A result of convergence for a mono-dimensional two-velocities lattice Boltzmann scheme

Filipa Caetano, François Dubois, Benjamin Graille

A RESULT OF CONVERGENCE FOR A MONO-DIMENSIONAL TWO-VELOCITIES LATTICE BOLTZMANN SCHEME

FILIPA CAETANO[1], FRANÇOIS DUBOIS[1,2], AND BENJAMIN GRAILLE[1,♮]

ABSTRACT. We consider a mono-dimensional two-velocities scheme used to approximate the solutions of a scalar hyperbolic conservative partial differential equation. We prove the convergence of the discrete solution toward the unique entropy solution by first estimating the supremum norm and the total variation of the discrete solution, and second by constructing a discrete kinetic entropy-entropy flux pair being given a continuous entropy-entropy flux pair of the hyperbolic system. We finally illustrate our results with numerical simulations of the advection equation and the Burgers equation.

1. INTRODUCTION

The lattice Boltzmann method is a numerical method which is largely used to simulate fluid dynamics equations, such as Navier Stokes, heat, acoustics equations, multi-phase and multi-component fluids (see Succi [24], Lallemand and Luo [18], and *e.g.* [17]). Its origin is in a discretized velocities version of the continuous Boltzmann equation (see Broadwell [5] and Gatignol [10]), with a specific collision kernel. The algorithm of the lattice Boltzmann method reads then as a fully discretized Boltzmann equation on a lattice.

The lattice Boltzmann method supposes that particles progress on a discrete cartesian lattice with a finite set of speeds. In one time step, each velocity allows particles to jump from one vertex of the lattice to another one. One iteration of the method can be described in two steps: a relaxation step, which is local to each vertex and which corresponds to the collision of the particles, followed by a transport step, which corresponds to the evolution of the particles on the lattice.

Despite the fact that they are widely used, mathematical numerical analysis of these kind of methods is far from being complete. In [21], a stability analysis is proposed by Rheinländer for a two velocities lattice Boltzmann scheme for the linear advection equation. In [16], Junk and Yang studied the convergence of approximation of smooth solutions for the incompressible Navier-Stokes equations. Concerning the linear one dimensional convection-diffusion equation, Dellacherie proves the convergence in L^∞ norm of the solutions of a two velocities lattice Boltzmann scheme [7].

In this contribution we consider a two velocities lattice Boltzmann scheme in one dimension, known in the domain as $\mathrm{D}_1\mathrm{Q}_2$, in order to approach a scalar conservation law. The purpose of our paper is to prove the convergence of the scheme towards the unique entropy solution of the hyperbolic equation, as the mesh size tends to zero, when the relaxation parameter of the scheme lies between 0 and 1. We point out a link with the relaxation system of Jin

♮ CORRESPONDING AUTHOR: benjamin.graille@u-psud.fr

[1] LABORATOIRE DE MATHÉMATIQUES D'ORSAY, UNIV. PARIS-SUD, CNRS, UNIVERSITÉ PARIS-SACLAY, 91405 ORSAY, FRANCE.

[2] CONSERVATOIRE NATIONAL DES ARTS ET MÉTIERS, LMSSC LABORATORY, PARIS, FRANCE

Date: May 29, 2019.

2010 *Mathematics Subject Classification.* 35L60, 65M08, 65M12, 76P99.

Key words and phrases. Lattice Boltzmann scheme, relaxation method, hyperbolic system.

and Xin [14], for which the equilibrium is described by the scalar conservation law. In [20] and in [23], Natalini and Serre obtained independently a rigorous proof of the convergence of the solutions of the Jin and Xin system towards the equilibrium, provided that a stability condition known as *sub-characteristic condition* is satisfied. This condition imposes that the solutions of the relaxed equation propagate with characteristic speed smaller than the one of the relaxation solutions. Concerning the numerical approximation of the solutions of scalar conservation laws based on the relaxation approximation of Jin and Xin, Aregba-Driollet and Natalini in [2] and Lattanzio and Serre in [19] proved, by considering two different discretizations of the relaxation system, the convergence of the finite volumes approximated solutions of the Jin and Xin system towards the entropy solution of the scalar conservation laws.

In this paper we will prove, by using finite volumes schemes techniques inspired by [6] such as total variation and L^∞ bounds, that the solutions of the $\mathsf{D}_1\mathsf{Q}_2$ scheme converge towards the numerical solution of a scalar hyperbolic conservation law, under the same stability condition explained above. We remark that in previous works, links between lattice Boltzmann and finite difference methods were done (see Junk [15] and [13]). Concerning non linear models, in [3], Boghosian, Love, and Yepez developed a two velocities entropic scheme for the viscous Burgers equation and studied its properties.

The plan of the paper is as follows. In section 2 we describe the $\mathsf{D}_1\mathsf{Q}_2$ scheme and recall its construction, by following the framework of d'Humières [8]. In section 3 we establish L^∞ and total variation discrete estimates. By using these estimates, in section 4 we prove the convergence of the $\mathsf{D}_1\mathsf{Q}_2$ scheme towards a weak solution of the scalar conservation law, in the case where the relaxation parameter of the scheme lies between 0 and 1. In section 5, by following Serre [23], Bouchut [4], and [9], we introduce numerical entropies for the $\mathsf{D}_1\mathsf{Q}_2$ scheme. Based on a L^1 estimation of the equilibrium gap independent of the mesh step established in section 3, we prove in section 6 the convergence of the $\mathsf{D}_1\mathsf{Q}_2$ scheme towards the unique entropy solution of the scalar conservation law. Finally, section 7 presents several numerical tests that illustrate the results previously obtained.

2. Description of the $\mathrm{D}_1\mathrm{Q}_2$ scheme

We consider the following mono-dimensional scalar conservation law

$$\partial_t u(t,x) + \partial_x \phi(u)(t,x) = 0, \qquad t > 0,\ x \in \mathbb{R}, \tag{1}$$

where the flux ϕ is a $\mathcal{C}^1$ function on $\mathbb{R}$, with the initial condition

$$u(0,x) = u^0(x), \qquad x \in \mathbb{R}. \tag{2}$$

It is well known that the Cauchy problem (1-2) possesses a unique entropy solution which belongs to $\mathsf{L}^\infty(]0,T[\times\mathbb{R})$, for all $T > 0$, and such that $u(t,\cdot) \in \mathsf{BV}(\mathbb{R})$, $\forall t > 0$, provided that the initial data $u^0 \in \mathsf{L}^\infty(\mathbb{R}) \cap \mathsf{BV}(\mathbb{R})$ (see *e.g.* [22], [12]).

In this contribution, a two-velocities lattice Boltzmann scheme is used to approximate the solution of this Cauchy problem.

2.1. The framework of d'Humières.

We use the notation proposed by d'Humières in [8] by considering $\mathcal{L} = \Delta x\mathbb{Z}$, a regular lattice in one dimension of space with typical mesh size Δx. The time step Δt is determined after the specification of the velocity scale λ by the relation:

$$\Delta t = \frac{\Delta x}{\lambda}.$$

For the scheme denoted by $\mathsf{D}_1\mathsf{Q}_2$, we introduce $\boldsymbol{v} = (-\lambda, \lambda)$ the set of the two velocities. The aim of the $\mathsf{D}_1\mathsf{Q}_2$ scheme is to compute a particle distributions vector $\boldsymbol{f} = (f_-, f_+)^{\mathrm{T}}$ on the lattice $\mathcal{L}$ at discrete values of time: it is a numerical scheme that is formally a discretization

in time, space and velocity, where only a finite number of velocities is considered (two, in our contribution), of the Boltzmann equation (even if it cannot be used to simulate this partial differential equation)

$$\partial_t f(t,x,c) + c\partial_x f(t,x,c) = \Omega(f),$$

with a specific collision operator $\Omega(f)$, whose effect is to relax the particle distributions f towards its equilibrium value f^{eq}.

The scheme splits into two phases for each time iteration: first, the relaxation phase that is local in space (corresponding to the consideration of the collision operator Ω), and second, the transport phase for which an exact characteristic method is used.

In the framework proposed by d'Humières [8], the relaxation phase reads as a linear relaxation toward the equilibrium that is diagonal into a peculiar base. The vectors of this base are called "moments", the terminology being taken from the kinetic theory. Moreover, the equilibrium is *a priori* a nonlinear function of the conservative variables.

In the following, we denote by $\boldsymbol{m} = (\mathrm{u}, \mathrm{v})^{\mathrm{T}}$ the moments defined for each space point $x \in \mathcal{L}$ and for each time t by

$$\mathrm{u} = f_- + f_+, \qquad \mathrm{v} = \lambda\left(-f_- + f_+\right).$$

The matrix of the moments $\boldsymbol{M}$ such that $\boldsymbol{m} = \boldsymbol{M}\boldsymbol{f}$ satisfies

$$\boldsymbol{M} = \begin{pmatrix} 1 & 1 \\ -\lambda & \lambda \end{pmatrix}, \qquad \boldsymbol{M}^{-1} = \begin{pmatrix} 1/2 & -1/(2\lambda) \\ 1/2 & 1/(2\lambda) \end{pmatrix}.$$

Let us now describe one time step of the scheme. The starting point is the particle distributions vector $\boldsymbol{f}(x,t)$ in $x \in \mathcal{L}$ at time t, the moments are computed by

$$\boldsymbol{m}(x,t) = \boldsymbol{M}\boldsymbol{f}(x,t). \tag{3}$$

The relaxation phase then reads

$$\mathrm{u}^\star(x,t) = \mathrm{u}(x,t), \qquad \mathrm{v}^\star(x,t) = \mathrm{v}(x,t) + s(\mathrm{v}^{\mathrm{eq}}(x,t) - \mathrm{v}(x,t)), \tag{4}$$

where s is the relaxation parameter and v^{eq} the second moment at equilibrium, which is a function of u. As a consequence, the first moment u is conserved during the relaxation phase. The relaxation parameter s is usually taken in $(0,2]$, for stability reasons.

The particle distributions are then computed after the relaxation phase by

$$\boldsymbol{f}^\star(x,t) = \boldsymbol{M}^{-1}\boldsymbol{m}^\star(x,t). \tag{5}$$

The transport phase finally reads

$$f_\epsilon(x, t+\Delta t) = f_\epsilon^\star(x - \epsilon\Delta x, t), \qquad \epsilon \in \{-,+\}. \tag{6}$$

In order to be consistent with Eq. (1), we impose that $\mathrm{v}^{\mathrm{eq}} = \phi(\mathrm{u})$ [13].

2.2. **A finite volume formalism for a relaxation system.** In order to study the convergence of the scheme, we rewrite it into a finite volume formalism in this section. We first introduce usual finite volume notations. We note $(x_j)_{j\in\mathbb{Z}}$ the sequence of the discrete points in space that make up the lattice $\mathcal{L}$ and $(t^n)_{n\in\mathbb{N}}$ the sequence of the discrete times, with

$$x_j = j\Delta x, \quad j \in \mathbb{Z}, \qquad t^n = n\Delta t, \quad n \in \mathbb{N}.$$

The scheme can be described either in terms of the particle distributions variables or in terms of the moments. It computes $(f_{-,j}^n, f_{+,j}^n)_{j\in\mathbb{Z},n\in\mathbb{N}}$ and $(\mathrm{u}_j^n, \mathrm{v}_j^n)_{j\in\mathbb{Z},n\in\mathbb{N}}$, which are respectively approximations of the local averages of the particle distributions and of the moments in each volume $[x_j, x_{j+1}]\times[t^n, t^{n+1}]$.

