0 & = & \|u_0\|_\infty \end{array} \right.$$

It leads to the estimate (17) and concludes the proof.

2.2 The semi-implicit scheme in the stiff case

For the semi-implicit scheme, we first prove that $u_{\mathcal{T},k}$ defined by (9), (11) and (13) is well defined and then we get an estimate on its L^∞-norm.

Proposition 2 *Under the hypotheses (3), (5) on the data, (7) on the mesh, (8) on the numerical fluxes and under the C.F.L. condition (12), there exists a unique solution* $\{u_p^n,\ p \in \mathcal{T},\ n \in \mathbb{N}\}$ *to the scheme (9), (11), such that :*

$$|u_p^n| \le (1+kD^1)^n(\|u_0\|_\infty + \frac{D^0}{D^1}) - \frac{D^0}{D^1},\ \forall n \in \mathbb{N},\ \forall p \in \mathcal{T}. \tag{19}$$

Therefore, for every $T > 0$*, there exists* $U(T) < +\infty$*, depending on* T*,* $\|u0\|_\infty$*,* D^0*,* D^1*, such that* $u_{\mathcal{T},k}$ *defined by (9), (11) and (13) satisfies :*

$$\|u_{\mathcal{T},k}\|_{L^\infty(\mathbb{R}^N \times [0,T[)} \le U(T). \tag{20}$$

Proof

The initial condition (9) defines $(u_p^0)_{p\in\mathcal{T}}$ uniquely with $|u_p^0| \le \|u_0\|_\infty$ for all $p \in \mathcal{T}$. Let $n \in \mathbb{N}^*$, we assume that $(u_p^n)_{p\in\mathcal{T}}$ is uniquely defined and that there exists U_n such that $|u_p^n| \le U_n$ for all $p \in \mathcal{T}$.

Let consider $H \in C^1(\mathbb{R},\mathbb{R})$ defined by $H(x) = x - kq(x)$; it satisfies $H(0) = 0$ and $H'(x) = 1 - kq'(x) \ge 1$. Therefore H is increasing and $|H(x)| \ge |x|$ for all $x \in \mathbb{R}$. Thus H is one-to-one and onto and it yields existence and uniqueness of $(u_p^{n+1})_{p\in\mathcal{T}}$ such that :

$$H(u_p^{n+1}) = G(u_p^n, u_q^n\ q \in N(p)) - \frac{1}{m(p)}\int_{t^n}^{t^{n+1}}\int_p \mathrm{div}_x F(x,t,u_p^n)dxdt,\ \forall p \in \mathcal{T}.$$

Moreover,

$$|u_p^{n+1}| \le (1+kD^1)U_n + kD^0.$$

Then, we deduce (19) from this inequality and (20) as a straightforward consequence of (19).

3 BV weak stability of the schemes

In this section, we prove some estimates on the discrete derivatives of the approximate solutions for both schemes. They are called BV-weak estimates.

In the sequel, we use the following bounds, which come from the hypotheses (7) on the mesh : $\mathrm{Card}\mathcal{T}_R \le Ch^{-N}$, $\mathrm{Card}\mathcal{E}_R^n \le Ch^{-N}$ and

$$\mathrm{Card}\ \{\sigma_{pq} \in B(0,R) \setminus B(0,R-h),\ (p,q) \in \mathcal{T}^2\} \le Ch^{1-N},$$

with C only depending on R, N and α.

3.1 The explicit scheme in the nonstiff case

Proposition 3 *Assume that the data satisfy (3), (4), the mesh (7), the numerical fluxes (8) and the time step the C.F.L. condition (12). The approximate solution is given by (9), (10) and (13). Let $T > 0$ and $R > h$, there exists C_{bv}, which only depends on F, u_0, q, M, α, ξ, R and T such that :*

$$\sum_{n=0}^{N_T} k \sum_{(p,q)\in\mathcal{E}_R^n} [\max_{u_q^n\le c\le d\le u_p^n}(F_{p,q}^n(d,c) - F_{p,q}^n(d,d))$$

$$+ \max_{u_q^n\le c\le d\le u_p^n}(F_{p,q}^n(d,c) - F_{p,q}^n(c,c))] \le \frac{C_{bv}}{\sqrt{h}} \tag{21}$$

and

$$\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} m(p)|u_p^{n+1} - u_p^n| \le \frac{C_{bv}}{\sqrt{h}}. \tag{22}$$

Proof

In this proof, we denote by $(C_i)_{i\in\mathbb{N}}$ some quantities that only depend on F, u_0, q, M, α, ξ, R and T.

In order to prove (21), we multiply the scheme (10) by $m(p)u_p^n$ and we sum the result over $n \in \{0,..,N_T\}$ and $p \in \mathcal{T}_R$. It yields :

$$B_1 + B_2 + B_3 = 0, \tag{23}$$

where

$$B_1 = \sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} m(p)u_p^n(u_p^{n+1} - u_p^n), \tag{24}$$

$$B_2 = \sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} k \sum_{q\in N(p)} u_p^n F_{p,q}^n(u_p^n, u_q^n), \tag{25}$$

$$B_3 = -\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} k m(p)u_p^n q_p^n. \tag{26}$$

By gathering the terms of B_2 by edges, we obtain B_{21} defined by :

$$B_{21} = \sum_{n=0}^{N_T}\sum_{(p,q)\in\mathcal{E}_R^n} k(u_p^n - u_q^n)F_{p,q}^n(u_p^n, u_q^n)$$

and that satisfies

$$|B_2 - B_{21}| \le C_1. \tag{27}$$

Let us introduce $\Psi^n_{p,q} \;:\; x \mapsto \int_0^x F^n_{p,q}(s,s)ds$. We rewrite B_{21} as :

$$B_{21} = B_{22} + B_{23} \tag{28}$$

where

$$\begin{aligned} B_{22} &= \sum_{n=0}^{N_T} \sum_{(p,q)\in\mathcal{E}_R^n} k \int_{u_q^n}^{u_p^n} \left(F^n_{p,q}(u_p^n, u_q^n) - F^n_{p,q}(s,s)\right) ds, \\ B_{23} &= \sum_{n=0}^{N_T} \sum_{(p,q)\in\mathcal{E}_R^n} k(\Psi^n_{p,q}(u_p^n) - \Psi^n_{p,q}(u_q^n)). \end{aligned}$$

As already proved in [1], we have :

$$\begin{aligned} B_{22} \geq \frac{1}{4M} \sum_{n=0}^{N_T} k \sum_{(p,q)\in\mathcal{E}_R^n} \Bigg(&\frac{1}{m(\sigma_{pq})} \max_{u_q^n\le c\le d\le u_p^n} (F^n_{p,q}(d,c) - F^n_{p,q}(d,d))^2 \\ &+ \frac{1}{m(\sigma_{pq})} \max_{u_q^n\le c\le d\le u_p^n} (F^n_{p,q}(d,c) - F^n_{p,q}(c,c))^2 \Bigg). \end{aligned} \tag{29}$$

In B_{23}, we can rearrange the sum on a sum on the cells of the mesh (instead of a sum on the edges). It gives :

$$B_{24} = \sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} k \sum_{q\in N(p)} \Psi^n_{p,q}(u_p^n)$$

and

$$|B_{23} - B_{24}| \leq C_2. \tag{30}$$

Thanks to the L^∞-stability and the hypotheses on $\mathrm{div}_x F$ and q, we have :

$$|B_3 + B_{24}| \leq C_3. \tag{31}$$

Let us now turn to an estimate of B_1.

$$B_1 = -\frac{1}{2} \sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} m(p)(u_p^{n+1} - u_p^n)^2 + \frac{1}{2} \sum_{p\in\mathcal{T}_R} m(p)(u_p^{N_T+1})^2 - \frac{1}{2} \sum_{p\in\mathcal{T}_R} m(p)(u_p^0)^2.$$

Applying Cauchy-Schwarz inequality in the scheme, we get

$$\begin{aligned} (u_p^{n+1} - u_p^n)^2 \leq \frac{2k^2}{m(p)^2} \left(\sum_{q\in N(p)} m(\sigma_{pq})\right) &\left(\sum_{q\in N(p)} \frac{1}{m(\sigma_{pq})} (F^n_{p,q}(u_p^n, u_q^n) - F^n_{p,q}(u_p^n, u_p^n))^2\right) \\ &+ \frac{2}{m(p)^2} \left(\int_{t^n}^{t^{n+1}} \int_p (\mathrm{div}_x F - q)(x,t,u_p^n) dx dt\right)^2. \end{aligned} \tag{32}$$

Therefore,

$$B_1 \geq -C_4 - \frac{(1-\xi)}{4M} \sum_{n=0}^{N_T} \sum_{(p,q)\in\mathcal{E}_R^n} k \left(\frac{1}{m(\sigma_{pq})} \max_{u_q^n \leq c \leq d \leq u_p^n} (F_{p,q}^n(d,c) - F_{p,q}^n(d,d))^2 \right.$$
$$\left. + \frac{1}{m(\sigma_{pq})} \max_{u_q^n \leq c \leq d \leq u_p^n} (F_{p,q}^n(d,c) - F_{p,q}^n(c,c))^2 \right). \tag{33}$$

Then, from (23), (27), (28), (29), (30), (31) and (33), we deduce :

$$\sum_{n=0}^{N_T} k \sum_{(p,q)\in\mathcal{E}_R^n} \frac{1}{m(\sigma_{pq})} \left(\max_{u_q^n \leq c \leq d \leq u_p^n} (F_{p,q}^n(d,c) - F_{p,q}^n(d,d))^2 \right.$$
$$\left. + \max_{u_q^n \leq c \leq d \leq u_p^n} (F_{p,q}^n(d,c) - F_{p,q}^n(c,c))^2 \right) \leq C_5$$

It yields the BV-weak estimate on the space derivatives (21) and, using again the scheme, we get the estimate on the time derivatives (22).

