HAL open science

NUMERICAL HIGH-FIELD LIMITS IN TWO-STREAM KINETIC MODELS AND 1D AGGREGATION EQUATIONS

Laurent Gosse, Nicolas Vauchelet

▶ **To cite this version:**

Laurent Gosse, Nicolas Vauchelet. NUMERICAL HIGH-FIELD LIMITS IN TWO-STREAM KINETIC MODELS AND 1D AGGREGATION EQUATIONS. SIAM Journal on Scientific Computing, 2016. hal-01233245

HAL Id: hal-01233245

https://hal.science/hal-01233245v1

Submitted on 24 Nov 2015

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

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

NUMERICAL HIGH-FIELD LIMITS IN TWO-STREAM KINETIC MODELS AND 1D AGGREGATION EQUATIONS

LAURENT GOSSE* AND NICOLAS VAUCHELET†

Abstract. Numerical resolution of two-stream kinetic models in strong aggregative setting is considered. To illustrate our approach, we consider an 1D kinetic model for chemotaxis in hydrodynamic scaling and the high field limit of the Vlasov-Poisson-Fokker-Planck system. A difficulty is that, in this aggregative setting, weak solutions of the limiting model blow up in finite time, therefore the scheme should be able to handle Dirac measures. It is overcome thanks to a careful discretization of the macroscopic velocity resulting of Vol'pert calculus: accordingly, a new well-balanced (WB) and asymptotic preserving (AP) numerical scheme is provided. Numerical simulations confirm a good behavior of solutions.

Key words. Asymptotic-Preserving and Well-Balanced scheme; Chemotaxis dynamics; Duality measure solutions; high-field limit; two-stream kinetic model; Vlasov-Poisson-Fokker-Planck system.

AMS subject classifications. 65M06, 92C17, 35L60.

1. Introduction. Two well-known 1D kinetic models, the Greenberg-Alt one of chemotaxis dynamics [25] and the attractive Vlasov-Poisson Fokker-Planck (VPFP), are numerically investigated in strong (high field) relaxing regime. Resulting macroscopic balance laws generally admit measure solutions involving Dirac atoms, which create havoc inside standard algorithms, because of their inherent *nonconservative products* [28]. Such singular limiting processes can be studied efficiently by means of *asymptotic preserving (AP)* [32], here based on *well-balanced (WB)* [21], schemes.

1.1. A two-stream chemotaxis model. The Cattaneo or Greenberg-Alt model [25, 12, 19, 26] describes the dynamics of the two distribution functions $f^{\pm}(t,x)$ of bacteria moving respectively in positive and negative directions, at time $t > 0$ and position $x \in \mathbb{R}$. Rescaling microscopic velocities to ± 1, it reads

$$\partial_t f^{\pm} \pm \partial_x f^{\pm} = \pm \frac{1}{\varepsilon}\left([\frac{1}{2} + \phi(\partial_x S)] f^- - [\frac{1}{2} - \phi(\partial_x S)] f^+ \right), \qquad x \in \mathbb{R}, \tag{1.1}$$

with initial/decay conditions $f^{\pm}(t=0,\cdot) = f_0^{\pm}$, $\lim_{x\to\pm\infty} f^{\pm}(t,x) = 0$. An equation holds on $S(t,x)$, the concentration of chemoattractant emitted by bacteria:

$$-\partial_{xx} S + S = \rho, \quad \rho = f^+ + f^-, \text{ i.e. } S(t,x) = \frac{1}{2}\exp(-|x|) * \rho(t,\cdot), \tag{1.2}$$

with the same type of decay at infinity, $\lim_{x\to\pm\infty} S(t,x) = 0$. Indeed, Young inequality yields $\|\partial_x S\|_{L^\infty} \le \|\partial_x K\|_{L^\infty} \|\rho\|_{L^1}$ with $K(x) = \frac{1}{2}e^{-|x|}$. In this system the density of cells is denoted ρ. The *turning kernels* $T^{\pm} := \frac{1}{2} \pm \phi(\partial_x S)$ correspond to bacteria reorientation rates, so $0 \le T^{\pm} \le 1$. Let $M = \|\rho\|_{L^1}$ the total (time-independent) mass, the latter condition on the turning kernel is satisfied provided a sub-characteristic

*IAC–CNR "MAURO PICONE" (SEZIONE DI ROMA), VIA DEI TAURINI, 19, 00185 ROME (ITALY) L.GOSSE@BA.IAC.CNR.IT

†SORBONNE UNIVERSITÉS, UPMC UNIV PARIS 06, UMR 7598, LABORATOIRE JACQUES-LOUIS LIONS, F-75005, PARIS, (FRANCE)
CNRS, UMR 7598, LABORATOIRE JACQUES-LOUIS LIONS, F-75005, PARIS,
INRIA-PARIS-ROCQUENCOURT, EPC MAMBA, DOMAINE DE VOLUCEAU, BP105, 78153 LE CHESNAY CEDEX, (FRANCE) VAUCHELET@ANN.JUSSIEU.FR

condition is met:

$$\sup_{v\in[-M/2,M/2]} |\phi(v)| \leq \frac{1}{2}. \tag{1.3}$$

Defining $J = f^+ - f^-$, (1.1) rewrites as a semi-linear relaxation system (see [2]),

$$\partial_t \rho + \partial_x J = 0, \qquad \partial_t J + \partial_x \rho = \frac{1}{\varepsilon}\big(2\phi(\partial_x S)\rho - J\big). \tag{1.4}$$

Thanks to (1.4), the model (1.1) appears to converge (at least formally), as $\varepsilon \to 0$, toward a continuity equation,

$$\partial_t \rho + \partial_x(a(\partial_x S)\rho) = 0, \qquad a(\partial_x S) = 2\phi(\partial_x S), \qquad -\partial_{xx} S + S = \rho, \tag{1.5}$$

belonging to the class of aggregation equations for which finite time blow-up of weak L^p-solutions for $p > 1$ was established (see e.g. [3]). Then Dirac deltas are produced in finite time within the macroscopic density ρ; at these atoms' location, the product $a(\partial_x S)\rho$ is not straightforwardly defined. Indeed, we infer from (1.2) that $\partial_x S$ is discontinuous at atoms' location.

1.2. A related Vlasov-Poisson-Fokker-Planck equation (VPFP). The VPFP is a fundamental system used *e.g.* in plasma physics: it models the evolution of a statistical distribution of electrons which dynamics are governed by both a self-consistent (gravitational or electrostatic) interaction and a Brownian motion induced by the underlying medium. For an attractive self-interaction, it reads [34]

$$\varepsilon(\partial_t f + v\partial_x f) + \partial_x \phi \cdot \partial_v f = \frac{1}{\theta}\partial_v(vf - \kappa\partial_v f), \qquad -\partial_{xx}\phi = \rho := \int_{\mathbb{R}} f(v)dv. \tag{1.6}$$

High field regime corresponds to letting $\varepsilon \to 0$. In the attractive case, like (1.1)-(1.2), an aggregation equation (for which blow-up of solutions usually occurs) emerges.

1.3. Organization of the manuscript. Although numerical approximation of kinetic system with stiff relaxation terms was thoroughly investigated in the last decade, there are only few contributions dealing with the strongly attractive case where the numerical scheme is meant to handle correctly the dynamics of aggregated matter, *i.e.* Dirac deltas. In [31], a fractional-step, asymptotic-preserving (AP) numerical scheme endowed with a careful discretization of the nonconservative product is proposed for the system (1.1)–(1.2). However this scheme is based on a time-splitting algorithm which is known to not have a good behavior for large time [2], i.e. does not satisfy the well-balanced property, in particular macroscopic flux for large time displays large peakons instead of being flat. This paper proposes a new numerical approach dealing with two-stream kinetic models in strongly aggregative setting. The WB property is obtained following the strategy of [21]. Obtaining an AP property relies on a careful discretization of Vol'pert calculus for BV functions.

The outline of the paper is as follows. In the next Section, we recall briefly the theoretical results of convergence towards aggregation equation as $\varepsilon \to 0$ for both models. Section 3 is devoted to a reminder of the strategy of [21] to derive WB scheme when the potential ($\partial_x S$ for (1.1) or E for (VPFP)) is assumed to be given. In §4, we develop our new strategy to build a WB and AP numerical scheme for the coupled models (see also [5, 9, 14]), the AP property of which is explained in §5. The main drawback of this strategy may be its CPU cost as it involves an iterative resolution of a nonlinear algebraic system (see Appendix A). Hence, we propose in §6 a simpler strategy for (1.1) based on hybridization [22]. Finally numerical comparisons between those algorithms are provided in §7.

2. Theoretical setups for both 1D kinetic models.

2.1. A two-stream model and its aggregation limit. A theory of measure solutions for equations of the type (1.5) was established in [30], along with convergence of kinetic solutions; besides, multi-dimensional aggregation equations were investigated in [7, 8]. In [30], (1.5) is treated as a conservative transport equation with a discontinuous velocity (see also [8, 10, 35, 36]). A convenient tool (in one dimension) to deal with such problems is provided by **duality solutions**, introduced in [4] (see also [36, 37]). Numerical investigation of these was carried out in [24].

DEFINITION 1. *The set* $(\rho, S) \in C(0,T;\mathcal{M}_b(\mathbb{R})) \times C(0,T;W^{1,\infty}(\mathbb{R}))$ *is a* **duality solution** *to (1.5) if there exists* $\alpha(t) \in L^1_{loc}(0,T)$ *such that, for* $0 < t_1 < t_2 < T$,

1. *the continuity equation holds* **in the sense of duality** *on* (t_1, t_2),
2. $\forall t \in (0,T)$, $\partial_x(a(\partial_x S))(t,\cdot) \le \alpha(t)$ *almost everywhere*[1] *and* $a(\partial_x S) \in L^\infty$,
3. *the elliptic equation (1.2) holds in a weak sense, for any* $\varphi \in C^1(\mathbb{R})$,

$$\forall t \in (0,T), \qquad \int_{\mathbb{R}} \partial_x S(t,x) \cdot \partial_x \varphi(x) + S(t,x) \cdot \varphi(x) \mathrm{d}x = \int_{\mathbb{R}} \rho(t,x)\varphi(x)\mathrm{d}x,$$

The terminology that a 1D continuity equation holds "in the sense of duality" explicitly refers to [4, §4.2] or [24, §2.2]: it essentially means that test-functions are chosen to be "reversible solutions" of the dual equation. The second point in Def. 1 is a one-sided Lipschitz condition. To secure uniqueness, a definition of $a(\partial_x S)\rho$, a **nonconservative product** [28] at places where $\rho(t,x)$ is a Dirac mass, is needed. Indeed at such location, the product of the discontinuous function $\partial_x S$ with the measure ρ is a priori not well defined.

From now on, we denote $\mathcal{M}_b(\mathbb{R})$ the space of local Borel measures on $\mathbb{R}$ with finite total variation; this space of measures is always endowed with the weak topology $\sigma(\mathcal{M}_b, C_0)$ where C_0 is the set of continuous functions that vanish at infinity.

DEFINITION 2. *Let* A *be such that* $A' = a$, *with* $A(0) = 0$. *Based on (1.2), the macroscopic flux is defined as follows,*

$$J(t,x) = -\partial_x A(\partial_x S) + a(\partial_x S) \cdot S. \tag{2.1}$$

The **Vol'Pert calculus** *(or chain rule for BV functions), again by* (1.2), *yields* $J = \widehat{a}_s \cdot \rho$, *where* $\widehat{a} = a(\partial_x S)$ *a.e.; the* **Vol'pert velocity** *is the quantity* $\widehat{a}_s$.

Thanks to both these definitions, the limiting process (1.1)–(1.2) is proved to rigorously converge as $\varepsilon \to 0$ toward the unique duality solution of (1.5):

THEOREM 1. *([30, Theorem 3.9 & 3.10]) Let* $0 \le f_0^\pm \in \mathcal{M}_b(\mathbb{R})$ *such that* $\int_{\mathbb{R}}(1+|x|^2)f_0^\pm(x)dx < \infty$ *and* $T > 0$:

- *there exists a unique duality solution on* $(0,T)$ *to (1.5)–(1.2) in the sense of Definition 1 such that,*

$$\partial_t \rho + \partial_x J = 0, \qquad \text{with } J \text{ defined in } (2.1);$$

- *the sequence* $f_\varepsilon^\pm, S_\varepsilon$ *of weak solutions to (1.1)–(1.2) converges weakly,*

$$\rho_\varepsilon := f_\varepsilon^+ + f_\varepsilon^- \rightharpoonup \rho, \text{ in } C([0,T],\mathcal{M}_b(\mathbb{R})), \qquad S_\varepsilon \rightharpoonup S, \text{ in } C([0,T],W^{1,\infty}(\mathbb{R})),$$

  where ρ, S *the unique duality solution of (1.5)–(1.2)–(2.1).*

Thus a challenging issue in the numerical handling of (1.1)–(1.2) is to unveil a scheme at the kinetic level, able to recover the flux definition (2.1) when $\varepsilon \to 0$.