According to (3), the particle distributions and the moments are linked by the relations

$$\begin{cases} \mathrm{u}_j^n = f_{-,j}^n + f_{+,j}^n, \\ \mathrm{v}_j^n = -\lambda f_{-,j}^n + \lambda f_{+,j}^n, \end{cases} \qquad \begin{cases} f_{-,j}^n = \frac{1}{2}\mathrm{u}_j^n - \frac{1}{2\lambda}\mathrm{v}_j^n, \\ f_{+,j}^n = \frac{1}{2}\mathrm{u}_j^n + \frac{1}{2\lambda}\mathrm{v}_j^n, \end{cases} \qquad j \in \mathbb{Z},\ n \in \mathbb{N}.$$

We denote by $(f_{-,j}^{n+1/2}, f_{+,j}^{n+1/2})$ (*resp.* $(\mathrm{u}_j^{n+1/2}, \mathrm{v}_j^{n+1/2})$) the approximated particle distributions functions (*resp.* moments) at time t^n after the relaxation step. By using these notations, according to (4) the relaxation step of the scheme reads

$$\mathrm{u}_j^{n+1/2} = \mathrm{u}_j^n, \qquad \mathrm{v}_j^{n+1/2} = (1-s)\mathrm{v}_j^n + s\phi(\mathrm{u}_j^n), \qquad j \in \mathbb{Z},\ n \in \mathbb{N}. \tag{7}$$

After the relaxation step, following (5), the particle distributions are then defined by

$$f_{-,j}^{n+1/2} = \tfrac{1}{2}\mathrm{u}_j^n - \tfrac{1}{2\lambda}\mathrm{v}_j^{n+1/2}, \qquad f_{+,j}^{n+1/2} = \tfrac{1}{2}\mathrm{u}_j^n + \tfrac{1}{2\lambda}\mathrm{v}_j^{n+1/2}, \qquad j \in \mathbb{Z},\ n \in \mathbb{N}, \tag{8}$$

and the transport phase (6) now reads

$$f_{+,j}^{n+1} = f_{+,j-1}^{n+1/2}, \qquad f_{-,j}^{n+1} = f_{-,j+1}^{n+1/2}, \qquad j \in \mathbb{Z},\ n \in \mathbb{N}. \tag{9}$$

We can then describe a complete step of the scheme in terms of the approximated distribution functions $(f_{-,j}^n, f_{+,j}^n)$ by

$$\begin{cases} f_{-,j}^{n+1} = (1-\frac{s}{2})f_{-,j+1}^n + \frac{s}{2}f_{+,j+1}^n - \frac{s}{2\lambda}\phi(f_{-,j+1}^n + f_{+,j+1}^n), \\ f_{+,j}^{n+1} = \frac{s}{2}f_{-,j-1}^n + (1-\frac{s}{2})f_{+,j-1}^n + \frac{s}{2\lambda}\phi(f_{-,j-1}^n + f_{+,j-1}^n), \quad j \in \mathbb{Z},\ n \in \mathbb{N}. \end{cases}$$

and in terms of the approximated moments $(\mathrm{u}_j^n, \mathrm{v}_j^n)$ by

$$\begin{cases} \mathrm{v}_j^{n+1/2} = (1-s)\mathrm{v}_j^n + s\phi(\mathrm{u}_j^n), \\ \mathrm{u}_j^{n+1} = \frac{1}{2}(\mathrm{u}_{j+1}^n + \mathrm{u}_{j-1}^n) - \frac{1}{2\lambda}(\mathrm{v}_{j+1}^{n+1/2} - \mathrm{v}_{j-1}^{n+1/2}), \\ \mathrm{v}_j^{n+1} = \frac{1}{2}(\mathrm{v}_{j+1}^{n+1/2} + \mathrm{v}_{j-1}^{n+1/2}) - \frac{\lambda}{2}(\mathrm{u}_{j+1}^n - \mathrm{u}_{j-1}^n), \quad j \in \mathbb{Z},\ n \in \mathbb{N}, \end{cases} \tag{10}$$

the discrete initial data $(\mathrm{u}_j^0, \mathrm{v}_j^0)_{j\in\mathbb{Z}}$ being taken as an approximation of $(u^0, \phi(u^0))$ that will be specified below.

For later purpose we rewrite the equation defining u_j^{n+1} in (10) as

$$\frac{\mathrm{u}_j^{n+1} - \mathrm{u}_j^n}{\Delta t} + \frac{\mathrm{v}_{j+1}^n - \mathrm{v}_{j-1}^n}{2\Delta x} - \lambda\frac{\mathrm{u}_{j+1}^n - 2\mathrm{u}_j^n + \mathrm{u}_{j-1}^n}{2\Delta x} = \frac{s}{2\lambda\Delta t}\Big[\big(\phi(\mathrm{u}_{j-1}^n) - \mathrm{v}_{j-1}^n\big) - \big(\phi(\mathrm{u}_{j+1}^n) - \mathrm{v}_{j+1}^n\big)\Big], \tag{11}$$

Let us now make a link between the $\mathsf{D}_1\mathsf{Q}_2$ scheme and a relaxation approximation of the scalar conservation law. For given Λ and $\varepsilon > 0$, let us consider the Jin and Xin relaxation system [14]:

$$\begin{cases} \partial_t u + \partial_x v = 0, \\ \partial_t v + \Lambda^2\partial_x u = \dfrac{1}{\varepsilon}(\phi(u) - v). \end{cases} \tag{12}$$

The $\mathsf{D}_1\mathsf{Q}_2$ scheme can be interpreted as a discretization of the relaxation system (12), by a splitting between the hyperbolic part—discretized with a Lax-Friedrichs scheme—and the relaxation part—discretized with an explicit Euler method—where first $\Lambda = \lambda$ and second the

relaxation parameter s, the time step Δt, and the Jin and Xin parameter ε are linked by the relation $\varepsilon = \Delta t/s$.

3. Total variation and L^∞ estimations

In this section we establish L^∞ and BV estimates for the numerical scheme (10). We begin this section by recalling the definitions of the functional spaces that we will use in the sequel and by establishing the assumptions we make in order to get these estimates.

3.1. **Notations and definitions.**

Let us introduce the following notations that will be useful in the sequel. If $\mathrm{w} = (\mathrm{w}_j)_{j\in\mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}$, we denote by $\delta_x \mathrm{w} := ((\delta_x \mathrm{w})_{j+1/2})_{j\in\mathbb{Z}}$ the sequence of $\mathbb{R}^{\mathbb{Z}}$ defined by

$$(\delta_x \mathrm{w})_{j+1/2} = \mathrm{w}_{j+1} - \mathrm{w}_j, \qquad j \in \mathbb{Z}.$$

If $\mathrm{w} = (\mathrm{w}^n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, we denote by $\delta^t \mathrm{w} := ((\delta^t \mathrm{w})^{n+1/2})_{n\in\mathbb{N}}$ the sequence of $\mathbb{R}^{\mathbb{N}}$ defined by

$$(\delta^t \mathrm{w})^{n+1/2} = \mathrm{w}^{n+1} - \mathrm{w}^n, \qquad n \in \mathbb{N}.$$

In order to investigate the convergence of the numerical solution, we recall some classical normed sub-spaces of $\mathbb{R}^{\mathbb{Z}}$ and the corresponding norms.

Definition 1 ($\ell^1(\mathbb{R}^{\mathbb{Z}})$ space, $\|\cdot\|_{\Delta x,\ell^1}$ norm)**.** Let $\mathrm{w} = (\mathrm{w}_j)_{j\in\mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}$ and $\Delta x > 0$ being given. We define the *sequential one norm* by

$$\|\mathrm{w}\|_{\Delta x,\ell^1} = \Delta x \sum_{j\in\mathbb{Z}} |\mathrm{w}_j|,$$

and the associated space $\ell^1(\mathbb{R}^{\mathbb{Z}})$ as

$$\ell^1(\mathbb{R}^Z) = \left\{ \mathrm{w} \in \mathbb{R}^{\mathbb{Z}} \,:\, \|\mathrm{w}\|_{\Delta x,\ell^1} < +\infty \right\}.$$

Definition 2 (sequential total variation)**.** Let $\mathrm{w} = (\mathrm{w}_j)_{j\in\mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}$. We define the *total variation in space* by

$$\mathsf{TV}(\mathrm{w}) = \sum_{j\in\mathbb{Z}} |(\delta_x \mathrm{w})_{j+1/2}|.$$

We recall below the classical Banach space of functions with bounded variation and the compactness Helly's theorem that we will use in the proof of our main result. To more details and to the proofs of these results, we address *e.g.* to [11] and to [1].

Definition 3 (Bounded variation space)**.** Let $\Omega \subset \mathbb{R}^n$ be an open set and $g \in \mathsf{L}^1_{\mathrm{loc}}(\Omega)$. The *total variation* of g over Ω denoted by $\mathsf{TV}_\Omega(g)$ is defined by

$$\mathsf{TV}_\Omega(g) = \sup_{\substack{\varphi\in\mathcal{C}^1_0(\Omega)\\ \|\varphi\|_\infty\le 1}} \int_\Omega g \,\mathrm{div}\,\varphi.$$

We denote by $\mathsf{BV}(\Omega)$ the subspace of $\mathsf{L}^1_{\mathrm{loc}}(\Omega)$ of functions with bounded total variation over Ω:

$$\mathsf{BV}(\Omega) = \left\{ g \in \mathsf{L}^1_{\mathrm{loc}}(\Omega) : \quad \mathsf{TV}_\Omega(g) < +\infty \right\}.$$

Proposition (Functional space)**.** *The space* $\mathsf{L}^1(\Omega) \cap \mathsf{BV}(\Omega)$ *is a Banach space for the norm*

$$\|g\| = \|g\|_{\mathsf{L}^1(\Omega)} + \mathsf{TV}_\Omega(g).$$

Theorem (Helly's sequential theorem)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with a Lipschitz boundary and let $(g_n)_{n\in\mathbb{N}}$ be a sequence which is bounded in $\mathsf{L}^1(\Omega)\cap\mathsf{BV}(\Omega)$. Then there exists a sub-sequence denoted by $(g_{\varphi(n)})_{n\in\mathbb{N}}$ and a function $g \in \mathsf{L}^1(\Omega) \cap \mathsf{BV}(\Omega)$ such that*

$$\begin{cases} g_{\varphi(n)} \longrightarrow g \text{ in } \mathsf{L}^1(\Omega) \text{ and a.e. } x \in \Omega, \\ \mathsf{TV}_\Omega(g) \leq \liminf\limits_{n\in\mathbb{N}} \mathsf{TV}(g_{\varphi(n)}). \end{cases}$$

In the sequel, we use the following notations for the sequences in $\mathbb{R}^{\mathbb{Z}}$

$$f_\pm^n = \left(f_{\pm,j}^n\right)_{j\in\mathbb{Z}}, \quad \mathrm{u}^n = \left(\mathrm{u}_j^n\right)_{j\in\mathbb{Z}}, \quad \mathrm{v}^n = \left(\mathrm{v}_j^n\right)_{j\in\mathbb{Z}}, \quad \phi(\mathrm{u}^n) = \left(\phi(\mathrm{u}_j^n)\right)_{j\in\mathbb{Z}},$$

for $n \in \mathbb{N}$, $\pm \in \{+,-\}$, and we use the following notations for the sequences in $\mathbb{R}^{\mathbb{N}}$

$$f_{\pm,j} = \left(f_{\pm,j}^n\right)^{n\in\mathbb{N}}, \quad \mathrm{u}_j = \left(\mathrm{u}_j^n\right)^{n\in\mathbb{N}}, \quad \mathrm{v}_j = \left(\mathrm{v}_j^n\right)^{n\in\mathbb{N}}, \quad \phi(\mathrm{u}_j) = \left(\phi(\mathrm{u}_j^n)\right)^{n\in\mathbb{N}},$$

for $j \in \mathbb{Z}$, $\pm \in \{+,-\}$.

Let us also introduce the functions

$$h^+(\xi) = \frac{\lambda\xi + \phi(\xi)}{2\lambda}, \qquad h^-(\xi) = \frac{\lambda\xi - \phi(\xi)}{2\lambda}, \qquad \xi \in \mathbb{R}. \tag{13}$$

The functions h^+ and h^- correspond to the equilibrium of the particle distributions functions f_+ and f_-. From relations (7) and (8), the relaxation phase can be rewritten as

$$f_{\pm,j}^{n+1/2} = (1-s)f_{\pm,j}^n + sh^\pm(\mathrm{u}_j^n), \quad j \in \mathbb{Z},\ n \in \mathbb{N}. \tag{14}$$

3.2. **Assumptions.** We will suppose throughout the paper that the initial condition u^0 belongs to $\mathsf{L}^\infty(\mathbb{R}) \cap \mathsf{BV}(\mathbb{R})$. We will then define

$$\alpha = \operatorname{ess\,inf} u^0, \qquad \beta = \operatorname{ess\,sup} u^0,$$
$$M = \max\left\{|\phi'(\xi)|, \text{ for } \alpha \leq \xi \leq \beta\right\}.$$

and we will make the following main assumptions concerning the numerical scheme (10).

Assumption 1. The relaxation parameter s lies in $(0,1]$.

Often in the applications, the value of the relaxation parameter s is chosen larger than 1 and often near 2. In the case of the linear $\mathsf{D}_1\mathsf{Q}_2$, the scheme remains numerically stable in a ℓ^2-sense for $s \in [0,2]$ and the first-order numerical diffusion term is proportional to $1/s - 1/2$: the choice $s = 2$ minimises then the error and the convergence rate is equal to 2 in that case (see [13]). For general lattice Boltzmann schemes, the "optimal" choices for the relaxation parameters are more complicated and are motivated by a combination of stability and accuracy reasons. However, these optimal parameters are generally larger than 1. The reason why we impose that s lies in $(0,1]$ is that the scheme has monotonicity properties in this case. And these properties are essential for dealing with weak-solutions according to our technique of estimates.

Assumption 2. The velocity of the scheme satisfies $\lambda = \Delta x/\Delta t \geq M$.

This assumption is known as the *sub-characteristic condition*. It states that the solutions of the equilibrium equation propagate with characteristic speed smaller that the one of the numerical scheme.

Assumption 3. The initial state is given by

$$\mathrm{u}_j^0 = \frac{1}{\Delta x}\int_{x_j}^{x_{j+1}} u^0(x)\,\mathrm{d}x, \qquad \mathrm{v}_j^0 = \phi(\mathrm{u}_j^0), \qquad j \in \mathbb{Z}.$$

This last assumption means that the initial values of the non-conserved moment are taken to the equilibrium. This choice is done more often than not as this second moment has to be a perturbation of the equilibrium state [18].

3.3. **Preliminary lemmas.** We begin by proving two lemmas that will be used throughout the proofs of the estimates and convergence of the numerical solutions. The first one concerns the total variation of the numerical initial data.

