Kinetic and Related Models
Vol. 18, No. 6, December 2025, pp. 989-1015
doi:10.3934/krm.2025012

FROM A NONLINEAR KINETIC EQUATION TO A VOLUME-EXCLUSION CHEMOTAXIS MODEL VIA ASYMPTOTIC PRESERVING METHODS

Gissell Estrada-Rodriguez[1], Diane Peurichard[2] and Xinran Ruan[*3]

[1]Mathematical Institute, University of Oxford, Oxford, U.K

[2]Sorbonne Université, Inria Paris - MUSCLEES project team, Université de Paris, CNRS, Laboratoire Jacques-Louis Lions UMR7598, Paris, France

[3]Capital Normal University, School of Mathematical Sciences, Beijing 100048, China

(Communicated by Alina Chertock)

Abstract. In this work we first prove, by formal arguments, that the diffusion limit of nonlinear kinetic equations, where both the *transport term and the turning operator are density-dependent*, leads to volume-exclusion chemotactic equations. We then numerically study this diffusive limit via an asymptotic preserving scheme based on a micro-macro decomposition. By properly discretizing the nonlinear term implicitly-explicitly in an upwind manner, the scheme produces accurate approximations also in the case of strong chemosensitivity. We show, via detailed calculations, that the scheme is asymptotic preserving and bound preserving and show numerically an energy dissipation property, which are essential for practical applications. We extend this scheme to two dimensional kinetic models and we validate its efficiency by means of 1D and 2D numerical experiments of pattern formation in biological systems.

1. **Introduction.** Chemotaxis is the mechanism by which cells and organisms adapt their movement in response to a chemical stimulus present in their environment. This phenomenon has been observed in many biological systems [3, 21, 22, 36].

The mathematical study of chemotaxis started from the seminal contributions of Patlak [35], Keller, Segel and Alt [3, 23, 24], where the authors introduced the celebrated Patlak-Keller-Segel-Alt (PKSA) model. This model was originally proposed for pattern formation in bacterial populations through an advection-diffusion system of two coupled parabolic equations describing the evolution of the cell density and the chemoattractant. The PKSA model has become the prevailing method for representing chemotactic behaviour in biological systems at population level (see Ref. [15] for a review about Keller-Segel type models). A drawback of studying biological systems directly at the continuous level, i.e. through partial differential equations (PDEs), is that the interactions comprised in such models can be

2020 *Mathematics Subject Classification.* Primary: 35Q70, 76M45, 35Q20, 92C17, 35K55, 35Q92, 92B05; Secondary: 35B36, 65M06.

Key words and phrases. Kinetic equations, asymptotic preserving (AP) methods, chemotaxis equations, diffusion approximation, velocity-jump processes.

*Corresponding author: Xinran Ruan.

difficult to interpret in terms of mechanisms at the cellular level. Therefore it is fundamental to explain the origin of these macroscopic models from a mechanistic/microscopic description of motion. In the case of the PKSA model, tremendous amount of works in this direction have been made by various authors (see reviews of Horstmann [16, 17]). Among other approaches, several authors showed that the chemotaxis models could be obtained from biased random walk approaches [3, 31], or as the parabolic limit of velocity-jump processes / transport models [8, 32, 33]. For instance in [33], the authors study the diffusion-limit of transport equations under a variety of external biases imposed on the turning operator. Depending on the strength of the bias, the authors show that this leads to anisotropic diffusion, drift term in the flux, or both. In particular, they show that the PKSA model only arises if the perturbation of the turning kernel is sufficiently small. They consider several examples of prototype models for chemotaxis of bacteria, of slime molds and of myxobacteria and present a general discussion of the derivation of diffusion approximations from various stochastic processes that model chemotaxis (see [32] and references therein). These works not only give insights into how the continuum PKSA model relates to models at the microscopic scale, but also enable to understand the link between different microscopic models and modeling assumptions.

