SPACE LOCALIZATION AND WELL-BALANCED SCHEMES FOR DISCRETE KINETIC MODELS IN DIFFUSIVE REGIMES*

LAURENT GOSSE† AND GIUSEPPE TOSCANI‡

Abstract. We derive and study Well-Balanced schemes for quasimonotone discrete kinetic models. By means of a rigorous localization procedure, we reformulate the collision terms as nonconservative products and solve the resulting Riemann problem whose solution is self-similar. The construction of an Asymptotic Preserving (AP) Godunov scheme is straightforward and various compactness properties are established within different scalings. At last, some computational results are supplied to show that this approach is realizable and efficient on concrete 2×2 models.

Key words. kinetic equations, diffusive relaxation schemes, nonconservative products.

AMS subject classifications. 65M06, 65M12, 35F25.

1. Introduction.

In this paper, we are interested in the numerical analysis of the forthcoming one-dimensional system of semilinear equations,

$$\partial_t f^\pm \pm \partial_x f^\pm = \mp G(f^+, f^-), \qquad x \in \mathbb{R}, \qquad t > 0, \tag{1}$$

in both rarefied and diffusive regimes, the latter being obtained through the transformation $t \to t/\varepsilon^2$, $x \to x/\varepsilon$, see (16). The unknowns $0 \le f^\pm$ are supposed to be at least *bounded variation (BV) functions*, [39], in the space variable.

One motivation comes from the study of classical Boltzmann models,

$$\partial_t f + \vec{\xi} \cdot \nabla f = Q(f, f), \tag{2}$$

where $0 \le f(t, x, \xi)$ stands for a density of particles moving with velocity $\vec{\xi}$ in the ambient space and $Q(f, f)$ is a collision operator satisfying some structural assumptions, see *e.g.* [7, 38]. Such a model relaxes under certain variables scale

$$\varepsilon \partial_t f + \vec{\xi} \cdot \nabla f = \frac{Q(f, f)}{\varepsilon}, \qquad \varepsilon \to 0+,$$

towards the incompressible Navier-Stokes system, [28], whereas for $\varepsilon \simeq 1$, it describes the flow of some cloudy bulk of rarefied molecules.

Coming back to our simplified model (16), we consider only particles moving with velocity ± 1. Therefore, the density f in (2) boils down to a two-component vector satisfying the system (1) and the collision term takes a simple form. In order to ensure some stability properties, namely $L^1(\mathbb{R})$-contraction, [24, 33, 38], we ask for the so-called *quasi-monotonicity* of the right-hand side which reads:

$$G \in C^1(\mathbb{R}^2);\ G(0,0) = 0, \qquad \partial_+ G \stackrel{def}{=} \frac{\partial G}{\partial f^+} \ge 0, \qquad \partial_- G \stackrel{def}{=} \frac{\partial G}{\partial f^-} < 0. \tag{3}$$

This matches essentially the standard hypotheses encountered in [8, 29, 30], with the notable exception of [31] in which compactness results are established by means of

*partially supported by the European Union "Asymptotic methods in kinetic theory" TMR project #ERBFMRXCT970157.

†Istituto per le Applicazioni del Calcolo (sezione di Bari), Via G. Amendola 122-I, 70126 Bari, ITALY. (l.gosse@area.ba.cnr.it).

‡Dipartimento di Matematica, Università degli Studi di Pavia, Via Ferrata 1, 27100 Pavia, ITALY. (toscani@dimat.unipv.it).

a different methodology. Our present objective is then to develop and study robust numerical processes for (1), (16), stable and reliable on the whole range $0 \leq \varepsilon \leq 1$. We aim also at establishing rigorous compactness properties for these schemes in the spirit of the former articles [8, 29].

Part of this program has already been reached by the authors of [17, 21, 22] who introduced the notion of *Asymptotic-Preserving (AP)* schemes relying on former works devoted to semilinear hyperbolic relaxation; in the present context, we refer to [30] and the quotations therein. Although the aforementioned schemes are computationally competitive, they still lack rigorous stability results as the stiffness strengthens in the diffusive limit; see however [19, 23]. On the other hand, the discretizations proposed in [2], Section 3.1, do not fully address the issue of stability as $\varepsilon \to 0$.

Thus we propose to work out this twofold objective making use of the so-called *Well-Balanced (WB)* schemes, [12, 15] (see also [4, 18, 26, 34]), built on a localization procedure already used in [1, 10, 11]. It relies on a rather simple idea: to reformulate the collision terms as a *nonconservative (NC) product*, [25], in order to be able to solve *exactly* self-similar Riemann problems inside a classical Godunov procedure, [9].

The concentration process for the right-hand side is carried out within Section 2 by means of uniform BV estimates and representation of weak-$\star$ limits of measures. Then in Section 3, we deduce in a rather straightforward way a Godunov scheme whose building blocks are nonconservative Riemann problems. Such a scheme turns out to be Well-Balanced in the rarefied regime (see Section 3.1) and Asymptotic-Preserving in the diffusive limit $\varepsilon \to 0$ (see Sections 3.2, 3.3). We give several compactness results under very reasonable CFL conditions of the type $\Delta t \simeq \max(\varepsilon h, h^2)$ (where $h, \Delta t$ stand for the space/time steps) by means of uniform BV bounds. Finally, we display some numerical results to illustrate various estimates in Section 4 on widely-used discrete kinetic models such as Carleman's, [6], or Ruijgrok-Wu's, [37].

2. A nonconservative reformulation for the kinetic model. In the present section, we are about to follow the canvas of [10] in order to reformulate (1) as an *homogeneous* but *nonconservative* weakly coupled 2×2 system.

2.1. Uniform BV estimates. We consider the Cauchy problem for $1 \geq \epsilon > 0$:

$$\partial_t f^{\pm} \pm \partial_x f^{\pm} = \mp G(f^+, f^-)\partial_x a^{\epsilon}, \qquad 0 \leq f^{\pm}(0,x) = f_0^{\pm}(x) \in L^1 \cap BV(\mathbb{R}). \tag{4}$$

We assume that a^ϵ is Lipschitz continuous for $\epsilon > 0$, more precisely:

$$a^\epsilon(x) = \begin{cases} jh, & \text{for} \quad x \in \left]jh, \left(j+\frac{1}{2}-\frac{\epsilon}{2}\right)h\right], \\ \frac{x}{\epsilon} + \left(j+\frac{1}{2}\right)h\left(1-\frac{1}{\epsilon}\right), & \text{for} \quad x \in \left]\left(j+\frac{1}{2}-\frac{\epsilon}{2}\right)h, \left(j+\frac{1}{2}+\frac{\epsilon}{2}\right)h\right], \\ (j+1)h, & \text{for} \quad x \in \left]\left(j+\frac{1}{2}+\frac{\epsilon}{2}\right)h, (j+1)h\right]. \end{cases} \tag{5}$$

This means that $a^{\epsilon=1}(x) = x$, $a^\epsilon \in BV_{loc}(\mathbb{R})$ uniformly in ϵ, $\partial_x a^\epsilon \geq 0$ and moreover, ($\mathbf{1}_A$ stands for the characteristic function of a set A)

$$a^\epsilon \stackrel{\epsilon \to 0}{\longrightarrow} \sum_{j\in\mathbb{Z}} jh\mathbf{1}_{](j-\frac{1}{2})h,(j+\frac{1}{2})h]}, \qquad h > 0.$$

From [24, 33], one deduces that the Cauchy problem for (4) is well-posed for any $\epsilon > 0$ but becomes ambiguous in the limit $\epsilon \to 0$ because a so-called *nonconservative product*, [25], appears on the right-hand side as $\partial_x a^\epsilon$ concentrates into a Dirac comb.

LEMMA 1. *Let $f_0^\pm \in L^1 \cap BV(\mathbb{R})$ have compact support; then the weak solutions to (4) $f^\pm$ belong to $BV_{loc}(\mathbb{R}_*^+ \times \mathbb{R})$ uniformly in ϵ.*

Proof. We split the proof into several steps for the sake of clarity.

(i)–From [33] and the fact that G is quasimonotone and $\partial_x a^\epsilon \geq 0$, one deduces a $L^1(\mathbb{R})$ contraction principle for any value $\epsilon > 0$: if $\tilde{f}_0^\pm \in L^1 \cap BV(\mathbb{R})$,

$$\forall t > 0, \qquad \partial_t \int_{\mathbb{R}} \Big(|f^+(t,x) - \tilde{f}^+(t,x)| + |f^-(t,x) - \tilde{f}^-(t,x)| \Big) dx \leq 0. \tag{6}$$

As $G(0,0) = 0$, the null solution satisfies trivially (4) and this ensures the L^1 stability and the positivity-preserving property. As we assumed that either $-\partial_- G$ (or $\partial_+ G$) is strictly positive, we can use the implicit function theorem to deduce that the equation $G(u,v) = 0$ admits as a unique solution a smooth curve $v = M(u)$, $M' \geq 0$, called the *Maxwellian distribution*. Therefore, using these curves as comparison functions inside (6) gives a maximum principle. More precisely, as in [38], the following domain

$$\Big[0, \|f_0^+\|_{L^\infty(\mathbb{R})}\Big] \times \Big[0, M(\|f_0^+\|_{L^\infty(\mathbb{R})})\Big]$$

is positively invariant for (4). But since (4) isn't translation invariant, (6) doesn't garantee the uniform $BV(\mathbb{R})$ stability.