3.2 The semi-implicit scheme in the stiff case

Proposition 4 *Assume that the data satisfy (3), (5), the mesh (7), the numerical fluxes (8) and the time step the C.F.L. condition (12). The approximate solution is given by (9), (11) and (13). Let $T > 0$ and $R > h$, there exists C_{bv}, which depends only on F, u_0, M, α, ξ, R, T and $\|q(u_{\mathcal{T},0})\|_{L^1(B(0,R))}$ such that :*

$$\sum_{n=0}^{N_T} k \sum_{(p,q)\in\mathcal{E}_R^n} [\max_{u_q^n \leq c \leq d \leq u_p^n} (F_{p,q}^n(d,c) - F_{p,q}^n(d,d))$$
$$+ \max_{u_q^n \leq c \leq d \leq u_p^n} (F_{p,q}^n(d,c) - F_{p,q}^n(c,c))] \leq \frac{C_{bv}}{\sqrt{h}} \tag{34}$$

and

$$\sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} m(p)|u_p^{n+1} - u_p^n| \leq \frac{C_{bv}}{\sqrt{h}}. \tag{35}$$

Proof

It is important to note that the quantity C_{bv} only depends on q through $q(u_{\mathcal{T},0})$. The mechanism of the proof is the same as the one of Proposition 3. We just have to take care about the influence of q in order to get a bound that does not depend on the derivative of q. We denote by $(C_i)_{i\in\mathbb{N}}$ some quantities that only depend on F, u_0, M, α, ξ, R, T and do not depend on q.

We multiply the scheme (11) by $m(p)u_p^n$ and we sum the result over $n \in \{0,..,N_T\}$ and $p \in \mathcal{T}_R$. It yields :

$$B_1 + B_2 + B_3 = B_4, \tag{36}$$

where B_1 and B_2 are respectively defined by (24) and (25) and

$$\begin{aligned} B_3 &= -\sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} km(p)u_p^n(q(u_p^{n+1}) - q(u_p^n)), \\ B_4 &= \sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} km(p)u_p^n q(u_p^n). \end{aligned}$$

The hypotheses (5) on q imply :

$$B_4 \le 0 \tag{37}$$

Following the proof of Proposition 3, we have :

$$\begin{aligned} B_2 \ge -C_1 + \frac{1}{4M}\sum_{n=0}^{N_T} k \sum_{(p,q)\in\mathcal{E}_R^n} \Bigg(&\frac{1}{m(\sigma_{pq})} \max_{u_q^n \le c \le d \le u_p^n} (F_{p,q}^n(d,c) - F_{p,q}^n(d,d))^2 \\ &+ \frac{1}{m(\sigma_{pq})} \max_{u_q^n \le c \le d \le u_p^n} (F_{p,q}^n(d,c) - F_{p,q}^n(c,c))^2 \Bigg). \end{aligned} \tag{38}$$

In order to estimate B_1, we introduce

$$B_{11} = -\frac{1}{2}\sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} m(p)(u_p^{n+1} - kq(u_p^{n+1}) - u_p^n)^2.$$

On one hand, we use the scheme as for (33) and :

$$\begin{aligned} B_{11} \ge -C_2 - \frac{(1-\xi)}{4M}\sum_{n=0}^{N_T} \sum_{(p,q)\in\mathcal{E}_R^n} k \Bigg(&\frac{1}{m(\sigma_{pq})} \max_{u_q^n \le c \le d \le u_p^n} (F_{p,q}^n(d,c) - F_{p,q}^n(d,d))^2 \\ &+ \frac{1}{m(\sigma_{pq})} \max_{u_q^n \le c \le d \le u_p^n} (F_{p,q}^n(d,c) - F_{p,q}^n(c,c))^2 \Bigg). \end{aligned} \tag{39}$$

On the other hand, we have :

$$(u_p^{n+1} - kq(u_p^{n+1}) - u_p^n)^2 = (u_p^{n+1} - u_p^n)^2 + k^2(q(u_p^{n+1}))^2 - 2kq(u_p^{n+1})(u_p^{n+1} - u_p^n).$$

Therefore,

$$-\frac{1}{2}\sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} m(p)(u_p^{n+1} - u_p^n)^2 = B_{11} + B_{12} + B_{13}$$

with

$$
\begin{aligned}
B_{12} &= \frac{1}{2}\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} m(p)k^2(q(u_p^{n+1}))^2,\\
B_{13} &= -\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} km(p)q(u_p^{n+1})(u_p^{n+1}-u_p^n).
\end{aligned}
$$

It is clear that $B_{12}\geq 0$ and a discrete integration by parts on B_{13} gives :

$$B_{13} = -B_3 + \sum_{p\in\mathcal{T}_R} km(p)u_p^0 q(u_p^0) - \sum_{p\in\mathcal{T}_R} km(p)u_p^{N_T+1}q(u_p^{N_T+1}).$$

The term $\sum_{p\in\mathcal{T}_R} km(p)u_p^0 q(u_p^0)$ occurs a dependence of C_{bv} on q. Finally :

$$
\begin{aligned}
B_1+B_3 \geq -C_3 - \frac{(1-\xi)}{4M}\sum_{n=0}^{N_T}\sum_{(p,q)\in\mathcal{E}_R^n} k\Bigg(&\frac{1}{m(\sigma_{pq})}\max_{u_q^n\leq c\leq d\leq u_p^n}(F_{p,q}^n(d,c)-F_{p,q}^n(d,d))^2\\
+&\frac{1}{m(\sigma_{pq})}\max_{u_q^n\leq c\leq d\leq u_p^n}(F_{p,q}^n(d,c)-F_{p,q}^n(c,c))^2\Bigg) - k\|u_0\|_\infty\|q(u_{\mathcal{T},0})\|_{L^1(B(0,R))}. \qquad (40)
\end{aligned}
$$

From (36), (37), (38) and (40), we deduce the BV-weak inequality (34). To prove the estimate on the time derivatives (35), we use the scheme (11). It yields :

$$
\begin{aligned}
\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} m(p)|u_p^{n+1}-u_p^n| \leq &\sum_{n=0}^{N_T} k\sum_{p\in\mathcal{T}_R}\sum_{q\in N(p)}|F_{p,q}^n(u_p^n,u_q^n)-F_{p,q}^n(u_p^n,u_p^n)|\\
&\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R}\int_{t^n}^{t^{n+1}}\int_p|\mathrm{div}_x F(x,t,u_p^n)|dxdt + \sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} km(p)|q(u_p^{n+1})| \qquad (41)
\end{aligned}
$$

We can bound the two first terms of the right hand side of (41) thanks to (34) and (19). But, to get the inequality (35) which does not depend on the derivative of q, we need an estimate of $\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} km(p)|q(u_p^{n+1})|$ that also does not depend on this quantity. To prove it, we multiply the scheme (11) by $m(p)sign(u_p^{n+1})$; it yields :

$$
\begin{aligned}
m(p)|u_p^{n+1}| \leq m(p)|u_p^n| + k&\sum_{q\in N(p)}|F_{p,q}^n(u_p^n,u_q^n)-F_{p,q}^n(u_p^n,u_p^n)|\\
&+\int_{t^n}^{t^{n+1}}\int_p|\mathrm{div}_x F(x,t,u_p^n)|dxdt - km(p)|q(u_p^{n+1})|.
\end{aligned}
$$

Summing over $n \in \{0,..,N_T\}$ and $p \in \mathcal{T}_R$, we get :

$$\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} km(p)|q(u_p^{n+1})| \leq \sum_{p\in\mathcal{T}_R} m(p)|u_p^0| + \sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R}\int_{t^n}^{t^{n+1}}\int_p |\mathrm{div}_x F(x,t,u_p^n)|dxdt$$
$$+\sum_{n=0}^{N_T} k \sum_{p\in\mathcal{T}_R}\sum_{q\in N(p)} |F_{p,q}^n(u_p^n,u_q^n) - F_{p,q}^n(u_p^n,u_p^n)|.$$

Therefore,

$$\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} km(p)|q(u_p^{n+1})| \leq C_4 + \frac{C_{bv}}{\sqrt{h}}. \tag{42}$$

From (41) and (42), we deduce the BV-weak inequality (35). It concludes the proof of Proposition 4.