The PKSA model has also been extensively studied in the literature from a theoretical viewpoint. Of particular interest is the tendency of solutions to exhibit finite-time blow-up [4, 8, 14, 30, 31]. It is now well established that in dimension $n \geq 2$, finite-time blow-up may occur if the initial condition exceeds some threshold. However, under the biologically relevant cases for aggregation to occur, initial conditions typically lie above this threshold. For these cases, the solutions of the PKSA model would correctly predict pattern formation but in the form of blow-up, i.e. the model will not be relevant for later stages following aggregation in real systems. Therefore, it is necessary to modify the model to allow pattern formation without blow-up. In this direction, the PKSA original model has been modified by various authors with the aim of improving its consistency with biological systems (see [15] for a comprehensive review). One example is the volume-exclusion chemotactic system (VEPKS) introduced by Hillen and Painter [34] to take into account the finite size of cells and volume limitations. In such models the chemotactic sensitivity (i.e. the term leading to cell aggregation) depends on both the chemical concentration in the medium and the local cell density, thus, the population directly modulates its own sensitivity response. The coupled system reads, in its parabolic-elliptic form as

$$\begin{gathered}\partial_t \rho - \nabla \cdot \Big(D_0(q(\rho) - \rho q'(\rho)) \nabla \rho - \chi_0 q(\rho) \rho \nabla c \Big) = h(\rho, c) \ , \qquad t \geq 0 \ , \mathbf{x} \in \Omega \subset \mathbb{R}^n \ , \\ D_c \Delta c + g(\rho, c) = 0 \ .\end{gathered} \tag{1}$$

Here, $\rho(t, \mathbf{x})$ is the cell density, $c(t, \mathbf{x})$ is the chemoattractant concentration and $q(\rho)$ is a function that describes the packing capacity of the cell aggregates. The diffusion coefficients for the cells and chemoattractant are D_0 and D_c, respectively, and χ_0 is the chemotactic sensitivity. The proliferation (or death) of the cells is described by $h(\rho, c)$, and the production and consumption of the chemoattractant is given by $g(\rho, c)$. The dependency of h and g on the cell density ρ enables to account for density-limited phenomena such as the logistic-growth function for h commonly used in the literature [28, 29]. Note that from (1), we recover the classical

PKSA model by taking $q(\rho) = 1$. It has been shown [5, 42, 43] that such volume-exclusion effects prevent blow-ups in finite time compared to the model without density effects (with $q(\rho) = 1$). The volume-exclusion chemotactic equation has been widely studied in the literature, from a modeling [34, 40], analytic [10, 13, 26, 42] and numerical perspectives [18]. A relevant example where volume-exclusion effects play a crucial role is in the modeling of glioblastoma (GBM) aggregates [38]. In particular, the system (1) was used in [2] to explain the mechanical changes at the cell level due to the presence of a chemical treatment. In [2], the cell's elasticity is modeled through the term $q(\rho)$ which incorporates the cell-cell interactions [41]. This function $q(\rho)$ was chosen to be $q(\rho) = 1 - (\rho/\bar{\rho})^\gamma$ where γ is a parameter that depends on the concentration of the treatment and $\bar{\rho}$ is the maximum cell density in each aggregate. For $\gamma = 1$ the cells were considered as solid particles while $\gamma > 1$ corresponded to semi-elastic particles that can squeeze into empty spaces.

In one dimension, Hillen and Painter [34] derived the VEPKS from a discrete-space random walk. The derivation begins with a master equation for a continuous-time and discrete space random walk model as in [31], but including the volume-filling effect by making the probability of making a jump depend on the availability of space into which cells can move. A natural question that arises is whether the VEPKS model can be obtained as the limit of a kinetic velocity-jump model as done for the PKSA model. In this paper, our first aim is to provide an interpretation of the volume-exclusion chemotactic system (1) as the diffusion limit of a kinetic velocity-jump model.

Inspired by the approach in [32, 33], we first show, by formal arguments, that the system (1) can be obtained in the diffusion limit of a nonlinear kinetic equation, provided that both the transport term and the turning operator are density-dependent (Theorem 2.1 in Section 2.2). The corresponding kinetic velocity-jump model we propose reads

$$\partial_t f + \mathbf{v}\cdot\nabla\big(F[\rho](t,\mathbf{x},\mathbf{v})f\big) = \lambda q(\rho)\big(-f + \rho T(\mathbf{v},\rho,\nabla c)\big) + r_0 f\Big(1-\frac{\rho}{\rho_{\max}}\Big)_+, \quad (2)$$