(ii)–Differentiating each equation in (4) with respect to time, multiplying by $\big(\text{sgn}(\partial_t f^+), \text{sgn}(\partial_t f^-)\big)^T$ and integrating on $x \in \mathbb{R}$, the same way one reaches:

$$\forall t > 0, \qquad \partial_t \int_{\mathbb{R}} \Big(|\partial_t f^+(t,x)| + |\partial_t f^-(t,x)| \Big) dx \leq 0. \tag{7}$$

This implies that $\partial_t f^\pm(t,.)$ are bounded measures on $\mathbb{R}$ and the same holds true for $G(f^+, f^-)\partial_x a^\epsilon$ by the L^∞ stability and (5); but by their very definition, one has also

$$|\partial_x f^\pm| - |G(f^+,f^-)\partial_x a^\epsilon| \leq |\partial_t f^\pm| \leq |\partial_x f^\pm| + |G(f^+,f^-)\partial_x a^\epsilon|. \tag{8}$$

So inside any interval $a < 0 < b$ large enough, one gets out of (7), (8) for any $t > 0$,

$$\int_a^b |\partial_x f^+(t,x)| + |\partial_x f^-(t,x)|.dx \leq \int_{\mathbb{R}} |\partial_x f_0^+| + |\partial_x f_0^-| + 4\|G(f^+,f^-)\|_{L^\infty} \int_a^b |\partial_x a^\epsilon|,$$

and this ensures the $BV_{loc}(\mathbb{R})$ stability for $t > 0$.

(iii)–It remains to check the L^1 modulus of continuity in the time variable. Thanks to the BV_{loc} bound, we deduce from the equations (4) that on the same interval, there holds for $t > s \geq 0$: (TV stands for the total variation in space)

$$\int_a^b |f^\pm(t,x) - f^\pm(s,x)|.dx \leq |t-s| \Big[TV(f_0^+) + TV(f_0^-) + 6\|G(f^+,f^-)\|_{L^\infty} \int_a^b |\partial_x a^\epsilon| \Big].$$

And this is enough to conclude the proof. □

Of course, by the classical Helly's compactness principle, we deduce that the sequence of weak solutions to (4) is relatively compact in the strong topology of $L^1_{loc}(\mathbb{R}_*^+ \times \mathbb{R})$ as $\epsilon \to 0$.

2.2. Limiting values of the right-hand side. In order to shed complete light on the limit system emanating from (4) as $\epsilon \to 0$, we must give a precise meaning to the ambiguous product appearing on its right-hand side. This can be done within the recent theory of *nonconservative products*, [25], which can be applied thanks to the uniform BV-bound established in the preceding section.

In order to reveal the nature of the limit for $G(f^+, f^-)\partial_x a^\epsilon$ in the weak-$\star$ topology of measures, we pick up a test function $\psi \in C_c^0(\mathbb{R}_*^+ \times \mathbb{R})$, that is, continuous and compactly supported and look at the behaviour of the sequence

$$\mathcal{I}^\epsilon = \int_{\mathbb{R}_*^+ \times \mathbb{R}} G(f^+, f^-)\partial_x a^\epsilon.\psi(t,x).dt.dx, \qquad \epsilon \to 0.$$

PROPOSITION 1. *Under the assumptions of Lemma 1, there holds as $\epsilon \to 0$,*

$$G(f^+, f^-)\partial_x a^\epsilon \overset{weak-\star}{\rightharpoonup} \mathcal{M}$$

$$\sum_{j\in\mathbb{Z}} h \left(\int_0^1 G(\bar{f}^+_{j+\frac12}, \bar{f}^-_{j+\frac12})(t,\xi).d\xi \right) \delta\left(x - \left(j + \frac12 \right) h \right),$$

where δ stands for the Dirac mass in $x = 0$ and the "microscopic profiles" $\bar{f}^\pm_{j+\frac12}$ satisfy the ordinary differential system

$$\partial_\xi \bar{f}^\pm_{j+\frac12} = -h.G(\bar{f}^+_{j+\frac12}, \bar{f}^-_{j+\frac12}), \qquad \xi \in [0,1], \tag{9}$$

with the initial data for $t \in \mathbb{R}^+$ and $x = \left(j + \frac12\right) h$, $j \in \mathbb{Z}$:

$$\bar{f}^+_{j+\frac12}(t, \xi = 0) = f^+(t, x - 0), \qquad \bar{f}^-_{j+\frac12}(t, \xi = 1) = f^+(t, x + 0). \tag{10}$$

We stress that the left/right values of $f^\pm(t,.)$ in (10) make sense thanks to the uniform BV-regularity.

Proof. We notice at once that thanks to the definition (5) of a^ϵ, we have

$$\mathcal{I}^\epsilon = \int_{\mathbb{R}_*^+} \sum_{j\in\mathbb{Z}} \int_{(j+\frac12-\frac{\epsilon}{2})h}^{(j+\frac12+\frac{\epsilon}{2})h} \frac{G(f^+, f^-)}{\epsilon} \psi(t,x).dx.dt.$$

Thus it is convenient to perform a rescaling of the space variable:

$$[0,1] \ni \xi = \frac{1}{h\epsilon}\left(x - \left(j + \frac12 - \frac{\epsilon}{2} \right) h \right); \quad x = \left(j + \frac12 - \frac{\epsilon}{2} \right) h + \xi h \epsilon \overset{\epsilon \to 0}{\to} \left(j + \frac12 \right) h.$$

Inside any stripe $\left] \left(j + \frac12 - \frac{\epsilon}{2}\right) h, \left(j + \frac12 + \frac{\epsilon}{2}\right) h \right]$, the unknowns $f^\pm$ satisfy the following semilinear boundary value problem for $\xi \in [0,1]$:

$$\begin{cases} \epsilon h \partial_t f^\pm \pm \partial_\xi f^\pm = \mp h.G(f^+, f^-), & t > 0, \\ f^+(t, \xi = 0) = f^+\left(t, \left(j + \frac12 - \frac{\epsilon}{2}\right) h\right), \\ f^-(t, \xi = 1) = f^-\left(t, \left(j + \frac12 + \frac{\epsilon}{2}\right) h\right), \\ f^\pm(t = 0, \xi) = f_0^\pm(\xi). \end{cases}$$

Its solution can be computed by the method of characteristics for $t \in [\tau_0, \tau_0 + \epsilon h]$:

$$\begin{array}{ll} \dot{\xi}^\pm_{\tau_0} = \frac{\pm 1}{\epsilon h}, & \xi^+_{\tau_0}(\tau_0) = 0, \ \xi^-_{\tau_0}(\tau_0) = 1; \\ \dot{f}^\pm = \frac{\mp 1}{\epsilon} G(f^+, f^-), & f^\pm(t) = f^\pm(t, \xi^\pm_{\tau_0}(t)). \end{array}$$

One sees therefore that along $\xi^{\pm}_{\tau_0}$, there holds:

$$\partial_\xi f^{\pm} = \dot{f}^{\pm}\left(\dot{\xi}^{\pm}_{\tau_0}\right)^{-1} = -h.G(f^+, f^-).$$

Since the system is semilinear, $\xi^{\pm}_{\tau_0}$ realize a diffeomophism from $[\tau_0, \tau_0 + \epsilon h]$ onto $[0,1]$ and has an inverse we note $\tau^{\pm}$; it satisfies

$$\forall \xi \in [0,1], \qquad \tau^+(0) = \tau^-(1) = \tau_0, \qquad \frac{d\tau^{\pm}}{d\xi} = \epsilon h \overset{\epsilon \to 0}{\to} 0.$$

This means in particular that for $\xi \in [0,1]$,

$$f^{\pm}(t) = f^{\pm}(t, \xi^{\pm}_{\tau_0}(t)) = f^{\pm}(\tau^{\pm}(\xi), \xi) \overset{\epsilon \to 0}{\to} f^{\pm}(\tau_0, \xi),$$

and keeps on satisfying the differential equation. It remains to rewrite:

$$\mathcal{I}^\epsilon = \int_{\mathbb{R}^+_*} \sum_{j \in \mathbb{Z}} \int_0^1 h.G(f^+, f^-)(t, \xi)\psi\left(\left(j + \frac{1}{2} - \frac{\epsilon}{2}\right)h + \xi h \epsilon\right).d\xi.dt.$$

We can invoke Lebesgue's dominated convergence theorem in order to pass to the limit $\epsilon \to 0$ in $\mathcal{I}^\epsilon$ and we are done. □

From now on, the meaning of the "distributions product" in

$$\partial_t f^{\pm} \pm \partial_x f^{\pm} = \mp \sum_{j \in \mathbb{Z}} h.G(f^+, f^-)\delta\left(x - \left(j - \frac{1}{2}\right)h\right), \tag{11}$$

is to be **always** understood following Proposition 1. In particular, this last result provides a unique way to solve the Riemann problem for (11) with three simple waves; two of them moving with velocity ± 1 associated with the convection process, and the static one rendering the action of the localized collision term. More precisely, if we supply four constant states at time $t = 0$ separated by a discontinuity in $x = \left(j - \frac{1}{2}\right)h$ $f^{\pm}_{L/R}$, the self-similar solution to (11) is given by:

$$\begin{cases} (f_L^+, f_L^-) & \text{for} \quad x - \left(j - \frac{1}{2}\right)h < -t, \\ (f_L^+, \tilde{f}^-) & \text{for} \quad -t < x - \left(j - \frac{1}{2}\right)h < 0, \\ (\tilde{f}^+, f_R^-) & \text{for} \quad 0 < x - \left(j - \frac{1}{2}\right)h < t, \\ (f_R^+, f_R^-) & \text{for} \quad x - \left(j - \frac{1}{2}\right)h > t, \end{cases} \tag{12}$$

where, according to the notation of Proposition 1,

$$\tilde{f}^+ = \bar{f}^+_{j-\frac{1}{2}}(t, \xi = 1) \text{ and } \tilde{f}^- = \bar{f}^-_{j-\frac{1}{2}}(t, \xi = 0).$$

Such a construction has already been succesfully used inside a numerical processing of the so-called *hyperbolic heat equations* in [13]. We also stress out that it's but a particular case of the "h–Riemann solvers" introduced in [1] within a different context.