Lemma 1. *The total variation of the discretized initial conserved moment* $\mathrm{u}^0 = (\mathrm{u}_j^0)_{j\in\mathbb{Z}}$ *is controlled by the total variation of the initial function* u^0*:*

$$\mathsf{TV}(\mathrm{u}^0) = \sum_{j\in\mathbb{Z}} |(\delta_x \mathrm{u}^0)_{j+1/2}| \leq \mathsf{TV}(u^0).$$

Proof. Using the definition of u_j^0 given in assumption 3, we have that

$$\begin{aligned}\mathsf{TV}(\mathrm{u}^0) &= \frac{1}{\Delta x}\sum_{j\in\mathbb{Z}}\left|\int_{x_j}^{x_{j+1}} \big(u^0(x+\Delta x) - u^0(x)\big)\,\mathrm{d}x\right| \\ &\leq \frac{1}{\Delta x}\sum_{j\in\mathbb{Z}}\int_{x_j}^{x_{j+1}} \Big|u^0(x+\Delta x) - u^0(x)\Big|\,\mathrm{d}x = \frac{1}{\Delta x}\int_{\mathbb{R}} \Big|u^0(x+\Delta x) - u^0(x)\Big|\,\mathrm{d}x.\end{aligned}$$

We will then prove that

$$\frac{1}{\Delta x}\int_{\mathbb{R}} \Big|u^0(x+\Delta x) - u^0(x)\Big|\,\mathrm{d}x \leq \mathsf{TV}(u^0). \tag{15}$$

Suppose first that $u^0 \in \mathcal{C}^1(\mathbb{R})$. In this case we have that

$$\mathsf{TV}(u^0) = \int_{\mathbb{R}} \Big|(u^0)'(x)\Big|\,\mathrm{d}x$$

and we get

$$\begin{aligned}\frac{1}{\Delta x}\int_{\mathbb{R}} \Big|u^0(x+\Delta x) - u^0(x)\Big|\,\mathrm{d}x &= \frac{1}{\Delta x}\int_{\mathbb{R}} \Big|\int_x^{x+\Delta x} (u^0)'(y)\,\mathrm{d}y\Big|\,\mathrm{d}x \\ &\leq \frac{1}{\Delta x}\int_{\mathbb{R}}\int_{\mathbb{R}} \Big|(u^0)'(y)\Big|\mathbb{1}_{[x,x+\Delta x]}(y)\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{\Delta x}\int_{\mathbb{R}} \Big|(u^0)'(y)\Big|\Delta x\,\mathrm{d}y = \mathsf{TV}(u^0).\end{aligned}$$

We now prove (15) for general $u^0 \in BV(\mathbb{R})$. To do so, we use the following result (see [1]): for $f \in \mathsf{BV}(\mathbb{R})$, there exists a sequence $(f_n)_n \in C^\infty(\mathbb{R})$ such that $f_n \to f$ in $\mathsf{L}^1_{\mathrm{loc}}(\mathbb{R})$ and $\mathsf{TV}(f_n) \to \mathsf{TV}(f)$.

We consider then such a sequence $(f_n)_n$ that converges toward u^0. We have, for all $M > 0$,

$$\frac{1}{\Delta x}\int_{-M}^{M} \Big|f_n(x+\Delta x) - f_n(x)\Big|\,\mathrm{d}x \leq \frac{1}{\Delta x}\int_{\mathbb{R}} \Big|f_n(x+\Delta x) - f_n(x)\Big|\,\mathrm{d}x \leq \mathsf{TV}(f_n), \tag{16}$$

since $f_n \in C^\infty(\mathbb{R})$. We have now, on the one hand, that

$$\frac{1}{\Delta x}\int_{-M}^{M} \Big|f_n(x+\Delta x) - f_n(x)\Big|\,\mathrm{d}x \xrightarrow[n\to\infty]{} \frac{1}{\Delta x}\int_{-M}^{M} \Big|u^0(x+\Delta x) - u^0(x)\Big|\,\mathrm{d}x,$$

since $f_n \to u^0$ in $\mathsf{L}^1_{\mathrm{loc}}(\mathbb{R})$. On the other hand, $\mathsf{TV}(f_n) \to \mathsf{TV}(u^0)$, so that, by passing (16) to the limit as $n \to \infty$, we obtain

$$\frac{1}{\Delta x}\int_{-M}^{M} \Big|u^0(x+\Delta x) - u^0(x)\Big|\,\mathrm{d}x \leq \mathsf{TV}(u^0).$$

By letting now $M \to \infty$, we obtain

$$\frac{1}{\Delta x}\int_{\mathbb{R}}\left|u^0(x+\Delta x)-u^0(x)\right|\mathrm{d}x \le \mathsf{TV}(u^0),$$

that ends the proof. ∎

The second lemma concerns some properties of the functions h^+ and h^- defined by (13).

Lemma 2. *Under assumption 2, we have :*

- *The functions h^+ and h^- are non decreasing over $[\alpha,\beta]$;*
- *For all $\xi \in \mathbb{R}$, $h^+(\xi)+h^-(\xi)=\xi$ and $(h^+)'(\xi)+(h^-)'(\xi)=1$.*

Proof. The functions h^+ and h^- are $\mathcal{C}^1(\mathbb{R})$ and their derivative is non negative over $[\alpha,\beta]$, under assumption 2. The second result of the lemma is a trivial consequence of the definition of the functions $h^\pm$. ∎

3.4. **Uniform bound estimates.** In this section we establish a maximum principle for the numerical solutions of the scheme (10).

Proposition 3 (maximum principle)**.** *Under assumptions 1, 2 and 3, we have*

$$\mathrm{u}_j^n \in [\alpha,\beta], \qquad f_{\pm,j}^n \in [h^\pm(\alpha),h^\pm(\beta)], \qquad j\in\mathbb{Z},\ n\in\mathbb{N}.$$

Proof. A recursive reasoning is done. Since $u^0 \in [\alpha,\beta]$, *a.e.* $x\in\mathbb{R}$, we have that $\mathrm{u}_j^0 \in [\alpha,\beta]$ for $j\in\mathbb{Z}$. By lemma 2, and since $\mathrm{v}_j^0=\phi(\mathrm{u}_j^0)$, we get

$$f_{\pm,j}^0 = h^\pm(\mathrm{u}_j^0) \in [h^\pm(\alpha),h^\pm(\beta)], \qquad j\in\mathbb{Z}.$$

We then assume that the three inclusions are true for a certain $n\in\mathbb{N}$. As $s\in(0,1]$, Eq. (14) implies that $f_{-,j}^{n+1/2}$ and $f_{+,j}^{n+1/2}$ are respectively convex linear combinations of $f_{-,j}^n$ and $h^-(\mathrm{u}_j^n)$, and of $f_{+,j}^n$ and $h^+(\mathrm{u}_j^n)$, so that $f_{-,j}^{n+1/2}\in[h^-(\alpha),h^-(\beta)]$ and $f_{+,j}^{n+1/2}\in[h^+(\alpha),h^+(\beta)]$, for $j\in\mathbb{Z}$. The transport phase just shift the distribution functions, so that the same inclusions yield for $f_{-,j}^{n+1}$ and $f_{+,j}^{n+1}$ for $j\in\mathbb{Z}$. Finally, we have

$$\mathrm{u}_j^{n+1} = f_{-,j}^{n+1}+f_{+,j}^{n+1} \in [h^-(\alpha)+h^+(\alpha),h^-(\beta)+h^+(\beta)] = [\alpha,\beta], \qquad j\in\mathbb{Z},$$

as $h^-(\xi)+h^+(\xi)=\xi$ for all $\xi\in\mathbb{R}$, by lemma 2. ∎

3.5. **Total variation estimates.** We establish now estimates on the total variation in space and in time of the numerical solutions.

Proposition 4 (Spatial total variation estimations)**.** *Under assumptions 1, 2 and 3, the particle distributions functions satisfy the total variation decreasing estimate*

$$\mathsf{TV}(f_+^{n+1})+\mathsf{TV}(f_-^{n+1}) \le \mathsf{TV}(f_+^n)+\mathsf{TV}(f_-^n), \qquad n\in\mathbb{N}.$$

Moreover, we have

$$\mathsf{TV}(f_+^n)+\mathsf{TV}(f_-^n) \le \mathsf{TV}(u^0), \qquad n\in\mathbb{N}, \tag{17}$$

$$\mathsf{TV}(\mathrm{u}^n) \le \mathsf{TV}(u^0), \qquad n\in\mathbb{N}, \tag{18}$$

$$\mathsf{TV}(\mathrm{v}^n) \le \lambda\mathsf{TV}(u^0), \qquad n\in\mathbb{N}. \tag{19}$$

Proof. First we remark that we have $\mathsf{TV}(\mathrm{u}^n) \leq \mathsf{TV}(f_-^n) + \mathsf{TV}(f_+^n)$, as $\mathrm{u}_j^n = f_{-,j}^n + f_{+,j}^n$. We first evaluate the total variation of the approximated particle distributions functions after the relaxation phase. We deduce from (14) that

$$(\delta_x f_\pm^{n+1/2})_{j+1/2} = (1-s)(\delta_x f_\pm^n)_{j+1/2} + s\big(h^\pm(\mathrm{u}_{j+1}^n) - h^\pm(\mathrm{u}_j^n)\big).$$

And we have, by performing a first order Taylor expansion,

$$\begin{aligned} h^\pm(\mathrm{u}_{j+1}^n) - h^\pm(\mathrm{u}_j^n) &= \frac{\lambda(\delta_x \mathrm{u}^n)_{j+1/2} \pm (\phi(\mathrm{u}_{j+1}^n) - \phi(\mathrm{u}_j^n))}{2\lambda} \\ &= \frac{(\lambda \pm \phi'(\xi_{j+1/2}^n))(\delta_x \mathrm{u}^n)_{j+1/2}}{2\lambda} = (h^\pm)'(\xi_{j+1/2}^n)(\delta_x \mathrm{u}^n)_{j+1/2}, \end{aligned}$$

where $\xi_{j+1/2}^n \in (\alpha, \beta)$ is such that $\phi(\mathrm{u}_{j+1}^n) - \phi(\mathrm{u}_j^n) = \phi'(\xi_{j+1/2}^n)(\mathrm{u}_{j+1}^n - \mathrm{u}_j^n)$. We then have, using that $(h^\pm)' \geq 0$ over $[\alpha, \beta]$ and that $s \in (0,1]$,

$$\left|(\delta_x f_\pm^{n+1/2})_{j+1/2}\right| \leq (1-s)\left|(\delta_x f_\pm^n)_{j+1/2}\right| + s(h^\pm)'(\xi_{j+1/2}^n)\left|(\delta_x \mathrm{u}^n)_{j+1/2}\right|.$$

Since $(h^+)' + (h^-)' = 1$, summing the above two inequalities over $j \in \mathbb{Z}$, for $\pm \in \{+,-\}$, yields

$$\mathsf{TV}(f_+^{n+1/2}) + \mathsf{TV}(f_-^{n+1/2}) \leq (1-s)\mathsf{TV}(f_+^n) + (1-s)\mathsf{TV}(f_-^n) + s\mathsf{TV}(\mathrm{u}^n),$$

and finally,

$$\mathsf{TV}(f_+^{n+1/2}) + \mathsf{TV}(f_-^{n+1/2}) \leq \mathsf{TV}(f_+^n) + \mathsf{TV}(f_-^n).$$

Concerning the transport phase, the total variation of the particle distributions functions after this phase is unchanged as they are defined as a translation of the particle distributions before the phase. We have thus

$$\mathsf{TV}(f_+^{n+1}) + \mathsf{TV}(f_-^{n+1}) = \mathsf{TV}(f_+^{n+1/2}) + \mathsf{TV}(f_-^{n+1/2}) \leq \mathsf{TV}(f_+^n) + \mathsf{TV}(f_-^n),$$

which allow us to conclude that

$$\mathsf{TV}(f_+^n) + \mathsf{TV}(f_-^n) \leq \mathsf{TV}(f_-^0) + \mathsf{TV}(f_+^0), \qquad \forall n \in \mathbb{N}. \tag{20}$$

We also have for the total variation of the conserved moment u^n

$$\mathsf{TV}(\mathrm{u}^n) \leq \mathsf{TV}(f_-^n) + \mathsf{TV}(f_+^n) \leq \mathsf{TV}(f_-^0) + \mathsf{TV}(f_+^0). \tag{21}$$

The definition of the initial numerical data in assumption 3 implies that

$$\left|(\delta_x f_\pm^0)_{j+1/2}\right| = \left|h^\pm(\mathrm{u}_{j+1}^0) - h^\pm(\mathrm{u}_j^0)\right| = (h^\pm)'(\xi_{j+1/2}^0)\left|(\delta_x \mathrm{u}^0)_{j+1/2}\right|.$$

By summing these two equalities over $j \in \mathbb{Z}$, for $\pm \in \{+,-\}$, we get $\mathsf{TV}(f_-^0) + \mathsf{TV}(f_+^0) = \mathsf{TV}(\mathrm{u}^0)$, which combined with (20) and with (21) gives (17) and (18). Since $|\mathrm{v}^n| \leq \lambda(|f_+^n| + |f_-^n|)$, we also obtain (19). ∎