where $f(t,\mathbf{x},\mathbf{v}) \geq 0$ is the phase space cell density, $\mathbf{x} \in \Omega \subset \mathbb{R}^n$ denotes the position, $\mathbf{v} \in V \subset \mathbb{R}^n$ is the velocity where V denotes the unit sphere $V = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| = 1\}$ and $t \in \mathbb{R}^+$ the time. Note that we have considered that cells proliferate according to a logistic-growth law $h(\rho,c) \equiv h(\rho) = r_0 f\Big(1-\frac{\rho}{\rho_{\max}}\Big)_+$ where $\rho_{\max}$ is the carrying capacity and r_0 is the proliferation rate. In (2), λ is the constant turning rate, with $1/\lambda$ giving a measure of the mean run length between velocity jumps, $T(\mathbf{v},\rho,\nabla c)$ gives the probability of a velocity jump to velocity $\mathbf{v}$, which accounts for cell preference to move up the chemotactic gradient ∇c, and the term $F[\rho](t,\mathbf{x},\mathbf{v})$ describes the anisotropic transport due to the density limited motion. The fact that the turning operator $T(\mathbf{v},\rho,c,\nabla c)$ depends on the gradient of chemical concentration ∇c is necessary to obtain the chemotactic behavior described by (2), as discussed thoroughly in [33]. For simplicity, we have assumed here that the turning rate only depends on ∇c but the derivation extends to more general choices $T(\mathbf{v},\rho,c,\nabla c)$, as briefly discussed in remark 2.2. While similar models have been considered in the literature, [32, 33], ours extends to higher dimensions and includes an extra nonlinearity in the transport term.

In model (2), it is noteworthy that cell motion is not only biased by external factors (leading to chemotactic motion) but also biased to account for density-limited motion/crowding effects. Model (2) shows that the minimal hypothesis

on the run-and-tumble process enabling to obtain the VEPKS system (1) in the diffusion limit are density-dependent transport term and turning kernel. The first hypothesis (density-dependent transport) is the most natural in terms of a modeling viewpoint, and expresses that cells will only be able to move in not already crowded regions. The second hypothesis (density-dependent turning kernel) is a more 'active' mechanism expressing that cells have less probability to turn towards crowding regions. Applied to bacteria, these hypothesis are consistent with experimental evidence showing that tightly packed colonies could collectively move faster by reducing the speed of their individual bacteria [27].

The second goal in this paper is to numerically investigate whether the relevant macroscopic volume exclusion equation corresponds to the underlying physical system described by the kinetic equation in the diffusion limit. These methods are widely known as asymptotic preserving schemes (AP) [12, 19] since they mimic the asymptotic behaviour of the kinetic equation when the scaling parameter approaches to zero and the mesh size and time steps are fixed. In this paper, we use a micro-macro decomposition of the unknown in the sense of [25] as detailed in Section 3. The finite difference discretization is explained in Section 4, where a proper implicit-explicit discretization of the nonlinear terms, including the chemosensitivity term, is applied to improve the efficiency and stability of our numerical scheme. This decomposition of the solution of the kinetic equation is analogous to a Chapman-Enskog expansion in the case of the classical Boltzmann equation. It uses the properties of the "collisional operator", which in our formulation describes the run and tumble movement of the individual. For the kinetic counterpart of the classical PKSA equation we refer to [7], where the authors used an odd-even splitting at the kinetic level and studied the blow-up behaviour of solutions. Other related works can be found in [9, 11, 20, 39].

Outline of the paper This paper is organized as follows. In Section 2 we introduce the new version of the kinetic velocity-jump model given by (2) and consider its diffusion limit using a Hilbert expansion method as in Ref. [33] under appropriate scaling assumptions for the turning operator. We show that the limiting system has a dissipating free energy and we derive an energy estimate used to validate the numerical method. In Section 3, we present the micro-macro decomposition of the kinetic model that we use to build the AP scheme presented, and analyzed various properties (positivity preserving, AP etc) in Section 4. Finally, we present the numerical experiments and draw conclusions in Section 5.