4 Entropy inequalities

In these section, we prove some entropy inequalities, (45), (54), satisfied by the approximate solution either in the nonstiff case or in the stiff case. These inequalities show in which sense the approximate solutions are close to the entropy solution : they differ from the definition of the entropy solution (6) only because of the error terms in the right hand side.

4.1 Discrete entropy inequalities

First, we establish entropy inequalities on each cell for both schemes. Lemma 1 deals with the explicit scheme for the nonstiff case while Lemma 2 deals with the semi-implicit scheme for the stiff case.

Lemma 1 *Assume the hypotheses (3), (4) on the data, (7) on the mesh, (8) on the numerical fluxes and the C.F.L. condition (12). Let $u_{\mathcal{T},k}$ be defined by (9), (10) and (13). Then, for all $\kappa \in \mathbb{R}$, $p \in \mathcal{T}$ and $n \in \mathbb{N}$, we have :*

$$|u_p^{n+1}-\kappa| - |u_p^n-\kappa| + \frac{k}{m(p)}\sum_{q\in N(p)}\Big((F_{p,q}^n(u_p^n\top\kappa, u_q^n\top\kappa) - F_{p,q}^n(u_p^n\bot\kappa, u_q^n\bot\kappa))$$
$$-(F_{p,q}^n(u_p^n\top\kappa, u_p^n\top\kappa) - F_{p,q}^n(u_p^n\bot\kappa, u_p^n\bot\kappa))\Big)$$
$$+\frac{k}{m(p)} sign(u_p^{n+1}-\kappa)\Big(\sum_{q\in N(p)} F_{p,q}^n(u_p^n,u_p^n) - m(p)q_p^n\Big) \leq 0. \tag{43}$$

Proof

Let $\kappa \in \mathbb{R}$, $x \mapsto |x-\kappa|$ is a convex function. Therefore, for all $a, b \in \mathbb{R}$, we have :

$$|b-\kappa| - |a-\kappa| - sign(a-\kappa)(b-a) \geq 0.$$

Applying this inequality to the scheme (10), we get :

$$|u_p^{n+1} - \kappa| \leq |u_p^n - \frac{k}{m(p)} \sum_{q \in N(p)} (F_{p,q}^n(u_p^n, u_q^n) - F_{p,q}^n(u_p^n, u_p^n)) - \kappa|$$
$$- \frac{k}{m(p)} sign(u_p^{n+1} - \kappa) \Big(\sum_{q \in N(p)} F_{p,q}^n(u_p^n, u_p^n) - m(p) q_p^n \Big).$$

But the monotony of the numerical fluxes implies under the CFL condition (see the discrete entropy inequality in [1]) :

$$|u_p^n - \frac{k}{m(p)} \sum_{q \in N(p)} (F_{p,q}^n(u_p^n, u_q^n) - F_{p,q}^n(u_p^n, u_p^n)) - \kappa| \leq |u_p^n - \kappa|$$
$$- \frac{k}{m(p)} \sum_{q \in N(p)} \Big((F_{p,q}^n(u_p^n \top \kappa, u_q^n \top \kappa) - F_{p,q}^n(u_p^n \bot \kappa, u_q^n \bot \kappa))$$
$$- (F_{p,q}^n(u_p^n \top \kappa, u_p^n \top \kappa) - F_{p,q}^n(u_p^n \bot \kappa, u_p^n \bot \kappa)) \Big).$$

It concludes the proof of (43).

Lemma 2 *Assume the hypotheses (3), (5) on the data, (7) on the mesh, (8) on the numerical fluxes and the C.F.L. condition (12). Let $u_{\mathcal{T},k}$ be defined by (9), (11) and (13). Then, for all $\kappa \in \mathbb{R}$, $p \in \mathcal{T}$ and $n \in \mathbb{N}$, we have :*

$$|u_p^{n+1} - \kappa| - |u_p^n - \kappa| + \frac{k}{m(p)} \sum_{q \in N(p)} \Big((F_{p,q}^n(u_p^n \top \kappa, u_q^n \top \kappa) - F_{p,q}^n(u_p^n \bot \kappa, u_q^n \bot \kappa))$$
$$- (F_{p,q}^n(u_p^n \top \kappa, u_p^n \top \kappa) - F_{p,q}^n(u_p^n \bot \kappa, u_p^n \bot \kappa)) \Big)$$
$$+ \frac{k}{m(p)} sign(u_p^{n+1} - \kappa) \Big(\sum_{q \in N(p)} F_{p,q}^n(u_p^n, u_p^n) - m(p) q(u_p^{n+1}) \Big) \leq 0. \tag{44}$$

Proof
The proof is the same as for Lemma 1.

4.2 Continuous entropy inequalities

We now prove that the approximate solutions defined by both schemes satisfy entropy inequalities similar to the definition of the entropy solution (6) but with some error terms. These error terms are expressed with the help of some measures.

For $\Omega = \mathbb{R}^N$ or $\Omega = \mathbb{R}^N \times \mathbb{R}_+$, we denote by $\mathcal{M}(\Omega)$ the set of measures on Ω, *i.e.* the set of positive continuous linear forms on $C_c(\Omega)$. If $\mu \in \mathcal{M}(\Omega)$, we set :

$$\langle \mu, g \rangle = \int_\Omega g d\mu, \text{ for all } g \in C_c(\Omega).$$

4.2.1 The explicit scheme for the nonstiff case

Theorem 3 *Assume the data satisfy (3), (4), the mesh (7), the numerical fluxes (8) and the time step (12). Let consider the approximate solution* $u_{\mathcal{T},k}$ *defined by the explicit scheme (9), (10) and (13). Then, there exist some measures* $\mu_{\mathcal{T}} \in \mathcal{M}(\mathbb{R}^N)$, $\mu_{\mathcal{T},k} \in \mathcal{M}(\mathbb{R}^N \times \mathbb{R}_+)$, $\xi_{\mathcal{T},k} \in \mathcal{M}(\mathbb{R}^N \times \mathbb{R}_+)$ *such that,* $\forall \kappa \in \mathbb{R}$, $\forall \varphi \in C_c^\infty(\mathbb{R}^N \times \mathbb{R}_+, \mathbb{R}_+)$ *:*

$$\begin{aligned}\int_{\mathbb{R}^N \times \mathbb{R}_+} \Big[& |u_{\mathcal{T},k} - \kappa| \varphi_t(x,t) + (F(x,t,u_{\mathcal{T},k} \top \kappa) - F(x,t,u_{\mathcal{T},k} \bot \kappa)).\nabla\varphi(x,t) \\ & -sign(u_{\mathcal{T},k} - \kappa)\Big(div_x(F(x,t,\kappa)) - q(x,t,u_{\mathcal{T},k})\Big)\varphi(x,t)\Big] dxdt \\ & + \int_{\mathbb{R}^N} |u_0(x) - \kappa| \varphi(x,0) dx \geq - \int_{\mathbb{R}^N} \varphi(x,0) d\mu_{\mathcal{T}}(x) \\ & - \int_{\mathbb{R}^N \times \mathbb{R}_+} (|\varphi_t(x,t)| + |\nabla\varphi(x,t)|) d\mu_{\mathcal{T},k}(x,t) - \int_{\mathbb{R}^N \times \mathbb{R}_+} \varphi(x,t) d\xi_{\mathcal{T},k}(x,t). \end{aligned} \tag{45}$$

Moreover, the measures $\mu_{\mathcal{T},k}$, $\xi_{\mathcal{T},k}$ *and* $\mu_{\mathcal{T}}$ *have the following properties :*

1. *For all* $R > 0$ *and* $T > 0$, *there exists* C_μ *that only depends on* F, u_0, q, M, α, ξ, R *and* T *such that*

$$\mu_{\mathcal{T},k}(B(0,R) \times [0,T]) \leq C_\mu \ (h + \sqrt{h}), \ \forall h < R. \tag{46}$$

2. *For all* $R > 0$ *and* $T > 0$, *there exists* C_ξ *depending only on* F, u_0, q, M, α, ξ, R *and* T *such that*

$$\xi_{\mathcal{T},k}(B(0,R) \times [0,T]) \leq C_\xi \ (h + \sqrt{h}), \ \forall h < R. \tag{47}$$

3. *The measure* $\mu_{\mathcal{T}}$ *is the measure of density* $|u_0 - u_{\mathcal{T},0}|$ *with respect to the Lebesgue measure. For all* $R > 0$, *we have :*

$$\lim_{h \to 0} (\mu_{\mathcal{T}}(B(0,R))) = 0. \tag{48}$$

   If $u_0 \in L^\infty \cap BV_{loc}(\mathbb{R}^N)$, *there exists* D_m *only depending on* u_0, α *and* R *such that :*

$$\mu_{\mathcal{T}}(B(0,R)) \leq C_\mu^0 \ h, \forall h < R. \tag{49}$$

Proof

The proof is similar to the proof of Theorem 1 in [1]. The main difference comes from the terms T_3 and T_3^* defined below. We also have to take care about the definition of $u_{\mathcal{T},k}$ which is different. In the proof, we denote by $(C_i)_{i\in\mathbb{N}}$ some quantities that only depend on F, u_0, q, M, α, ξ, R and T.