In order to control the total variation of the numerical approximation in time and space variables, we prove now uniform total time variation estimates for the approximated solutions. To do so, we estimate the quantity

$$\sum_{j\in\mathbb{Z}} \left|(\delta^t f_{+,j})^{n+1/2}\right| + \left|(\delta^t f_{-,j})^{n+1/2}\right|.$$

Proposition 5 (total variation in time estimates)**.** *Under assumptions 1, 2 and 3, the particle distributions functions satisfy the estimate*

(22)
$$\sum_{j\in\mathbb{Z}} \left|(\delta^t f_{+,j})^{n+1/2}\right| + \left|(\delta^t f_{-,j})^{n+1/2}\right| \leq \sum_{j\in\mathbb{Z}} \left|(\delta^t f_{+,j})^{n-1/2}\right| + \left|(\delta^t f_{-,j})^{n-1/2}\right| \leq 2\mathsf{TV}(u^0), \qquad n\in\mathbb{N}.$$

Moreover we have

$$\sum_{j\in\mathbb{Z}} \left|(\delta^t \mathrm{u}_j)^{n+1/2}\right| \leq 2\mathsf{TV}(u^0), \qquad \sum_{j\in\mathbb{Z}} \left|(\delta^t \mathrm{v}_j)^{n+1/2}\right| \leq 2\lambda\mathsf{TV}(u^0), \qquad n\in\mathbb{N}. \tag{23}$$

Proof. Suppose that

$$\sum_{j\in\mathbb{Z}} \left|(\delta^t f_{+,j})^{n-1/2}\right| + \left|(\delta^t f_{-,j})^{n-1/2}\right| < +\infty.$$

Since $\mathrm{u} = f_+ + f_-$, then we also have

$$\sum_{j\in\mathbb{Z}} \left|(\delta^t \mathrm{u}_j)^{n-1/2}\right| < +\infty.$$

From (9) and (14), we have $f^{n+1}_{\pm,j} = f^{n+1/2}_{\pm,j\mp1} = (1-s) f^n_{\pm,j\mp1} + s h^{\pm}(\mathrm{u}^n_{j\mp1})$. We obtain then

$$\begin{aligned}(\delta^t f_{\pm,j})^{n+1/2} &= (1-s)(\delta^t f_{\pm,j\mp1})^{n-1/2} + s\left(h^{\pm}(\mathrm{u}^n_{j\mp1}) - h^{\pm}(\mathrm{u}^{n-1}_{j\mp1})\right)\\ &= (1-s)(\delta^t f_{\pm,j\mp1})^{n-1/2} + s\frac{\lambda \pm \phi'(\xi^{n-1/2}_{j\mp1})}{2\lambda}(\delta^t \mathrm{u}_{j\mp1})^{n-1/2},\end{aligned}$$

where we used that, for $j\in\mathbb{Z}$,

$$h^{\pm}(\mathrm{u}^n_j) - h^{\pm}(\mathrm{u}^{n-1}_j) = (h^{\pm})'(\xi^{n-1/2}_j)(\delta^t \mathrm{u}_j)^{n-1/2} = \frac{\lambda \pm \phi'(\xi^{n-1/2}_j)}{2\lambda}(\delta^t \mathrm{u}_j)^{n-1/2},$$

with $\xi^{n-1/2}_j$ lying between u^n_j and u^{n-1}_j.

Assumptions 1 and 2 yield

$$\left|(\delta^t f_{\pm,j})^{n+1/2}\right| \leq (1-s)\left|(\delta^t f_{\pm,j\mp1})^{n-1/2}\right| + s\frac{\lambda \pm \phi'(\xi^{n-1/2}_{j\mp1})}{2\lambda}\left|(\delta^t \mathrm{u}_{j\mp1})^{n-1/2}\right|,$$

By summing the above two inequalities over $j\in\mathbb{Z}$, for $\pm\in\{+,-\}$, we get

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \left(\left|(\delta^t f_{+,j})^{n+1/2}\right| + \left|(\delta^t f_{-,j})^{n+1/2}\right|\right) &\leq (1-s)\sum_{j\in\mathbb{Z}} \left(\left|(\delta^t f_{+,j})^{n-1/2}\right| + \left|(\delta^t f_{-,j})^{n-1/2}\right|\right)\\ &\quad + s\sum_{j\in\mathbb{Z}} \left|(\delta^t \mathrm{u}_j)^{n-1/2}\right|.\end{aligned}$$

Since we have $|\mathrm{u}| \leq |f_+| + |f_-|$, we get

$$\sum_{j\in\mathbb{Z}} \left(\left|(\delta^t f_{+,j})^{n+1/2}\right| + \left|(\delta^t f_{-,j})^{n+1/2}\right|\right) \leq \sum_{j\in\mathbb{Z}} \left(\left|(\delta^t f_{+,j})^{n-1/2}\right| + \left|(\delta^t f_{-,j})^{n-1/2}\right|\right).$$

We conclude then, by reasoning recursively, that

$$\sum_{j\in\mathbb{Z}} \left(\left|(\delta^t f_{+,j})^{n+1/2}\right| + \left|(\delta^t f_{-,j})^{n+1/2}\right|\right) \leq \sum_{j\in\mathbb{Z}} \left(\left|(\delta^t f_{+,j})^{1/2}\right| + \left|(\delta^t f_{-,j})^{1/2}\right|\right), \ \forall n\in\mathbb{N}. \tag{24}$$

Now, we have, by using assumption 3,

$$(\delta^t f_{+,j})^{1/2} = |f^1_{+,j} - f^0_{+,j}| = |(1-s)f^0_{+,j-1} + sh^+(\mathrm{u}^0_{j-1}) - f^0_{+,j}| \\ = |h^+(\mathrm{u}^0_{j-1}) - h^+(\mathrm{u}^0_j)| \le |\mathrm{u}^0_{j-1} - \mathrm{u}^0_j|,$$

and, in the same way,

$$(\delta^t f_{-,j})^{1/2} \le |\mathrm{u}^0_{j+1} - \mathrm{u}^0_j|,$$

which, combined with (24), implies (22). Since $|\mathrm{u}^n| \le |f^n_+| + |f^n_-|$ and $|\mathrm{v}^n| \le \lambda(|f^n_+| + |f^n_-|)$, we also obtain (23). ∎

3.6. **Estimation of the equilibrium gap.** We aim to prove that, at the discrete level, the equilibrium gap $\phi(u) - v$ tends to 0, as the discretization step tends to 0. The purpose of the next proposition is then to estimate the quantity $\phi(\mathrm{u}^n_j) - \mathrm{v}^n_j$.

Proposition 6 (estimation of the discrete equilibrium gap)**.** *Under assumptions 1, 2 and 3, the equilibrium gap is bounded, that is*

$$\|\phi(\mathrm{u}^n) - \mathrm{v}^n\|_{\Delta x,\ell^1} \le \frac{2\lambda\Delta x}{s}\mathsf{TV}(u^0), \qquad n \in \mathbb{N}.$$

Proof. We have to estimate

$$\|\phi(\mathrm{u}^n) - \mathrm{v}^n\|_{\Delta x,\ell^1} = \Delta x \sum_{j\in\mathbb{Z}} |\phi(\mathrm{u}^n_j) - \mathrm{v}^n_j|.$$

For $n = 0$, we have $\|\phi(\mathrm{u}^0) - \mathrm{v}^0\|_{\Delta x,\ell^1} = 0$. Then, we use the relations $\phi(\mathrm{u}^{n+1}_j) - \phi(\mathrm{u}^n_j) = \phi'(\zeta^{n+1/2}_j)(\mathrm{u}^{n+1}_j - \mathrm{u}^n_j)$ with $\zeta^{n+1/2}_j \in (\alpha,\beta)$ and $\mathrm{v}^{n+1/2}_j = (1-s)\mathrm{v}^n_j + s\phi(\mathrm{u}^n_j)$. We have

$$\begin{aligned}
\phi(\mathrm{u}^{n+1}_j) - \mathrm{v}^{n+1}_j &= \phi(\mathrm{u}^{n+1}_j) - \phi(\mathrm{u}^n_j) + \phi(\mathrm{u}^n_j) - \mathrm{v}^{n+1/2}_j + \mathrm{v}^{n+1/2}_j - \mathrm{v}^{n+1}_j \\
&= (1-s)\big(\phi(\mathrm{u}^n_j) - \mathrm{v}^n_j\big) + \phi'(\zeta^{n+1/2}_j)(\mathrm{u}^{n+1}_j - \mathrm{u}^n_j) - (\mathrm{v}^{n+1}_j - \mathrm{v}^{n+1/2}_j) \\
&= (1-s)\big(\phi(\mathrm{u}^n_j) - \mathrm{v}^n_j\big) + \phi'(\zeta^{n+1/2}_j)\big(f^{n+1/2}_{+,j-1} + f^{n+1/2}_{-,j+1} - f^{n+1/2}_{+,j} - f^{n+1/2}_{-,j}\big) \\
&\qquad - \lambda\big(f^{n+1/2}_{+,j-1} - f^{n+1/2}_{-,j+1} - f^{n+1/2}_{+,j} + f^{n+1/2}_{-,j}\big) \\
&= (1-s)\big(\phi(\mathrm{u}^n_j) - \mathrm{v}^n_j\big) + \big(\phi'(\zeta^{n+1/2}_j) - \lambda\big)\big(f^{n+1/2}_{+,j-1} - f^{n+1/2}_{+,j}\big) \\
&\qquad + \big(\phi'(\zeta^{n+1/2}_j) + \lambda\big)\big(f^{n+1/2}_{-,j+1} - f^{n+1/2}_{-,j}\big).
\end{aligned}$$

By assumptions 1 and 2, we obtain

$$|\phi(\mathrm{u}^{n+1}_j) - \mathrm{v}^{n+1}_j| \le |1-s|\,|\phi(\mathrm{u}^n_j) - \mathrm{v}^n_j| + 2\lambda\left(|f^{n+1/2}_{+,j-1} - f^{n+1/2}_{+,j}| + |f^{n+1/2}_{-,j+1} - f^{n+1/2}_{-,j}|\right).$$

As a consequence of proposition 4, summing the above inequalities over $j \in \mathbb{Z}$ yields

$$\|\phi(\mathrm{u}^{n+1}) - \mathrm{v}^{n+1}\|_{\Delta x,\ell^1} \le |1-s|\|\phi(\mathrm{u}^n) - \mathrm{v}^n\|_{\Delta x,\ell^1} + 2\lambda\Delta x\mathsf{TV}(\mathrm{u}^0)$$

as $\mathsf{TV}(f^{n+1/2}_+) + \mathsf{TV}(f^{n+1/2}_-) \le \mathsf{TV}(\mathrm{u}^0)$. A recursive reasoning then implies that

$$\|\phi(\mathrm{u}^n) - \mathrm{v}^n\|_{\Delta x,\ell^1} \le 2\lambda\Delta x\mathsf{TV}(\mathrm{u}^0)\sum_{k=0}^{n}|1-s|^k \le \frac{2\lambda\Delta x}{s}\mathsf{TV}(\mathrm{u}^0), \qquad n \in \mathbb{N},$$

as $s \in (0,1]$ by assumption 1. ∎

4. Convergence of the numerical scheme

This section is devoted to proving the convergence of the numerical solution towards a weak solution of the nonlinear Cauchy problem (1-2). This result is obtained as a consequence of the spatial and temporal estimates proved in section 3. The compactness of the numerical sequences is used in the space of the functions with bounded variation and is obtained as a consequence of Helly's theorem.

As usual in the context of finite volume schemes, we seek an approximated solution of the form

$$\mathrm{u}_{\Delta t,\Delta x}(t,x) = \sum_{j\in\mathbb{Z}}\sum_{n\in\mathbb{N}} \mathrm{u}_j^n \, \mathbb{1}_{[t^n,t^{n+1})} \, \mathbb{1}_{[x_j,x_{j+1})},$$
$$\mathrm{v}_{\Delta t,\Delta x}(t,x) = \sum_{j\in\mathbb{Z}}\sum_{n\in\mathbb{N}} \mathrm{v}_j^n \, \mathbb{1}_{[t^n,t^{n+1})} \, \mathbb{1}_{[x_j,x_{j+1})}.$$

We begin by proving the convergence of the sequence $(\mathrm{u}_{\Delta t,\Delta x}, \mathrm{v}_{\Delta t,\Delta x})$, as the space-meshing Δx and the time-step Δt tend to 0, towards a function $(\bar{u},\bar{v})$, where $\bar{u}$ is a weak solution of the Cauchy problem (1-2) and $\bar{v} = \phi(\bar{u})$.