2. Volume-exclusion kinetic equation: Macroscopic limit.

In this section, we analyze the diffusion limit of the following transport equation

$$\partial_t f + \mathbf{v}\cdot\nabla\big(F[\rho](t,\mathbf{x},\mathbf{v})f\big) = \lambda q(\rho)\big(-f+\rho T(\mathbf{v},\rho,\nabla c)\big) + r_0 f\Big(1-\frac{\rho}{\rho_{\max}}\Big)_+ .$$

The turning kernel is supposed to be independent on the previous velocity of the jumping particle, and only dependent on the new velocity $\mathbf{v}$, the chemical concentration is given by $c(t,\mathbf{x})$ and the density of particles by $\rho(t,\mathbf{x}) = \int_V f(t,\mathbf{x},\mathbf{v})\,\mathrm{d}\mathbf{v}$ near position $\mathbf{x}$. The function $q(\rho)$ is the probability for a cell to find space at its neighbouring locations, and we assume that only a finite number of cells, $\bar\rho$, can be accommodated at any site. We will therefore consider functions $q(\rho)$ such that

$$q(\bar\rho) = 0 \quad \text{and} \quad q(\rho) \geq 0 \quad \text{for all} \;\; 0 \leq \rho \leq \bar\rho \; .$$

We suppose that the turning kernel T integrates to 1 in the velocity variable,

$$\int_V T(\mathbf{v}, \rho, \nabla c)\, \mathrm{d}\mathbf{v} = 1 \ .$$

In the volume-exclusion approach, following the lines of [15], we assume that the probability of making a jump depends upon the availability of space into which a cell can move. To this aim, we suppose that cells can only make a turn in directions where space is available, and we choose the turning operator T to be

$$T(\mathbf{v}, \rho, \nabla c) = \tilde{c}(t, \mathbf{x})\psi(\mathbf{v}, \nabla c) q(\rho(t, \mathbf{x} + \alpha \mathbf{v})) \ ,$$

where α is a small parameter, $\psi(\mathbf{v}, \nabla c)$ will be defined later in (6), and $\tilde{c}(t, \mathbf{x})$ is a normalisation factor given by

$$\tilde{c}(t, \mathbf{x}) = \frac{1}{\displaystyle\int_V \psi(\mathbf{v}, \nabla c) q(\rho(t, \mathbf{x} + \alpha \mathbf{v}))\, \mathrm{d}\mathbf{v}} \ .$$

Note that under these assumptions, particles will only make a turn if (i) they are not already trapped in a high density region (where they stop) and (ii) only in directions where the density of cells is not already too large.

In order to take into account density limited motion, we will also consider that cells are only transported to non-overcrowded regions, and choose for the transport term

$$F[\rho](t, \mathbf{x}, \mathbf{v}) = q\big(\rho(t, \mathbf{x} + \alpha \mathbf{v})\big) \ .$$

Note that $F[\rho](t, \mathbf{x}, \mathbf{v}) = 0$ when $\rho(t, \mathbf{x}) = \bar{\rho}$, meaning that when the maximum capacity of an aggregate is reached, cells cannot move in that direction.

We consider that the saturation on both the turning operator T and the density-dependent transport F, is given by the same function q. This is because cells will sense crowded regions in the same way either when they are going to choose the next direction of movement, or when they are in the running phase. This is of course a modeling assumption in our case, but in general, different functions can be considered.

2.1. **Diffusion scaling.** Following the lines of [15], we aim to obtain a macroscopic limit by choosing space and time scales on which there are many velocity jumps in one order of time, but small net displacements on this time scale. To this aim, we define the dimensionless velocity, space and time variables as

$$\mathbf{u} = \frac{\mathbf{v}}{s} \ , \quad \xi = \frac{\mathbf{x}}{L} \ , \quad \tau = \frac{t}{\sigma} \ ,$$

where s is the characteristic speed, L the characteristic length scale and σ yet to be determined. Equation (2) now writes,

$$\frac{1}{\sigma}\partial_\tau \tilde{f} + \frac{s}{L}\mathbf{u} \cdot \nabla_\xi (F[\tilde{\rho}](\tau, \xi, \mathbf{u})\tilde{f}) = \lambda q(\tilde{\rho})(-\tilde{f} + \tilde{\rho} T(\mathbf{u}, \tilde{\rho}, \nabla c)) + \frac{1}{\sigma} r_0 \tilde{f}\Big(1 - \frac{\tilde{\rho}}{\rho_{\max}}\Big)_+ \ , \quad (3)$$

where $\tilde{f}(\tau, \xi, \mathbf{u}) = f(\sigma\tau, L\xi, s\mathbf{u})$, $\tilde{\rho}(\tau, \xi) = \rho(\sigma t, L\xi)$ and therefore

$$F[\tilde{\rho}](\tau, \xi, \mathbf{u}) = q(\tilde{\rho}(\tau, \xi + \alpha \frac{s}{L}\mathbf{u})) \ ,$$