Let $\varphi \in C_c^\infty(\mathbb{R}^N \times \mathbb{R}_+, \mathbb{R}_+)$ and $\kappa \in \mathbb{R}$. Let $T > 0$ and $R > 0$ such that $\varphi(x,t) \neq 0$ implies $|x| \leq R - h$ and $t \in [0,T]$. Let us multiply (43) by $\displaystyle\int_{t^n}^{t^{n+1}} \int_p \varphi(x,t)dxdt$ and sum the result for all $p \in \mathcal{T}_R$ and $n \in \{0, .., N_T\}$. It yields :

$$T_1 + T_2 + T_3 \leq 0 \tag{50}$$

with

$$T_1 = \sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} \frac{|u_p^{n+1} - \kappa| - |u_p^n - \kappa|}{k} \int_{t^n}^{t^{n+1}} \int_p \varphi(x,t)dxdt,$$

$$\begin{aligned} T_2 = \sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} \frac{1}{m(p)} \int_{t^n}^{t^{n+1}} \int_p \varphi(x,t)dxdt \sum_{q\in N(p)} &(F_{p,q}^n(u_p^n \top \kappa, u_q^n \top \kappa) - F_{p,q}^n(u_p^n \top \kappa, u_p^n \top \kappa) \\ &- F_{p,q}^n(u_p^n \bot \kappa, u_q^n \bot \kappa) + F_{p,q}^n(u_p^n \bot \kappa, u_p^n \bot \kappa)) \end{aligned}$$

and

$$T_3 = \sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} \frac{1}{m(p)} \int_{t^n}^{t^{n+1}} \int_p \varphi(x,t)dxdt\ sign(u_p^{n+1} - \kappa) \left(\sum_{q\in N(p)} F_{p,q}^n(u_p^n, u_p^n) - m(p)q_p^n \right).$$

Now, we compare each of these three terms with a part of the left hand side of (45). Let introduce :

$$\begin{aligned} T_1^* &= -\int_{\mathbb{R}^N\times\mathbb{R}_+} |u_{\mathcal{T},k}(x,t) - \kappa| \varphi_t(x,t)dxdt - \int_{\mathbb{R}^N} |u_0(x) - \kappa|\varphi(x,0)dx, \\ T_2^* &= -\int_{\mathbb{R}^N\times\mathbb{R}_+} (F(x,t,u_{\mathcal{T},k}(x,t)\top\kappa) - F(x,t,u_{\mathcal{T},k}(x,t)\bot\kappa)).\nabla\varphi(x,t)dxdt \\ &\quad - \int_{\mathbb{R}^N\times\mathbb{R}_+} (\mathrm{div}_x F(x,t,u_{\mathcal{T},k}\top\kappa) - \mathrm{div}_x F(x,t,u_{\mathcal{T},k}\bot\kappa))\varphi(x,t)dxdt \\ T_3^* &= \int_{\mathbb{R}^N\times\mathbb{R}_+} sign(u_{\mathcal{T},k} - \kappa)(\mathrm{div}_x F - q)(x,t,u_{\mathcal{T},k})\varphi(x,t)dxdt. \end{aligned}$$

Comparison of T_1 and T_1^*

Using the definitions of $u_{\mathcal{T},k}$, (13), and of $u_{\mathcal{T},0}$, (14), in T_1^*, we get :

$$|T_1 - T_1^*| \leq \sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} |u_p^{n+1} - u_p^n| \int_{t^n}^{t^{n+1}} \int_p |\varphi_t(x,t)|dxdt$$

$$+\int_{\mathbb{R}^N} |u_0(x) - u_{\mathcal{T},0}(x)|\varphi(x,0)dx,$$

which rewrites :

$$|T_1 - T_1^*| \le \int_{\mathbb{R}^N\times\mathbb{R}_+} |\varphi_t(x,t)| d\lambda_{\mathcal{T},k}(x,t) + \int_{\mathbb{R}^N} \varphi(x,0) d\mu_{\mathcal{T}}(x). \tag{51}$$

The measure $\lambda_{\mathcal{T},k}$ will contribute to the measure $\mu_{\mathcal{T},k}$. Thanks to the BV-weak estimate on the discrete time derivatives (22), we have :

$$\lambda_{\mathcal{T},k}(B(0,R)\times[0,T]) \le TC_{bv}\sqrt{h} \quad \forall R>0,\ \forall T>0.$$

Comparison of T_2 and T_2^*

In order to use what has already been done in [1], we introduce :

$$\overline{T_2^*} = -\int_{\mathbb{R}^N\times\mathbb{R}_+} (F(x,t,\overline{u_{\mathcal{T},k}}(x,t)\top\kappa) - F(x,t,\overline{u_{\mathcal{T},k}}(x,t)\bot\kappa)).\nabla\varphi(x,t)dxdt$$
$$-\int_{\mathbb{R}^N\times\mathbb{R}_+} (\mathrm{div}_x F(x,t,\overline{u_{\mathcal{T},k}}\top\kappa) - \mathrm{div}_x F(x,t,\overline{u_{\mathcal{T},k}}\bot\kappa))\varphi(x,t)dxdt,$$

where $\overline{u_{\mathcal{T},k}}(x,t) = u_p^n\ \forall x\in p,\ \forall t\in[t^n,t^{n+1}[$.
Indeed, because of the regularity of F (hypotheses (3 *iii*) and (3 *iv*)), we have :

$$|T_2^* - \overline{T_2^*}| \le C_1 \sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R} |u_p^{n+1} - u_p^n| \int_{t^n}^{t^{n+1}}\int_p (|\nabla\varphi|(x,t) + \varphi(x,t))dxdt.$$

Furthermore, as in [1] (Theorem 1), we prove the existence of $\nu_{\mathcal{T},k}\in\mathcal{M}(\mathbb{R}^N\times\mathbb{R}_+)$ such that

$$|T_2 - \overline{T_2^*}| \le \int_{\mathbb{R}^N\times\mathbb{R}_+} \Big(|\varphi_t(x,t)| + |\nabla\varphi(x,t)|\Big) d\nu_{\mathcal{T},k}(x,t)$$

and

$$\nu_{\mathcal{T},k}(B(0,R)\times[0,T]) \le C_2(h+\sqrt{h}) \quad \forall R>0,\ \forall T>0.$$

The measure $\nu_{\mathcal{T},k}$ contains the space derivatives of the approximate solution and the last inequality is due to the BV-weak estimate (21).
Then, we get :

$$|T_2 - T_2^*| \le C_1 \int_{\mathbb{R}^N\times\mathbb{R}_+} \Big(|\nabla\varphi(x,t)| + \varphi(x,t)\Big) d\lambda_{\mathcal{T},k}(x,t)$$
$$+\int_{\mathbb{R}^N\times\mathbb{R}_+} \Big(|\varphi_t(x,t)| + |\nabla\varphi(x,t)|\Big) d\nu_{\mathcal{T},k}(x,t). \tag{52}$$

Comparison of T_3 and T_3^*

The terms T_3 and T_3^* rewrite :

$$T_3^* = \sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} \int_{t^n}^{t^{n+1}} \int_p sign(u_p^{n+1} - \kappa)(\mathrm{div}_x F - q)(x, t, u_p^{n+1})\varphi(x, t)dxdt$$

and

$$T_3 = \sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} \int_{t^n}^{t^{n+1}} \int_p sign(u_p^{n+1} - \kappa)(\frac{1}{km(p)} \int_{t^n}^{t^{n+1}} \int_p (\mathrm{div}_x F - q)(y, s, u_p^n)dyds)\varphi(x, t)dxdt.$$