[1] "almost everywhere" is always with respect to the Lebesgue measure.

2.2. High field limit for a Vlasov-Poisson-Fokker-Planck system. We switch now to (1.6), and its high field limit obtained by letting $\varepsilon \to 0$:

$$\varepsilon(\partial_t f + v\partial_x f) + E\partial_v f = \frac{1}{\theta}\partial_v(vf - \kappa\partial_v f), \tag{2.2}$$

$$-\partial_{xx}\phi = \rho, \qquad E = \partial_x \phi. \tag{2.3}$$

In this attractive case, concentration occurs together with Dirac masses within the limiting problem. Considering only the first two moments of equation (2.2) leads to,

$$\partial_t \rho + \partial_x J = 0, \qquad \partial_t J + \partial_x \int_{\mathbb{R}} v^2 f(v)\, dv = \frac{1}{\theta\varepsilon}(\theta\rho E - J),$$

with standard notations $\rho = \int_{\mathbb{R}} f(v)dv$ and $J = \int_{\mathbb{R}} vf(v)\, dv$. Letting formally $\varepsilon \to 0$, a scalar conservation law emerges,

$$J = \theta\rho E, \qquad \partial_t \rho + \theta\partial_x(\rho E) = 0. \tag{2.4}$$

As (2.3) implies that $\rho = -\partial_x E$, provided that the product $\rho E = -\frac{1}{2}\partial_x E^2$ is correctly defined, an integration in space implies that E solves a Burgers-Hopf equation:

$$\partial_t E + \frac{\theta}{2}\partial_x E^2 = 0, \qquad F(E) = \frac{\theta}{2}|E|^2. \tag{2.5}$$

This formal computation has been rigorously studied in [34] (see also [29]):

THEOREM 2. *(see [34, Theorems 1 & 2] and [29, Theorem 3.2]) Let a sequence of nonnegative initial data $0 \le f_0^\varepsilon$ be such that:*

- *$\int_{\mathbb{R}^2}(1 + |x|^2 + |v|^2) f_0^\varepsilon dxdv$, is uniformly bounded with respect to $\varepsilon > 0$,*
- *corresponding $\rho_0^\varepsilon \rightharpoonup \rho_0$ in $\mathcal{M}_+(\mathbb{R})$-weak$\star$.*

Then, there exists an unique duality solution (ρ, E), $\rho \ge 0$ to the system (2.3)–(2.4) with initial data ρ_0 and where the product is defined by $\rho E = -\frac{1}{2}\partial_x E^2$. It satisfies $\rho = -\partial_x E$ where E is the unique entropy solution of (2.5). Moreover, the sequence $(f_\varepsilon, E_\varepsilon)$ from (2.2)–(2.3) weakly converges to this unique solution, in the sense

$$\rho_\varepsilon = \int_{\mathbb{R}} f_\varepsilon(v)\, dv \rightharpoonup \rho, \ \text{in } \mathcal{D}'([0,T] \times \mathbb{R}) \qquad \text{as } \varepsilon \to 0.$$

Thus an interesting issue in the numerical resolution of the high field limit of VPFP is to design a numerical scheme, for which the numerical solution converges to the one computed by solving the Burgers equation (2.5). Numerical simulation of VPFP systems was studied by several authors, see *e.g.* [16, 33, 38, 42, 11]. However, up to our knowledge, the numerical resolution of the high field limit has still not been studied and in particular no numerical comparisons with solutions of (2.5) are available. In fact, specific difficulties occur, if $\varepsilon \ll \Delta x$, because of the nonconservative product appearing in the definition of the macroscopic flux.

3. Derivation of WB schemes based on scattering matrices. Hereafter we shall always work on a uniform Cartesian computational grid, defined by the parameters $\Delta t, \Delta x > 0$, with $\Delta t \le \Delta x$; in standard notation, numerical approximations of kinetic densities read $f^\pm(t^n, x_j) \simeq f^\pm_{j,n}$ for $x_j = j\Delta x$, $j \in \mathbb{Z}$ and $t^n = n\Delta t$, $n \in \mathbb{N}$.

3.1. Chemotaxis model (1.1). According to [21], a convenient manner of treating a weakly nonlinear model like (1.1)–(1.2) consists in building a Godunov scheme based on an exact Riemann solver for the slightly modified problem,

$$\partial_t f^\pm \pm \partial_x f^\pm = \pm \sum_{j\in\mathbb{Z}} \frac{\Delta x}{2\varepsilon} \Big([1 + a^n_{j-\frac12}] f^- - [1 - a^n_{j-\frac12}] f^+ \Big) \cdot \delta(x - x_{j-\frac12}), \tag{3.1}$$

where $a^n_{j-\frac12}$ is a "consistent discretization" of $2\phi(\partial_x S)(t^n, x_{j-\frac12})$: this amounts to decouple the original hyperbolic/elliptic model into a position-dependent linear 2×2 system at each time-step. A justification comes from the explicit form of the solution of (1.2), a convolution product, which supposedly reacts "slowly" to perturbations in ρ (this idea traces back to Gummel [27]). Assuming that the form of $a^n_{j-\frac12}$ is known, the Riemann problem arising at each interface $x_{j-\frac12}$ gives rise to 3 elementary waves (or normal modes): 2 usual linear convective ones, and a steady-state discontinuity linked to the Dirac mass, across which unknowns jump according to a Boundary-Value problem (BVP) for the stationary equations of (1.1), for any $j, n \in \mathbb{Z}\times\mathbb{N}$,

$$2\varepsilon\partial_x \bar f^\pm(x) = [1 + a^n_{j-\frac12}]\bar f^- - [1 - a^n_{j-\frac12}]\bar f^+, \;\; \bar f^+(0) = f^+_{j-1,n}, \;\; \bar f^-(\Delta x) = f^-_{j,n}, \tag{3.2}$$

which are equivalent to the "macroscopic stationary equations",

$$\partial_x \bar J = 0, \qquad \varepsilon\partial_x \bar\rho(x) = a^n_{j-\frac12}\bar\rho - \bar J, \qquad x \in (0, \Delta x). \tag{3.3}$$

Accordingly, the collision term is numerically handled as a countable collection of "local scattering centers" [18], each of which inducing a 2×2 scattering matrix, $\mathcal{S}^n_{j-\frac12}$ relating incoming states (or "inflow data" of (3.2)) to outgoing ones:

$$\begin{pmatrix} f^+_{j-\frac12,n} \\ f^-_{j-\frac12,n} \end{pmatrix} := \begin{pmatrix} \bar f^+(\Delta x) \\ \bar f^-(0) \end{pmatrix} = \mathcal{S}^n_{j-\frac12} \begin{pmatrix} f^+_{j-1,n} \\ f^-_{j,n} \end{pmatrix}. \tag{3.4}$$

The Divergence Theorem furnishes the expression of the resulting Godunov scheme,

$$f^+_{j,n+1} = f^+_{j,n} - \frac{\Delta t}{\Delta x}(f^+_{j,n} - f^+_{j-\frac12,n}), \qquad f^-_{j,n+1} = f^-_{j,n} + \frac{\Delta t}{\Delta x}(f^-_{j+\frac12,n} - f^-_{j,n}),$$

which rewrites, in a vectorial form,

$$\begin{pmatrix} f^+_{j,n+1} \\ f^-_{j-1,n+1} \end{pmatrix} = (1 - \frac{\Delta t}{\Delta x}) \begin{pmatrix} f^+_{j,n} \\ f^-_{j-1,n} \end{pmatrix} + \frac{\Delta t}{\Delta x}\mathcal{S}^n_{j-\frac12} \begin{pmatrix} f^+_{j-1,n} \\ f^-_{j,n} \end{pmatrix}. \tag{3.5}$$

LEMMA 3.1. *Suppose that for all j, n, the matrix $\mathcal{S}^n_{j-\frac12}$ is left-stochastic and the CFL restriction, $\Delta t \le \Delta x$, holds. Then the scheme (3.5) dissipates the L^1 norm,*

$$\forall n \in \mathbb{N}, \qquad \sum_{j\in\mathbb{Z}} \Delta x(|f^+_{j,n+1}| + |f^-_{j,n+1}|) \le \sum_{j\in\mathbb{Z}} \Delta x(|f^+_{j,n}| + |f^-_{j,n}|).$$

For nonnegative initial data $(f^\pm_{j,n=0})_{j\in\mathbb{Z}}$, it preserves both positivity and total mass.

Proof. The dissipation is a direct consequence of the fact $\|\mathcal{S}^n_{j-\frac12}\|_{\ell^1} = 1$ in (3.5). The positivity-preserving property results from convex combination arguments. □

At this point, we left behind the issue of defining a value of $a^n_{j-\frac12}$, able to compute correctly products of the type $a^n_{j-\frac12}\cdot f^\pm$, likely to become singular as $\varepsilon \to 0$. Without a specific treatment, the Vol'Pert product (2.1) is usually not recovered numerically, as shown in our Fig. 7.1.

3.2. VPFP model. As above, let an uniform grid in the space variable be x_j, $j \in \mathbb{Z}$, together with a discretization v_k, $k \in \mathbb{Z}$, in the velocity domain. One starts again with stationary equations and incoming boundary conditions, in (x_{j-1}, x_j):

$$\begin{aligned} &\varepsilon v \partial_x f = \tfrac{1}{\theta}\partial_v\Big((v-\theta E)f - \kappa\partial_v f\Big), \\ &f(0,v) = f_{j-1}(v), \qquad f(\Delta x, -v) = f_j(-v), \qquad \text{for } v>0. \end{aligned} \tag{3.6}$$

At this early stage, let the field be a constant value E, given. Plugging the usual "separated variables" ansatz $f(x,v) = \exp(-\lambda x)\varphi_\lambda(v)$, one recovers a Sturm-Liouville problem, [6, 21, Chapter 12.3]. To the null eigenvalue $\lambda = 0$, are associated two non-damped modes $\Psi_0^\pm$ ("diffusion solutions" in the sense of [15]), and f expands into

$$f(x,v) = \gamma^+\Psi_0^+(x,v) + \gamma^-\Psi_0^-(x,v) + \sum_{\ell\in\mathbb{N}^*}\gamma_\ell^\pm\Psi_\ell^\pm(x,v),$$

Coefficients $\gamma_\ell^\pm$ in this full-range expansion follow from incoming boundary conditions in (3.6), therefore they depend only on incoming boundary condition and are constants. In particular, for $\ell > 0$, damped modes $\Psi_\ell^\pm$ may be considered as boundary layers, hence negligible away from the boundaries. Hereafter, we shall restrict ourselves to the diffusion solutions, $\ell = 0$, so $f(x,v) \approx \gamma^+\Psi_0^+(x,v) + \gamma^-\Psi_0^-(x,v)$. Beside simplicity, a reason for such a simplification is that none of the damped modes $\Psi_\ell^\pm$ carries net macroscopic flux, as is seen from the relation, [21, Remark 12.4]

$$\int_{\mathbb{R}} v\Psi_{\pm\ell}(0,v)\Psi_{\pm\ell'}(0,v)\ \exp((v-\theta E)^2/2\kappa)dv = C\cdot\delta_{\ell,\ell'}, \qquad \ell,\ell'\in\mathbb{Z}^2,$$

inside which one picks a space-homogeneous mode $\Psi_0^\pm$, explicitly given in [21, p. 251]:

- when $E > 0$,

$$\Psi_0^+(x,v) = \exp(-\frac{(v-|\theta E|)^2}{2\kappa}); \quad \Psi_0^-(x,v) = \exp(-\frac{|\theta E|^2}{2\kappa})\exp(\frac{Ex}{\varepsilon\kappa} - \frac{v^2}{2\kappa}); \tag{3.7}$$

- when $E < 0$,

$$\Psi_0^+(x,v) = \exp(-\frac{|\theta E|^2}{2\kappa})\exp(\frac{Ex}{\varepsilon\kappa} - \frac{v^2}{2\kappa}); \quad \Psi_0^-(x,v) = \exp(-\frac{(v+|\theta E|)^2}{2\kappa}). \tag{3.8}$$