where the parameter α carries a dimension of time. We estimate the diffusion coefficient as the product of the characteristic speed times the distance traveled between velocity jumps, giving $D \approx \mathcal{O}(\frac{s^2}{\lambda})$, and we deduce the characteristic diffusion time on the length scale L by $\tau_{diff} \approx \frac{L^2\lambda}{s^2}$. The characteristic drift

time is defined by $\tau_{drift} = \frac{L}{s}$ and we assume that the space scale is such that $\tau_{run} = \frac{1}{\lambda} \ll \tau_{drift} \ll \tau_{diff}$. We therefore introduce a small parameter $\varepsilon \ll 1$ and ensure that $\tau_{run} = \mathcal{O}(1)$, $\tau_{drift} = \mathcal{O}(\frac{1}{\varepsilon})$ and $\tau_{diff} = \mathcal{O}(\frac{1}{\varepsilon^2})$ by choosing the time and space scales to be $L \approx \mathcal{O}(\frac{s}{\varepsilon})$ and $\sigma = \tau_{diff}$. Finally, the parameter α follows the fastest time scale, i.e $\tilde{\alpha} = \frac{1}{\lambda}\alpha$ and without loss of generality we set $\tilde{\alpha} = \lambda = 1$. Now equation (3) becomes (dropping the tildes and, with a slight abuse of notations, going back to $\mathbf{x}, \mathbf{v}$ and t for the dimensionless quantities $\xi, \mathbf{u}$ and τ),

$$\varepsilon^2 \partial_t f^\varepsilon + \varepsilon \mathbf{v} \cdot \nabla (F_\varepsilon[\rho](t,\mathbf{x},\mathbf{v}) f^\varepsilon) = q(\rho)(-f^\varepsilon + \rho T_\varepsilon(\mathbf{v},\rho,\nabla c)) + \varepsilon^2 r_0 f^\varepsilon \Big(1 - \frac{\rho}{\rho_{\max}}\Big)_+ , \quad (4)$$

where

$$T_\varepsilon(\mathbf{v},\rho,\nabla c) = \frac{\psi_\varepsilon(\mathbf{v},\nabla c) q(\rho(t,\mathbf{x}+\varepsilon\mathbf{v}))}{\displaystyle\int_V \psi_\varepsilon(\mathbf{v},\nabla c) q(\rho(t,\mathbf{x}+\varepsilon\mathbf{v}))\,\mathrm{d}\mathbf{v}} , \qquad F_\varepsilon[\rho](t,\mathbf{x},\mathbf{v}) = q(\rho(t,\mathbf{x}+\varepsilon\mathbf{v})) . \quad (5)$$

We denote by $\langle . \rangle$ the integration in the $\mathbf{v}$ variable, $\langle u \rangle = \int_V u(\mathbf{v}) d\mathbf{v}$, and we will consider that $\langle \psi_\varepsilon \rangle = \int_V \psi_\varepsilon(\mathbf{v},\nabla c)\,\mathrm{d}\mathbf{v} = 1$, where $\psi_\varepsilon(\mathbf{v},\nabla c)$ is a non-negative and decreasing function in ∇c. This modeling choice amounts to consider that cells are less likely to tumble when the chemical gradient increases. We then consider that the dependency of the turning operator on the chemical gradient ∇c happens as a perturbation of magnitude ε in the following way

$$\psi_\varepsilon(\mathbf{v},\nabla c) = \psi_0(\mathbf{v}) + \varepsilon \psi_1(\mathbf{v},\nabla c) . \quad (6)$$

In order to recover the volume-exclusion Keller-Segel equation, we will assume that ψ_0 is radially symmetric and that the perturbation $\psi_1(\mathbf{v},\nabla c)$ depends linearly on the chemical gradient ∇c. These assumptions lead to the following hypotheses:

Hypothesis 1.

$$\int_V \psi_0(\mathbf{v})\,\mathrm{d}\mathbf{v} = 1 , \quad \int_V \psi_1(\mathbf{v},\nabla c)\,\mathrm{d}\mathbf{v} = 0 , \quad \text{(H1)}$$

Hypothesis 2.