Theorem 7 (convergence towards a weak solution)**.** *Under assumptions 1, 2 and 3, there exist sub-sequences of* $\mathrm{u}_{\Delta t,\Delta x}$, $\mathrm{v}_{\Delta t,\Delta x}$ *(still denoted by* $\mathrm{u}_{\Delta t,\Delta x}$, $\mathrm{v}_{\Delta t,\Delta x}$*) and functions* $\bar{u}$, $\bar{v}$ *with*

$$\bar{u},\bar{v} \in \mathsf{L}^\infty(\mathbb{R}^+\times\mathbb{R}) \cap \mathsf{BV}([0,T]\times\mathbb{R}),$$

for all $T > 0$*, such that*

$$(\mathrm{u}_{\Delta t,\Delta x}, \mathrm{v}_{\Delta t,\Delta x}) \xrightarrow[\Delta t,\Delta x\to 0]{} (\bar{u},\bar{v}) \quad \mathsf{L}^1_{loc}([0,+\infty[\times\mathbb{R}) \times \mathsf{L}^1_{loc}([0,+\infty[\times\mathbb{R}).$$

We have in addition that $\bar{u}$ *is a weak solution of (*1-2*) in* $[0,+\infty[\times\mathbb{R}$*, and that* $\bar{v} = \phi(\bar{u})$*.*

Proof. We have to prove that the families of functions $\mathrm{u}_{\Delta t,\Delta x}$ and $\mathrm{v}_{\Delta t,\Delta x}$ are bounded in $\mathsf{L}^\infty(\Omega) \cap \mathsf{BV}(\Omega)$ for all bounded open subset Ω of $\mathbb{R}^+\times\mathbb{R}$.

Proposition (3) implies that $\mathrm{u}_{\Delta t,\Delta x}$ and $\mathrm{v}_{\Delta t,\Delta x}$ are uniformly bounded in $\mathsf{L}^\infty(\mathbb{R}^+\times\mathbb{R})$.

Let Ω be a bounded open subset of $\mathbb{R}^+\times\mathbb{R}$, such that $\Omega \subseteq [0,T]\times[-M,M]$, for some $T > 0$ and $M > 0$, and let $N, J \in \mathbb{N}$ such that $(N-1)\Delta t \le T < N\Delta t$, $(J-1)\Delta x \le M < J\Delta x$. We have then that

$$\mathsf{TV}_\Omega(\mathrm{u}_{\Delta t,\Delta x}) \le \Delta t \sum_{n=0}^{N}\sum_{j=-J}^{J} \left|\mathrm{u}_{j+1}^n - \mathrm{u}_j^n\right| + \Delta x \sum_{n=0}^{N}\sum_{j=-J}^{J} \left|\mathrm{u}_j^{n+1} - \mathrm{u}_j^n\right|.$$

From Proposition 4 we obtain that

$$\Delta t \sum_{n=0}^{N}\sum_{j=-J}^{J} \left|\mathrm{u}_{j+1}^n - \mathrm{u}_j^n\right| \le \Delta t \sum_{n=0}^{N} \mathsf{TV}(u^0) \le N\Delta t \mathsf{TV}(u^0) \le (T+\Delta t)\mathsf{TV}(u^0),$$

which is bounded as $\Delta t \to 0$.

On the other hand, as a consequence of Proposition 5, we have

$$\Delta x \sum_{n=0}^{N}\sum_{j=-J}^{J} \left|\mathrm{u}_j^{n+1} - \mathrm{u}_j^n\right| \le \Delta x \sum_{n=0}^{N} 2\mathsf{TV}(u^0)$$
$$= \frac{\Delta x}{\Delta t}(N+1)\Delta t 2\mathsf{TV}(u^0) \le 2\lambda(T+\Delta t)\mathsf{TV}(u^0),$$

which is also bounded as $\Delta t \to 0$.

We obtain for $\mathrm{v}_{\Delta t,\Delta x}$ a similar result. These estimations imply in particular that the set $\{(\mathrm{u}_{\Delta t,\Delta x}, \mathrm{v}_{\Delta t,\Delta x})\}_{\Delta t,\Delta x}$ remains bounded in $\mathsf{L}^\infty(\Omega) \cap \mathsf{BV}(\Omega)$, and this is valid for any bounded

set $\Omega \subseteq \mathbb{R}^+ \times \mathbb{R}$. The compactness of $\mathsf{BV}(\Omega) \cap \mathsf{L}^\infty(\Omega)$ in $\mathsf{L}^1(\Omega)$ (Helly's theorem) implies that there exists a sub-sequence of $(\mathrm{u}_{\Delta t,\Delta x}, \mathrm{v}_{\Delta t,\Delta x})$ and functions $(\bar{u}, \bar{v})$ satisfying the conditions of the Theorem.

Proposition 6 now implies that

$$\Big\| \phi(\mathrm{u}_{\Delta t,\Delta x}) - \mathrm{v}_{\Delta t,\Delta x} \Big\|_{\mathsf{L}^\infty([0,T];\mathsf{L}^1_{\mathrm{loc}}(\mathbb{R}))} \xrightarrow[\Delta t,\Delta x\to 0]{} 0.$$

Since $(\mathrm{u}_{\Delta t,\Delta x}, \mathrm{v}_{\Delta t,\Delta x}) \longrightarrow (\bar{u}, \bar{v})$, we get $\phi(\bar{u}) = \bar{v}$, a.e. $x \in \mathbb{R}$, for all $t > 0$.

In order to prove that $\bar{u}$ is a weak solution of (1-2), let us consider $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^+ \times \mathbb{R})$ and put

$$\varphi_j^n = \varphi(t^n, x_j), \qquad \varphi_{\Delta t,\Delta x}(t,x) = \sum_{n\in\mathbb{N}} \sum_{j\in\mathbb{Z}} \varphi_j^n \mathbb{1}_{[t^n,t^{n+1})}(t) \mathbb{1}_{[x_j,x_{j+1})}(x).$$

We multiply both sides of (11) by $\Delta t \Delta x \varphi_j^n$, we sum over $n \in \mathbb{N}$ and $j \in \mathbb{Z}$, do a discrete integration by parts, and pass to the limit as $\Delta t, \Delta x \to 0$. The first term on the left-hand side reads

$$\begin{aligned}
&\Delta t\Delta x \sum_{n\geq 1} \sum_{j\in\mathbb{Z}} \mathrm{u}_j^n \frac{\varphi_j^{n-1} - \varphi_j^n}{\Delta t} - \Delta t\Delta x \sum_{j\in\mathbb{Z}} \frac{\mathrm{u}_j^0 \varphi_j^0}{\Delta t} \\
&\qquad = \int_{\Delta t}^{+\infty} \int_{\mathbb{R}} \mathrm{u}_{\Delta t,\Delta x}(t,x) \frac{\varphi_{\Delta t,\Delta x}(t-\Delta t,x) - \varphi_{\Delta t,\Delta x}(t,x)}{\Delta t} \,\mathrm{d}x\,\mathrm{d}t \\
&\qquad\qquad - \int_{\mathbb{R}} \mathrm{u}_{\Delta t,\Delta x}(0,x) \varphi_{\Delta t,\Delta x}(0,x) \,\mathrm{d}x \\
&\qquad\qquad \xrightarrow[\Delta t,\Delta x\to 0]{} - \int_0^{+\infty} \int_{\mathbb{R}} \bar{u}(t,x) \partial_t \varphi(t,x) \,\mathrm{d}x\,\mathrm{d}t - \int_{\mathbb{R}} u^0(x) \varphi(0,x) \,\mathrm{d}x.
\end{aligned}$$

By a similar reasoning we get for the second term on the left-hand side

$$\begin{aligned}
&\Delta t\Delta x \sum_{n\geq 0} \sum_{j\in\mathbb{Z}} \mathrm{v}_j^n \frac{\varphi_{j-1}^n - \varphi_{j+1}^n}{2\Delta x} \\
&\qquad = \int_0^{+\infty} \int_{\mathbb{R}} \mathrm{v}_{\Delta t,\Delta x}(t,x) \frac{\varphi_{\Delta t,\Delta x}(t,x-\Delta x) - \varphi_{\Delta t,\Delta x}(t,x+\Delta x)}{2\Delta x} \,\mathrm{d}x\,\mathrm{d}t \\
&\qquad \xrightarrow[\Delta t,\Delta x\to 0]{} - \int_0^{+\infty} \int_{\mathbb{R}} \bar{v}(t,x) \partial_x \varphi(t,x) \,\mathrm{d}x\,\mathrm{d}t = - \int_0^{+\infty} \int_{\mathbb{R}} \phi(\bar{u}(t,x)) \partial_x \varphi(t,x) \,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}$$

Concerning the third term, it can be written as

$$\int_0^{+\infty} \int_{\mathbb{R}} \mathrm{u}_{\Delta t,\Delta x}(t,x) \Big(\frac{\varphi_{\Delta t,\Delta x}(t,x-\Delta x) - \varphi_{\Delta t,\Delta x}(t,x)}{2\Delta x} - \frac{\varphi_{\Delta t,\Delta x}(t,x) - \varphi_{\Delta t,\Delta x}(t,x+\Delta x)}{2\Delta x} \Big) \,\mathrm{d}x\,\mathrm{d}t,$$

which vanishes as $\Delta t, \Delta x \to 0$. Let us now treat the right-hand side of (11). After integration by parts, we get

$$\begin{aligned}
&s\Delta t\Delta x \sum_{n\in\mathbb{N}} \sum_{j\in\mathbb{Z}} \big(\phi(\mathrm{u}_j^n) - \mathrm{v}_j^n \big) \frac{\varphi_{j+1}^n - \varphi_{j-1}^n}{2\Delta x} \\
&\qquad = s \int_0^{+\infty} \int_{\mathbb{R}} \big(\phi(\mathrm{u}_{\Delta t,\Delta x}) - \mathrm{v}_{\Delta t,\Delta x} \big) \frac{\varphi_{\Delta t,\Delta x}(t,x+\Delta x) - \varphi_{\Delta t,\Delta x}(t,x-\Delta x)}{2\Delta x} \,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}$$

Let $T > 0$, $M > 0$, such that $\mathrm{supp}(\varphi) \subseteq [0,T] \times [-M, M]$. Since

$$\left| \int_0^{+\infty} \int_{\mathbb{R}} \left(\phi(\mathrm{u}_{\Delta t,\Delta x}) - \mathrm{v}_{\Delta t,\Delta x}\right) \frac{\varphi_{\Delta t,\Delta x}(t, x+\Delta x) - \varphi_{\Delta t,\Delta x}(t, x-\Delta x)}{2\Delta x} \,\mathrm{d}x\,\mathrm{d}t \right|$$
$$\leq \iint_{\mathrm{supp}(\varphi)} \left|\phi(\mathrm{u}_{\Delta t,\Delta x}) - \mathrm{v}_{\Delta t,\Delta x}\right| \|\partial_x \varphi\|_\infty \,\mathrm{d}x\,\mathrm{d}t$$
$$\leq \|\partial_x \varphi\|_\infty \|\phi(\mathrm{u}_{\Delta t,\Delta x}) - \mathrm{v}_{\Delta t,\Delta x}\|_{\mathrm{L}^\infty([0,T];\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})},$$

we conclude, as a consequence of Proposition 6, that this last term tends towards 0 when Δt and Δx go to 0. We conclude then that $\bar{u}$ is a weak solution of (1). ∎

5. Entropies and numerical entropy estimates

In this section we aim to establish discrete entropy estimates for the numerical scheme (10). To do so, we will use the relaxation entropies introduced in [23] in order to construct numerical entropies for the scheme. We will also make a link between these relaxation entropies and a kinetic decomposition of the dual entropy for the nonlinear conservation law, introduced in [4, 9].

5.1. **Entropy.**

Let us consider an entropy-entropy flux pair (η, q), with η strictly convex, for the hyperbolic scalar conservation law (1),

$$\eta'(u)\phi'(u) = q'(u), \qquad \eta'' > 0.$$

For given λ, we introduce the homogeneous Jin and Xin relaxation system ([14]):

$$\begin{cases} \partial_t u + \partial_x v = 0, \\ \partial_t v + \lambda^2 \partial_x u = 0, \end{cases} \tag{25}$$

which we can write in an equivalent way in the characteristic variables (f_+, f_-):

$$\begin{cases} \partial_t f_+ + \lambda \partial_x f_+ = 0, \\ \partial_t f_- - \lambda \partial_x f_- = 0. \end{cases} \tag{26}$$

Following Serre ([23]), we define an entropy couple (E, Q) for system (25) as follows. Let us first define the couple $(\mathrm{e}^+, \mathrm{e}^-)$ by

$$\mathrm{e}^\pm(f) = \frac{\lambda\eta \pm q}{2\lambda}((h^\pm)^{-1}(f)), \qquad f \in [h^\pm(\alpha), h^\pm(\beta)],$$

and the couple $(\tilde{\eta}, \tilde{q})$ by

$$(\tilde{\eta}, \tilde{q})(f_-, f_+) = \left(\mathrm{e}^+(f_+) + \mathrm{e}^-(f_-), \lambda\mathrm{e}^+(f_+) - \lambda\mathrm{e}^-(f_-)\right), \qquad f_\pm \in [h^\pm(\alpha), h^\pm(\beta)].$$

We remark then that the couple $(\tilde{\eta}, \tilde{q})$ is an entropy-flux pair for the system (26) in the set $[h^+(\alpha), h^+(\beta)] \times [h^-(\alpha), h^-(\beta)]$.