Accordingly, (3.6) reduces to a two-stream system by restricting the velocity space to a pair $\{-v_1, v_1\}$ (belonging to a Gauss-Hermite quadrature), so that $\gamma^\pm \in \mathbb{R}^2$ in

$$f(x,\pm v_1) = \gamma^+\Psi_0^+(x,\pm v_1) + \gamma^-\Psi_0^-(x,\pm v_1), \tag{3.9}$$

are uniquely determined by the data of incoming boundary conditions,

$$\begin{aligned} f(0,v_1) &= f_{j-1}(v_1) = \gamma^+\Psi_0^+(0,v_1) + \gamma^-\Psi_0^-(0,v_1), \\ f(\Delta x,-v_1) &= f_j(-v_1) = \gamma^+\Psi_0^+(\Delta x,-v_1) + \gamma^-\Psi_0^-(\Delta x,-v_1), \end{aligned}$$

through an easily invertible 2×2 linear system. Denoting

$$M = \begin{pmatrix} \Psi_0^+(0,v_1) & \Psi_0^-(0,v_1) \\ \Psi_0^+(\Delta x,-v_1) & \Psi_0^-(\Delta x,-v_1)\end{pmatrix}, \qquad \begin{pmatrix}\gamma^+\\ \gamma^-\end{pmatrix} = M^{-1}\begin{pmatrix} f_{j-1}(v_1)\\ f_j(-v_1)\end{pmatrix},$$

corresponding outward values of f follow thanks to (3.9),

$$f(\Delta x, v_1) = \gamma^+ \Psi_0^+(\Delta x, v_1) + \gamma^- \Psi_0^-(\Delta x, v_1),$$
$$f(0, -v_1) = \gamma^+ \Psi_0^+(0, -v_1) + \gamma^- \Psi_0^-(0, -v_1),$$

which can be rephrased as

$$\widetilde{M} = \begin{pmatrix} \Psi_0^+(\Delta x, v_1) & \Psi_0^-(\Delta x, v_1) \\ \Psi_0^+(0, -v_1) & \Psi_0^-(0, -v_1) \end{pmatrix}, \quad \begin{pmatrix} f(\Delta x, v_1) \\ f(0, -v_1) \end{pmatrix} = \widetilde{M} M^{-1} \begin{pmatrix} f_{j-1}(v_1) \\ f_j(-v_1) \end{pmatrix}. \tag{3.10}$$

Tedious computations, using expressions (3.7)–(3.8), lead to

$$\widetilde{M} M^{-1} = \begin{pmatrix} s_1 & s_2 \\ 1 - s_1 & 1 - s_2 \end{pmatrix},$$

with

$$s_1 = \frac{1 - e^{-2|\theta E| \frac{v}{\kappa}}}{1 - e^{-2|\theta E| \frac{v}{\kappa} - |E| \frac{\Delta x}{\varepsilon \kappa}}}, \quad s_2 = \frac{1 - e^{-2|E| \frac{\Delta x}{\varepsilon \kappa}}}{1 - e^{-2|\theta E| \frac{v}{\kappa} - |E| \frac{\Delta x}{\varepsilon \kappa}}}, \quad \text{if } E > 0;$$
$$s_1 = \frac{e^{-|E| \frac{\Delta x}{\varepsilon \kappa}} (1 - e^{-2|\theta E| \frac{v}{\kappa}})}{1 - e^{-2|\theta E| \frac{v}{\kappa} - |E| \frac{\Delta x}{\varepsilon \kappa}}}, \quad s_2 = \frac{e^{-2|\theta E| \frac{v}{\kappa}} (1 - e^{-2E \frac{\Delta x}{\varepsilon \kappa}})}{1 - e^{-2|\theta E| \frac{v}{\kappa} - |E| \frac{\Delta x}{\varepsilon \kappa}}}, \quad \text{if } E < 0.$$

The issue of determining a reliable discretization of the Poisson force field has been left apart since we assumed in all this section that the value E was given.

4. Scattering algorithm involving discontinuous velocities. When ρ is endowed with atoms, one must get rid of standard discretizations (like the ones in [19, 21]) because they produce wrong dynamics, as illustrated in Fig. 7.1 or in [40].

4.1. Derivation of the numerical process for (1.1). Here, the resolution of this stationary problem (3.2) proceeds by splitting into 2 contributions, one coming from the left with velocity $a^n_{j-\frac{1}{2},L}$ and another coming from the right with velocity $a^n_{j-\frac{1}{2},R}$, the precise values of which will be given later in §4.2.

- The first step consists in replacing the BVP (3.2) by two Cauchy problems on $(0, \Delta x)$. The flux $\bar{J}$ defined in (3.3) is constant and can be decomposed as $\bar{J} = J_L + J_R$ by means of two constants J_L and J_R, standing for left and right fluxes, respectively. Accordingly, $\bar{\rho}(x) = \rho_L(x) + \rho_R(x)$ such that $\rho(0) = \rho_L(0)$ and $\rho(\Delta x) = \rho_R(\Delta x)$. Thus, $\rho_R(0) = 0$ and $\rho_L(\Delta x) = 0$.
- The BVP (3.2) is reformulated as two Cauchy problems on $(0, \Delta x)$:

$$\varepsilon \frac{d}{dx} \rho_L = a_{j-\frac{1}{2},L} \rho_L - J_L, \qquad \rho_L(\Delta x) = 0, \tag{4.1}$$
$$\varepsilon \frac{d}{dx} \rho_R = a_{j-\frac{1}{2},R} \rho_R - J_R, \qquad \rho_R(0) = 0. \tag{4.2}$$

Moreover,

$$J = J_L + J_R, \qquad \rho(0) = \rho_L(0), \qquad \rho(\Delta x) = \rho_R(\Delta x). \tag{4.3}$$

The unknowns of this system are J_L, J_R, $f^-(0)$ and $f^+(\Delta x)$.

LEMMA 4.1. *For any $j, n \in \mathbb{Z}\times\mathbb{N}$, the scattering matrix $\mathcal{S}^n_{j-1/2}$ has the expression,*

$$\mathcal{S}^n_{j-1/2} = \begin{pmatrix} \dfrac{2c_L}{1-c_R+c_L} & \dfrac{1+c_R+c_L}{1-c_R+c_L} \\ \dfrac{1-c_R-c_L}{1-c_R+c_L} & \dfrac{-2c_R}{1-c_R+c_L} \end{pmatrix}, \tag{4.4}$$

with the signed coefficients ($a^n_{j-\frac12,L}$ and $a^n_{j-\frac12,R}$ may be equal) :

$$c_R = \frac{a^n_{j-\frac12,R}}{1-\exp(a^n_{j-\frac12,R}\Delta x/\varepsilon)} \le 0, \qquad c_L = \frac{a^n_{j-\frac12,L}}{1-\exp(-a^n_{j-\frac12,L}\Delta x/\varepsilon)} \ge 0. \tag{4.5}$$

Under the discrete sub-characteristic condition,

$$\sup_{j,n\in\mathbb{Z}\times\mathbb{N}} |a^n_{j-\frac12,L/R}| \le 1, \tag{4.6}$$

the matrix $\mathcal{S}^n_{j-1/2}$ is left-stochastic.

Proof. The quantities J_R, J_L, $a^n_{j-\frac12,R}$ and $a^n_{j-\frac12,L}$ are constant across each interface $x_{j-1/2}$, i.e. on $(0,\Delta x)$, so both (4.1), (4.2) are solved explicitly:

$$a_{j-\frac12,L}\,\rho_L(0) = J_L\big(1-e^{-a^n_{j-\frac12,L}\,\Delta x/\varepsilon}\big), \quad a_{j-\frac12,R}\,\rho_R(\Delta x) = J_R\big(1-e^{a^n_{j-\frac12,R}\,\Delta x/\varepsilon}\big).$$

Hence, using the notations of (4.5),

$$J_L = c_L(f^+_{j-1} + \bar f^-(0)), \qquad J_R = c_R(\bar f^+(\Delta x) + f^-_j). \tag{4.7}$$

Moreover, since the total flux $J = \bar f^+ - \bar f^-$ is constant, $\bar f^+(\Delta x) - f^-_j = f^+_{j-1} - \bar f^-(0)$ which gives a 2×2 system of unknowns $\bar f^-(0), \bar f^+(\Delta x)$

$$\bar f^+(\Delta x) - f^-_j = f^+_{j-1} - \bar f^-(0), \qquad \bar f^+(\Delta x) - f^-_j = J_L + J_R. \tag{4.8}$$

Recombining equations (4.8) with (4.7) gives

$$\begin{aligned} \bar f^+(\Delta x) + \bar f^-(0) &= f^+_{j-1} + f^-_j \\ \bar f^+(\Delta x) - f^-_j &= c_R\big(f^-_j + \bar f^+(\Delta x)\big) + c_L\big(f^+_{j-1} + \bar f^-(0)\big), \end{aligned}$$

which rewrites as,

$$\begin{aligned} \bar f^-(0)\big(1-c_R+c_L\big) &= -2c_R f^-_j + \big(1-c_R-c_L\big) f^+_{j-1}, \\ \bar f^+(\Delta x)\big(1-c_R+c_L\big) &= f^-_j\big(1+c_R+c_L\big) + 2c_L f^+_{j-1}. \end{aligned}$$

These equalities can be written in matrix form as in (4.4). At last, under the sub-characteristic condition, one checks easily that $c_L \ge 0$, $c_R \le 0$ and $c_L + c_R \le 1$, so that the sum of each column's (nonnegative) entries in the scattering matrix $\mathcal{S}^n_{j-1/2}$ equals 1, that is $\mathcal{S}^n_{j-1/2}$ is left-stochastic. □

4.2. Specific handling of the asymptoticly singular product. As formerly explained, careful values $a^n_{j-\frac12,L/R}$ must be devised in order to recover the correct behavior as $\varepsilon \to 0$. Accordingly, when put altogether, both values ρ_L and ρ_R at each interface induce sequences $(\rho_{j-\frac12,L})_j$ and $(\rho_{j-\frac12,R})_j$, (we drop the time index n)

$$\forall j \in \mathbb{Z}, \qquad \rho_{j-\frac12,L} = f^+_{j-1} + f^-_{j-\frac12}\,, \qquad \rho_{j-\frac12,R} = f^+_{j-\frac12} + f^-_j\,.$$

Then from the expression of $\mathcal{S}^n_{j-1/2}$ given in Lemma 4.1 we deduce,

$$\rho_{j-\frac{1}{2},L} = \frac{2(1-c_R)}{1-c_R+c_L} f^+_{j-1} - \frac{2c_R}{1-c_R+c_L} f^-_j \,, \tag{4.9}$$

$$\rho_{j-\frac{1}{2},R} = \frac{2c_L}{1-c_R+c_L} f^+_{j-1} + \frac{2(1+c_L)}{1-c_R+c_L} f^-_j \,. \tag{4.10}$$

These densities induce corresponding potentials S_L and S_R solving each elliptic problem, $-\partial_{xx}S_L + S_L = \rho_L$ and $-\partial_{xx}S_R + S_R = \rho_R$. Standard centered finite differences furnish approximations $S_{j-\frac{1}{2},L}$ and $S_{j-\frac{1}{2},R}$, respectively,

$$-\frac{S_{j+\frac{1}{2},L/R} - 2S_{j-\frac{1}{2},L/R} + S_{j-\frac{3}{2},L/R}}{\Delta x^2} + S_{j-\frac{1}{2},L/R} = \rho_{j-\frac{1}{2},L/R}\,,. \tag{4.11}$$

Centered discretizations of partial derivatives of $S_{L/R}$ read

$$\partial_x S_{j,L} = \frac{S_{j+\frac{1}{2},L} - S_{j-\frac{1}{2},L}}{\Delta x} \qquad \text{and} \qquad \partial_x S_{j,R} = \frac{S_{j+\frac{1}{2},R} - S_{j-\frac{1}{2},R}}{\Delta x}\,.$$

Yet, left and right values at any interface $x_{j-\frac{1}{2}}$ of the macroscopic velocity are computed thanks to the Vol'pert discretization :

$$a_{j-\frac{1}{2},L} = \begin{cases} \dfrac{A(\partial_x S_{j,L}) - A(\partial_x S_{j-1,L})}{\partial_x S_{j,L} - \partial_x S_{j-1,L}}, & \text{if} \quad \partial_x S_{j,L} \neq \partial_x S_{j-1,L}\,; \\ 0, & \text{else}\,; \end{cases} \tag{4.12}$$

$$a_{j-\frac{1}{2},R} = \begin{cases} \dfrac{A(\partial_x S_{j,R}) - A(\partial_x S_{j-1,R})}{\partial_x S_{j,R} - \partial_x S_{j-1,R}}, & \text{if} \quad \partial_x S_{j,R} \neq \partial_x S_{j-1,R}\,; \\ 0, & \text{else}\,; \end{cases} \tag{4.13}$$

We notice that according to assumption (1.3), the macroscopic velocities above satisfy the sub-characteristic condition (4.6) given in Lemma 4.1. In practice, computing velocities $a_{j-\frac{1}{2},L/R}$ consists in iteratively solving a nonlinear system for $(S_{j-\frac{1}{2},L})_{j\in\mathbb{Z}}$ and $(S_{j-\frac{1}{2},R})_{j\in\mathbb{Z}}$: namely, equations (4.5)-(4.9)-(4.10)-(4.11)-(4.12)-(4.13).

See our Appendix A for some implementations details.