$$\int_V \mathbf{v}\psi_0(\mathbf{v})\,\mathrm{d}\mathbf{v} = 0 , \quad \psi_1(\mathbf{v},\nabla c) = \phi(\mathbf{v}) \cdot \nabla c . \quad \text{(H2)}$$

2.2. **Macroscopic model.** In this section we prove the following theorem:

Theorem 2.1. *(formal) The limit $\varepsilon \to 0$ of f^ε solving (4)-(6) together with Hypotheses H1 and H2 is $f^0 = \rho(t,\mathbf{x})\psi_0(\mathbf{v})$, where ρ solves*

$$\partial_t \rho - \nabla \cdot \big(D_0\big(q(\rho) - \rho q'(\rho)\big)\nabla\rho - \chi_0 \rho q(\rho) \nabla c\big) = r_0 \rho \Big(1 - \frac{\rho}{\rho_{\max}}\Big)_+ , \quad (7)$$

with the diffusion coefficient D_0 and the chemotactic sensitivity parameter χ_0 given by

$$D_0 = \langle (\mathbf{v} \otimes \mathbf{v})\psi_0(\mathbf{v}) \rangle \qquad \text{and} \qquad \chi_0 = \langle \mathbf{v} \otimes \phi(\mathbf{v}) \rangle . \quad (8)$$

Proof. We first expand the transport quantity F_ε and the turning operator T_ε given by (5). For $\varepsilon \ll 1$, we have

$$F_\varepsilon[\rho] = q(\rho(t,\mathbf{x}+\varepsilon\mathbf{v})) = q(\rho(t,\mathbf{x})) + \varepsilon q'(\rho(t,\mathbf{x}))\mathbf{v} \cdot \nabla\rho(t,\mathbf{x}) + \mathcal{O}(\varepsilon^2) , \quad (9)$$

where $q'(\rho) = \frac{\mathrm{d}q}{\mathrm{d}\rho}$. Introducing this expansion in the expression for T_ε, we write

$$\tilde{c}_\varepsilon(t,\mathbf{x}) = \frac{1}{q(\rho)\langle\psi_\varepsilon\rangle}\Big(1 - \varepsilon \frac{q'(\rho)}{q(\rho)} \nabla\rho \cdot \frac{\langle \mathbf{v}\psi_\varepsilon \rangle}{\langle\psi_\varepsilon\rangle}\Big) + \mathcal{O}(\varepsilon^2) ,$$

where $\langle\psi_\varepsilon\rangle = \int_V \psi_\varepsilon(\mathbf{v},\nabla c)\,\mathrm{d}\mathbf{v}$ and $\langle\mathbf{v}\psi_\varepsilon\rangle = \int_V \mathbf{v}\psi_\varepsilon(\mathbf{v},\nabla c)\,\mathrm{d}\mathbf{v}$ denote the first and second moments of ψ_ε. Finally, we obtain

$$T_\varepsilon(\mathbf{v},\rho,\nabla c) = \frac{\psi_\varepsilon(\mathbf{v},\nabla c)}{\langle\psi_\varepsilon\rangle} + \varepsilon\frac{q'(\rho)}{q(\rho)}\frac{\psi_\varepsilon(\mathbf{v},\nabla c)}{\langle\psi_\varepsilon\rangle}\nabla\rho\cdot\left(\mathbf{v} - \frac{\langle\mathbf{v}\psi_\varepsilon\rangle}{\langle\psi_\varepsilon\rangle}\right) + \varepsilon^2\bar{T} + \mathcal{O}(\varepsilon^3)\ ,$$

where $\langle\bar{T}\rangle = 0$. Note that the error terms contained in $\mathcal{O}(\varepsilon^2)$ integrate to 0 in the velocity variable. Using (6), the turning kernel T_ε writes

$$T_\varepsilon(\mathbf{v},\rho,\nabla c) = \psi_0(\mathbf{v}) + \varepsilon\Big(\psi_1(\mathbf{v},\nabla c) + \frac{q'(\rho)}{q(\rho)}(\mathbf{v}\cdot\nabla\rho)\psi_0(\mathbf{v})\Big) + \varepsilon^2\bar{T} + \mathcal{O}(\varepsilon^3)\ , \quad (10)$$

where the $\mathcal{O}(\varepsilon^3)$-term is such that $\langle\mathcal{O}(\varepsilon^3)\rangle = 0$. We also note that using Hypotheses H1 and H2 we have $\langle T_\varepsilon\rangle = 1$ which describes the conservation of individuals during the velocity reorientation. We now consider a second order regular expansion of f^ε in ε,