The definition of $u_{\mathcal{T},k}$ by (13) is crucial at this point. It permits that the *sign* function has the same argument in T_3 and in T_3^* and therefore, the hypotheses of regularity on $\mathrm{div}_x F$ (3 *iv*) and q (4 *ii*) imply :

$$|T_3 - T_3^*| \leq C_3 \sum_{n=0}^{N_T} \sum_{p\in\mathcal{T}_R} |u_p^{n+1} - u_p^n| \int_{t^n}^{t^{n+1}} \int_p \varphi(x, t)dxdt$$
$$+C_4(h + k) \int_{\mathbb{R}^N\times\mathbb{R}_+} \varphi(x, t)dxdt.$$

Then we define $\pi_{\mathcal{T},k} \in \mathcal{M}(\mathbb{R}^N \times \mathbb{R}_+)$ by its action on $C_c(\mathbb{R}^N \times \mathbb{R}_+)$:

$$\langle \pi_{\mathcal{T},k}, \varphi\rangle = \langle \lambda_{\mathcal{T},k}, C_3\varphi\rangle + C_4(h + k) \int_{\mathbb{R}^N\times\mathbb{R}_+} \varphi(x, t)dxdt.$$

This measure verifies (47) and we have :

$$|T_3 - T_3^*| \leq \int_{\mathbb{R}^N\times\mathbb{R}_+} \varphi(x, t)d\xi_{\mathcal{T},k}(x, t). \tag{53}$$

Finally, by setting $\mu_{\mathcal{T},k} = C_1\lambda_{\mathcal{T},k} + \nu_{\mathcal{T},k}$ and $\xi_{\mathcal{T},k} = C_1\lambda_{\mathcal{T},k} + \pi_{\mathcal{T},k}$, we deduce (45) from (50), (51), (52) and (53).

4.2.2 The semi-implicit scheme for the stiff case

Theorem 4 *Assume the data satisfy (3), (5), the mesh (7), the numerical fluxes (8) and the time step (12). Let consider the approximate solution $u_{\mathcal{T},k}$ defined by the semi-implicit scheme (9), (11) and (13). Then, there exist some measures $\mu_{\mathcal{T}} \in \mathcal{M}(\mathbb{R}^N)$, $\mu_{\mathcal{T},k} \in \mathcal{M}(\mathbb{R}^N \times \mathbb{R}_+)$, $\xi_{\mathcal{T},k} \in \mathcal{M}(\mathbb{R}^N \times \mathbb{R}_+)$ such that, $\forall\kappa \in \mathbb{R}$, $\forall\varphi \in C_c^\infty(\mathbb{R}^N \times \mathbb{R}_+, \mathbb{R}_+)$:*

$$\int_{\mathbb{R}^N\times\mathbb{R}_+} \Big[|u_{\mathcal{T},k} - \kappa|\varphi_t(x, t) + (F(x, t, u_{\mathcal{T},k}\top\kappa) - F(x, t, u_{\mathcal{T},k}\bot\kappa)).\nabla\varphi(x, t)$$

$$-sign(u_{\mathcal{T},k}-\kappa)\Big(div_x(F(x,t,\kappa))-q(x,t,u_{\mathcal{T},k})\Big)\varphi(x,t)\Big]dxdt$$

$$+\int_{\mathbb{R}^N}|u_0(x)-\kappa|\varphi(x,0)dx \geq -\int_{\mathbb{R}^N}\varphi(x,0)d\mu_{\mathcal{T}}(x)$$

$$-\int_{\mathbb{R}^N\times\mathbb{R}_+}(|\varphi_t(x,t)|+|\nabla\varphi(x,t)|)d\mu_{\mathcal{T},k}(x,t)-\int_{\mathbb{R}^N\times\mathbb{R}_+}\varphi(x,t)d\xi_{\mathcal{T},k}(x,t). \tag{54}$$

The measure $\mu_{\mathcal{T}}$ is still the measure of density $|u_0-u_{\mathcal{T},0}|$ with respect to the Lebesgue measure and satisfies the properties (48) and (49). The measures $\mu_{\mathcal{T},k}$ and $\xi_{\mathcal{T},k}$ also verify (46) and (47) but with C_μ and C_ξ depending only on F, u_0, M, α, ξ, R, T and $\|q(u_{\mathcal{T},0})\|_{L^1(B(0,R))}$.

Proof

The proof is similar to the one of Theorem 3. The only difference is in the comparison of T_3 and T_3^*. For the scheme (11), we have :

$$T_3-T_3^*=\sum_{n=0}^{N_T}\sum_{p\in\mathcal{T}_R}\int_{t^n}^{t^{n+1}}\int_p sign(u_p^{n+1}-\kappa)\varphi(x,t)$$

$$[\mathrm{div}_x F(x,t,u_p^{n+1})-\frac{1}{km(p)}\int_{t^n}^{t^{n+1}}\int_p \mathrm{div}_x F(y,s,u_p^{n+1})dyds]dxdt$$

There is no more q in this expression and q appears in the estimates (46) and (47) on $\mu_{\mathcal{T},k}$ and $\xi_{\mathcal{T},k}$ only through the quantity C_{bv} providing from the BV-weak estimates (34) and (35). Therefore, for the semi-implicit scheme in the stiff case, the estimates on $\mu_{\mathcal{T},k}$ and $\xi_{\mathcal{T},k}$ only depends on q through $\|q(u_{\mathcal{T},0})\|_{L^1(B(0,R))}$.

5 Error estimates

5.1 A preliminary lemma

The error estimates given in Theorem 1 and Theorem 2 are straightforward consequences of the continous entropy inequalities (45) and (54) and of the following Lemma.

Lemma 3 *Assume the data satisfy (3) and either (4) or (5). Let $u_0\in L^\infty\cap BV_{loc}(\mathbb{R}^N)$. Let $\tilde{u}\in L^\infty(\mathbb{R}^N\times[0,T[)$ for all $T>0$. We suppose that there exist $\mu\in\mathcal{M}(\mathbb{R}^N\times\mathbb{R}_+)$, $\xi\in\mathcal{M}(\mathbb{R}^N\times\mathbb{R}_+)$ and $\mu_0\in\mathcal{M}(\mathbb{R}^N)$ such that, for all $\kappa\in\mathbb{R}$, for all $\varphi\in C_c^\infty(\mathbb{R}^N\times\mathbb{R}_+,\mathbb{R}_+)$:*

$$\int_{\mathbb{R}^N\times\mathbb{R}_+}[|\tilde{u}(x,t)-\kappa|\varphi_t(x,t)+(F(x,t,\tilde{u}(x,t)\top\kappa)-F(x,t,\tilde{u}(x,t)\perp\kappa)).\nabla\varphi(x,t)]dxdt$$

$$-\int_{\mathbb{R}^N\times\mathbb{R}_+}sign(\tilde{u}(x,t)-\kappa)(div_xF(x,t,\kappa)-q(x,t,\tilde{u}(x,t)))\varphi(x,t)dxdt$$

$$
+\int_{\mathbb{R}^N} |u_0(x)-\kappa|\varphi(x,0)dx \geq -\int_{\mathbb{R}^N} \varphi(x,0)d\mu_0(x)
$$

$$
(55) \qquad -\int_{\mathbb{R}^N\times\mathbb{R}_+} (|\varphi_t(x,t)|+|\nabla\varphi(x,t)|)d\mu(x,t) - \int_{\mathbb{R}^N\times\mathbb{R}_+} \varphi(x,t)d\xi(x,t)
$$

Let $u \in L^\infty(\mathbb{R}^N \times [0,T[)$ for all $T>0$ be the unique entropy solution to (1), satisfying : for all $\kappa \in \mathbb{R}$, for all $\varphi \in C_c^\infty(\mathbb{R}^N \times \mathbb{R}_+, \mathbb{R}_+)$,

$$
\int_{\mathbb{R}^N\times\mathbb{R}_+} [|u(y,s)-\kappa|\varphi_s(y,s) + (F(y,s,u(y,s)\top\kappa) - F(y,s,u(y,s)\bot\kappa)).\nabla\varphi(y,s)]dyds
$$

$$
-\int_{\mathbb{R}^N\times\mathbb{R}_+} sign(u(y,s)-\kappa)(div_x F(y,s,\kappa) - q(y,s,u(y,s)))\varphi(y,s)
$$

$$
(56) \qquad +\int_{\mathbb{R}^N} |u_0(y)-\kappa|\varphi(y,0)dy \geq 0.
$$

Then, for every compact subset $E \subset \mathbb{R}^N \times \mathbb{R}_+$, there exists R and T depending only on E, F, u, $\tilde{u}$, u_0 and C_1, C_2, depending only on R, T, F, q, u, $\tilde{u}$ and u_0 such that :

$$
\int_E |\tilde{u}(x,t)-u(x,t)|dxdt \leq C_1\Big(\mu_0(B(0,R)) + \mu(B(0,R)\times[0,T])
$$

$$
(57) \qquad +(\mu(B(0,R)\times[0,T]))^{\frac{1}{2}} + \xi(B(0,R)\times[0,T])\Big) + C_2(\mu(B(0,R)\times[0,T]))^{\frac{1}{2}}
$$

Moreover, if the source term is stiff, what means that q verifies (5), C_1 does not depend on q and C_2 depends on q only through $\|q(u_0)\|_{L^1(B(0,R+1))}$.