4.3. A similar handling of (VPFP). In a quite similar way, there holds:

LEMMA 4.2. *For any $j, n \in \mathbb{Z}\times\mathbb{N}$, the scattering matrix $\mathcal{S}^n_{j-1/2}$ still has the expression (4.4), yet with more involved coefficients*

$$\begin{aligned} c_L &= \frac{\tanh\left(\frac{v_1}{\kappa}|\theta E_{j-\frac{1}{2},L}|\right)}{\omega_1(1-\exp(-|E_{j-\frac{1}{2},L}|\frac{\Delta x}{\varepsilon\kappa}))}\left(\mathbf{1}_{E_{j-\frac{1}{2},L}>0} + e^{-|E_{j-\frac{1}{2},L}|\frac{\Delta x}{\varepsilon\kappa}}\mathbf{1}_{E_{j-\frac{1}{2},L}<0}\right) \geq 0, \\ c_R &= \frac{-\tanh\left(\frac{v_1}{\kappa}|\theta E_{j-\frac{1}{2},R}|\right)}{\omega_1(1-\exp(-|E_{j-\frac{1}{2},R}|\frac{\Delta x}{\varepsilon\kappa}))}\left(e^{-|E_{j-\frac{1}{2},R}|\frac{\Delta x}{\varepsilon\kappa}}\mathbf{1}_{E_{j-\frac{1}{2},R}>0} + \mathbf{1}_{E_{j-\frac{1}{2},R}<0}\right) \leq 0, \end{aligned} \tag{4.14}$$

where we recall that v_1 is the first node in the Gauss-Hermite quadrature and ω_1 the associated weight. The matrix $\mathcal{S}^n_{j-1/2}$ is left-stochastic.

Again, an appropriate discretization of the macroscopic force fields $E_{j-\frac{1}{2},L}$ and $E_{j-\frac{1}{2},R}$ will be provided in the forthcoming §4.4.

Proof. From (3.10), it comes,

$$f_{j-1}(v_1) - f(\Delta x, -v_1) = f(0, v_1) - f_j(-v_1) = s_1 f_{j-1}(v_1) + (s_2 - 1) f_j(-v_1),$$

so that the corresponding flux reads,

$$J_{j-1/2} = v_1 \omega_1 (s_1 f_{j-1}(v_1) + (s_2 - 1) f_j(-v_1)). \tag{4.15}$$

Following the idea in §4, we introduce the following macroscopic densities

$$\rho_{j-\frac{1}{2},L} = \omega_1 (f_{j-1}(v_1) + f(0, -v_1)), \qquad \rho_{j-\frac{1}{2},R} = \omega_1 (f(\Delta x, v_1) + f_j(-v_1)). \tag{4.16}$$

From equation (3.10), one deduces

$$\begin{cases} \rho_{j-\frac{1}{2},L} = & \omega_1((2 - s_1) f_{j-1}(v_1) + (1 - s_2) f_j(-v_1)), \\ \rho_{j-\frac{1}{2},R} = & \omega_1 (s_1 f_{j-1}(v_1) + (1 + s_2) f_j(-v_1)). \end{cases} \tag{4.17}$$

Combining (4.15) and (4.17), and after straightforward computations,

$$J_{j-\frac{1}{2}} = v_1 \left(\frac{s_1}{1 - s_1 + s_2} \rho_{j-\frac{1}{2},L} + \frac{s_2 - 1}{1 - s_1 + s_2} \rho_{j-\frac{1}{2},R} \right). \tag{4.18}$$

Following §4, we define the flux generated by the left boundary $J_{j-\frac{1}{2},L}$ as the flux obtained for $E = E_{j-\frac{1}{2},L}$, such that $\rho_{j-\frac{1}{2},L} = \rho_{j-1} = \omega_1 (f_{j-1}(v_1) + f(0, -v_1))$ and $\rho_{j-\frac{1}{2},R} = 0$. Respectively, we define the flux generated by the right boundary $J_{j-\frac{1}{2},R}$ as the one obtained for $E = E_{j-\frac{1}{2},R}$, such that $\rho_{j-\frac{1}{2},R} = \rho_j = \omega_1 (f(\Delta x, v_1) + f_j(-v_1))$ and $\rho_{j-\frac{1}{2},L} = 0$. Replacing s_1 and s_2 by their respective values in (4.18) yield,

$$J_{j-\frac{1}{2},L} = v_1 \omega_1 \cdot c_L \rho_{j-1}, \qquad J_{j-\frac{1}{2},R} = v_1 \omega_1 \cdot c_R \rho_j, \tag{4.19}$$

where c_L and c_R are given in (4.14). The total flux reads $J_{j-1/2} = J_{j-1/2,L} + J_{j-1/2,R}$, so that $(f(0, -v_1), f(\Delta x, v_1))$ solve the following 2×2 system:

$$\begin{aligned} f(\Delta x, v_1) - f_j(-v_1) &= f_{j-1}(v_1) - f(0, -v_1), \\ v_1 \omega_1 \cdot (f^+(\Delta x) - f_j^-) &= J_{j-1/2,L} + J_{j-1/2,R}. \end{aligned}$$

The first equation expresses flux conservation. Recombining and using (4.19),

$$\begin{aligned} f(\Delta x, v_1) + f(0, -v_1) &= f_{j-1}(v_1) + f_j(-v_1), \\ f(\Delta x, v_1) - f_j(-v_1) &= c_L \big(f_j(-v_1) + f(\Delta x, v_1) \big) + c_R \big(f_{j-1}(v_1) + f(0, -v_1) \big). \end{aligned}$$

This is the same system as the one in the proof of Lemma 4.1 so we are done. □

4.4. Discretization of Vol'pert product in (VPFP). The scope is now the discretization of the macroscopic force fields E_L and E_R, as used in Lemma 4.2. Again, it heavily relies on Vol'pert calculus so as to handle correctly the nonconservative product ρE at the singular limit. By definition (4.16),

$$\rho_{j-\frac{1}{2},L} = \omega_1 \cdot (f_{j-1}(v_1) + f(0, -v_1)), \qquad \rho_{j-\frac{1}{2},R} = \omega_1 \cdot (f_j(-v_1) + f(\Delta x, v_1)).$$

Using the scattering matrix (4.4) to compute $f(0, -v_1)$ and $f(\Delta x, v_1)$, it comes

$$\begin{aligned} \rho_{j-\frac{1}{2},L} &= \omega_1 \left(\frac{2(1 - c_R)}{1 - c_R + c_L} f_{j-1}(v_1) - \frac{2 c_R}{1 - c_R + c_L} f_j(-v_1) \right), \\ \rho_{j-\frac{1}{2},R} &= \omega_1 \left(\frac{2 c_L}{1 - c_R + c_L} f_{j-1}(v_1) - \frac{2(1 + c_L)}{1 - c_R + c_L} f_j(-v_1) \right), \end{aligned} \tag{4.20}$$

so that potentials $\phi_{j-\frac{1}{2},L}$ and $\phi_{j-\frac{1}{2},R}$ follow by solving finite-differences:

$$-\frac{\phi_{j+\frac{1}{2},L/R} - 2\phi_{j-\frac{1}{2},L/R} + \phi_{j-\frac{3}{2},L/R}}{\Delta x^2} = \rho_{j-\frac{1}{2},L/R}, \qquad \text{for } j \in \mathbb{Z}. \tag{4.21}$$

Finally, our discretization of the electric field is defined by

$$E_{j-\frac{1}{2},L/R} = \frac{1}{2}(E_{j-1,L/R} + E_{j,L/R}), \quad \text{where } E_{j,L/R} = \frac{\phi_{j+\frac{1}{2},L/R} - \phi_{j-\frac{1}{2},L/R}}{\Delta x}. \tag{4.22}$$

In summary, knowing $f_{j-1}(v_1)$ and $f_j(-v_1)$, the discrete force fields E_L and E_R are computed by solving the nonlinear problem (4.14)-(4.20)-(4.21)-(4.22). Finally,

$$\begin{pmatrix} f_j^{n+1}(v_1) \\ f_{j-1}^{n+1}(-v_1) \end{pmatrix} = (1 - v_1 \frac{\Delta t}{\Delta x}) \begin{pmatrix} f_j^n(v_1) \\ f_{j-1}^n(-v_1) \end{pmatrix} + v_1 \frac{\Delta t}{\Delta x}\, \mathcal{S}_{j-1/2}^{\Delta x} \begin{pmatrix} f_{j-1}^n(v_1) \\ f_j^n(-v_1) \end{pmatrix}, \tag{4.23}$$

where $\mathcal{S}_{j-1/2}^{\Delta x}$ is given in Lemma 4.2. Obviously, the CFL restriction of Lemma 3.1 should be changed to $v_1 \Delta t \leq \Delta x$.

5. Consistency and asymptotic properties.

5.1. Hydrodynamic behavior (AP property) of (1.1). Standard notation for both the macroscopic density and flux at the cell's center reads

$$\forall j, n \in \mathbb{Z} \times \mathbb{N}, \qquad \rho_j^n = f_{j,n}^+ + f_{j,n}^-, \qquad J_j^n = f_{j,n}^+ - f_{j,n}^-.$$

PROPOSITION 5.1. *In (5.2) the flux $J_{j+1/2}^n$ satisfies, as $\varepsilon \to 0$: $\forall\, j, n \in \mathbb{Z} \times \mathbb{N}$,*

$$\boxed{J_{j-1/2}^n \to \max\left(0, a_{j-\frac{1}{2},L}^n\right) \rho_{j-\frac{1}{2},L}^n + \min\left(0, a_{j-\frac{1}{2},R}^n\right) \rho_{j-\frac{1}{2},R}^n\,.} \tag{5.1}$$

Proof. Notice that we get from (4.8)

$$f_{j-\frac{1}{2},n}^+ - f_{j,n}^- = J_{j-1/2}^n = f_{j-1,n}^+ - f_{j-\frac{1}{2},n}^-,$$

that can be inserted into (3.5) in order to produce,

$$\begin{array}{rcl} f_{j,n+1}^+ & = & f_{j,n}^+ - \frac{\Delta t}{\Delta x}\big(f_{j,n}^+ - f_{j,n}^- - J_{j-1/2}^n\big) \\ f_{j-1,n+1}^- & = & f_{j-1,n}^- - \frac{\Delta t}{\Delta x}\big(f_{j-1,n}^- - f_{j-1,n}^+ + J_{j-1/2}^n\big) \end{array}$$

Adding both lines, the second being shifted at x_j (instead of x_{j-1}), we get :

$$\rho_j^{n+1} = \rho_j^n - \frac{\Delta t}{\Delta x}\big(J_{j+1/2}^n - J_{j-1/2}^n\big), \tag{5.2}$$

where, using (4.8),

$$J_{j-\frac{1}{2}}^n = c_L(f_{j-1,n}^+ + f_{j-\frac{1}{2}}^-) + c_R(f_{j-\frac{1}{2}}^+ + f_{j,n}^-).$$

According to the definitions (4.9)-(4.10), this rewrites

$$J_{j-\frac{1}{2}}^n = c_L\, \rho_{j-\frac{1}{2},L}^n + c_R\, \rho_{j+\frac{1}{2},R}^n\,.$$

By (4.5), as $\varepsilon \to 0$, $c_R \to \min(0, a_{j-\frac{1}{2},R})$ and $c_L \to \max(0, a_{j-\frac{1}{2},L})$. □

The asymptotic scheme (5.1)–(5.2) is reminiscent of the upwind schemes studied in [24] (see §4.2): conventional (centered) discretizations like the ones used in [31] yield a more diffusive Lax-Friedrichs scheme in the hydrodynamical limit. Our discretization of the macroscopic velocity using Vol'pert calculus is consistent with nonconservative products as in Definition 2. Indeed, both (4.11) and (4.12) yield,

$$a^n_{j-\frac{1}{2},L}\,\rho^n_{j-\frac{1}{2},L} = -\frac{A(\partial_x S^n_{j,L}) - A(\partial_x S^n_{j-1,L})}{\Delta x} + a^n_{j-\frac{1}{2},L} S^n_{j-\frac{1}{2},L},$$

and the analogue for $a^n_{j-\frac{1}{2},R}\,\rho^n_{j-\frac{1}{2},R}$. In other words, this definition of the macroscopic velocity is a discretization of the Vol'pert calculus (chain rule for BV functions) :

$$-\partial_x(A(\partial_x S_L)) = \widehat{a}_L(\rho_L - S_L), \qquad \text{with} \qquad \widehat{a}_L = a(\partial_x S) \quad \text{a.e.} \ ,$$

where we systematically use the fact that $-\partial_{xx} S_L + S_L = \rho_L$. If $\rho(t,\cdot)$ is continuous, left/right values $a^n_{j-\frac{1}{2},L/R}$ are very close to each other, so this construction is consistent with a standard finite-differences approximation $a^n_{j-\frac{1}{2}} = \phi(\frac{S^n_j - S^n_{j-1}}{\Delta x})$.