2.3. Uniqueness via $L^1(\mathbb{R})$ contraction. Thanks to the dissipative structure of (4), it is straightforward to establish uniqueness for the singular problem (11) as soon as it has been given a rigorous sense within the theory of distributions.

PROPOSITION 2. *Under the assumptions of Lemma 1, let $f_0^{\pm}, \tilde{f}_0^{\pm}$ be two sets of initial data for (11). It holds for all $t > 0$:*

$$\|f^+(t,.) - \tilde{f}^+(t,.)\|_{L^1(\mathbb{R})} + \|f^-(t,.) - \tilde{f}^-(t,.)\|_{L^1(\mathbb{R})} \leq \|f_0^+ - \tilde{f}_0^+\|_{L^1(\mathbb{R})} + \|f_0^- - \tilde{f}_0^-\|_{L^1(\mathbb{R})}.$$

In particular, the weak solution in the sense of Proposition 1 to (11), $f_0^\pm \in L^1 \cap BV(\mathbb{R})$ with compact support is unique and belongs to $L^\infty(\mathbb{R}^+; BV(\mathbb{R})) \cap Lip(\mathbb{R}^+; L^1(\mathbb{R}))$.

Proof. We proceed by approximation and come back to (4) for which the contraction property (6) holds uniformly in ϵ. It remains to integrate it in time and make use of the compactness results to conclude. □

REMARK 1. *We would like to mention at this level that the present construction provides an alternative route to the compactness results of [10] concerning the convex scalar balance law whose right-hand side concentrates like (4), (5)*

$$\partial_t u + \partial_x f(u) = k(x) g(u) \partial_x a^\epsilon, \qquad 0 \le u(t=0,x) = u_0(x) \in L^1 \cap BV(\mathbb{R}).$$

Indeed, if it is assumed that $0 \le k \in C_c^0(\mathbb{R})$, $g' \le 0$ and $g(\bar{u}) = 0$ for a $\bar{u} > 0$, then the $L^1(\mathbb{R})$ contraction principle, [24], holds uniformly in ϵ and the interval

$$\Big[0, \max(\bar{u}, \|u_0\|_{L^\infty(\mathbb{R})})\Big]$$

is positively invariant since its endpoints give rise as initial data to sub/super-solutions of the associated homogeneous conservation law. Therefore, it is easy to follow the lines of Lemma 1 to derive

$$\partial_t \left(\int_{\mathbb{R}} |\partial_t u|(t,x).dx \right) \le 0,$$

which gives in turn as long as k has compact support,

$$\int_{\mathbb{R}} |\partial_x f(u)|(t,x).dx \le \int_{\mathbb{R}} |\partial_x f(u_0)| + 2\|k(x)g(u)\|_{L^\infty} \int_{Supp(k)} |\partial_x a^\epsilon|,$$

together with a time Lipschitz-modulus of continuity. If moreover, a non–resonance assumption $f' \ge c > 0$ holds, it turns out that this is enough to establish a uniform $BV_{loc}(\mathbb{R}_*^+ \times \mathbb{R})$ bound for u as $\epsilon \to 0$. Therefore we recover (part of) the conclusion of Lemma 7 in [10].

We close this section introducing $\mathbf{S}$, the *solution operator* for (11):

PROPOSITION 3. *There exists a unique "nonconservative contraction semigroup" $\mathbf{S}$ whose domain is $L^1 \cap BV(\mathbb{R})$ and such that any trajectory $0 < t \mapsto \mathbf{S}(t) f_0^\pm$ coincides with the unique weak solution to (11), $f^\pm(t=0,.) = f_0^\pm$, in the sense of distributions.*

3. Derivation and convergence of well-balanced schemes. Roughly speaking, we are about to develop and study Godunov schemes, [9], for (1) relying on solving elementary Riemann problems for (11) whose self-similar solution is given by (12). More precisely, given a time-step $\Delta t > 0$ and a mesh-size $h > 0$, we can define a computational cartesian grid. The cells read for all $j, n \in \mathbb{Z} \times \mathbb{N}$,

$$C_j = \left] \left(j - \frac{1}{2}\right) h, \left(j + \frac{1}{2}\right) h \right], \qquad I^n = \Big[n\Delta t, (n+1)\Delta t \Big[.$$

Let $\mathcal{P}^h$ be the standard L^2 projector on piecewise-constant functions:

$$\begin{array}{rccc} \mathcal{P}^h : & L^1 \cap BV(\mathbb{R}) & \to & L^1 \cap BV(\mathbb{R}) \\ & \varphi & \mapsto & \left(\displaystyle\int_{C_j} \frac{\varphi(x)}{h}.dx \right)_{j \in \mathbb{Z}} \end{array}$$

It remains to discretize the initial data as follows: we define a piecewise constant approximation $f_h^{\pm}(t=0,.)$ by taking the pointwise values

$$\forall j \in \mathbb{Z}, \qquad f_{j,0}^{\pm} = f_0^{\pm}(jh),$$

and this makes sense thanks to the BV regularity of the considered functions. Our well-balanced Godunov scheme reads therefore:

$$f_h^{\pm}(t,.) = \mathbf{S}(t - n\Delta t) \circ \Big(\mathcal{P}^h \circ \mathbf{S}(\Delta t)\Big)^n f_h^{\pm}(t=0,.), \tag{13}$$

where n stands for the integer part of $t/\Delta t$. Therefore Riemann problems for (11) are to be solved at the endpoints of each C_j, $j \in \mathbb{Z}$.

3.1. The rarefied regime. We focus first on (1) in its hyperbolic scaling. Using the divergence theorem, one sees that the Godunov scheme (13) generates the following values as $n \in \mathbb{N}$, $j \in \mathbb{Z}$:

$$f_{j,n+1}^{+} = f_{j,n}^{+} - \frac{\Delta t}{h}\Big(f_{j,n}^{+} - f_{j-\frac{1}{2},n}^{+}\Big), \qquad f_{j,n+1}^{-} = f_{j,n}^{-} + \frac{\Delta t}{h}\Big(f_{j+\frac{1}{2},n}^{-} - f_{j,n}^{-}\Big). \tag{14}$$

The values at the borders of each cell C_j are given by the generalized jump relations (9), (10). Thus the upwind scheme (14) rewrites:

$$\begin{aligned} f_{j,n+1}^{+} &= f_{j,n}^{+} - \frac{\Delta t}{h}\Big(f_{j,n}^{+} - f_{j-1,n}^{+}\Big) - \Delta t \int_0^1 G(\bar f_{j-\frac{1}{2}}^{+}, \bar f_{j-\frac{1}{2}}^{-})(n\Delta t, \xi).d\xi, \\ f_{j,n+1}^{-} &= f_{j,n}^{-} + \frac{\Delta t}{h}\Big(f_{j+1,n}^{-} - f_{j,n}^{-}\Big) + \Delta t \int_0^1 G(\bar f_{j+\frac{1}{2}}^{+}, \bar f_{j+\frac{1}{2}}^{-})(n\Delta t, \xi).d\xi. \end{aligned} \tag{15}$$

This highlights the ability of this scheme to preserve **exactly** the steady-state curves of (1) since by their very definition, they satisfy, cf. (9), (10), for all $j \in \mathbb{Z}$:

$$\begin{aligned} f_{j,0}^{+} - f_{j-1,0}^{+} &= -h \int_0^1 G(\bar f_{j-\frac{1}{2}}^{+}, \bar f_{j-\frac{1}{2}}^{-})(0,\xi).d\xi, \\ f_{j+1,0}^{-} - f_{j,0}^{-} &= -h \int_0^1 G(\bar f_{j+\frac{1}{2}}^{+}, \bar f_{j+\frac{1}{2}}^{-})(0,\xi).d\xi. \end{aligned}$$

The following stability result is easily established:

LEMMA 2. *Let $f_0^{\pm} \in L^1 \cap BV(\mathbb{R})$; under the hyperbolic CFL condition $\Delta t \le h$, the approximate solutions $f_h^{\pm}$ obtained from (13), (14) satisfy:*

$$TV(f_h^{+}(t,.)) + TV(f_h^{-}(t,.)) \le \exp\Big(2\frac{t}{h}\big(\exp(Lip(G)h) - 1\big)\Big)\Big[TV(f_0^{+}) + TV(f_0^{-})\Big],$$

where $Lip(G)$ stands for the Lipschitz constant of G.