Proof

Here, we prove the estimate (57) only for stiff source terms; it is the most difficult case. Indeed, if we want to explicit the dependence of C_1 and C_2 upon q, we have to watch all the terms including q. Therefore, the hypothesis of monotony of q is crucial.

In the case where the source term is nonstiff, we would only use the lipschitzianity of q and then get some quantities C_1 and C_2 in the error estimate (57) which depend on the Lipschitz constants of q.

We can split the proof into two steps :

1. For all $\psi \in C_c^\infty(\mathbb{R}^N \times \mathbb{R}_+, \mathbb{R}_+)$, we set :

   $$K = \{(x,t) \in \mathbb{R}^N \times \mathbb{R}_+ \; ; \; \psi(x,t) \neq 0\} \text{ and } K_0 = \{x \in \mathbb{R}^N \; ; \; \psi(x,0) \neq 0\},$$

   and

   $$U_{max} = \max\{\|u\|_{L^\infty(K)}, \; \|\tilde{u}\|_{L^\infty(K)}, \; \|u_0\|_{L^\infty(K_0)}\}.$$

We prove that, for all $\psi \in C_c^\infty(\mathbb{R}^N \times \mathbb{R}_+, \mathbb{R}_+)$, there exists C depending only on ψ, F, u_0 and U_{max} such that :

$$\begin{aligned}
&\int_{\mathbb{R}^N\times\mathbb{R}_+} |\tilde{u}(x,t) - u(x,t)|\psi_t(x,t)dxdt \\
&\quad + \int_{\mathbb{R}^N\times\mathbb{R}_+} (F(x,t,\tilde{u}(x,t)\top u(x,t)) - F(x,t,\tilde{u}(x,t)\bot u(x,t))).\nabla\psi(x,t)dxdt \geq \\
&\qquad -C\left\{\mu_0(K_0) + (\|q(u_0)\|_{L^\infty(K_0)} + 1)(\mu(K))^{\frac{1}{2}} + \mu(K) + \xi(K)\right\}.
\end{aligned} \tag{58}$$

2. Then, we deduce (57).

Step 1

Let us prove (58). We introduce two functions $\rho_N \in C_c^\infty(\mathbb{R}^N, \mathbb{R})$ and $\bar{\rho}_1 \in C_c^\infty(\mathbb{R}, \mathbb{R})$ that satisfy :

$$\left\{\begin{array}{ll}
\text{Supp }(\rho_N) \subset \{x \in \mathbb{R}^N;\ |x| \leq 1\}, & \text{Supp }(\bar{\rho}_1) \subset [-1, 0], \\
\rho_N(x) \geq 0,\ \forall x \in \mathbb{R}^N, & \bar{\rho}_1(x) \geq 0,\ \forall x \in \mathbb{R}, \\
\displaystyle\int_{\mathbb{R}^N} \rho_N(x)dx = 1, & \displaystyle\int_{\mathbb{R}} \bar{\rho}_1(x)dx = 1.
\end{array}\right\} \tag{59}$$

We define the regularizing functions $\rho_{N,r} : x \to r^N\rho_N(rx)$ and $\bar{\rho}_{1,r} : x \to r\bar{\rho}_1(rx)$ for all $r \geq 1$. Let $\psi \in C_c^\infty(\mathbb{R}^N \times \mathbb{R}_+, \mathbb{R}_+)$, we set :

$$\varphi(x,t,y,s) = \psi(x,t)\rho_{N,r}(x-y)\bar{\rho}_{1,r}(t-s).$$

We apply (55) with $\varphi(.,.,y,s)$ and $\kappa = u(y,s)$ and we integrate with respect to y on $\mathbb{R}^N$ and s on $\mathbb{R}_+$, which yields :

$$D_1 + D_2 + D_3 + D_4 + D_5 + D_6 \geq -E, \tag{60}$$

with

$$\begin{aligned}
D_1 &= \int_{(\mathbb{R}^N\times\mathbb{R}_+)^2} |\tilde{u}(x,t) - u(y,s)|\psi_t(x,t)\rho_{N,r}(x-y)\bar{\rho}_{1,r}(t-s)dxdtdyds, \\
D_2 &= \int_{(\mathbb{R}^N\times\mathbb{R}_+)^2} |\tilde{u}(x,t) - u(y,s)|\psi(x,t)\rho_{N,r}(x-y)\bar{\rho}'_{1,r}(t-s)dxdtdyds, \\
D_3 &= \int_{(\mathbb{R}^N\times\mathbb{R}_+)^2} (F(x,t,\tilde{u}(x,t)\top u(y,s)) - F(x,t,\tilde{u}(x,t)\bot u(y,s))). \\
&\qquad\qquad \nabla\psi(x,t)\rho_{N,r}(x-y)\bar{\rho}_{1,r}(t-s)dxdtdyds, \\
D_4 &= \int_{(\mathbb{R}^N\times\mathbb{R}_+)^2} (F(x,t,\tilde{u}(x,t)\top u(y,s)) - F(x,t,\tilde{u}(x,t)\bot u(y,s))).
\end{aligned}$$

$$\nabla\rho_{N,r}(x-y)\psi(x,t)\bar\rho_{1,r}(t-s)dxdtdyds,$$

$$D_5 = -\int_{(\mathbb{R}^N\times\mathbb{R}_+)^2} sign(\tilde u(x,t)-u(y,s))(\mathrm{div}_x F(x,t,u(y,s))-q(x,t,\tilde u(x,t)))\,\psi(x,t)\rho_{N,r}(x-y)\bar\rho_{1,r}(t-s)dxdtdyds,$$

$$D_6 = \int_{\mathbb{R}^N\times\mathbb{R}^N\times\mathbb{R}_+} |u_0(x)-u(y,s)|\psi(x,0)\rho_{N,r}(x-y)\bar\rho_{1,r}(-s)dxdyds$$

and

$$\begin{aligned} E = \int_{(\mathbb{R}^N\times\mathbb{R}_+)^2} &(|\rho_{N,r}(x-y)(\psi_t(x,t)\bar\rho_{1,r}(t-s)+\psi(x,t)\bar\rho'_{1,r}(t-s))| \\ &+|\bar\rho_{1,r}(t-s)(\nabla\psi(x,t)\rho_{N,r}(x-y)+\psi(x,t)\nabla\rho_{N,r}(x-y))|)d\mu(x,t)dyds \\ &+\int_{(\mathbb{R}^N\times\mathbb{R}_+)^2}\psi(x,t)\rho_{N,r}(x-y)\bar\rho_{1,r}(t-s)d\xi(x,t)dyds \\ &+\int_{\mathbb{R}^N\times\mathbb{R}_+\times\mathbb{R}^N}|\psi(x,0)\rho_{N,r}(x-y)\bar\rho_{1,r}(-s)|d\mu_0(x)dyds. \end{aligned}\tag{61}$$

In the sequel, we denote by $(C_i)_{i\in\mathbb{N}}$ some quantities depending only on ψ, F and U_{max}. The properties of the functions ρ_N, $\bar\rho_1$ (59) imply :

$$E \le C_1\Big((r+1)\mu(K)+\xi(K)+\mu_0(K_0)\Big).\tag{62}$$

Then, we set :

$$\begin{aligned} \varepsilon(r,K,u) &= \sup\{\int_K |u(x,t)-u(x+\eta,t+\tau)|dxdt,\ |\eta|\le\frac{1}{r},\ 0\le\tau\le\frac{1}{r}\}, \\ \varepsilon(r,K_0,u_0) &= \sup\{\int_{K_0}|u_0(x)-u_0(x+\eta)|dxdt,\ |\eta|\le\frac{1}{r}\} \end{aligned}$$

Using (59), we also get the following estimates :

$$|D_1-\int_{\mathbb{R}^N\times\mathbb{R}_+}|u(x,t)-\tilde u(x,t)|\psi_t(x,t)dxdt| \le \|\psi_t\|_\infty\varepsilon(r,K,u),\tag{63}$$

and

$$\begin{aligned} |D_3-\int_{\mathbb{R}^N\times\mathbb{R}_+}(F(x,t,\tilde u(x,t)\top u(x,t))-F(x,t,\tilde u(x,t)\bot u(x,t)).\nabla\psi(x,t)dxdt| \\ \le C_2\|\nabla\psi\|_\infty\varepsilon(r,K,u). \end{aligned}\tag{64}$$