$$f^\varepsilon(t,\mathbf{x},\mathbf{v}) = f^0(t,\mathbf{x},\mathbf{v}) + \varepsilon f^1(t,\mathbf{x},\mathbf{v}) + \varepsilon^2 f^2(t,\mathbf{x},\mathbf{v}) + \mathcal{O}(\varepsilon^3)\ ,$$

where $\int_V f^\varepsilon(t,\mathbf{x},\mathbf{v})\,\mathrm{d}\mathbf{v} = \int_V f^0(t,\mathbf{x},\mathbf{v})\,\mathrm{d}\mathbf{v} = \rho(t,\mathbf{x})$, therefore $\int_V f^i(t,\mathbf{x},\mathbf{v})\,\mathrm{d}\mathbf{v} = 0$, $\forall i \geq 1$. Introducing this ansatz in (4), we obtain

$$\begin{aligned}
&\varepsilon^2\partial_t f^0 + \varepsilon\mathbf{v}\cdot\nabla\big(q(\rho)\big(f^0 + \varepsilon f^1\big) + \varepsilon q'(\rho)(\mathbf{v}\cdot\nabla\rho)f^0\big)\\
&\quad = q(\rho)\Big[-f^0 + \rho\psi_0(\mathbf{v}) + \varepsilon\Big(\big(-f^1 + \rho\Big(\psi_1(\mathbf{v},\nabla c) + \frac{q'(\rho)}{q(\rho)}(\mathbf{v}\cdot\nabla\rho)\psi_0(\mathbf{v}) + \varepsilon^2\bar{T}\Big)\Big)\\
&\qquad + \varepsilon^2 r_0 f^0\Big(1 - \frac{\rho}{\rho_{\max}}\Big)_+ + \mathcal{O}(\varepsilon^3)\Big]\ .
\end{aligned}$$

Identifying the different equations in powers of ε, we obtain

$$\varepsilon^0 : f^0(t,\mathbf{x},\mathbf{v}) = \rho(t,\mathbf{x})\psi_0(\mathbf{v})\ , \quad (11)$$

$$\varepsilon^1 : \mathbf{v}\cdot\nabla(q(\rho)f^0) = q(\rho)\left(-f^1 + \rho\left(\psi_1(\mathbf{v},\nabla c) + \frac{q'(\rho)}{q(\rho)}(\mathbf{v}\cdot\nabla\rho)\psi_0(\mathbf{v})\right)\right)\ , \quad (12)$$

$$\varepsilon^2 : \partial_t f^0 + \mathbf{v}\cdot\nabla\Big(q(\rho)f^1 + q'(\rho)(\mathbf{v}\cdot\nabla\rho)f^0\Big) = r_0 f^0\Big(1 - \frac{\rho}{\rho_{\max}}\Big)_+ + \bar{T}\ . \quad (13)$$

Integrating (13) with respect to $\mathbf{v}\in V$ and noticing that $\langle\bar{T}\rangle = 0$, using Hypothesis H1, we get

$$\partial_t\rho + \nabla\cdot\big(q(\rho)\langle\mathbf{v}f^1\rangle + \rho q'(\rho)\langle(\mathbf{v}\otimes\mathbf{v})\psi_0\rangle\nabla\rho\big) = r_0\rho\Big(1 - \frac{\rho}{\rho_{\max}}\Big)_+\ . \quad (14)$$

Next, after replacing f^0 by its expression (Eq. 11), we multiply (12) by $\mathbf{v}$ and we integrate again with respect to $\mathbf{v}$ to obtain

$$q(\rho)\langle\mathbf{v}f^1\rangle = -\nabla\cdot\big(\rho q(\rho)\langle(\mathbf{v}\otimes\mathbf{v})\psi_0\rangle\big) + \rho q(\rho)\langle\mathbf{v}\psi_1\rangle + \rho q'(\rho)\langle(\mathbf{v}\otimes\mathbf{v})\psi_0\rangle\nabla\rho\ . \quad (15)$$