Let us now define the couple (E, Q) by

$$E(u, v) = \mathrm{e}^+\left(\frac{\lambda u + v}{2\lambda}\right) + \mathrm{e}^-\left(\frac{\lambda u - v}{2\lambda}\right),$$
$$Q(u, v) = \lambda\mathrm{e}^+\left(\frac{\lambda u + v}{2\lambda}\right) - \lambda\mathrm{e}^-\left(\frac{\lambda u - v}{2\lambda}\right),$$

for (u, v) such that $(\lambda u + v)/(2\lambda) \in [h^+(\alpha), h^+(\beta)]$, $(\lambda u - v)/(2\lambda) \in [h^-(\alpha), h^-(\beta)]$. Then, the couple (E, Q) is an entropy-flux pair for the system (25), which satisfies

$$(E, Q)(u, \phi(u)) = (\eta, q)(u), \quad \forall u \in [\alpha, \beta].$$

Let us now state some properties of the entropies that will be useful to establish entropy estimates for the numerical solutions.

Due to the assumption (2), we can easily check the following lemma.

Lemma 8. *Under assumption* (2)*, the functions* e^+ *and* e^- *satisfy*

$$\mathrm{e}^{\pm\prime}(f) = \eta'((h^\pm)^{-1}(f)), \quad \mathrm{e}^{\pm\prime\prime}(f) > 0, \quad f \in [h^\pm(\alpha), h^\pm(\beta)].$$

Furthermore we have

$$\mathrm{e}^{+\prime}(h^+(u)) = \mathrm{e}^{-\prime}(h^-(u)) = \eta'(u).$$

Proof. The result of the lemma follows immediately from the definition of $\mathrm{e}^\pm$ and from the relation $\eta'\phi' = q'$. Since $h^{\pm\prime}(\xi) = \frac{\lambda \pm \phi'(\xi)}{2\lambda}$, for $\xi \in \mathbb{R}$, we have

$$\mathrm{e}^{\pm\prime}(f) = \left(\eta' \frac{\lambda \pm \phi'}{2\lambda}\right)((h^\pm)^{-1}(f)) \frac{1}{(h^\pm)'((h^\pm)^{-1}(f))} = \eta'((h^\pm)^{-1}(f)).$$

■

5.2. **Dual entropy.** In this paragraph we make a link between the relaxation entropies $\mathrm{e}^+, \mathrm{e}^-$ defined previously with the kinetic decomposition of the dual entropy introduced by Bouchut in [4] and by one of the authors in [9] for the nonlinear conservation law (1).

Let us introduce the entropy variable ϑ for the scalar conservation law:

$$\vartheta = \eta'(u).$$

We define the dual entropy $\vartheta \longmapsto \eta^\star(\vartheta)$ according to

$$\eta^\star(\vartheta) = \sup_w \big(\vartheta w - \eta(w)\big).$$

We then have

$$\eta^\star(\vartheta) = \vartheta(\eta')^{-1}(\vartheta) - \eta\big((\eta')^{-1}(\vartheta)\big) = \vartheta u - \eta(u),$$

$$\frac{\mathrm{d}\eta^\star}{\mathrm{d}\vartheta}(\vartheta) = (\eta')^{-1}(\vartheta) = u.$$

The dual entropy flux $\vartheta \longmapsto q^\star(\vartheta)$ is then defined according to

$$q^\star(\vartheta) = \vartheta\phi\big((\eta')^{-1}(\vartheta)\big) - q\big((\eta')^{-1}(\vartheta)\big) = \vartheta\phi(u) - q(u).$$

We have the following differential

$$\frac{\mathrm{d}q^\star}{\mathrm{d}\vartheta}(\vartheta) = \phi\big((\eta')^{-1}(\vartheta)\big) = \phi(u).$$

Following [9], let us introduce two convex functions $\vartheta \longmapsto \psi_\pm^\star(\vartheta)$ that satisfy the kinetic decomposition of the dual entropy

$$\psi_+^\star(\vartheta) + \psi_-^\star(\vartheta) = \eta^\star(\vartheta), \qquad \lambda\psi_+^\star(\vartheta) - \lambda\psi_-^\star(\vartheta) = q^\star(\vartheta), \tag{27}$$

and the equilibrium functions $u \longmapsto f_\pm^{\mathrm{eq}}(u)$ according to the relation

$$f_\pm^{\mathrm{eq}}(u) = \frac{\mathrm{d}\psi_\pm^\star}{\mathrm{d}\vartheta}\big(\eta'(u)\big) = \frac{\mathrm{d}\psi_\pm^\star}{\mathrm{d}\vartheta}(\vartheta).$$

We can then define a kinetic hyperbolic system

$$\begin{cases} \partial_t f_+ + \lambda\partial_x f_+ = \dfrac{1}{\tau}\big(f_+^{\mathrm{eq}}(f_+ + f_-) - f_+\big), \\ \partial_t f_- - \lambda\partial_x f_- = \dfrac{1}{\tau}\big(f_-^{\mathrm{eq}}(f_+ + f_-) - f_-\big). \end{cases} \tag{28}$$

We introduce the duals $f \longmapsto \psi_\pm(f)$ of the functions $\psi_\pm^\star$ defined in (27):

$$\psi_\pm(f) = \sup_\vartheta \big(f\vartheta - \psi_\pm^\star(\vartheta) \big).$$

Then

$$\psi_\pm'(f) = \left(\frac{\mathrm{d}\psi_\pm^\star}{\mathrm{d}\vartheta} \right)^{-1} (f).$$

We know that such a framework is able to put in evidence a "H-theorem" [9]. Just multiply each equation of (28) by $\psi_\pm'(f_\pm)$; then

$$\partial_t \big(\psi_+(f_+) + \psi_-(f_-) \big) + \partial_x \big(\lambda\psi_+(f_+) - \lambda\psi_-(f_-) \big) \leq 0.$$

The natural question is to make a link between this framework and the tools introduced in this contribution, *id est* to make a link between $h^\pm$ and $f_\pm^{\mathrm{eq}}$ and between $\mathrm{e}^\pm$ and $\psi_\pm$. We have the following proposition.

Proposition 9 (Duality and entropy decomposition)**.** *We have*

$$f_\pm^{\mathrm{eq}}(\xi) = h^\pm(\xi), \qquad \xi \in \mathbb{R},$$
$$\psi_\pm(f) = \mathrm{e}^\pm(f), \qquad f \in [h^\pm(\alpha), h^\pm(\beta)].$$

Proof. We first have

$$\psi_\pm^\star(\vartheta) = \frac{\lambda\eta^\star(\vartheta) \pm q^\star(\vartheta)}{2\lambda}, \qquad \frac{\mathrm{d}\psi_\pm^\star}{\mathrm{d}\vartheta}(\vartheta) = \frac{\lambda u \pm \phi(u)}{2\lambda} = h^\pm(u) \quad \text{with } \vartheta = \eta'(u).$$

Then, we can identify $f_\pm^{\mathrm{eq}} = h^\pm$.

At the optimum value ϑ that define $\psi_\pm(f)$, we have $f = \big(\psi_\pm^\star\big)'(\vartheta) = h^\pm(u)$ with $\eta'(u) = \vartheta$. Then

$$\begin{aligned}
\psi_\pm(f) &= f\vartheta - \frac{1}{2}\Big(\eta^\star(\vartheta) \pm \frac{1}{\lambda} q^\star(\vartheta) \Big) = f\vartheta - \frac{1}{2}\Big(\vartheta u - \eta(u) \pm \frac{1}{\lambda}\big(\vartheta\phi(u) - q(u) \big) \Big) \\
&= \vartheta\Big(f - \frac{1}{2}\Big(u \pm \frac{1}{\lambda}\phi(u) \Big) \Big) + \frac{1}{2}\Big(\eta(u) \pm \frac{1}{\lambda} q(u) \Big) \\
&= \vartheta(f - h^\pm(u)) + \frac{1}{2}\Big(\eta \pm \frac{1}{\lambda} q \Big)\Big((h^\pm)^{-1}(f) \Big) = \mathrm{e}^\pm(f),
\end{aligned}$$

and the proof is established. ∎

5.3. **Numerical entropies estimates.** In this section we construct numerical entropies and corresponding numerical entropy-fluxes for the numerical scheme and we prove the dissipation of the numerical entropy.

Inspired by [19], let us define the numerical entropies for the numerical scheme by

$$E_j^{n+1/2} = \mathrm{e}^+(f_{+,j}^{n+1/2}) + \mathrm{e}^-(f_{-,j}^{n+1/2}), \qquad j \in \mathbb{Z},\ n \in \mathbb{N}, \tag{29}$$
$$Q_{j+1/2}^{n+1/2} = \lambda\mathrm{e}^+(f_{+,j}^{n+1/2}) - \lambda\mathrm{e}^-(f_{-,j+1}^{n+1/2}), \qquad j \in \mathbb{Z},\ n \in \mathbb{N}.$$

Note that the numerical entropies are defined for each time step after the relaxation phase and before the transport phase. This is essential in order to obtained the estimates. Let us now define the numerical entropy production

$$\mathcal{E}_j^n = \frac{E_j^{n+1/2} - E_j^{n-1/2}}{\Delta t} + \frac{Q_{j+1/2}^{n-1/2} - Q_{j-1/2}^{n-1/2}}{\Delta x}, \qquad j \in \mathbb{Z},\ n \in \mathbb{N}^*. \tag{30}$$

We begin by proving that the entropy production has a sign.

Proposition 10. *Under assumptions 1 and 2, we have*

$$\mathcal{E}_j^n \leq 0, \qquad j \in \mathbb{Z},\ n \in \mathbb{N}^*.$$

Proof. We develop $\mathcal{E}_j^n$. Since we have $\lambda = \Delta x/\Delta t$, we obtain

$$\mathcal{E}_j^n = \frac{\mathrm{e}^+(f_{+,j}^{n+1/2}) - \mathrm{e}^+(f_{+,j-1}^{n-1/2})}{\Delta t} + \frac{\mathrm{e}^-(f_{-,j}^{n+1/2}) - \mathrm{e}^-(f_{-,j+1}^{n-1/2})}{\Delta t}.$$

By using one time step of the scheme, we have

$$f_{\pm,j}^{n+1/2} = f_{\pm,j\mp1}^{n-1/2} \pm \frac{s}{2\lambda}\big(\phi(\mathrm{u}_j^{n-1/2}) - \mathrm{v}_j^{n-1/2}\big) = f_{\pm,j\mp1}^{n-1/2} \pm \frac{s}{2\lambda(1-s)}\big(\phi(\mathrm{u}_j^{n+1/2}) - \mathrm{v}_j^{n+1/2}\big),$$

the convexity of $\mathrm{e}^\pm$ implies that

$$\mathrm{e}^\pm(f_{\pm,j}^{n+1/2}) \mp \frac{s}{2\lambda(1-s)}\big(\phi(\mathrm{u}_j^{n+1/2}) - \mathrm{v}_j^{n+1/2}\big)\mathrm{e}^{\pm\prime}(f_{\pm,j}^{n+1/2}) \leq \mathrm{e}^\pm(f_{\pm,j\mp1}^{n-1/2}).$$

We have thus that

$$\frac{\mathrm{e}^\pm(f_{\pm,j}^{n+1/2}) - \mathrm{e}^\pm(f_{\pm,j\mp1}^{n-1/2})}{\Delta t} \leq \pm\frac{s}{2\lambda(1-s)}\big(\phi(\mathrm{u}_j^{n+1/2}) - \mathrm{v}_j^{n+1/2}\big)\mathrm{e}^{\pm\prime}(f_{\pm,j}^{n+1/2}).$$

Hence we get

$$\mathcal{E}_j^n \leq \frac{s}{2\lambda(1-s)}\big(\phi(\mathrm{u}_j^{n+1/2}) - \mathrm{v}_j^{n+1/2}\big)\big(\mathrm{e}^{+\prime}(f_{+,j}^{n+1/2}) - \mathrm{e}^{-\prime}(f_{-,j}^{n+1/2})\big). \tag{31}$$

Now from

$$f_{\pm,j}^{n+1/2} = h^\pm(\mathrm{u}_j^{n+1/2}) \pm \frac{1}{2\lambda}\big(\mathrm{v}_j^{n+1/2} - \phi(\mathrm{u}_j^{n+1/2})\big)$$

it follows that

$$\mathrm{e}^{\pm\prime}(f_{\pm,j}^{n+1/2}) = \mathrm{e}^{\pm\prime}(h^\pm(\mathrm{u}_j^{n+1/2})) \pm \frac{1}{2\lambda}\big(\mathrm{v}_j^{n+1/2} - \phi(\mathrm{u}_j^{n+1/2})\big)\mathrm{e}^{\pm\prime\prime}(\xi_j^{\pm,n+1/2}),$$

where $\xi_j^{\pm,n+1/2}$ lie between $f_{\pm,j}^{n+1/2}$ and $h^\pm(\mathrm{u}_j^{n+1/2})$. Combining with (31) and using Lemma 8, we conclude that

$$\mathcal{E}_j^n \leq -\frac{s}{4\lambda^2(1-s)}\big(\phi(\mathrm{u}_j^{n+1/2}) - \mathrm{v}_j^{n+1/2}\big)^2\big(\mathrm{e}^{+\prime\prime}(\xi_j^{+,n+1/2}) + \mathrm{e}^{-\prime\prime}(\xi_j^{-,n+1/2})\big) \leq 0,$$

and the proof is established. ∎

6. Convergence towards the entropic solution

In this section we establish the final convergence result, by using the discrete entropy estimates obtained at the previous section. We will prove that the weak solution of the nonlinear Cauchy problem (1-2) given by Theorem 7, obtained as the limit $\bar{u}$ of the numerical scheme, is indeed the unique entropic solution of the Cauchy problem (1-2).