Proof. We start from (15) and follow the proof of Lemma 10 in [10]: by the classical theory of differential equations and a linearization of G, one gets the following inequalities:

$$
\begin{aligned}
|f^+_{j,n+1} - f^+_{j-1,n+1}| &\leq |f^+_{j,n} - f^+_{j-1,n}| \left(1 - \frac{\Delta t}{h}\right) \\
&+ \frac{\Delta t}{h}\Big(1 + \big(\exp(Lip(G)h) - 1\big)\Big)|f^+_{j-1,n} - f^+_{j-2,n}| \\
&+ \frac{\Delta t}{h}\big(\exp(Lip(G)h) - 1\big)|f^-_{j+1,n} - f^-_{j,n}|, \\
|f^-_{j+1,n+1} - f^-_{j,n+1}| &\leq |f^-_{j+1,n} - f^-_{j,n}| \left(1 - \frac{\Delta t}{h}\right) \\
&+ \frac{\Delta t}{h}\Big(1 + \big(\exp(Lip(G)h) - 1\big)\Big)|f^-_{j+2,n} - f^-_{j+1,n}| \\
&+ \frac{\Delta t}{h}\big(\exp(Lip(G)h) - 1\big)|f^+_{j,n} - f^+_{j-1,n}|.
\end{aligned}
$$

It remains to add these two inequalities, to sum on $j \in \mathbb{Z}$ to derive:

$$
\sum_{j\in\mathbb{Z}} \Big(|f^+_{j,n+1} - f^+_{j-1,n+1}| + |f^-_{j+1,n+1} - f^-_{j,n+1}|\Big) \leq
$$
$$
\Big(1 + \frac{2\Delta t}{h}\big(\exp(Lip(G)h) - 1\big)\Big) \sum_{j\in\mathbb{Z}} \Big(|f^+_{j,n} - f^+_{j-1,n}| + |f^-_{j+1,n} - f^-_{j,n}|\Big).
$$

And this is enough to conclude since $BV(\mathbb{R}) \subset L^\infty(\mathbb{R})$. □

REMARK 2. *This proof is a direct adaptation to (1) of the one given for Lemma 10 in [10]. We mention here that there exists however a much quicker way to establish this former compactness result relying on [27]. Indeed, one can consider any nonhomogeneous scalar balance law endowed with a localized right-hand side,*

$$
\partial_t u + \partial_x f(u) - g(u)\partial_x a = 0, \qquad \partial_t a = 0,
$$

as an elementary but nonconservative 2×2 *Temple system whose wave curves are the level sets of the strong Riemann invariants (in the notation of [10]):*

$$
a \text{ and } w(u,a) = \phi^{-1} \circ (\phi(u) - a), \qquad \phi'(u) = \frac{f'(u)}{g(u)}.
$$

Therefore Lemmas 3.1 and 3.2 in [27] imply that the Godunov scheme decreases the total variation of $w(t,.)$. *This entails control on the* $u(t,.)$ *variable in the case where*

$$
0 < c \leq \partial_u w(u,a) \leq C < +\infty,
$$

and this is a consequence of the non-resonance assumption $f'(u) \neq 0$. *Thus strong* L^1_{loc} *compactness for the scalar "well-balanced" scheme follows. It does not seem that the same shortcut applies here in order to shrink the proof of Lemma 2.*

By standard arguments, we can establish strong L^1_{loc} compactness for $f^\pm_h$ as $h \to 0$ relying on the bound stated in Lemma 2.

3.2. The diffusive regime: BV stability. We move now to the study of numerical approximations to (1) in its *diffusive scaling*, that is to say:

$$
\partial_t f^\pm \pm \frac{1}{\varepsilon}\partial_x f^\pm = \mp\frac{1}{\varepsilon^2}G(f^+, f^-), \qquad 0 < \varepsilon < 1, \qquad x \in \mathbb{R}, \qquad t > 0. \tag{16}
$$

In this perspective, the so-called *hyperbolic heat equations* treated in [13] correspond to the special case $G(f^+, f^-) = f^+ - f^-$. We assume also that the Maxwellian distribution is given by $M(f) = f$, that is to say:

$$f^+ = f^- \qquad \Rightarrow \qquad G(f^+, f^-) = 0.$$

In this setting and for any $\varepsilon > 0$, the previous technique relying on a localization of the source term onto a Dirac comb still applies, but the differential system (9), (10) in Proposition 1 has to be rescaled

$$(17) \qquad \forall j \in \mathbb{Z}, \qquad \partial_\xi \bar{f}^{\pm}_{j+\frac{1}{2}} = \frac{-h}{\varepsilon} G(\bar{f}^+_{j+\frac{1}{2}}, \bar{f}^-_{j+\frac{1}{2}}),$$

and the stability result given in Lemma 2 becomes obsolete because of both the unrealistic restriction $\Delta t \leq \varepsilon h$ and the fact that $Lip(G)/\varepsilon$ can be made arbitrarily big. It is therefore of interest to consider the *macroscopic variables* which read

$$(18) \qquad \rho = f^+ + f^-, \qquad J = \frac{f^+ - f^-}{\varepsilon},$$

and within which the system (17) rewrites:

$$(19) \quad \partial_\xi J = 0, \qquad \partial_\xi \rho = -\frac{2h}{\varepsilon} G\left(\frac{1}{2}(\rho + \varepsilon J), \frac{1}{2}(\rho - \varepsilon J)\right) \stackrel{def}{=} -2hA(\rho, J, \varepsilon J).$$

Indeed, the precise form of A can be revealed relying on the mean-value theorem:

$$A(\rho, J, \varepsilon J) = \frac{1}{\varepsilon} \underbrace{G\left(\frac{\rho}{2}, \frac{\rho}{2}\right)}_{=0} + \frac{J}{2}(\partial_+ G - \partial_- G)\left(\frac{\rho + \theta\varepsilon J}{2}, \frac{\rho - \theta\varepsilon J}{2}\right),$$

for some $\theta \in [0, 1]$. As ρ and J realize also the first two moments of the discrete kinetic model (1), it can be expected that they satisfy the semilinear hyperbolic system

$$(20) \qquad \partial_t \rho + \partial_x J = 0, \qquad \varepsilon^2 \partial_t J + \partial_x \rho = -2A(\rho, J, \varepsilon J),$$

which has been shown recently to exhibit diffusive asymptotics as $\varepsilon \to 0$. More precisely, by the implicit function theorem, one goes formally from (20) to

$$J = -B(\rho, \partial_x \rho), \qquad \partial_t \rho = \partial_x \Big(B(\rho, \partial_x \rho)\Big),$$

and this limiting behaviour holds rigorously for instance if:

- $A(\rho, J, \varepsilon J) = A(\rho, J) = \rho^\alpha J$ and $B(\rho, \partial_x \rho) = \frac{1}{2}\rho^{-\alpha}\partial_x \rho$, $\alpha < -1$, [29]
- $A(\rho, J, \varepsilon J) = J - \frac{1}{2}\big(\rho^2 + (\varepsilon J)^2\big)$ and $B(\rho, \partial_x \rho) = -\frac{1}{2}(\rho^2 - \partial_x \rho)$, [8].

Some other results are available in different contexts, see for instance [5, 31].

The main point of the so-called *Asymptotic-Preserving* (AP) schemes, [17, 21], is to capture these features numerically as $\varepsilon \to 0$ with a fixed (and reasonable!) $h > 0$.

Following Proposition 1, we integrate (19) on $\xi \in [0, 1]$ and, inverting (18), we find within the notation of (12) and with $\bar{\rho} = \bar{f}^+ + \bar{f}^-$:

$$(21) \quad \Phi(J; f_L^+, f_R^-) \stackrel{def}{=} (2f_R^- + \varepsilon J) - (2f_L^+ - \varepsilon J) + 2h \int_0^1 A(\bar{\rho}, J, \varepsilon J).d\xi = 0.$$

We plan to apply the implicit function theorem to Φ since:

$$\partial_J \Phi = 2\varepsilon + 2h \int_0^1 \partial_J A(\bar{\rho}, J, \varepsilon J).d\xi, \qquad \partial_J A = \frac{1}{2}(\partial_+ G - \partial_- G) > 0.$$

Therefore, the solution of the equation (21) in J is given by a smooth *flux function*:

$$\begin{array}{cccc} F: & (\mathbb{R}_*^+)^2 & \to & \mathbb{R} \\ & (f_L^+, f_R^-) & \mapsto & J = F(f_L^+, f_R^-). \end{array} \tag{22}$$

Moreover, we know that:

$$\nabla F = \left(\frac{-1}{\partial_J \Phi}\right) \nabla \Phi.$$

Therefore we propose the following definition:

DEFINITION 1. *We say that the flux function F (22) is* **monotone** *if $F(0,0) = 0$ and it is increasing (resp. decreasing) with respect to its first (resp. second) variable:*

$$\partial_+ F \geq 0, \qquad \partial_- F \leq 0.$$

And we aim at treating (16) by means of the modified (partly implicit) numerical scheme one gets out of (13), (14), (15):

$$\begin{aligned} f_{j,n+1}^+ &= f_{j,n}^+ - \frac{\Delta t}{\varepsilon h}\left(f_{j,n+1}^+ - f_{j,n+1}^-\right) + \frac{\Delta t}{h} F(f_{j-1,n}^+, f_{j,n}^-), \\ f_{j,n+1}^- &= f_{j,n}^- + \frac{\Delta t}{\varepsilon h}\left(f_{j,n+1}^+ - f_{j,n+1}^-\right) - \frac{\Delta t}{h} F(f_{j,n}^+, f_{j+1,n}^-). \end{aligned} \tag{23}$$

In sharp contrast with (15), this emphasizes the consistency of such a discretization with a diffusive asymptotic behaviour for $\rho_{j,n} = f_{j,n}^+ + f_{j,n}^-$. Of course, we keep on using the notation $f_h^\pm$ for the piecewise constant numerical approximations to (16) generated by (23).