5.2. Burgers-Hopf limit of (VPFP). Since the scattering matrix is very similar, computations in §5.1 with $\rho_j = \omega_1 \cdot (f_j(v_1) + f_j(-v_1))$, ω_1 being the weight associated to v_1 in e.g. (4.15) or (4.16), yield now

$$\rho^{n+1}_j = \rho^n_j - \frac{\Delta t}{\Delta x}(J^n_{j+\frac{1}{2}} - J^n_{j-\frac{1}{2}}), \qquad J^n_{j-\frac{1}{2}} = c_L \rho^n_{j-\frac{1}{2},L} + c_R \rho^n_{j-\frac{1}{2},R}.$$

From (4.14) a suffucient condition for consistency to hold is $v_1\theta/\kappa \ll 1$. Accordingly, when $\varepsilon \to 0$,

$$c_L \to \max\left(0, \frac{\tanh(\frac{v_1\theta}{\kappa} E_{j-\frac{1}{2},L})}{\omega_1}\right) \ \text{ and } \ c_R \to \min\left(0, \frac{\tanh(\frac{v_1\theta}{\kappa} E_{j-\frac{1}{2},R})}{\omega_1}\right),$$

and this leads to an asymptotic behavior similar to the one in Proposition 5.1.

Remark 1. *Thanks to this discretization of the field,* (4.22)-(4.21) *imply that*

$$\boxed{E_{j-\frac{1}{2},L/R} \cdot \rho_{j-\frac{1}{2},L/R} = \frac{1}{2\Delta x}\Big(\big(E_{j,L/R}\big)^2 - \big(E_{j-1,L/R}\big)^2\Big),}$$

the right hand side being a discretization of $\frac{1}{2}|\partial_x E|^2$. *Hence our definition of the discrete velocity field is consistent with the product* ρE *as defined in Theorem 2. Moreover, it correctly handles that product even if* ρ *contains Dirac atoms.*

6. Hybrid scheme mixing scattering and time-splitting. As the previous scheme's CPU cost mainly comes from solving the "nonlinear system" (4.5)-(4.9)-(4.10)-(4.11)-(4.12)-(4.13), we now follow [22] and select a discretization of the macroscopic velocity such that $a^n_{j-\frac{1}{2}}$ depends only on a sole interface value $\rho^n_{j-\frac{1}{2}}$. A time-splitting pre-processing, when applied to VPFP system leads to macroscopic equations, where κ disappeared while integrating in the v variable,

$$\partial_t \rho = 0, \qquad \partial_t J = \frac{1}{\theta\varepsilon}(\theta\rho E - J).$$

6.1. A simpler handling of the nonconservative product. To proceed, we first propose a discretization based on the Vol'pert calculus by setting :

$$\forall j \in \mathbb{Z}, \qquad a^n_{j-\frac{1}{2}} = \frac{A(\partial_x S^n_j) - A(\partial_x S^n_{j-1})}{\partial_x S^n_j - \partial_x S^n_{j-1}}, \qquad \text{if } \partial_x S^n_j \neq \partial_x S^n_{j-1}, \tag{6.1}$$

where the centered divided difference reads $\partial_x S^n_j = (S^n_{j+\frac{1}{2}} - S^n_{j-\frac{1}{2}})/\Delta x$. The potential $S^n_{j-\frac{1}{2}}$ being solution of the discretized elliptic problem,

$$\forall n \in \mathbb{N}, \qquad -\frac{S^n_{j+\frac{1}{2}} - 2S^n_{j-\frac{1}{2}} + S^n_{j-\frac{3}{2}}}{\Delta x^2} + S^n_{j-\frac{1}{2}} = \rho^n_{j-\frac{1}{2}}. \tag{6.2}$$

From (6.1)-(6.2), we deduce that

$$(a \cdot \rho)^n_{j-\frac{1}{2}} = -\frac{A(\partial_x S^n_j) - A(\partial_x S^n_{j-1})}{\Delta x} + a^n_{j-\frac{1}{2}} S^n_{j-\frac{1}{2}}. \tag{6.3}$$

This value is consistent with the non-conservative product $a \cdot \rho$ when the density contains Dirac atoms. However, the "interface value" $\rho^n_{j-\frac{1}{2}}$ must be made explicit.

6.2. A time-splitting (TS) pre-processing procedure. Hereafter we shall use the shorthand notation, $\mathcal{L}(f^\pm) = \big(1+a\big)f^- - \big(1-a\big)f^+$. Moreover, for $0 < \varepsilon \ll 1$ small enough, we define two complementary parameters ε_{WB} and ε_{TS} such that,

$$\frac{1}{\varepsilon} = \frac{1}{\varepsilon_{WB}} + \frac{1}{\varepsilon_{TS}}, \qquad \varepsilon_{TS} = \frac{\varepsilon \cdot \varepsilon_{WB}}{\max(0, \varepsilon_{WB} - \varepsilon)} \in (0, +\infty].$$

Accordingly the kinetic system (1.1) is treated by decomposing the collision term into

$$\partial_t f^\pm \mp \frac{\mathcal{L}(f^\pm)}{\varepsilon_{TS}} = \mp \partial_x f^\pm \pm \frac{\mathcal{L}(f^\pm)}{\varepsilon_{WB}} \tag{6.4}$$

The left part of (6.4) is intended to be handled by time-splitting; the right part, by applying a scattering matrix as in the previous section. At each interface $x_{j-\frac{1}{2}}$, its "incoming states" are modified by means of the mass-preserving differential equation,

$$\frac{d}{dt}\widetilde{f}^\pm(t) = \pm\frac{1}{\varepsilon_{TS}}\mathcal{L}\big(\widetilde{f}^\pm(t)\big), \qquad x \in (x_{j-1}, x_j), t \in (t^n, t^{n+1}).$$

The unknowns of this system are the distribution functions (f^+_{j-1}, f^-_j) entering into (x_{j-1}, x_j) (or equivalently, approaching in the sense of Glimm the discontinuity at $x_{j-\frac{1}{2}}$), so it rewrites as, (we now drop the time index n since there is no ambiguity)

$$\frac{d\widetilde{f}^+_{j-1}(t)}{dt} = \frac{1}{2\varepsilon_{TS}}\big(a_{j-\frac{1}{2}}\widetilde{\rho}_{j-\frac{1}{2}} - \widetilde{J}_{j-\frac{1}{2}}\big), \qquad \frac{d\widetilde{f}^-_j(t)}{dt} = -\frac{1}{2\varepsilon_{TS}}\big(a_{j-\frac{1}{2}}\widetilde{\rho}_{j-\frac{1}{2}} - \widetilde{J}_{j-\frac{1}{2}}\big),$$

where we switched to the macroscopic variables,

$$\widetilde{\rho}_{j-\frac{1}{2}}(t) = \widetilde{f}^+_{j-1}(t) + \widetilde{f}^-_j(t), \qquad \text{and} \qquad \widetilde{J}_{j-\frac{1}{2}}(t) = \widetilde{f}^+_{j-1}(t) - \widetilde{f}^-_j(t). \tag{6.5}$$

Adding and taking the difference, the later system is equivalent to

$$\frac{d}{dt}\widetilde{\rho}_{j-\frac{1}{2}}(t) = 0, \qquad \frac{d}{dt}\widetilde{J}_{j-\frac{1}{2}}(t) = \frac{1}{\varepsilon_{TS}}\Big(a_{j-\frac{1}{2}} \cdot \widetilde{\rho}_{j-\frac{1}{2}} - \widetilde{J}_{j-\frac{1}{2}}(t)\Big),$$

so $\widetilde{\rho}_{j-\frac{1}{2}}$ is a constant, implying that $a_{j-\frac{1}{2}}$ is a constant, too. Then we can solve exactly the second equation and get

$$\widetilde{J}_{j-\frac{1}{2}}(t+\Delta t) = a_{j-\frac{1}{2}}\widetilde{\rho}_{j-\frac{1}{2}} + \big(\widetilde{J}_{j-\frac{1}{2}}(t) - a_{j-\frac{1}{2}}\widetilde{\rho}_{j-\frac{1}{2}}\big)e^{-\Delta t/\varepsilon_{TS}}.$$

So, $\widetilde{f}^+_{j-1} := \widetilde{f}^+_{j-1}(t+\Delta t)$ and $\widetilde{f}^-_j := \widetilde{f}^+_j(t+\Delta t)$ solve a linear algebraic system:

$$\begin{cases} \widetilde{f}^+_{j-1} + \widetilde{f}^-_j &= f^+_{j-1} + f^-_j, \\ \widetilde{f}^+_{j-1} - \widetilde{f}^-_j &= a_{j-\frac{1}{2}}(f^+_{j-1} + f^-_j) + \big(f^+_{j-1} - f^-_j - a_{j-\frac{1}{2}}(f^+_{j-1} + f^-_j)\big)e^{-\Delta t/\varepsilon_{TS}}. \end{cases}$$

After straightforward computations, a matrix formulation comes out:

$$\begin{pmatrix} \widetilde{f}^+_{j-1} \\ \widetilde{f}^-_j \end{pmatrix} = O^{\Delta t}_{j-\frac{1}{2}} \begin{pmatrix} f^+_{j-1} \\ f^-_j \end{pmatrix}, \tag{6.6}$$

where we used the notation,

$$O^{\Delta t}_{j-\frac{1}{2}} = \frac{1}{2}\begin{pmatrix} 1 + a_{j-\frac{1}{2}} + (1 - a_{j-\frac{1}{2}})e^{-\Delta t/\varepsilon_{TS}} & (1 + a_{j-\frac{1}{2}})\big(1 - e^{-\Delta t/\varepsilon_{TS}}\big) \\ (1 - a_{j-\frac{1}{2}})\big(1 - e^{-\Delta t/\varepsilon_{TS}}\big) & 1 - a_{j-\frac{1}{2}} + (1 + a_{j-\frac{1}{2}})e^{-\Delta t/\varepsilon_{TS}} \end{pmatrix}.$$

Observe that when $\varepsilon_{TS} \to 0$, it simply reduces to,

$$\widetilde{f}^+_{j-1} = \frac{1 + a_{j-\frac{1}{2}}}{2}\big(f^+_{j-1} + f^-_j\big), \qquad \widetilde{f}^-_j = \frac{1 - a_{j-\frac{1}{2}}}{2}\big(f^+_{j-1} + f^-_j\big). \tag{6.7}$$

6.3. Complementary well-balanced (WB) discretization. In a second step, we consider the right hand side part of (6.4) and solve

$$\partial_t f^\pm = \mp\partial_x f^\pm \pm \frac{\mathcal{L}(f^\pm)}{\varepsilon_{WB}}$$

We proceed in the same way as in §4.1 so we will not repeat similar arguments. Stationary equations at each interface $x_{j-\frac{1}{2}}$, with incoming boundary conditions on (x_{j-1}, x_j) read, in macroscopic variables,

$$\varepsilon\partial_x\rho = a_{j-\frac{1}{2}}\rho - J, \qquad \rho = f^+ + f^-, \qquad J = f^+ - f^-.$$

As usual, such an equation can be solved explicitly,

$$J = a_{j-\frac{1}{2}}\Big(\frac{\rho_{j-1} - e^{-a_{j-\frac{1}{2}}\Delta x/\varepsilon_{WB}}\rho_j}{1 - e^{-a_{j-\frac{1}{2}}\Delta x/\varepsilon_{WB}}}\Big). \tag{6.8}$$

Notations are now adapted to a definition of $a_{j-\frac{1}{2}}$ continuous across the interface:

$$c_R = \frac{a_{j-\frac{1}{2}}}{1 - e^{a_{j-\frac{1}{2}}\Delta x/\varepsilon_{WB}}}, \qquad c_L = \frac{a_{j-\frac{1}{2}}}{1 - e^{-a_{j-\frac{1}{2}}\Delta x/\varepsilon_{WB}}}, \tag{6.9}$$

notice the difference with (4.5). Recombining, we derive a formulation similar to (3.4),

$$\begin{pmatrix} f^+_{j,L} \\ f^-_{j-1,R} \end{pmatrix} = \begin{pmatrix} f^+(\Delta x) \\ f^-(0) \end{pmatrix} = \mathcal{S}^{\Delta x}_{j-\frac{1}{2}} \begin{pmatrix} f^+_{j-1} \\ f^-_j \end{pmatrix},$$

where, with the notations (4.5), the scattering matrix $\mathcal{S}^{\Delta x}_{j-\frac{1}{2}}$ is identical to $\mathcal{S}_{j-\frac{1}{2}}$, which expression is given in (4.4), so the Godunov scheme is a slight variation of (3.5),

$$\begin{pmatrix} f_j^+ \\ f_{j-1}^- \end{pmatrix}^{n+1} = \begin{pmatrix} \widetilde{f}_j^+ \\ \widetilde{f}_{j-1}^- \end{pmatrix}^n - \frac{\Delta t}{\Delta x}\left[\begin{pmatrix} \widetilde{f}_j^+ \\ \widetilde{f}_{j-1}^- \end{pmatrix}^n - \mathcal{S}^{\Delta x}_{j-\frac{1}{2}}\begin{pmatrix} \widetilde{f}_{j-1}^+ \\ \widetilde{f}_j^- \end{pmatrix}^n\right]. \tag{6.10}$$