Theorem 11 (convergence result)**.** *Let u be a weak solution of the Cauchy problem (*1-2*) given by Theorem 7. Then we have that u is the unique entropic solution of (*1-2*).*

Proof. Let (η, q) be an entropy-entropy-flux pair for (1), with η strictly convex and let $\mathcal{E}_j^n$ be defined by (30). Let us also consider $\psi \in \mathcal{D}(]0,+\infty[\times\mathbb{R})$, $\psi \geq 0$ and put

$$\psi_j^n = \psi(t^n, x_j), \qquad \psi_{\Delta t,\Delta x}(t,x) = \sum_{n\in\mathbb{N}}\sum_{j\in\mathbb{Z}} \mathbb{1}_{[t^n,t^{n+1})}(t)\mathbb{1}_{[x_j,x_{j+1})}(x)\psi_j^n, \qquad j \in \mathbb{Z},\ n \in \mathbb{N}.$$

The result of Proposition 10 implies that

$$\Delta t\Delta x\mathcal{E}_j^n\psi_j^n \leq 0, \qquad j \in \mathbb{Z},\ n \in \mathbb{N}^*,$$

and by summing over $n \in \mathbb{N}^*$ and over $j \in \mathbb{Z}$, we get

$$\Delta t \Delta x \sum_{j\in\mathbb{Z}} \sum_{n\in\mathbb{N}^*} \psi_j^n \left(\frac{E_j^{n+1/2} - E_j^{n-1/2}}{\Delta t} + \frac{Q_{j+1/2}^{n-1/2} - Q_{j-1/2}^{n-1/2}}{\Delta x} \right) \leq 0.$$

We now do a discrete integration by parts. We obtain

$$\Delta t \Delta x \left(\sum_{j\in\mathbb{Z}} \sum_{n\in\mathbb{N}} E_j^{n+1/2} \frac{\psi_j^n - \psi_j^{n+1}}{\Delta t} + \sum_{j\in\mathbb{Z}} \sum_{n\in\mathbb{N}^*} Q_{j+1/2}^{n-1/2} \frac{\psi_j^n - \psi_{j+1}^n}{\Delta x} \right) \leq 0.$$

By using the definition (29) of the numerical entropies $E_j^{n+1/2}$ and $Q_{j+1/2}^{n-1/2}$, the above inequality writes as

$$\sum_{j\in\mathbb{Z}} \sum_{n\in\mathbb{N}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \left(e^+(f_{+,j}^{n+1/2}) + e^-(f_{-,j}^{n+1/2}) \right) \frac{\psi(t^n, x_j) - \psi(t^{n+1}, x_j)}{\Delta t} \, \mathrm{d}t \, \mathrm{d}x$$
$$+ \sum_{j\in\mathbb{Z}} \sum_{n\in\mathbb{N}^*} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \left(\lambda e^+(f_{+,j}^{n+1/2}) - \lambda e^-(f_{-,j+1}^{n+1/2}) \right) \frac{\psi(t^n, x_j) - \psi(t^n, x_{j+1})}{\Delta x} \, \mathrm{d}t \, \mathrm{d}x \leq 0.$$

We now use that $f_{+,j}^{n+1} = f_{+,j-1}^{n+1/2}$, $f_{-,j}^{n+1} = f_{-,j+1}^{n+1/2}$, for all $j \in \mathbb{Z}$ and $n \in \mathbb{N}$. We obtain

$$\sum_{j\in\mathbb{Z}} \sum_{n\in\mathbb{N}} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \left(e^+(f_{+,j+1}^{n+1}) + e^-(f_{-,j-1}^{n+1}) \right) \frac{\psi(t^n, x_j) - \psi(t^{n+1}, x_j)}{\Delta t} \, \mathrm{d}t \, \mathrm{d}x$$
$$+ \sum_{j\in\mathbb{Z}} \sum_{n\in\mathbb{N}^*} \int_{x_j}^{x_{j+1}} \int_{t^n}^{t^{n+1}} \left(\lambda e^+(f_{+,j+1}^{n+1}) - \lambda e^-(f_{-,j}^{n+1}) \right) \frac{\psi(t^n, x_j) - \psi(t^n, x_{j+1})}{\Delta x} \, \mathrm{d}t \, \mathrm{d}x$$
$$= \int_0^{+\infty} \int_{\mathbb{R}} \left(e^+(f_{\Delta t,\Delta x}^+(t + \Delta t, x + \Delta x)) + e^-(f_{\Delta t,\Delta x}^-(t + \Delta t, x - \Delta x)) \right)$$
$$\times \left(\frac{\psi_{\Delta t,\Delta x}(t, x) - \psi_{\Delta t,\Delta x}(t + \Delta t, x)}{\Delta t} \right) \mathrm{d}t \, \mathrm{d}x$$
$$+ \int_{\Delta t}^{+\infty} \int_{\mathbb{R}} \left(\lambda e^+(f_{\Delta t,\Delta x}^+(t + \Delta t, x + \Delta x)) - \lambda e^-(f_{\Delta t,\Delta x}^-(t + \Delta t, x)) \right)$$
$$\times \left(\frac{\psi_{\Delta t,\Delta x}(t, x) - \psi_{\Delta t,\Delta x}(t, x + \Delta x)}{\Delta x} \right) \mathrm{d}t \, \mathrm{d}x \leq 0.$$

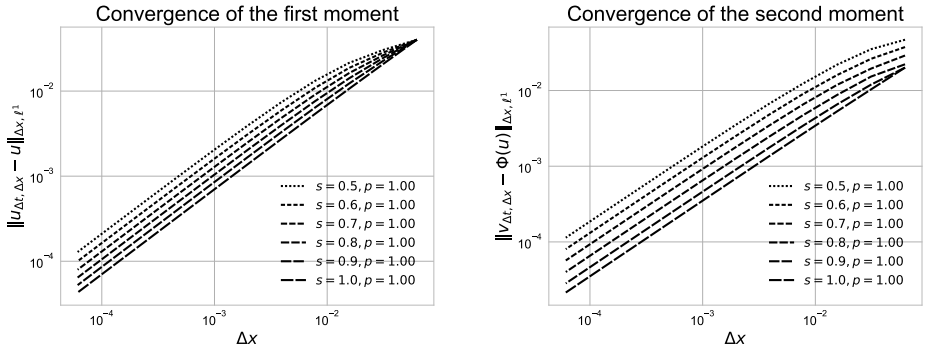
Following the definition of (e^+, e^-) and since $(f_{\Delta t,\Delta x}^+, f_{\Delta t,\Delta x}^-) \xrightarrow[\Delta t,\Delta x\to 0]{} (h^+(u), h^-(u))$, by passing to the limit as $\Delta t,\ \Delta x \to 0$, we obtain

$$\int_0^{+\infty} \int_{\mathbb{R}} \left(\eta(u) \frac{\partial \psi}{\partial t} + q(u) \frac{\partial \psi}{\partial x} \right) \mathrm{d}x \, \mathrm{d}t \geq 0,$$

and the result of the theorem follows. ∎

7. Numerical illustrations

In this section, numerical simulations are given in order to illustrate the theoretical results of the previous sections. Two models are investigated: the advection equation with a constant velocity and the Burger's equation, both considering regular then discontinuous initial conditions. In particular, the numerical rate of convergence and the entropy production are computed in these cases.



(a) First moment u *with* (I_A) *(b) Second moment* v *with* (I_A)

Fig. 1. Error in norm ℓ^1 of the numerical solution of the advection equation (32) according to the space steps Δx for several relaxation parameter values $s \in \{0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$. p is the convergence rate.

Two initial conditions are systematically chosen: first a regular function

$$(I_A) \qquad u^0(x) = \begin{cases} 0 & \text{if} \quad x \le x_L - \delta, \\ \frac{1}{2} + \frac{(x-x_L)(3\delta^2-(x-x_L)^2)}{4\delta^3} & \text{if} \quad x_L - \delta \le x \le x_L + \delta, \\ 1 & \text{if} \quad x_L + \delta \le x \le x_R + \delta, \\ \frac{1}{2} - \frac{(x-x_R)(3\delta^2-(x-x_R)^2)}{4\delta^3} & \text{if} \quad x_R - \delta \le x \le x_R + \delta, \\ 0 & \text{if} \quad x_R + \delta \le x, \end{cases}$$

and second a discontinuous function

$$(I_B) \qquad u^0(x) = \begin{cases} 0 & \text{if} \quad x \le x_L, \\ 1 & \text{if} \quad x_L \le x \le x_R, \\ 0 & \text{if} \quad x_R \le x. \end{cases}$$

Numerically, we took $x_L = 1/4$, $x_R = 3/4$, $\delta = 1/10$. Note that the second case (I_B) can be obtained from the first case (I_A) in the limit δ goes to 0.

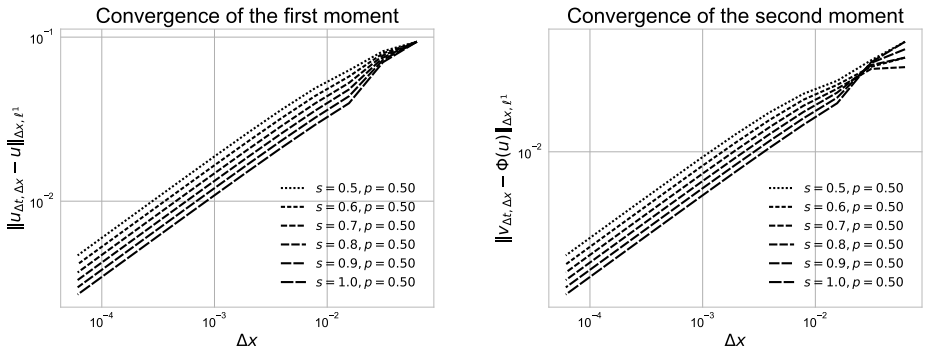
7.1. **Constant velocity advection equation.** The first studied model corresponds to the advection equation at velocity a

$$(32) \qquad \begin{cases} \partial_t u(t,x) + a\partial_x u(t,x) = 0, & x \in \mathbb{R}, t > 0, \\ u(0,x) = u^0(x), & x \in \mathbb{R}, \end{cases}$$

where the velocity a is taken equal to 0.75. The exact solution is well known for both initial conditions (I_A) and (I_B) and reads $u(t,x) = u^0(x - at)$.

The convergence rates of the numerical solution can be read in Figure 1: the error on the first moment u and on the second moment v converges toward 0 with a convergence rate equal to 1 for regular solutions (Figure 1a and 1b) and equal to 0.5 for discontinuous solutions (Figure 1c and 1d). As expected, the error is even larger when the relaxation parameter s is small, this numerical results staying true for s lying in $[1, 2]$ as observed in [13] for the first moment in other test cases.

We now present the results on the numerical entropy production. As the system (32) is linear, the exact entropy production is zero. In Figure 2, the local in space numerical entropy



(c) First moment u *with* (I_B) *(d) Second moment* v *with* (I_B)

Fig. 1. Error in norm ℓ^1 of the numerical solution of the advection equation (32) according to the space steps Δx for several relaxation parameter values $s \in \{0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$. p is the convergence rate.

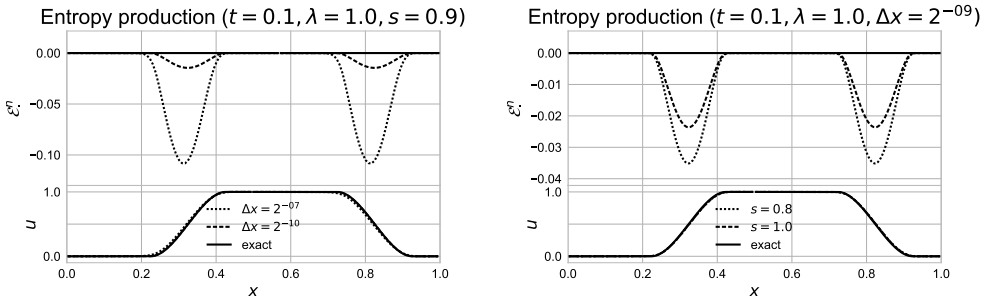


Fig. 2. Local in space numerical entropy production of the advection equation with initial condition $(I_A)$for varying space steps Δx at left and for varying relation parameters s at right. The numerical and exact solutions at $t = 0.1$ are shown bellow and the corresponding entropy productions above.