LEMMA 3. *Let $0 \leq f_0^\pm \in L^1 \cap BV(\mathbb{R})$; if the flux function F is monotone and under the parabolic CFL condition $(\Delta t + \varepsilon h) Lip(F) \leq h$, one has for all $t > 0$, $\varepsilon > 0$:*

- $\|f_h^+(.,t)\|_{L^1(\mathbb{R})} + \|f_h^-(.,t)\|_{L^1(\mathbb{R})} \leq \|f_0^+\|_{L^1(\mathbb{R})} + \|f_0^-\|_{L^1(\mathbb{R})}$,
- $TV(f_h^+(.,t)) + TV(f_h^-(.,t)) \leq TV(f_0^+) + TV(f_0^-)$.

and the scheme (23) is positivity preserving.

The aforementioned CFL condition is said to be *parabolic* because $Lip(F)$ is $O(h^{-1})$ and doesn't blow up as $\varepsilon \to 0$. It means also in most cases that $\varepsilon \leq O(h)$; in this sense, it completes the picture with (14) which is stable in the complementary range of parameters.

REMARK 3. *We stress that there exist many cases of interest for which the monotonicity of F can be established rigorously. For instance if $A(\rho, J) = k(\rho)J$, $k > 0$ [29], one may use a different functional:*

$$\tilde{\Phi}(J; f_L^+, f_R^-) \stackrel{def}{=} \phi(2f_R^- + \varepsilon J) - \phi(2f_L^+ - \varepsilon J) + 2hJ = 0, \qquad \phi'(\rho) = \frac{1}{k(\rho)}. \tag{24}$$

Proof. For ease of writing, we denote $a = 1 + \frac{\Delta t}{\varepsilon h}$, $b = \frac{\Delta t}{\varepsilon h}$. The system (23) can be explicitly solved and this is a desirable feature according to [17]:

$$f^+_{j,n+1} = \frac{a}{a+b}\left(f^+_{j,n} + \frac{\Delta t}{h}F(f^+_{j-1,n}, f^-_{j,n})\right) + \frac{b}{a+b}\left(f^-_{j,n} - \frac{\Delta t}{h}F(f^+_{j,n}, f^-_{j+1,n})\right),$$
$$f^-_{j,n+1} = \frac{b}{a+b}\left(f^+_{j,n} + \frac{\Delta t}{h}F(f^+_{j-1,n}, f^-_{j,n})\right) + \frac{a}{a+b}\left(f^-_{j,n} - \frac{\Delta t}{h}F(f^+_{j,n}, f^-_{j+1,n})\right).$$

In order to control the L^1 norm, we linearize F around $(0,0)$:

$$F(f^+_{j-1,n}, f^-_{j,n}) = \partial_+ F(\xi_{j-\frac{1}{2},n})f^+_{j-1,n} + \partial_- F(\xi_{j-\frac{1}{2},n})f^-_{j,n}.$$

The monotonicity property of the flux function together with the CFL restriction give:

$$\begin{aligned}
|f^+_{j,n+1}| \leq\ & \frac{1}{a+b}\Big[|f^+_{j,n}|\Big(a - \frac{b\Delta t}{h}\partial_+F(\xi_{j+\frac{1}{2},n})\Big) + |f^+_{j-1,n}|\frac{a\Delta t}{h}\partial_+F(\xi_{j-\frac{1}{2},n}) \\
&+ |f^-_{j,n}|\Big(b + \frac{a\Delta t}{h}\partial_-F(\xi_{j-\frac{1}{2},n})\Big) - |f^-_{j+1,n}|\frac{b\Delta t}{h}\partial_-F(\xi_{j+\frac{1}{2},n})\Big] \\
|f^-_{j,n+1}| \leq\ & \frac{1}{a+b}\Big[|f^+_{j,n}|\Big(b - \frac{a\Delta t}{h}\partial_+F(\xi_{j+\frac{1}{2},n})\Big) + |f^+_{j-1,n}|\frac{b\Delta t}{h}\partial_+F(\xi_{j-\frac{1}{2},n}) \\
&+ |f^-_{j,n}|\Big(a + \frac{b\Delta t}{h}\partial_-F(\xi_{j-\frac{1}{2},n})\Big) - |f^-_{j+1,n}|\frac{a\Delta t}{h}\partial_-F(\xi_{j+\frac{1}{2},n})\Big].
\end{aligned}$$

Such a convex combination ensures the positivity-preserving property for (23). Adding the two inequalities leads to:

$$\sum_{j\in\mathbb{Z}} h\Big(|f^+_{j,n+1}| + |f^-_{j,n+1}|\Big) \leq \sum_{j\in\mathbb{Z}} h\Big(|f^+_{j,n}| + |f^-_{j,n}|\Big).$$

The decay in time of the total variation in space is shown by similar arguments. □

This stability result is already enough to ensure strong compactness for the numerical approximations $f^\pm_h$ generated by (23) as $h \to 0$ as long as the relaxation parameter ε remains strictly positive.

3.3. The diffusive regime: limiting behaviour. We are now interested in the behaviour of (23) as $\varepsilon \to 0$. The forthcoming result completes Lemma 3:

LEMMA 4. *Under the hypotheses of Lemma 3, we suppose* $\varepsilon\|\tilde{F}\|_{L^\infty} < 1$ *in the following decomposition which holds for* F*, uniformly in* $h \geq 0$*:*

$$\begin{aligned}
F(f^+_h, f^-_h) &= (f^+_h - f^-_h)\tilde{F}(f^+_h, f^-_h) + g(f^+_h, f^-_h), \\
\tilde{F} \in C^1(\mathbb{R}^2),\ \tilde{F}(f^+_h&, f^-_h) \in L^\infty(\mathbb{R}),\ g(f^+_h, f^-_h) \in L^1(\mathbb{R}).
\end{aligned} \tag{25}$$

Then one has the estimates uniformly in $\varepsilon \geq 0$ *(where* C *is an absolute constant):*

- $\|f^+(t,.) - f^-(t,.)\|_{L^1(\mathbb{R})} \leq \|f^+_0 - f^-_0\|_{L^1(\mathbb{R})}$
  $+C\varepsilon\Big(\|g\|_{L^1(\mathbb{R})} + hLip(F)\big(TV(f^+_0) + TV(f^-_0)\big)\Big)$,
- $\|f^+(t,.) - f^+(s,.)\|_{L^1(\mathbb{R})} + \|f^-(t,.) - f^-(s,.)\|_{L^1(\mathbb{R})} \leq \sqrt{|t-s|}\times$
  $\left[\frac{2}{\varepsilon}\|f^+_0 - f^-_0\|_{L^1(\mathbb{R})} + hLip(F)(1+C)\Big(TV(f^+_0) + TV(f^-_0)\Big) + C\|g\|_{L^1(\mathbb{R})}\right].$

The condition $\varepsilon\|\tilde{F}\|_{L^\infty} < 1$ roughly means that $\varepsilon < O(h)$; this kind of restriction has already been encountered in [14] in order to show compactness in the context of another relaxation problem. The technical assumption (25) will be checked in the numerical examples later on: it expresses the fact that the Maxwellian distribution for (16), (1) should be given by $f^+ = f^-$ (in the case of the Goldstein-Taylor model, the decomposition is trivial, $\tilde{F} \equiv \frac{1}{h+\varepsilon}$, $g \equiv 0$, [13]).

Proof. We keep on using the same notations; from (23), we get

$$(1+2b)(f^+_{j,n+1} - f^-_{j,n+1}) = (f^+_{j,n} - f^-_{j,n}) + \frac{2\Delta t}{h} F(f^+_{j,n}, f^-_{j,n})$$
$$+ \frac{\Delta t}{h}\Big[F(f^+_{j-1,n}, f^-_{j,n}) + F(f^+_{j,n}, f^-_{j+1,n}) - 2F(f^+_{j,n}, f^-_{j,n})\Big].$$

We use now the decomposition (25) in order to get:

$$(1+2b)|f^+_{j,n+1} - f^-_{j,n+1}| \leq |f^+_{j,n} - f^-_{j,n}|\Big(1 + \frac{2\Delta t}{h}\|\tilde{F}\|_{L^\infty}\Big)$$
$$+ \frac{2\Delta t}{h}|g(f^+_{j,n}, f^-_{j,n})| + \frac{\Delta t}{h} Lip(F)\Big[|f^+_{j-1,n} - f^+_{j,n}| + |f^-_{j+1,n} - f^-_{j,n}|\Big].$$

Thus an elementary computation shows that:

$$\alpha \stackrel{def}{=} \frac{1 + \frac{2\Delta t}{h}\|\tilde{F}\|_{L^\infty}}{1 + \frac{2\Delta t}{\varepsilon h}} < 1 \qquad \Leftrightarrow \qquad \varepsilon\|\tilde{F}\|_{L^\infty} < 1.$$

This implies that

$$\begin{aligned} \|f^+(t,.) - f^-(t,.)\|_{L^1(\mathbb{R})} \leq \; & \|f^+_0 - f^-_0\|_{L^1(\mathbb{R})} + \\ & \tfrac{\varepsilon}{1-\alpha}\Big(\|g\|_{L^1(\mathbb{R})} + h.Lip(F)\big[TV(f^+_0) + TV(f^-_0)\big]\Big), \end{aligned}$$

and the control on the Maxwellian distribution follows.