From (6.9), one sees that, when $\varepsilon_{WB} \to 0$,

$$c_R \to \min(0, a_{j-\frac{1}{2}}) \le 0 \qquad \text{and} \qquad c_L \to \max(0, a_{j-\frac{1}{2}}) \ge 0,$$

and so the scheme (6.10) becomes, with $\widetilde{f}^{\pm}$ given by (6.6),

$$\left\{\begin{array}{rcl} f_j^{+,n+1} & = & \widetilde{f}_j^{+,n} - \frac{\Delta t}{\Delta x}\big(\widetilde{f}_j^{+,n} - \widetilde{f}_j^{-,n}\big) + \frac{\Delta t}{\Delta x}\Big(\frac{2c_R}{1-c_R+c_L}\widetilde{f}_j^{-,n} + \frac{2c_L}{1-c_R+c_L}\widetilde{f}_{j-1}^{+,n}\Big), \\ f_{j-1}^{-,n+1} & = & \widetilde{f}_{j-1}^{-,n} - \frac{\Delta t}{\Delta x}\big(\widetilde{f}_{j-1}^{-,n} - \widetilde{f}_{j-1}^{+,n}\big) - \frac{\Delta t}{\Delta x}\Big(\frac{2c_R}{1-c_R+c_L}\widetilde{f}_j^{-,n} + \frac{2c_L}{1-c_R+c_L}\widetilde{f}_{j-1}^{+,n}\Big). \end{array}\right. \tag{6.11}$$

6.4. Asymptotic behavior of the resulting scheme as $\varepsilon \to 0$. Let,

$$\widetilde{J}^n_{j-\frac{1}{2}} = \frac{2c_R}{1-c_R+c_L}\widetilde{f}_j^{-,n} + \frac{2c_L}{1-c_R+c_L}\widetilde{f}_{j-1}^{+,n}.$$

PROPOSITION 6.1. *When both $\varepsilon_{WB} \to 0$ and $\varepsilon_{TS} \to 0$, $\forall j, n \in \mathbb{Z} \times \mathbb{N}$,*

$$\rho^{n+1}_{j-\frac{1}{2}} = \big(1 - \frac{\Delta t}{\Delta x}\big)\rho^n_{j-\frac{1}{2}} + \frac{\Delta t}{2\Delta x}\big(\rho^n_{j-3/2} + \rho^n_{j+\frac{1}{2}}\big) - \frac{\Delta t}{2\Delta x}\big(a^n_{j+\frac{1}{2}}\rho^n_{j+\frac{1}{2}} - a^n_{j-3/2}\rho^n_{j-3/2}\big). \tag{6.12}$$

Proof. Letting $\varepsilon_{WB} \to 0$ we deduce from the definitions (6.9) that

$$\widetilde{J}^n_{j-\frac{1}{2}} \to \frac{2\min(0, a_{j-\frac{1}{2}})}{1 - \min(0, a_{j-\frac{1}{2}}) + (\max(0, a_{j-\frac{1}{2}})}\widetilde{f}_j^{-,n} + \frac{2\max(0, (a_{j-\frac{1}{2}})}{1 - \min(0, a_{j-\frac{1}{2}}) + \max(0, (a_{j-\frac{1}{2}})}\widetilde{f}_{j-1}^{+,n}. \tag{6.13}$$

Then, letting $\varepsilon_{TS} \to 0$ we use (6.7) to pass to the limit in (6.13),

$$\widetilde{J}^n_{j-\frac{1}{2}} \to \min(0, a_{j-\frac{1}{2}})\rho^n_{j-\frac{1}{2}} + \max(0, a_{j-\frac{1}{2}})\rho^n_{j-\frac{1}{2}} = a_{j-\frac{1}{2}}\rho^n_{j-\frac{1}{2}}. \tag{6.14}$$

From (6.11), we deduce that at the limit $\varepsilon \to 0$,

$$\begin{array}{rcl} f_j^{+,n+1} & = & \Big(1 - \frac{\Delta t}{\Delta x}\Big)\widetilde{f}_j^{+,n} + \frac{\Delta t}{\Delta x}\widetilde{f}_j^{-,n} + \frac{\Delta t}{\Delta x}a_{j-\frac{1}{2}}\rho^n_{j-\frac{1}{2}}, \\ f_{j-1}^{-,n+1} & = & \Big(1 - \frac{\Delta t}{\Delta x}\Big)\widetilde{f}_{j-1}^{-,n} + \frac{\Delta t}{\Delta x}\widetilde{f}_{j-1}^{+,n} - \frac{\Delta t}{\Delta x}a_{j-\frac{1}{2}}\rho^n_{j-\frac{1}{2}}, \end{array}$$

and using (6.7),

$$\begin{array}{rcl} f_j^{+,n+1} & = & \Big(1 - \frac{\Delta t}{\Delta x}\Big)\widetilde{f}_j^{+,n} + \frac{\Delta t}{2\Delta x}\big(1 + a_{j-\frac{1}{2}}\big)\rho^n_{j-\frac{1}{2}}, \\ f_{j-1}^{-,n+1} & = & \Big(1 - \frac{\Delta t}{\Delta x}\Big)\widetilde{f}_{j-1}^{-,n} + \frac{\Delta t}{2\Delta x}\big(1 - a_{j-\frac{1}{2}}\big)\rho^n_{j-\frac{1}{2}}. \end{array}$$

It easily follows that,

$$\begin{array}{rcl} f_{j-1}^{+,n+1} & = & \Big(1 - \frac{\Delta t}{\Delta x}\Big)\widetilde{f}_{j-1}^{+,n} + \frac{\Delta t}{2\Delta x}\big(1 + a_{j-3/2}\big)\rho^n_{j-3/2}, \\ f_j^{-,n+1} & = & \Big(1 - \frac{\Delta t}{\Delta x}\Big)\widetilde{f}_j^{-,n} + \frac{\Delta t}{2\Delta x}\big(1 - a_{j+\frac{1}{2}}\big)\rho^n_{j+\frac{1}{2}}. \end{array}$$

Summing these equalities and recalling $\widetilde{\rho}^n_{j-\frac{1}{2}} = \rho^n_{j-\frac{1}{2}}$, gives the conclusion. □

We recognize in (6.12) a Lax-Friedrichs discretization of the aggregation equation, on a staggered grid, though, since it acts on the nodal variables $\rho^n_{j+\frac{1}{2}}$ as defined in (6.5). From these discrete data (on each staggered interval (x_{j-1}, x_j)), a piecewise constant approximation is built for which convergence toward the unique duality solution of the aggregation equation was proved in [31], as $\Delta t \leq \Delta x \to 0$.

7. Numerical validations and benchmarks. In this Section, practical comparisons between previously introduced schemes for the simulation of both systems (1.1)–(1.2) and (1.6) are achieved.

7.1. Concentrations in the two-stream chemotaxis model. The problem is posed on a computational domain $x \in (-2.5, 2.5)$ with an uniform Cartesian grid of 2^8 nodes (the CFL restriction is taken as $\Delta t = 0.9\Delta x$), $\Delta x \simeq 0.02$. A regularized, but sharply varying, Heaviside-type function reads $a(x) = 2\phi(x) = \frac{2}{\pi}\arctan(50x)$: it satisfies the sub-characteristic condition (1.3), too. Initial data was chosen as,

$$\rho^0(x) = e^{-10(x-1.25)^2} + 0.8e^{-20x^2} + e^{-10(x+1)^2}.$$

The kinetic equation (1.1) is complemented with specular reflection, whereas the elliptic equation (1.2), with Neumann (zero-flux) boundary conditions. These parameters correspond to the ones in [31], where a centered time-splitting scheme was set up, that we intend to compare to ours. Before going into details, one may question the

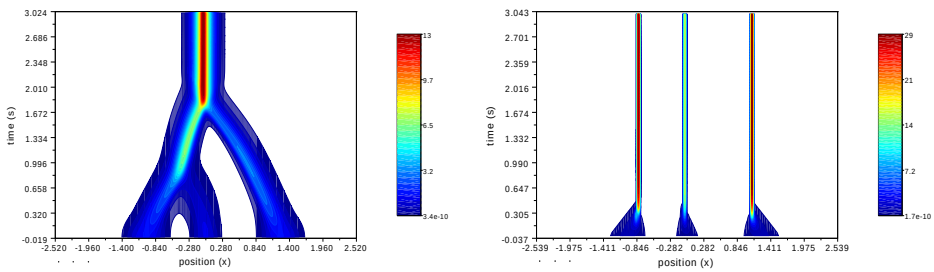

FIG. 7.1. *Macroscopic density ρ with $\varepsilon = 0.1$ (left), $\varepsilon = 0.005$ (right): aggregates' numerical velocity are flawed when $\varepsilon \ll \Delta x$ without a specific treatment.*

importance of the careful treatment of nonconservative products that is advocated. Especially (see e.g. [41]) the kinetic equation, for bounded tumbling kernel, propagates in time the smoothness of its initial data. Hence, at the kinetic level, all quantities involved are well-defined and there seems to be no *nonconservative product*. However, to illustrate the importance of having a good definition of the macroscopic velocity at the numerical level, even in this kinetic realm, the algorithm of §4 is implemented with $a^n_{j-1/2,L} = a^n_{j-1/2,R} = a\big((S^n_{j+1} - S^n_j)/\Delta x\big)$: numerical results are displayed in Figure 7.1. As previously observed by Twarogowska on a related model [40], the dynamics of ρ are correct for $\varepsilon = 0.1$ (left) but spurious for $\varepsilon = 0.005$ (right). Indeed, for $\varepsilon = 0.1$, the macroscopic density forms a single aggregate in finite time. Oppositely, when $\varepsilon = 0.005$, 3 aggregates quickly appear, but remain static and neither interact nor glue altogether, as expected. In particular, there seems to be non continuity as ε is decreased so that $\varepsilon \ll \Delta x$. Even in a kinetic setup, the concentration process

is so strong that, without an appropriate (numerical) handling of nonconservative products, numerical results are flawed, unless $\Delta x \ll \varepsilon$, an unacceptable restriction. This is a consequence of the non-uniqueness of solution for the aggregation equation when the nonconservative product is not defined correctly (we refer the reader to [30, Section 5] where counter examples of uniqueness are given).

7.1.1. Comparison with centered time-splitting scheme. Numerical results obtained by means of the process described in §4 are now compared to the ones obtained by the time-splitting scheme in [31], see Fig. 7.2 for $\varepsilon = 10^{-5}$. The top line of Fig. 7.2 displays the outcome of [31] with same parameters. As ex-

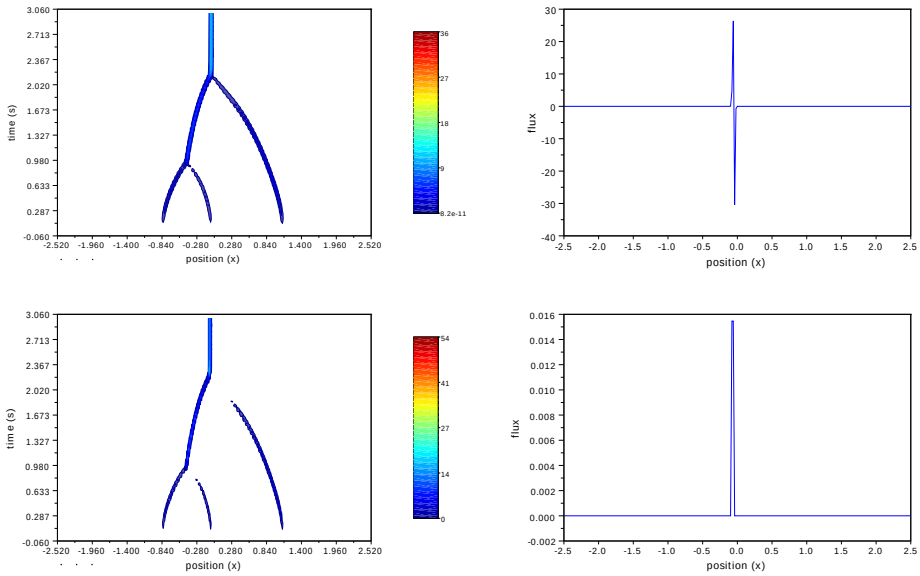

FIG. 7.2. *Macroscopic ρ (left) and J (right), $\varepsilon = 10^{-5}$: time-split (top) and WB (bottom).*

pected, the behavior of ρ is (roughly) the same for both numerical schemes. The 3 initial bumps quickly collapse into numerical Dirac masses. Then, dynamics after blowup see these 3 atoms aggregate into one. As noticed in [31] (see also [8] in several dimensions), this suggests that two different time-scales co-exist: a first one corresponding to initial blowups into individual atoms, and a second one, to these atoms interacting with each other. Let us stress that our new algorithm produces a stronger aggregation numerical process, as the maximal value of ρ is 54 instead of 36 for centered time-splitting. Moreover, a big difference shows up in the macroscopic flux J: at equilibrium, a comparison is displayed in Fig. 7.2 (right). For the time-splitting algorithm, it varies strongly across the Dirac mass, whereas the order of magnitude of the WB flux is $\simeq 10^{-2}$. Yet, taking mirrored initial data, *i.e.* $\rho^0(x) = e^{-10(x+1.25)^2} + 0.8e^{-20x^2} + e^{-10(x-1)^2}$, produces just opposite values (see Fig. 7.3). Other initial data, $\rho^0(x) = \frac{1}{2}(e^{-10(x-0.7)^2} + e^{-10(x+0.7)^2})$, allows to generate an equilibrium endowed with a macroscopic flux with an order of magnitude of 10^{-10} (see Fig. 7.4), which is remarkable. Boundary conditions probably have an influence on J, even if the Dirac mass stabilizes far from the edges of the domain.