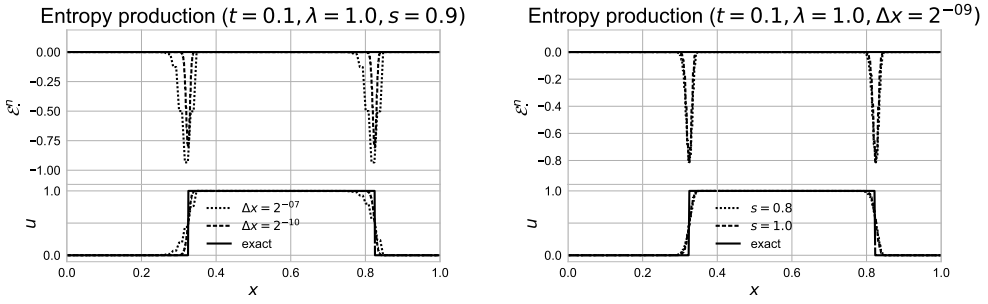
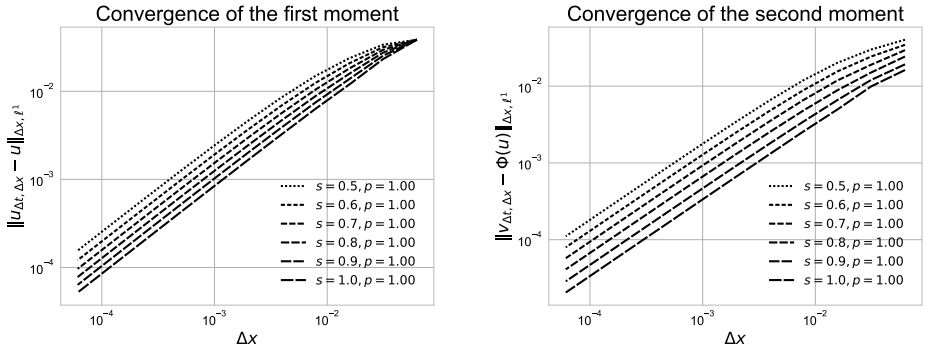


Fig. 3. Local in space numerical entropy production of the advection equation with initial condition $(I_B)$for varying space steps Δx at left and for varying relation parameters s at right. The numerical and exact solutions at $t = 0.1$ are shown bellow and the corresponding entropy productions above.



(a) First moment u with (I_A) *(b) Second moment v with (I_A)*

Fig. 4. Error in norm ℓ^1 of the numerical solution of the Burgers equation (33) according to the space steps Δx for several relaxation parameter values $s \in \{0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$. p is the convergence rate.

production $\mathcal{E}_j^n$, $j \in \mathbb{Z}$, is shown for the regular initial condition (I_A). This production is negative as proved in Proposition 10, located where the solution is not constant, and is smaller for small space step Δx and for large relaxation parameter s. In Figure 3, the entropy production is shown for the discontinuous initial condition (I_B). This production is also negative and located on the discontinuities. The ℓ^1-norm is decreasing when the space step Δx is lower or when the relaxation parameter s is larger.

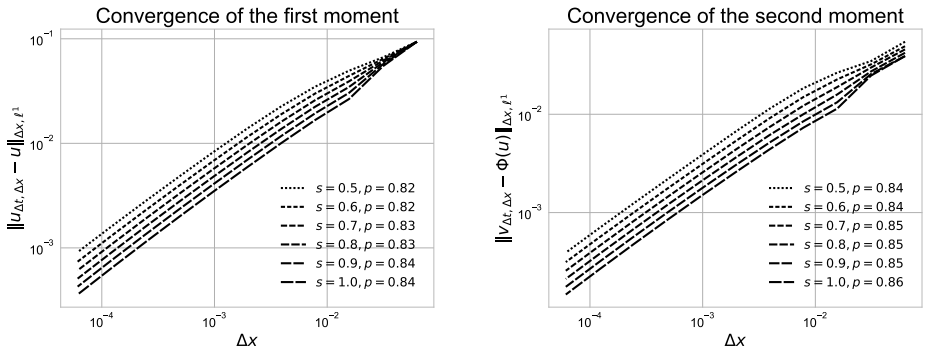
7.2. **Burger's equation.** The second studied model corresponds to the Burger's equation

$$\begin{cases} \partial_t u(t,x) + \partial_x(\frac{1}{2}u^2(t,x)) = 0, & x \in \mathbb{R}, t > 0, \\ u(0,x) = u^0(x), & x \in \mathbb{R}. \end{cases} \tag{33}$$

The exact solution for both initial conditions can be computed: for the regular function (I_A) with the theory of the characteristics until the shock appears and for the discontinuous function (I_B) with the combination of a shock and a rarefaction wave.

The convergence rates of the numerical solution can be read in Figure 4: the error on the first moment u and on the second moment v converges toward 0. For regular solutions, the convergence rate is equal to 1 (Figure 4a and 4b) and, for discontinuous solutions, it is about 0.8 (Figure 4c and 4d). As expected, the error is even larger when the relaxation parameter s is small, this numerical results staying true for s lying in $[1, 2]$ as observed in [13] for the first moment in other test cases. Note that we cannot explain the better convergence rate of this nonlinear equation for discontinuous solutions as previously shown in [13].

We now present the results on the numerical entropy production. For the system (33), the exact entropy production is zero but on the decreasing discontinuities where it is a Dirac measure multiplied by a negative value. In Figure 5, the local in space numerical entropy production $\mathcal{E}_j^n$, $j \in \mathbb{Z}$, is shown for the regular initial condition (I_A). This production is negative as proved in Proposition 10, located where the solution is not constant, and is smaller for small space step Δx and for large relaxation parameter s. In Figure 6, the entropy production is shown for the discontinuous initial condition (I_B). This production is also negative and essentially located on the discontinuity (a smaller contribution is located on the rarefaction wave, this contribution being decreasing when the space step Δx is smaller or when the relaxation parameter s is larger).



(c) First moment u with (I_B) *(d) Second moment v with* (I_B)

Fig. 4. Error in norm ℓ^1 of the numerical solution of the Burgers equation (33) according to the space steps Δx for several relaxation parameter values $s \in \{0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$. p is the convergence rate.

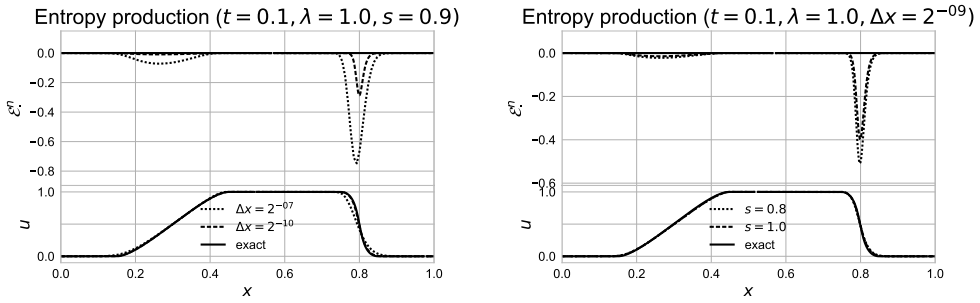


Fig. 5. Local in space numerical entropy production of the Burgers equation with initial condition $(I_A)$for varying space steps Δx at left and for varying relation parameters s at right. The numerical and exact solutions at $t = 0.1$ are shown bellow and the corresponding entropy productions above.

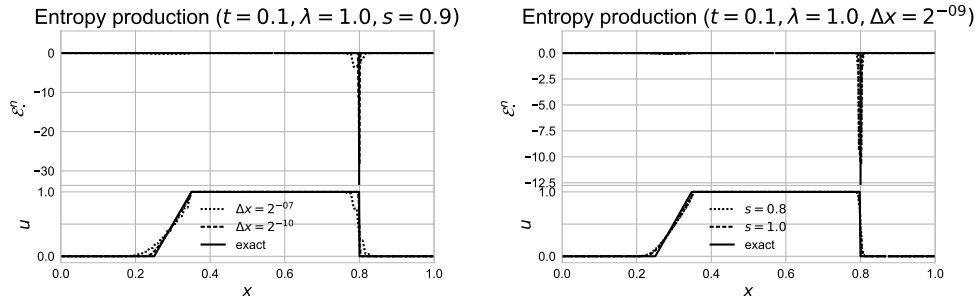


Fig. 6. Local in space numerical entropy production of the Burgers equation with initial condition $(I_B)$for varying space steps Δx at left and for varying relation parameters s at right. The numerical and exact solutions at $t = 0.1$ are shown bellow and the corresponding entropy productions above.

7.3. **Concluding remarks.** These numerical illustrations confirm the theoretical results of the convergence of the numerical solution for the ℓ^1-norm for both moments u and v. Moreover, the discrete entropy production is non-positive and seems to converge towards the exact entropy production.

8. Conclusion

In this paper, we prove, by using techniques based on finite volume methods—in particular on relaxation methods—, a convergence result for a nonlinear hyperbolic one dimensional scalar conservation law. More precisely, we prove the convergence of the numerical solutions of the $\mathsf{D}_1\mathsf{Q}_2$ scheme towards the unique entropy solution of the scalar conservation law. We prove in addition a numerical entropy estimation. Our results are based on convexity properties of the scheme.

Future works could be dedicated to extend the present results to a wide range of relaxation parameters. Moreover, other popular lattice Boltzmann schemes, like a scheme with three velocities in one space dimension, should also be studied. Of course, general results concerning the convergence of more complex lattice Boltzmann schemes remains an open question.

Acknowledgement

The authors are grateful to Frédéric Lagoutière and to Jean-François Babadjian for their explanations concerning bounded variation functions. The simulations have been obtained with the `python` package `pylbm` (`https://pylbm.readthedocs.io/en/latest/`).

References

1. L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, vol. 254, Clarendon Press Oxford, 2000.
2. D. Aregba-Driollet and R. Natalini, *Convergence of relaxation schemes for conservation laws*, Applicable Analysis **61** (1996), no. 1-2, 163–193.
3. B. M. Boghosian, P. Love, and J. Yepez, *Entropic lattice Boltzmann model for Burgers's equation*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **362** (2004), no. 1821, 1691–1701.
4. F. Bouchut, *Entropy satisfying flux vector splittings and kinetic BGK models*, Numer. Math. **94** (2003), no. 4, 623–672.
5. J. E. Broadwell, *Shock structure in a simple discrete velocity gas*, The Physics of Fluids **7** (1964), no. 8, 1243–1247.
6. F. Caetano, *Sur certains problèmes de linéarisation et de couplage pour les systèmes hyperboliques non linéaires*, Ph.D. thesis, Université Paris 6, 2006.
7. S. Dellacherie, *Construction and analysis of lattice Boltzmann methods applied to a 1D convection-diffusion equation*, Acta Applicandae Mathematicae **131** (2014), no. 1, 69–140.
8. D. d'Humières, *Generalized Lattice-Boltzmann equations*, Rarefied Gas Dynamics: Theory and simulation, vol. 159, AIAA Progress in astronomics and aeronautics, 1992, pp. 450–458.
9. F. Dubois, *Stable lattice Boltzmann schemes with a dual entropy approach for monodimensional nonlinear waves*, Comput. Math. Appl. **65** (2013), no. 2, 142–159.
10. R. Gatignol, *Théorie cinétique des gaz à répartition discrète de vitesses*, Springer-Verlag, Berlin-New York, 1975, Lecture Notes in Physics, Vol. 36.
11. E. Giusti, *Minimal surfaces and functions of bounded variation*, Monographs in Mathematics, vol. 80, Birkhäuser Verlag, Basel, 1984.
12. E. Godlewski and P.-A. Raviart, *Hyperbolic systems of conservation laws*, Mathématiques & Applications (Paris) [Mathematics and Applications], Ellipses, Paris, 1991.
13. B. Graille, *Approximation of mono-dimensional hyperbolic systems: A lattice Boltzmann scheme as a relaxation method*, J. Comp. Physics **266** (2014), 74–88.
14. S. Jin and Z. Xin, *The relaxation schemes for systems of conservation laws in arbitrary space dimensions*, Comm. Pure Appl. Math **48** (1995), 235–277.
15. M. Junk, *A Finite Difference Interpretation of the Lattice Boltzmann Method*, Numer. Methods Partial Differ. Equations **17** (2001), 383–402.

16. M. Junk and Z. Yang, *Convergence of lattice Boltzmann methods for Navier–Stokes flows in periodic and bounded domains*, Numerische Mathematik **112** (2009), no. 1, 65–87.
17. T. Krüger, H. Kusumaatmaja, A. Kuzmin, O. Shardt, G. Silva, and E. Viggen, *The lattice Boltzmann method*, vol. 10, Springer, 2017.
18. P. Lallemand and L.-S. Luo, *Theory of the lattice Boltzmann method: dispersion, dissipation, isotropy, Galilean invariance, and stability*, Phys. Rev. E **61** (2000), 6546–6562.
19. C. Lattanzio and D. Serre, *Convergence of a relaxation scheme for hyperbolic systems of conservation laws*, Numerische Mathematik **88** (2001), no. 1, 121–134.
20. R. Natalini, *Convergence to equilibrium for the relaxation approximations of conservation laws*, Communications on pure and applied mathematics **49** (1996), no. 8, 795–823.
21. M. K. Rheinländer, *Stability and multiscale analysis of an advective lattice Boltzmann scheme*, Progress in Computational Fluid Dynamics (PCFD) **8** (2008), no. 1-4, 56–68.
22. D. Serre, *Systèmes de lois de conservation. I*, Diderot Editeur, Paris, 1996.
23. ———, *Relaxations semi-linéaire et cinétique des systèmes de lois de conservation*, Ann. Inst. H. Poincaré Anal. Non Linéaire **17** (2000), no. 2, 169–192.
24. S. Succi, *Lattice Boltzmann 2038*, EPL (Europhysics Letters) **109** (2015), no. 5, 50001.