Concerning the L^1-modulus of continuity in time, following [13], we rewrite the first equation in (23) as

$$f^+_{j,n+1} - f^+_{j,n} + \frac{\Delta t}{\varepsilon h}(f^+_{j,n+1} - f^+_{j,n}) = \frac{\Delta t}{\varepsilon h}(f^-_{j,n+1} - f^-_{j,n}) - \frac{\Delta t}{\varepsilon h}(f^+_{j,n} - f^-_{j,n}) + \frac{\Delta t}{h} F(f^+_{j-1,n}, f^-_{j,n}),$$

in order to derive for all $j \in \mathbb{Z}$, $n \in \mathbb{N}$,

$$|f^+_{j,n+1} - f^+_{j,n}|\Big(1 + \frac{\Delta t}{\varepsilon h}\Big) - \frac{\Delta t}{\varepsilon h}|f^-_{j,n+1} - f^-_{j,n}| \leq \frac{\Delta t}{\varepsilon h}|f^+_{j,n} - f^-_{j,n}| + \frac{\Delta t}{h}|F(f^+_{j-1,n}, f^-_{j,n})|,$$

together with a similar expression for the f^- variable. Therefore, we arrive at

$$|f^+_{j,n+1} - f^+_{j,n}| + |f^-_{j,n+1} - f^-_{j,n}| \leq \frac{2\Delta t}{\varepsilon h}|f^+_{j,n} - f^-_{j,n}| +$$
$$\frac{\Delta t}{h} Lip(F)\Big[|f^+_{j-1,n} - f^+_{j,n}| + |f^-_{j+1,n} - f^-_{j,n}|\Big],$$

and we are done with the second inequality of Lemma 4 just summing on $j \in \mathbb{Z}$ and noticing that under a parabolic CFL condition, $\frac{\Delta t}{h} = O(h) = O(\sqrt{\Delta t})$. □

As a consequence of Lemma 4, strong compactness as $\varepsilon \to 0$, $h > 0$ fixed follows as soon as one provides a so-called "well-prepared" initial datum, that is to say:

$$\|f^+_0 - f^-_0\|_{L^1(\mathbb{R})} = O(\varepsilon). \tag{26}$$

Such an initialization for (16) cancels any kind of *initial layer* which would destroy the Hölder time regularity of the process (see also [35] for a similar remark) in $t = 0+$.

With obvious notation, we can deduce easily the asymptotic behaviour of the scheme (23) as $\varepsilon \to 0$; adding the two equations, we derive

$$\rho_{j,n+1} = \rho_{j,n} + \frac{\Delta t}{h}\Big(F\big(\rho_{j-1,n}/2, \rho_{j,n}/2\big) - F\big(\rho_{j,n}/2, \rho_{j+1,n}/2\big)\Big) + \mathcal{R}_{j,n}, \tag{27}$$

and the estimates from Lemma 4 ensure that the remaining term is of the order of ε in L^1 as a consequence of the smoothness of F. This has already been evidenced by explicit computations with the Goldstein-Taylor model in [13].

REMARK 4. *At this point, we underline that relying on the monotonicity of F, (25) can be easily fulfilled using the mean-value theorem:*

$$F(f^+, f^-) \quad = \quad (f^+ - f^-)\partial_+ F(\xi) + div(F)(\xi) f^-, \qquad \xi \in (\mathbb{R}^+)^2 \tag{28}$$

Since $f^\pm(t,.) \in L^1(\mathbb{R})$, what we need is just $div(F) = O(1)$ uniformly in $h \geq 0$. It turns out that this corresponds to a splitting of the flux function F between a diffusive and a convective part when looking at the scheme (27) since it rewrites:

$$\rho_{j,n+1} = \rho_{j,n} + \frac{\Delta t}{2h}\partial_+ F(\xi_{j,n})\Big(\rho_{j+1,n} - 2\rho_{j,n} + \rho_{j-1,n}\Big) - \frac{\Delta t}{2h} div(F)(\xi_{j,n})\Big(\rho_{j+1,n} - \rho_{j,n}\Big).$$

One readily checks that this scheme is L^∞-stable; its L^1 and BV stability are just particular cases of Lemma 3.

As a byproduct of Lemmas 3 and 4, we can state the following theorem which deals with the cases investigated by Lions and Toscani in [29]:

THEOREM 1. *Assume $0 \leq f_0^\pm \in L^1 \cap BV(\mathbb{R})$ are initial data for (16) with $G(f^+, f^-) = (f^+ + f^-)^\alpha (f^+ - f^-)$, $\alpha \leq 0$ and satisfy (26). Then as $h, \varepsilon \to 0$ under the prescribed CFL conditions, the sequence $f_h^\pm$ generated by the scheme (23) is relatively compact in the strong topology of $L^1_{loc}(\mathbb{R}_*^+ \times \mathbb{R})$. In particular, $\rho_h = f_h^+ + f_h^-$ converges towards the unique solution in the sense of distributions to:*

$$\partial_t \rho = \frac{1}{2}\partial_x\Big(\rho^{-\alpha}\partial_x \rho\Big), \qquad \rho(t=0,.) = 2f_0^+ = 2f_0^-.$$

Proof. We are precisely in position to use the modified functional $\tilde{\Phi}$ in (24) since in the notation of Remark 3, $k(\rho) = \rho^\alpha \geq 0$. Thus the flux function F exists and is monotone in the sense of Definition 1. Moreover, the Maxwellian distribution is $f^+ = f^-$ and we can take $g \equiv 0$ in (25) since by the mean-value theorem, we get some expression for the flux function from (24):

$$F(f_L^+, f_R^-) = \frac{f_L^+ - f_R^-}{h.k(\zeta) + \varepsilon}, \qquad \zeta \in \mathbb{R}^+. \tag{29}$$

The conclusion thus follows from Lemma 4 and (27). □

We left behind the cases $0 < \alpha \leq 1$ (fast diffusion equations) as $\tilde{\Phi}$ can become singular if $\rho = 0$; this isn't an issue in practical computations, see Section 4.3.

4. Numerical experiments. We illustrate in this section the results previously stated by means of some numerical runs of increasing difficulty after [13].

4.1. The porous media equation. We simulate here the so-called Barenblatt's problem, [3], consisting in finding a particular solution to the porous media equation whose analytical expression is known. More precisely, we select $A(\rho, J) = \frac{J}{4\rho}$ in (20) which gives $\alpha = -1$ in Theorem 1 and we look for:

$$\rho(t,x) = \frac{1}{r(t)}\left(1 - \left(\frac{x}{r(t)}\right)^2\right)\mathbf{1}_{|x|\leq r(t)}, \qquad r(t) = \big(12(1+t)\big)^{\frac{1}{3}}.$$

In this case, the equation (24) is solved explicitly and the flux function is given by:

$$F(f_L^+, f_R^-) = \frac{(f_L^+)^2 - (f_R^-)^2}{(h/4) + \varepsilon(f_L^+ + f_R^-)}.$$

Its partial derivatives have constant signs whatever values take ε, h since the scheme (23) preserves positivity and:

$$\partial_+ F(f_L^+, f_R^-) = \frac{2hf_L^+ + 4\varepsilon(f_L^+ + f_R^-)^2}{\big(h + 4\varepsilon(f_L^+ + f_R^-)\big)^2}, \quad \partial_- F(f_L^+, f_R^-) = -\frac{2hf_R^- + 4\varepsilon(f_L^+ + f_R^-)^2}{\big(h + 4\varepsilon(f_L^+ + f_R^-)\big)^2}.$$

Notice that since we deal with a quadratic nonlinearity, the value of ζ in (29) is simply given by an arithmetic average. Numerical results are shown in Figure 1 in time $T = 3$ with the parameters $h = 0.15$ and $\Delta t = 0.01$. On the right, the absolute error between (23) and the exact solution is displayed as a function of ε; it stalls below a certain value as the value of h becomes a limiting factor.