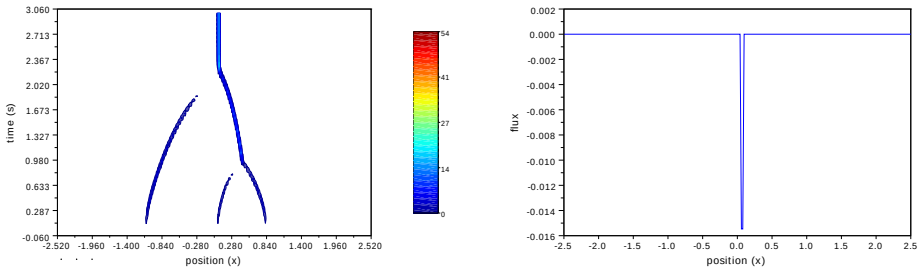

FIG. 7.3. *Macroscopic density ρ (left) and flux J (right): mirror data and opposite fluxes.*

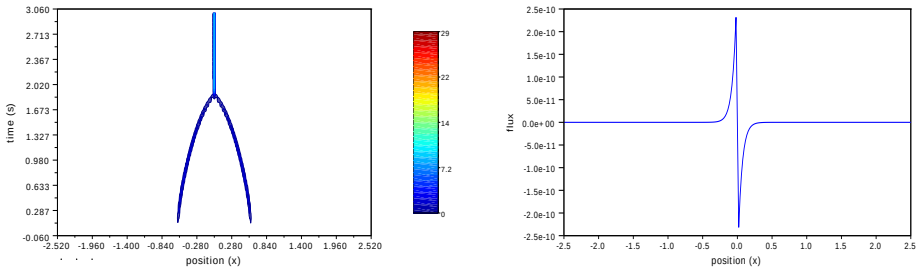

FIG. 7.4. *Macroscopic density ρ (left) and flux J (right): WB scheme and symmetric data.*

7.1.2. Hybrid scheme vs. time-splitting scheme. Figure 7.5 displays a comparison between the hybrid scheme of §4 and the time-splitting one of [31]. Both algorithms recover a Lax-Friedrichs kind of discretization when $\varepsilon \to 0$. Then numerical diffusion is stronger than in previous subsection. Although the equilibrium flux is piked with both schemes, numerical values for the hybrid algorithm are twice smaller. Its CPU cost is lighter since no iterative resolution of a nonlinear problem involved. More precisely the CPU time to obtain Fig. 7.5 (bottom) is 15 s, whereas the CPU time to run the simulation on SciLab software of Fig. 7.2 (bottom) is 65.5 s. However, even if we recover a similar dynamics for the macroscopic density, the hybrid scheme cannot generate macroscopic fluxes of such high quality as in Fig. 7.4.

7.2. Numerical results for VPFP model. The computational domain in space is now $x \in (-4, 4)$ gridded with 2^8 nodes; for the discretization of the velocity domain, a Gauss-Hermite quadrature rule is chosen, $v_1 = \sqrt{\kappa}$ with a weight $\omega_1 = \sqrt{\kappa\pi/2}$. The time-step is such that $v_1 \Delta t = 0.95\Delta x$, $\Delta x = 1/32$. Finally, $\theta = 1$, $\kappa = 0.5$ and the same iterative procedure (as described in Appendix A) to solve iteratively the nonlinear problem is used. Fig. 7.6 displays the results obtained with the initial data $\rho^0(x) = \frac{1}{2}(e^{-5(x-2)^2} + e^{-5(x+2)^2})$ for $\varepsilon = 0.1$ (top) and for $\varepsilon = 10^{-5}$ (bottom). The order of magnitude of the flux is 10^{-10}, which is negligible; this confirms that the well-balanced property is numerically satisfied, even in presence of a Dirac atom in ρ. Theorem 2 states that, if $\varepsilon \ll 1$, the field E approaches a solution of Burgers-Hopf equation (2.5). Accordingly, to numerically check this property, it is interesting to compare numerical results obtained with our algorithm with the outcome of a standard Godunov scheme on the Burgers law. Such a comparison is provided in Fig. 7.7, where 4 successive snapshots for $\varepsilon = 10^{-5} \ll \Delta x$ are plotted: a

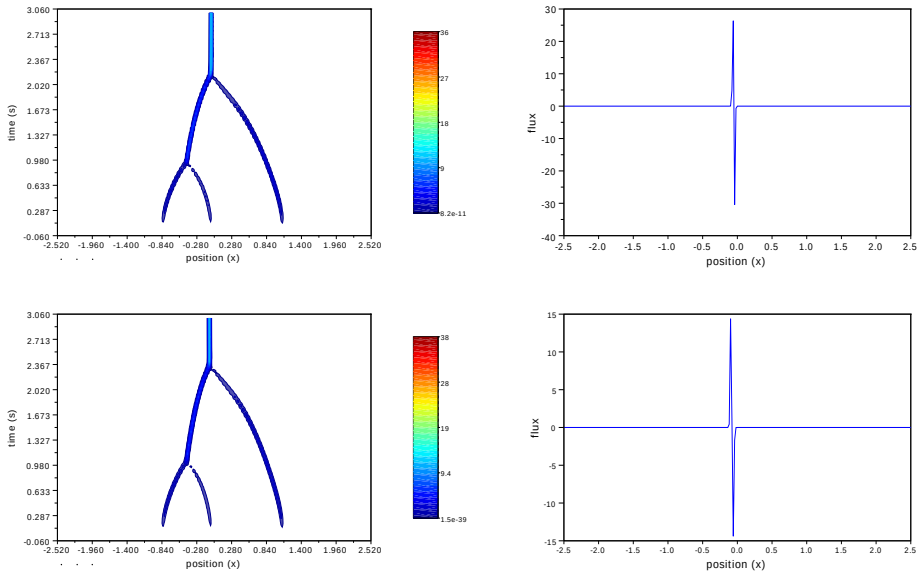

FIG. 7.5. *Macroscopic densities ρ (left) and fluxes J (right): centered (top), hybrid (bottom).*

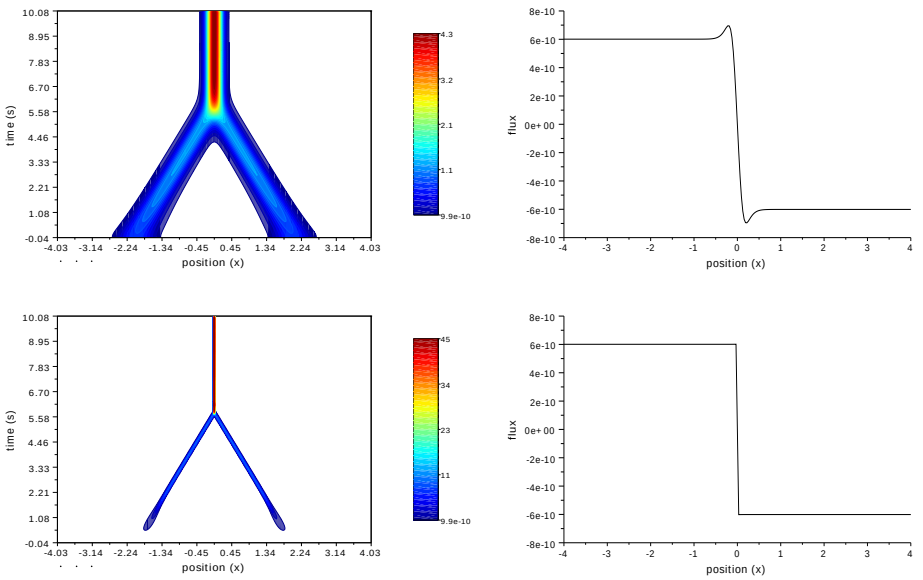

FIG. 7.6. *Macroscopic densities ρ (left), fluxes J (right) for $\varepsilon = 0.1$ (top), $\varepsilon = 10^{-5}$ (bottom).*

satisfying behavior is observed.

Finally, similar results are displayed in Fig. 7.8 and 7.9 where we chose a non-symmetric initial data: $\rho^0(x) = 0.75e^{-5(x-1.5)^2} + 0.4e^{-5(x+0.5)^2} + 0.85e^{-5(x+2.5)^2}$. In this numerical simulation we have taken $\varepsilon = 10^{-5}$. The order of magnitude of the flux at final time in Fig. 7.8-right is 10^{-6} which is very small, even in the presence of a

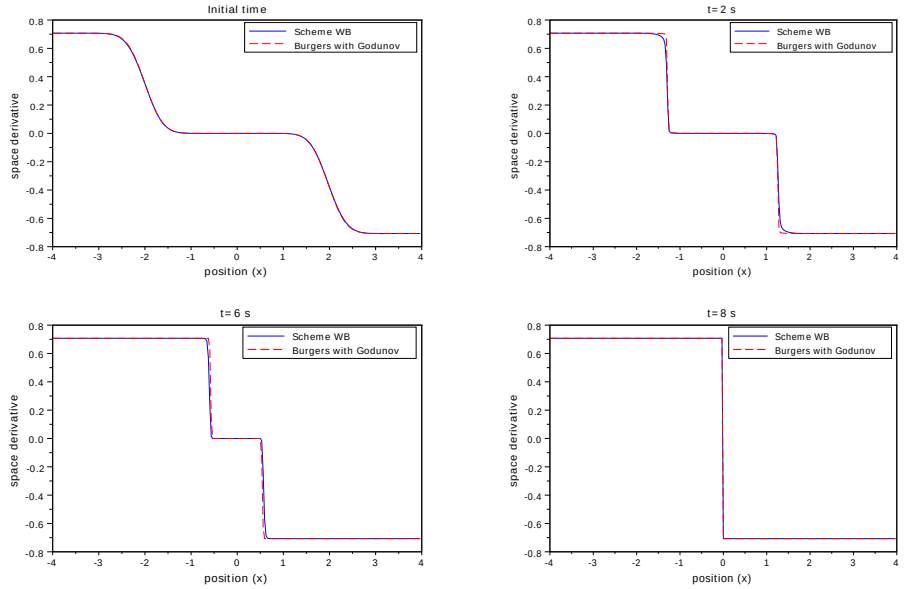

FIG. 7.7. *Comparison* $E = \partial_x \phi$ *for* $\varepsilon = 10^{-5}$ *vs. a Godunov scheme on Burgers equation* (2.5).

Dirac delta in the density (see Fig. 7.8-left). A comparison with the Godunov scheme for the Burgers-Hopf equation in Fig. 7.9 illustrates numerically the good behaviour of the scheme when $\varepsilon \ll 1$.

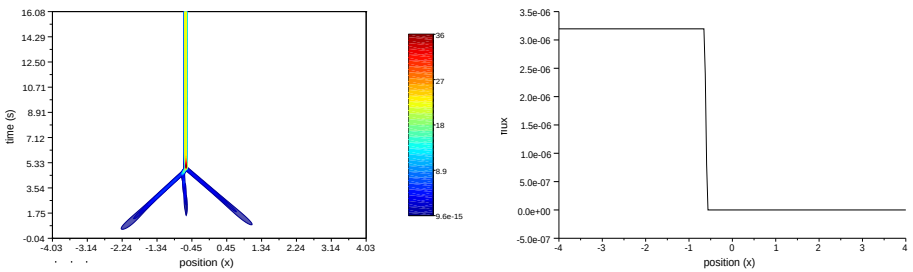

FIG. 7.8. *Macroscopic densities* ρ *(left), fluxes* J *(right) for* $\varepsilon = 10^{-5}$ *with a non-symmetric initial data.*

8. Conclusion and outlook. We developped in this paper a new strategy to deal with numerical resolution of two-stream kinetic models in strongly aggregative setting. The main difficulty is due to the finite-time blow-up of solutions to the limiting model, which involves the treatment of possible Dirac deltas at the limit. This implies a nonconservative product which should be treated with care, here we use a careful discretization of the Vol'pert calculus as explained in §4. We propose two strategies, a hybrid scheme and a full well-balanced scheme. Although the full well-balanced scheme has a large computational cost, it has the advantage to produce a numerical flux of high quality even at the locations of Dirac deltas in the macroscopic density.