Moreover, we can also write (56) with $\varphi(x,t,.,.)$ and $\kappa=\tilde u(x,t)$ and integrate with respect to x and t, which yields :

$$-D_2-D_{40}-D_{50}\ge 0,\tag{65}$$

with

$$\begin{aligned}
D_{40} &= \int_{(\mathbb{R}^N\times\mathbb{R}_+)^2} (F(y,s,u(y,s)\top\tilde{u}(x,t)) - F(y,s,u(y,s)\bot\tilde{u}(x,t))). \\
&\qquad\qquad\qquad \nabla\rho_{N,r}(x-y)\psi(x,t)\bar{\rho}_{1,r}(t-s)dxdydtds \\
D_{50} &= -\int_{(\mathbb{R}^N\times\mathbb{R}_+)^2} sign(\tilde{u}(x,t)-u(y,s))(\mathrm{div}_x F(y,s,\tilde{u}(x,t)) - q(y,s,u(y,s)) \\
&\qquad\qquad\qquad \psi(x,t)\rho_{N,r}(x-y)\bar{\rho}_{1,r}(t-s)dxdtdyds.
\end{aligned}$$

In order to find a lower bound of $-D_2-D_4-D_5$, we compare D_4 to D_{40} and D_5 to D_{50}. If we replace $u(y,s)$ by $u(x,t)$ in $D_{40}-D_4$, we obtain, after having integrated by parts with respect to y :

$$\begin{aligned}
D_{41} = \int_{(\mathbb{R}^N\times\mathbb{R}_+)^2} &(\mathrm{div}_x F(y,s,u(x,t)\top\tilde{u}(x,t)) - \mathrm{div}_x(F(y,s,u(x,t)\bot\tilde{u}(x,t))) \\
&\qquad\qquad \psi(x,t)\rho_{N,r}(x-y)\bar{\rho}_{1,r}(t-s)dxdydtds
\end{aligned}$$

and, thanks to the regularity of F,

$$|D_{40}-D_4-D_{41}| \le C_3\|\psi\|_\infty.\varepsilon(r,K,u) \tag{66}$$

Furthermore, we can rewrite $D_{50}-D_5 = D_{51}+D_{52}+D_{53}$ with :

$$\begin{aligned}
D_{51} &= -\int_{(\mathbb{R}^N\times\mathbb{R}_+)^2} (\mathrm{div}_x F(x,t,\tilde{u}(x,t)\top u(y,s)) - \mathrm{div}_x F(x,t,\tilde{u}(x,t)\bot u(y,s)) \\
&\qquad\qquad\qquad \psi(x,t)\rho_{N,r}(x-y)\bar{\rho}_{1,r}(t-s)dxdtdyds, \\
D_{52} &= \int_{(\mathbb{R}^N\times\mathbb{R}_+)^2} (sign(\tilde{u}(x,t)-u(y,s))(\mathrm{div}_x F(x,t,\tilde{u}(x,t)) - \mathrm{div}_x F(y,s,\tilde{u}(x,t))) \\
&\qquad\qquad\qquad \psi(x,t)\rho_{N,r}(x-y)\bar{\rho}_{1,r}(t-s)dxdtdyds, \\
D_{53} &= \int_{(\mathbb{R}^N\times\mathbb{R}_+)^2} (sign(\tilde{u}(x,t)-u(y,s))(q(y,s,u(y,s)) - q(x,t,\tilde{u}(x,t))) \\
&\qquad\qquad\qquad \psi(x,t)\rho_{N,r}(x-y)\bar{\rho}_{1,r}(t-s)dxdtdyds
\end{aligned}$$

Under the hypotheses (5), the source term q only depends on its third argument and is nonincreasing. Therefore, the term D_{53} has the great sign : $D_{53}\ge 0$. Moreover :

$$|D_{51}+D_{41}| \le C_4\|\psi\|_\infty\Big(\varepsilon(r,K,u)+\frac{1}{r}\Big) \tag{67}$$

and

$$|D_{52}| \le \frac{C_5\|\psi\|_\infty}{r} \tag{68}$$

Therefore, the inequalities (65), (66), (67), (68) lead to :

$$-D_2-D_4-D_5 \ge -C_6\|\psi\|_\infty(\frac{1}{r}+\varepsilon(r,K,u)) \tag{69}$$

It remains to control D_6. In order to get some information about this term, we rewrite (56) with

$$\varphi(y,s) = \psi(x,0)\rho_{N,r}(x-y)\int_s^\infty \bar{\rho}_{1,r}(-\tau)d\tau$$

and $\kappa = u_0(x)$ and we integrate with respect to x. We get :

$$-D_6 \geq D_7 + D_8 + D_9 + D_{10}$$

where

$$D_7 = \int_{\mathbb{R}^N}\int_{\mathbb{R}^N\times\mathbb{R}_+} (F(y,s,u(y,s)\top u_0(x)) - F(y,s,u(y,s)\perp u_0(x))). \\ \nabla\rho_{N,r}(x-y)\psi(x,0)\int_s^\infty \bar{\rho}_{1,r}(-\tau)d\tau dxdyds,$$

$$D_8 = \int_{\mathbb{R}^N}\int_{\mathbb{R}^N\times\mathbb{R}_+} sign(u(y,s)-u_0(x)) \ \mathrm{div}_x F(y,s,u_0(x)) \\ \psi(x,0)\rho_{N,r}(x-y)\int_s^\infty \bar{\rho}_{1,r}(-\tau)d\tau dxdyds,$$

$$D_9 = -\int_{\mathbb{R}^N}\int_{\mathbb{R}^N\times\mathbb{R}_+} sign(u(y,s)-u_0(x)) \ q(u(y,s)) \\ \psi(x,0)\rho_{N,r}(x-y)\int_s^\infty \bar{\rho}_{1,r}(-\tau)d\tau dxdyds,$$

$$D_{10} = \int_{(\mathbb{R}^N)^2} |u_0(x)-u_0(y)|\psi(x,0)\rho_{N,r}(x-y)dxdy.$$

In D_7, we can replace $u_0(x)$ by $u_0(y)$ and integrate by parts with respect to x. It yields :

$$|D_7| \leq C_7(\frac{1}{r} + \varepsilon(r,K_0,u_0))$$

For D_8 and D_{10}, we obtain directly the upper bounds :

$$|D_8| \leq \frac{C_8}{r} \ \text{ and } \ |D_{10}| \leq \|\psi(.,0)\|_\infty \varepsilon(r,K_0,u_0)$$

In D_9, we add and substract $q(u_0)(x)$. Using the monotony of q, we get :

$$D_9 \geq -\frac{\|q(u_0)\|_{L^1(K_0)}\|\psi(.,0)\|_\infty}{r}.$$

Then,

$$-D_6 \geq -C_9\Big(\frac{1}{r}+\varepsilon(r,K_0,u_0)\Big)-\frac{\|q(u_0)\|_{L^1(K_0)}\|\psi(.,0)\|_\infty}{r} \tag{70}$$

and, from (60), (62), (63), (64), (69), (70), we deduce :

$$\int_{\mathbb{R}^N\times\mathbb{R}_+}|\tilde{u}(x,t)-u(x,t)|\psi_t(x,t)dxdt$$
$$+\int_{\mathbb{R}^N\times\mathbb{R}_+}(F(x,t,\tilde{u}(x,t)\top u(x,t))-F(x,t,\tilde{u}(x,t)\bot u(x,t))).\nabla\psi(x,t)dxdt \geq$$
$$-C_{10}\Big((r+1)\mu(K)+\xi(K)+\mu_0(K_0)+\frac{1}{r}+\varepsilon(r,K,u)+\varepsilon(r,K_0,u_0)+\frac{\|q(u_0)\|_{L^1(K_0)}}{r}\Big).$$

But, if $u_0\in BV_{loc}(\mathbb{R}^N)$, there exist C_{K_0,u_0} and $C_{K,u_0,F}$ such that :

$$\varepsilon(r,K_0,u_0)\leq\frac{C_{K_0,u_0}}{r},\quad \varepsilon(r,K,u)\leq\frac{C_{K,u_0,F}}{r}.$$

Then, we get the estimate (58) by taking $r=\dfrac{1}{\sqrt{\mu(K)}}$ (or $r\to\infty$ if $\mu(K)=0$) in the last inequality.