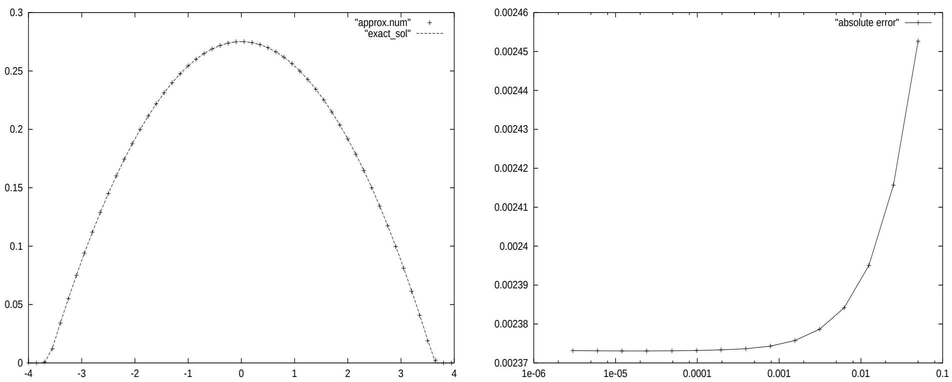

FIG. 1. *Numerical results for (23) on Barenblatt's problem in* $T = 3$*:* $3.10^{-6} \leq \varepsilon \leq 5.10^{-2}$.

4.2. The advection-diffusion equation. We move on now to another equation which has been investigated from the relaxation point of view in [20, 5, 2]:

$$\partial_t \rho + \partial_x \rho = \frac{1}{2}\partial_{xx}\rho, \qquad x \in \mathbb{R}, \qquad t > 0. \tag{30}$$

Therefore, we apply the same program relying on (20) together with the right-hand side $A(\rho, J) = J - \rho$. Once again, we are able to solve (19) and the flux function F is monotone since it comes out:

$$F(f_L^+, f_R^-) = 2\frac{\exp(2h)f_L^+ - f_R^-}{\exp(2h) - 1 + \varepsilon(1 + \exp(2h))}.$$

By means of Lemma 3, compactness holds in $h \geq 0$ under the CFL condition

$$\frac{2(\Delta t + \varepsilon h)}{1 - \exp(-2h)} \leq h \qquad \Rightarrow \qquad 0 \leq \Delta t = O(h^2),\ \varepsilon < h,$$

and one can check therefore that for (28),

$$div(F) = \frac{2}{1 + \varepsilon\left(\frac{\exp(2h)+1}{\exp(2h)-1}\right)} = O(1), \qquad 0 \leq h \leq 1.$$

Moreover concerning (25) and Lemma 4, we can choose $\tilde{F} = \partial_+ F$ and we get:

$$\varepsilon \|\tilde{F}\|_{L^\infty} < 1 \quad \Leftrightarrow \quad \varepsilon < h < 1.$$

Therefore, gathering inside Lemmas 3, 4 and Remark 4, we obtain an analogue of Theorem 1 for the equation (30):

THEOREM 2. *Assume* $0 \leq f_0^{\pm} \in L^1 \cap BV(\mathbb{R})$ *are initial data for (18), (20) with* $A(\rho, J) = J - \rho$ *and satisfy (26). Then as* $h, \varepsilon \to 0$ *under the prescribed conditions, the sequence* $f_h^{\pm}$ *generated by the scheme (23) is relatively compact in the strong topology of* $L^1_{loc}(\mathbb{R}^+_* \times \mathbb{R})$. *In particular,* $\rho_h = f_h^+ + f_h^-$ *converges towards the unique solution to (30),* $\rho(t = 0, .) = 2f_0^+ = 2f_0^-$.

We close this paragraph displaying some numerical results illustrating our statements. We choose some Maxwellian initial data $f_0^{\pm}(x) = \mathbf{1}_{x<5}$ to compute the solution of (20) in $x \in [0, 10]$ for $T = 2$. Of course, the exact solution of (30) is given by $\rho(t,x) = 1 - \text{erf}\big((x - t - 5)/\sqrt{2t}\big)$. We took $h = 0.2$ and $\Delta t = 0.02$: see Figure 2.

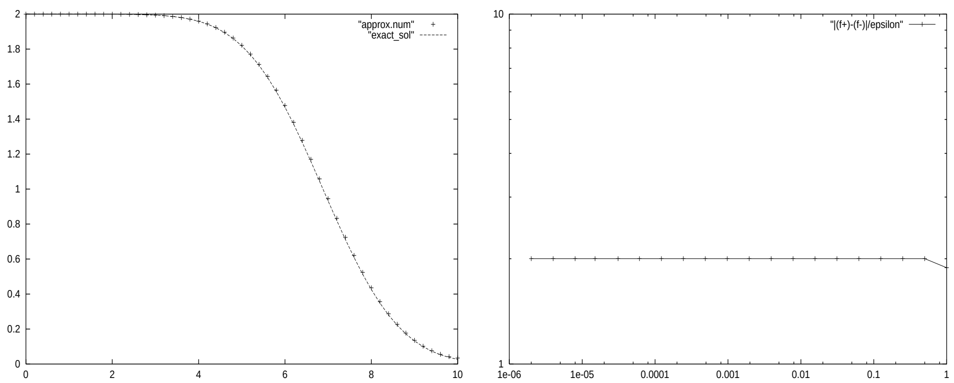

FIG. 2. *Numerical results for (23), (30) on a Riemann's problem:* $2.10^{-6} \leq \varepsilon \leq 1$.

4.3. The Carleman model. We consider the so-called Carleman's model, [6], which corresponds to the choice $A(\rho, J) = \rho J$, that is to say $\alpha = 1$ in Theorem 1. This presents a difficulty as the asymptotic behaviour is singular if $\rho = 0$; anyway, we succeeded in simulating the following *initial-boundary value problem* by means of the scheme (23):

$$\begin{aligned} \partial_t \rho = \frac{1}{2}\partial_{xx}\Big(\ln(\rho)\Big), \qquad & \rho(0,.) = 1 + \mathbf{1}_{x>0.5}; \\ \rho(., x = 0) = 1.5, \qquad & \rho(., x = 1) = 2.5. \end{aligned} \tag{31}$$

The computational domain is $x \in [0,1]$ and we took $h = 0.03$, $\Delta t = 0.001$ in order to produce the results displayed in Figure 3. In this case, the flux function cannot be known analytically but it turns out from (24) that the equation

$$F(f_L^+, f_R^-) = J = \frac{1}{2h} \ln \left(\frac{2f_L^+ - \varepsilon J}{2f_R^- + \varepsilon J} \right)$$

can be easily solved by means of a fixed point algorithm if ε is small enough. We compared our results with those generated by a standard second order centered scheme for (31): the absolute error between them is of the order of ε as announced in (27).

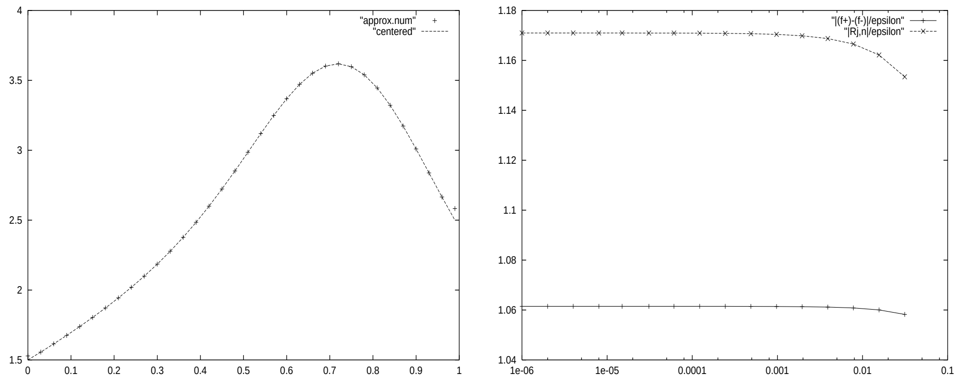

FIG. 3. *Transient regime ($T = 0.01$) for Carleman's model on the BVP (31):* $10^{-6} \leq \varepsilon \leq 3.10^{-2}$.

4.4. The Ruijgrok-Wu model. We close this section devoted to numerical tests with the Ruijgrok-Wu's model of the Boltzmann equation, [8, 29, 30], which relaxes under a diffusive scaling of variables towards the viscous Burgers equation, [16]. More precisely, we aim at reproducing with (23) the smooth solution of the initial-boundary value problem:

$$\begin{aligned} \partial_t \rho + \rho \partial_x \rho &= \frac{1}{2} \partial_{xx} \rho, \qquad \rho(0,.) = 2(2 - \mathbf{1}_{x>0}); \\ \rho(., x = -1) &= 4, \qquad \rho(., x = 1) = 2, \end{aligned} \tag{32}$$

by means of (20) with the right-hand side $A(\rho, J, \varepsilon J) = J - \frac{\varepsilon^2}{2} J^2 - \frac{1}{2}\rho^2$.

Several features makes it difficult: the first one is that, according to [8, 30], this 2×2 system isn't unconditionally quasi-monotone. Indeed, this property holds only in case the initial data satisfy $f_0^- \leq 1/2\varepsilon$; this has to be taken into account when starting the simulation. Moreover, the differential equation (19) cannot be always solved analytically and we integrated it approximately through the mid-point rule. Thus the flux function we used comes out of a fixed point algorithm on the following equation which is deduced from (21):

$$J = \frac{1}{2h} \left[(2f_L^+ - \varepsilon J)\Big(1 + \frac{h}{2}(2f_L^+ - \varepsilon J)\Big) - (2f_R^- + \varepsilon J)\Big(1 - \frac{h}{2}(2f_R^- + \varepsilon J)\Big) + \varepsilon h J^2 \right].$$

This choice led to the results displayed on Figure 4 with the parameters $h = 0.05$, $\Delta t = 0.001$. This demonstrates the fact that an approximate treatment of the jump relation (19) still gives a good outcome on a nonlinear problem and the Maxwellian estimate of Lemma 4 is kept.