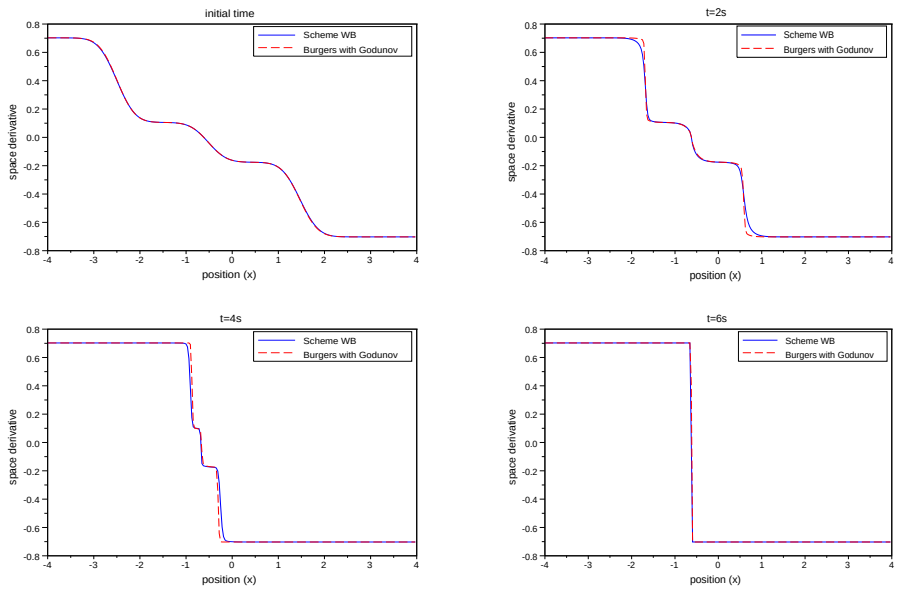

FIG. 7.9. *Comparison $E = \partial_x \phi$ for $\varepsilon = 10^{-5}$ vs. a Godunov scheme on Burgers equation* (2.5) *for a non symetric initial data.*

A drawback inherent to the present study lies in its limitation to two-stream kinetic models, for which the number of macroscopic and microscopic variables is equal. For the VPFP model, such a reduction implies that the limiting macroscopic velocity is only an hyperbolic tangent (see §5.2). Accordingly, the resulting numerical process may underestimate this velocity for large fields E (or θ). This discrepancy finds its origin into an Hermite quadrature involving only two discrete velocities. So, it appears logical to try to extend our present results toward continuous (at least, larger) velocity models for which strong aggregation occurs, see e.g. [13, 17, 20, 39]. Relying either on "Caseology" methods [1] or on approximations [21, §10.3-4], it is still possible to derive scattering matrices $S^n_{j-\frac{1}{2}}$ and to set up a scheme of the form (3.5). However, such a roadmap will again prove inefficient if confronted to strong confinement and its corresponding singular limits (for which nonconservative products appear), as illustrated on Fig.7.1.

Acknowledgment. The second author thanks the Istituto per le Applicazioni del Calcolo (CNR, Rome) where part of this work was carried (november 2014) and partial support from French "ANR blanche" project Kibord: ANR-13-BS01-0004.

Appendix A. Nonlinear iterative process. The algorithm, described in §4, involves the resolution of a nonlinear problem. Let us assume that $(f^+_{j,n})_j$ and $(f^-_{j,n})_j$ are known at time t_n. Then, at time t_{n+1}, the distribution functions are computed thanks to (3.5) with the scattering matrix (4.4)-(4.5), provided the macroscopic velocities $a^n_{j-1/2,L/R}$ are known. To do so, we implement the following process:

1. For $(S_{j-\frac{1}{2},L})_j$ and $(S_{j-\frac{1}{2},R})_j$ given, $a_{j-\frac{1}{2},L}$ and $a_{j-\frac{1}{2},R}$ are computed thanks to the relations (4.12) and (4.13) respectively.
2. Then with (4.5) and (4.9)-(4.10), right/left densities $(\rho_{j-\frac{1}{2},L})_j$ and $(\rho_{j-\frac{1}{2},R})_j$ are computed.

3. Solving Poisson equation (4.11) yields updated values $(\widetilde{S}_{j-\frac{1}{2},L})_j$, $(\widetilde{S}_{j-\frac{1}{2},R})_j$.
4. Let the error be $\epsilon = \|S_{j-\frac{1}{2},L} - \widetilde{S}_{j-\frac{1}{2},L}\|_\infty + \|S_{j-\frac{1}{2},R} - \widetilde{S}_{j-\frac{1}{2},R}\|_\infty$. While this quantity is bigger than a threshold value ϵ_r, we set $S_{j-\frac{1}{2},L} = \widetilde{S}_{j-\frac{1}{2},L}$, $S_{j-\frac{1}{2},R} = \widetilde{S}_{j-\frac{1}{2},R}$ and come back to step 1. In our examples, $\epsilon_r = 10^{-3}$.
5. If the error ϵ is less than ϵ_r, we set $S^n_{j-\frac{1}{2},L} = \widetilde{S}_{j-\frac{1}{2},L}$, $S^n_{j-\frac{1}{2},R} = \widetilde{S}_{j-\frac{1}{2},R}$, and compute $a^n_{j-\frac{1}{2},L/R}$ with (4.12) and (4.13). Then $(f^+_{j,n+1})_j$ and $(f^-_{j,n+1})_j$ follow with (3.5) where the scattering matrix is given by (4.4)-(4.5)

REFERENCES

[1] R.E. Aamodt, K.M. Case, *Useful identities for half-space problems in linear transport theory*, Ann. Physics **21** (1963) 284–301.

[2] D. Amadori, L. Gosse, *Stringent error estimates for one-dimensional, space-dependent* 2×2 *relaxation systems*, Ann. IHP (C) Nonlinear Anal. (2015), DOI: 10.1016/j.anihpc.2015.01.001

[3] A.L. Bertozzi, J.A. Carrillo, T. Laurent, *Blow-up in multidimensional aggregation equation with mildly singular interaction kernels*, Nonlinearity **22** (2009), 683–710.

[4] F. Bouchut, F. James, *One-dimensional transport equations with discontinuous coefficients*, Nonlinear Analysis TMA, **32** (1998), n° 7, 891–933.

[5] G. Bretti, R. Natalini and M. Ribot, A hyperbolic model of chemotaxis on a network: a numerical study, ESAIM: M2AN 48 231–258 (2014).

[6] M.A. Burschka, U.M. Titulaer; *The Kinetic Boundary Layer for the Fokker-Planck Equation: A Brownian Particle in a Uniform Field.* Physica **112A**, 315–330 (1982).

[7] J. A. Carrillo, M. DiFrancesco, A. Figalli, T. Laurent, D. Slepčev, *Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations*, Duke Math. J. **156** (2011), 229–271.

[8] J. A. Carrillo, F. James, F. Lagoutière, N. Vauchelet, *The Filippov characteristic flow for the aggregation equation with mildly singular potentials*, submitted.

[9] F. Chalub, P. Markowich, B. Perthame, and C. Schmeiser, *Kinetic models for chemotaxis and their drift-diffusion limits*, Monats. Math. **142** (2004) 123–141.

[10] L. Derbel and P.-E. Jabin, The set of concentration for some hyperbolic models of chemotaxis. J. Hyperbolic Differ. Equ. 4(2) 331-349 (2007).

[11] B. Despres, Ch. Buet, *The structure of well-balanced schemes for Friedrichs systems with linear relaxation.* Submitted.

[12] Y. Dolak, T. Hillen, *Cattaneo Models for Chemosensitive Movement: Numerical Solution and Pattern Formation*, J. Math. Biol. **46** (2003) 153–170.

[13] Y. Dolak, C. Schmeiser, *Kinetic models for chemotaxis: Hydrodynamic limits and spatio-temporal mechanisms*, J. Math. Biol. **51** (2005) 595–615.

[14] F. Filbet, P. Laurencot, B. Perthame, Derivation of hyperbolic models for chemosensitive Movement, J. Math. Biology, 50 (2) 189–207 (2005).

[15] N.J. Fisch, M. Kruskal, *Separating variables in two-way diffusion equations*, J.Math. Phys. **21** (1980) 740–750.

[16] J.C.M. Fok, B. Guo, T. Tang, *Combined Hermite spectral-finite difference method for the Fokker-Planck equation*, Math. Comput. **71** (2001) 1497–1528.

[17] R.M. Ford, P.T. Cummings, *On the relationship between cell balance equations for chemotactic cell populations*, SIAM J. Applied Math. **52** (1992) 1426–1441.

[18] J. Glimm, D.H. Sharp, *An S-matrix theory for classical nonlinear physics*, Found. Phys. **16** (1986), 125–141.

[19] L. Gosse, *Asymptotic-Preserving and Well-Balanced scheme for the 1D Cattaneo model of chemotaxis movement in both hyperbolic and diffusive regimes*, J. Math. Anal. Applic. **388** (2012) 964–983.

[20] L. Gosse, *A well-balanced scheme for kinetic models of chemotaxis derived from one-dimensional local forward-backward problems*, Math. Biosci. **242** (2013) 117–128.

[21] L. Gosse, **Computing Qualitatively Correct Approximations of Balance Laws**, SIMAI Springer Series **2** Springer (2013) - ISBN 978-88-470-2891-3.

[22] L. Gosse, *A well-balanced scheme able to cope with hydrodynamic limits for linear kinetic models*, Applied Math. Letters **42** (2015) 15–21.

[23] L. Gosse, *A well-balanced processing of a two-dimensional Vlasov-Poisson Fokker-Planck model*

of population dynamics, in preparation.

[24] L. Gosse, F. James, *Numerical approximations of one-dimensional linear conservation equations with discontinuous coefficients*, Math. Comput. **69** (2000) 987–1015.

[25] J. Greenberg, W. Alt, *Stability results for a diffusion equation with functional shift approximating a chemotaxis model*, Trans. Amer. Math. Soc. **300** (1987) 235–258.

[26] F. Guarguaglini, C. Mascia, R. Natalini, M. Ribot, *Global stability of constant states and qualitative behavior of solutions to a one dimensional hyperbolic model of chemotaxis*, Discrete Contin. Dyn. Syst. Ser. B **12** (2009) 39–76.

[27] H. K. Gummel, *A Self -Consistent Iterative Scheme for One-Dimensional Steady State Transistor Calculations*, IEEE Trans. Electron. Devices **11** (1964) 455–465.

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

[29] F. James, N. Vauchelet, *A remark on duality solutions for some weakly nonlinear scalar conservation laws*, C. R. Acad. Sci. Paris, Sér. I, **349** (2011), 657-661.

[30] F. James, N. Vauchelet, *Chemotaxis: from kinetic equations to aggregate dynamics*, Nonlinear Differ. Equ. Appl. (NoDEA) **20** (2013) 101–127.

[31] F. James, N. Vauchelet, *Numerical methods for one-dimensional aggregation equations*, SIAM J. Num. Anal., **53** no 2 (2015), 895–916.

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

[33] S. Jin, Li Wang, *An asymptotic preserving scheme for the Vlasov-Poisson-Fokker-Planck system in the high field regime*, Acta Mathematica Scientia **31** (2011) 2219–2232.

[34] J. Nieto, F. Poupaud, J. Soler, *High field limit for the Vlasov-Poisson-Fokker-Planck system*, Arch. Rat. Mech. Anal. **158** (2001) 29–59.

[35] B. Perthame, A.-L. Dalibard, Existence of solutions of the hyperbolic Keller-Segel model, Trans. Amer. Math. Soc. 361, 2319-2335, 2009.

[36] F. Poupaud, Diagonal Defect Measures, adhesion dynamics and Euler equation, Methods and Applications of Analysis, 2002.

[37] F. Poupaud and M. Rascle, Measure solutions to the linear multi-dimensional transport equation with non smooth coefficients. Comm. PDE 22, 337-358 (1997).

[38] J. Schaeffer, *A difference scheme for the Vlasov-Poisson-Fokker-Planck system*, SIAM J. Numer. Anal., 35 (1998) 1149–1175.

[39] D. W. Stroock, *Some stochastic processes which arise from a model of the motion of a bacterium*, Probab. Theory Related Fields **28** (1974) 305–315.

[40] Monika Twarogowska, **Numerical approximation and analysis of mathematical models arising in cells movement**, Ph.D. Thesis, Univ. dell'Aquila (february 2012).

[41] N. Vauchelet, *Numerical simulation of a kinetic model for chemotaxis*, Kinetic & Related Models **3** (2010) 501–526.

[42] S. Wollman, E. Ozizmir, *A deterministic particle method for the Vlasov-Fokker-Planck equation in one dimension*, J. Comput. Applied Math. **213** (2008) 316–365.