Step 2

This step has already been proved in [1]. Let E be a compact set of $\mathbb{R}^N\times\mathbb{R}_+$. Let

$$U=\max\{\|u\|_{L^\infty(E)},\|\tilde{u}\|_{L^\infty(E)}\}.$$

We set $\omega=V_{[-U,U]}$. Let $R>0$ and $T\in]0,\dfrac{R}{\omega}[$ such that :

$$E\subset\bigcup_{0\leq t\leq T}(B(0,R-\omega t)\times\{t\}).$$

We consider a smooth function $\rho\in C_c^\infty(\mathbb{R}^+,[0,1])$ verifying :

$$\begin{aligned}&\rho(r)=1 \text{ if } r\in[0,R],\\&\rho(r)=0 \text{ if } r\in[R+1,+\infty[,\\&\rho'(r)\leq 0 \text{ for all } r\in\mathbb{R}^+\end{aligned}$$

and we define $\psi\in C_c^\infty(\mathbb{R}^N\times\mathbb{R}_+,\mathbb{R}^+)$ by

$$\psi(x,t)=\begin{cases}\rho(|x|+\omega t)\dfrac{T-t}{T} & \text{for } x\in\mathbb{R} \text{ and } t\in[0,T],\\[2ex] 0 & \text{for } x\in\mathbb{R} \text{ and } t\geq T.\end{cases} \tag{71}$$

and apply (58) with ψ as test function. It gives :

$$\int_E |u(x,t)-\tilde{u}(x,t)|dxdt \le CT\Big(\mu_0(B(0,R+1))+\mu(B(0,R+1)\times[0,T])$$
$$+(\|q(u_0)\|_{L^1(B(0,R+1))}+1)(\mu(B(0,R+1)\times[0,T]))^{\frac{1}{2}}+\xi(B(0,R+1)\times[0,T])\Big)$$

and concludes the proof of Lemma 3.

5.2 Proof of Theorem 1 and Theorem 2

The error estimate given in Theorem 1 (*resp.* Theorem 2) is a straightforward consequence of the continuous entropy inequality (45) (*resp.* (2)) and of Lemma 3.
Indeed, thanks to Theorem 3 and Theorem 4, we know that $u_{\mathcal{T},k}$ defined by (9), (10) or (11) and (13) satisfies the assumptions made on $\tilde{u}$ in Lemma 3 with $\mu_0=\mu_{\mathcal{T}}$, $\mu=\mu_{\mathcal{T},k}$ and $\xi=\xi_{\mathcal{T},k}$. Therefore, in the nonstiff case as in the stiff case, Lemma 3 implies :

$$\int_E |u_{\mathcal{T},k}(x,t)-u(x,t)|dxdt \le C_1\Bigg(\mu_{\mathcal{T}}(B(0,R))+\mu_{\mathcal{T},k}(B(0,R)\times[0,T])$$
$$+(\mu_{\mathcal{T},k}(B(0,R)\times[0,T]))^{\frac{1}{2}}+\xi_{\mathcal{T},k}(B(0,R)\times[0,T])\Bigg)+C_2(\mu_{\mathcal{T},k}(B(0,R)\times[0,T]))^{\frac{1}{2}}$$

Using the properties of $\mu_{\mathcal{T}}$, $\mu_{\mathcal{T},k}$ and $\xi_{\mathcal{T},k}$, we get either (15) or (16).

5.3 Some remarks on the results

We want here to discuss the error estimate in the stiff case. We succeeded in cutting out the Lipschitz constant of q in the error estimate, but finally it remains $\|q(u_0)\|_{L^1}$ and $\|q(u_{\mathcal{T},0})\|_{L^1}$.
If we just know that q satisfies (5) and is of the form $q(s)=\frac{1}{\delta}g(s)$, we may assume $\|g(u_0)\|_{L^1}\le G\delta$ and $\|g(u_{\mathcal{T},0})\|_{L^1}\le G\delta$ in order to get an error estimate independant of δ. It means that the result holds for very small initial data. It is for instance the assumption made by H. J. Schroll and R. Winther in their paper on finite difference schemes for the nonhomogeneous conservation law [20]. The dependence of the error estimate on $\|q(u_0)\|_{L^1}$ also remains in the paper by T. Tang [21] concerning operator-splitting methods.
But you can also find some cases where the smallness of u_0 is not required. There are very particular cases. We propose here the example given by T. Tang in [21] :

$$q(s)=\exp(-\frac{s}{\delta})-1 \text{ with } u_0\ge 0 \text{ on } \mathbb{R}^N.$$

With such q and u_0, $\|q(u_0)\|_{L^1}$ and $\|q(u_{\mathcal{T},0})\|_{L^1}$ are of order 1 with respect to δ.

References

[1] C. Chainais-Hillairet, *Finite volume schemes for a nonlinear hyperbolic equation. Convergence towards the entropy solution and error estimate.*, M^{2}AN **33** (1999), n° 1, 129–156.

[2] A. Chalabi, *Stable upwind schemes for hyperbolic conservation laws with source terms*, IMA J. Numer. Anal. **12** (1992), n° 2, 217–241.

[3] A. Chalabi, *On convergence of numerical schemes for hyperbolic conservation laws with stiff source terms*, Math. Comp. **66** (1997), n° 218, 527–545.

[4] S. Champier, *Error estimates for the approximate solution of a nonlinear hyperbolic equation with source term given by finite volume scheme*, Preprint de l'UMR 5585 (1997), n° 258.

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

[6] B. Cockburn, F. Coquel, P. LeFloch, *An error estimate for finite volume methods for multidimensional conservation laws*, Math. of Comp. **63** (1994), 77–103.

[7] B. Cockburn, F. Coquel, P. LeFloch, *Convergence of the finite volume method for multidimensional conservation laws*, SIAM J. Numer. Anal. **32** (1995), n° 3, 687–705.

[8] M. G. Crandall, A. Majda, *Monotone difference approximations for scalar conservation laws*, Math. Comp. **34** (1980), n° 149, 1–21.

[9] R. Eymard, T. Gallouët, M. Ghilani, R. Herbin, *Error estimates for the approximate solutions of a nonlinear hyperbolic equation given by some finite volume schemes*, I.M.A. Journal of Numer. Anal. **18** (1998), 563–594.

[10] R. Eymard, T. Gallouët, R. Herbin, *The finite volume method*, Ph. Ciarlet and J.L. Lions ed., to appear in the Handbook of Numerical Analysis.

[11] R. Eymard, T. Gallouët, R. Herbin, *Convergence of a finite volume scheme for a nonlinear hyperbolic equation*, Proceedings of the Third colloquium on numerical analysis, Elsevier (D. Bainov, V. Covachev, eds.), 1995, pp. 61–70.

[12] D. Kröner, M. Rokyta, *Convergence of upwind finite volume schemes for scalar conservation laws in two dimensions*, SIAM J. Numer. Anal. **31** (1994), n° 2, 324–343.

[13] S.N. Kruskov, *First order quasilinear equations with several space variables*, Math. USSR. Sb. **10** (1970), 217–243.

[14] N.N. Kuznetsov, *Accuracy of some aproximate methods for computing the weak solutions of a first order quasi-linear equation*, USSR Comp. Math. and Math. Phys. **16** (1976), 105–119.

[15] R. J. LeVeque, H. C. Yee, *A study of numerical methods for hyperbolic conservation laws with stiff source terms*, J. Comput. Phys. **86** (1990), n° 1, 187–210.

[16] T. P. Liu, *Quasilinear hyperbolic systems*, Comm. Math. Phys. **68** (1979), n° 2, 141–172.

[17] T. P. Liu, *Nonlinear stability and instability of transonic flows through a nozzle*, Comm. Math. Phys. **83** (1982), n° 2, 243–260.

[18] R. B. Pember, *Numerical methods for hyperbolic conservation laws with stiff relaxation. I. Spurious solutions*, SIAM J. Appl. Math. **53** (1993), n° 5, 1293–1330.

[19] R. B. Pember, *Numerical methods for hyperbolic conservation laws with stiff relaxation. II. Higher-order Godunov methods*, SIAM J. Sci. Comput. **14** (1993), n° 4, 824–859.

[20] H. J. Schroll, R. Winther, *Finite-difference schemes for scalar conservation laws with source terms*, IMA J. Numer. Anal. **16** (1996), n° 2, 201–215.

[21] T. Tang, *Convergence analysis for operator-splitting methods applied to conservation laws with stiff source terms*, SIAM J. Numer. Anal. **35** (1998), n° 5, 1939–1968.

[22] T. Tang, Z.-H. Teng, *Error bounds for fractional step methods for conservation laws with source terms*, SIAM J. Numer. Anal. **32** (1995), n° 1, 110–127.

[23] T. Tang, Z.-H. Teng, *Monotone difference schemes for two-dimensional nonhomogeneous conservation laws*, Recent advances in differential equations (Kunming, 1997) (Harlow), Longman, Harlow, 1998, pp. 229–243.

[24] J.-P. Vila, *Convergence and error estimate in finite volume schemes for general multidimensional conservation laws*, M^2AN **28** (1994), 267–285.