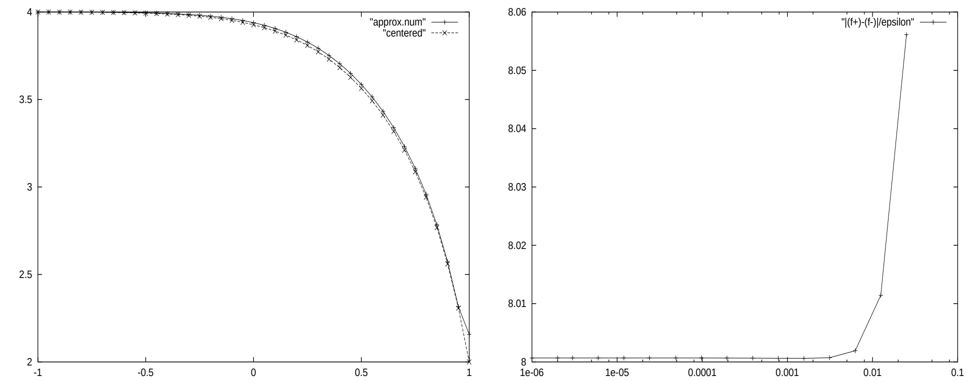

FIG. 4. *Transient regime ($T = 0.3$) for Ruijgrok-Wu's model on (32):* $10^{-6} \leq \varepsilon \leq 3.10^{-2}$.

5. Conclusion. We presented in this paper a study of a Well-Balanced scheme for discrete 2×2 kinetic models. This scheme is endowed with several interesting properties as it preserves steady-state curves in the rarefied regime (1) and is Asymptotic-Preserving in the sense of [17] as $\varepsilon \to 0$ in (16). Moreover, these statements can be rigorously established in many significant situations like for instance the cases investigated in [29]. From this point of view, this work can be seen as an extension of [10, 11] to relaxation problems with a diffusive behaviour. The present approach may be further developed in several directions; first, multidimensional problems could be considered, then more complex asymptotics can also be tackled in the spirit of [31]. In any case, intermediate scalings considered in [8, 32] (the so-called *incompressible Euler limits*) relaxing to conservation laws can be handled by a suitable modification of (17) inside the scheme (23).

REFERENCES

[1] D. Amadori, L. Gosse, G. Guerra, *Global BV entropy solutions and uniqueness for hyperbolic systems of balance laws*, Arch. Rational Mech. Anal. **162** (2002), 327–366.

[2] D. Aregba-Driollet, R. Natalini, S.Q. Tang, *Diffusive kinetic explicit schemes for nonlinear degenerate parabolic systems*, Quaderno IAC 26/2000, to appear in Math. Comp.

[3] G.I. Barenblatt, *On some steady motion of a liquid or a gas in a porous medium*, Prikl. Mat. Mekh. **16** (1952) p.67.

[4] R. Botschorijvili, B. Perthame, A. Vasseur, *Equilibrium schemes for scalar conservation laws with stiff source terms*, to appear in Math. Comp.

[5] F. Bouchut, F.R. Guarguaglini and R. Natalini, *Diffusive BGK Approximations for Nonlinear Multidimensional Parabolic Equations*, Indiana Univ. Math. J. **49** (2000), 723-749

[6] T. Carleman, *Problèmes mathématiques de la théorie cinétique des gaz*, Almquist Wiksells (1957)

[7] C. Cercignani, R. Illner, M. Pulvirenti, *The mathematical theory of dilute gases*, Applied Mathematical Sciences, **106**. Springer-Verlag, New York, 1994.

[8] E. Gabetta, B. Perthame, *Scaling limits for the Ruijgrok-Wu model of the Boltzmann equation*, Math. Mod. Appl. Sci. **24** (2001) 949–967.

[9] S.K. Godunov, *Finite difference schemes for numerical computation of solutions of the equations of fluid dynamics*, Mat. USSR Sbornik **47** (1959) 271–306

[10] L. Gosse, *Localization effects and measure source terms in numerical schemes for balance laws*, Math. Comp. **71** (2002), 553–582.

[11] L. Gosse, *Time-splitting schemes and measure source terms for a quasilinear relaxing system*, to appear in Math. Models Methods Appl. Sci.

[12] L. Gosse, *A well-balanced scheme using non-conservative products designed for hyperbolic systems of conservation laws with source terms*, Math. Models Methods Appl. Sci. **11** (2001)

339–365

[13] L. Gosse, G. Toscani, *An asymptotic-preserving well-balanced scheme for the hyperbolic heat equations*, C.R. Math. Acad. Sci. Paris **334** (2002) 337–342.

[14] L. Gosse, A.E. Tzavaras, *Convergence of relaxation schemes to the equations of elastodynamics*, Math. Comp. **70** (2001), 555–577

[15] J. Greenberg, A.Y. LeRoux, *A well balanced scheme for the numerical processing of source terms in hyperbolic equations*, SIAM J. Numer. Anal. **33** (1996) 1 – 16.

[16] E. Hopf, *The partial differential equation* $u_t + uu_x = \mu u_{xx}$, Comm. Pure Appl. Math. **3** (1950) 201 – 230.

[17] S. Jin, *Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations*, SIAM J. Sci. Comput. **21** (1999), 441–454.

[18] S. Jin, *A steady-state capturing method for hyperbolic systems with geometrical source terms* , Math. Mod. Numer. Anal. **35** (2001) 631 – 645.

[19] S. Jin, F. Golse, C.D. Levermore, *The convergence of numerical transfer schemes in diffusive regimes I: The discrete-ordinate method*, SIAM J. Numer. Anal. **36** (1999), 1333–1369.

[20] S. Jin, H. Liu, *Diffusion limit of a hyperbolic system with relaxation*, Methods Appl. Anal. **5** (1998), 317–334.

[21] S. Jin, L. Pareschi, G. Toscani, *Diffusive relaxation schemes for multiscale discrete-velocity kinetic equations*, SIAM J. Numer. Anal. **35** (1998), 2405–2439

[22] A. Klar, *An asymptotic induced scheme for nonstationary transport equations in the diffusive limit*, SIAM J. Numer. Anal. **35** (1998) 1073–1094

[23] A. Klar, A. Unterreiter, *Uniform stability of a finite difference scheme for transport equations in diffusive regimes*, SIAM J. Numer. Anal. **40** (2002) 891–913.

[24] S.N. Kružkov, *First order quasilinear equations in several independant space variables*, Mat. USSR Sbornik **81** (1970), 228–255.

[25] Ph. LeFloch, A.E. Tzavaras, *Representation of weak limits and definition of nonconservative products*, SIAM J. Math. Anal. **30** (1999), 1309 – 1342.

[26] R.J. LeVeque, *Balancing source terms and flux gradients in high resolution Godunov methods: the quasi steady wave propagation algorithm*, J. Comp. Phys. **146** (1998) 346 – 365.

[27] R.J. LeVeque, B. Temple, *Stability of Godunov's method for a class of* 2×2 *systems of conservation laws*, Trans. A.M.S. **288** (1985) 115–123.

[28] P.L. Lions, *Mathematical topics in fluid mechanics* **I** (incompressible models), Oxford Univ. Press.

[29] P.L. Lions, G. Toscani, *Diffusive limit for finite velocity Boltzmann kinetic models*, Rev. Mat. Iberoamericana **13** (1997), 473–513.

[30] H. Liu, J. Wang, G. Warnecke, *Convergence of a splitting scheme applied to the Ruijgrok-Wu model of the Boltzmann equation*, J. Comp. Appl. Math. **134** (2001) 343–367.

[31] P. Marcati, A. Milani, *The one-dimensional Darcy's law as the limit of a compressible Euler flow*, J. Differential Equations **84** (1990), 129–147

[32] G. Naldi, L. Pareschi, *Numerical schemes for hyperbolic systems of conservation laws with stiff diffusive relaxation* SIAM J. Numer. Anal. **37** (2000) 1246–1270

[33] R. Natalini, B. Hanouzet, *Weakly coupled systems of quasilinear hyperbolic equations*, Diff. Integ. Equations **9** (1997) 1279–1292

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

[35] B. Perthame, E. Tadmor, *A kinetic equation with kinetic entropy functions for scalar conservation laws*, Comm. Math. Phys. **136** (1991) 501 – 517.

[36] T. Platkowski, R. Illner, *Discrete velocity models of the Boltzmann equation: a survey on the mathematical aspects of the theory.* SIAM Rev. **30** (1988), no. 2, 213–255.

[37] W. Ruijgrok, T.T. Wu, *A completely solvable model of the nonlinear Boltzmann equation*, Physica A **113** (1982) 401–416.

[38] A.E. Tzavaras, *On the mathematical theory of fluid dynamic limits to conservation laws* In: Advances in Mathematical Fluid Mechanics, J. Malek, J. Nečas and M. Rokyta, eds.; Springer, New York, (2000) 192–222

[39] A.I. Vol'Pert, *Spaces BV and quasilinear equations*, Mat. USSR Sbornik **2** (1967) 225–267