Substituting (15) into (14) we get

$$\begin{aligned}
\partial_t\rho + \nabla\cdot\Big[&-\langle(\mathbf{v}\otimes\mathbf{v})\psi_0\rangle\nabla(\rho q(\rho)) + \rho q(\rho)\langle\mathbf{v}\psi_1\rangle\\
&+ 2\rho q'(\rho)\langle(\mathbf{v}\otimes\mathbf{v})\psi_0\rangle\nabla\rho\Big] = r_0\rho\Big(1 - \frac{\rho}{\rho_{\max}}\Big)_+\ .
\end{aligned}$$

Noting that $\nabla(q(\rho)\rho) = q'(\rho)\rho\nabla\rho + q(\rho)\nabla\rho$ and using Hypothesis H2 for the perturbation ψ_1, we finally arrive to the volume-exclusion Keller-Segel model (7) together with (8). □

This macroscopic equation describes the volume-exclusion chemotactic motion associated with the so-called *squeezing probability* $q(\rho)$. Depending on the choice of this function we can consider the cells either as solid blocks, for the case $q(\rho) = 1-\frac{\rho}{\bar{\rho}}$, where $\bar{\rho}$ is the maximum cell density in each aggregate, or as semi-elastic entities for $q(\rho) = 1 - \left(\frac{\rho}{\bar{\rho}}\right)^{\gamma}$ (see [41]). In the following section we show that equation (7) admits an energy functional decreasing in time.

Remark 2.2. As discussed in [33], the assumption (H2), which consists in assuming that the perturbation ψ_1 is linearly dependent on ∇c, is necessary to recover the chemotactic term in the limit $\epsilon \to 0$. We note here that the derivation extends easily to perturbations of the form $\psi_1(\mathbf{v}, c, \nabla c) = \tilde{\phi}(\mathbf{v}, c) \cdot \nabla c$. In this case, notice that we would obtain a non constant chemotactic intensity $\chi_0(c) = \langle v \otimes \phi(v, c)\rangle$ in the limit.

2.3. **Energy dissipation in the macroscopic model.** Thus far, we could perform the derivation without specifying the equation for the chemoattractant c. From the remaining of the paper, we choose the production and consumption of chemoattractant $g(\rho, c)$ in Eq. (2) to be linear in both variables $g(\rho, c) = \rho - c$. Therefore, c solves:

$$\Delta c + \rho - c = 0. \tag{16}$$

Under proper assumptions, we prove that the macroscopic volume-exclusion Keller-Segel model (7) has the property of energy dissipation, which will be a key feature to be preserved in numerical methods. For simplicity, we consider the case without proliferation, i.e. $r_0 = 0$ in (7).

Following a gradient flow approach to energy in the sense of Refs. [1, 5, 6], we start by defining $H(\rho) = \frac{D_0}{\chi_0} \ln\left(\frac{\rho}{q(\rho)}\right)$. Then the volume-exclusion Keller-Segel model (7) can be reformulated as

$$\partial_t \rho + \nabla \cdot (\chi_0 q(\rho)\rho(-H'(\rho)\nabla\rho + \nabla c)) = 0 \ , \tag{17}$$

where $H'(\rho) = D_0 \frac{q(\rho) - \rho q'(\rho)}{\chi_0 q(\rho)\rho}$. The energy functional of the model can be given by

$$\mathcal{E}(t) = \int \Phi(\rho)\, \mathrm{d}\mathbf{x} - \frac{1}{2}\int \rho c\, \mathrm{d}\mathbf{x} \ , \tag{18}$$

where

$$\Phi'(\rho) = H(\rho) \ . \tag{19}$$

Proposition 2.3. *Suppose that $c(t, \mathbf{x})$ and $\rho(t, \mathbf{x})$ solve Eqs. (16)-(17). We have*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) = -\int \chi_0 q(\rho)\rho |\nabla(H - c)|^2\, \mathrm{d}\mathbf{x} \leq 0 \ , \tag{20}$$

where the energy functional $\mathcal{E}(t)$ is given by (18).

Proof. Multiplying (17) by $H - c$, we get

$$\begin{aligned}(H - c)\partial_t \rho &= (H - c)\nabla \cdot (\chi_0 q(\rho)\rho \nabla(H - c)) \\ &= \frac{1}{2}\nabla \cdot \left(\chi_0 q(\rho)\rho \nabla(H - c)^2\right) - \chi_0 q(\rho)\rho |\nabla(H - c)|^2 \ ,\end{aligned}$$

where we used the relation $H'(\rho)\nabla\rho - \nabla c = \nabla(H - c)$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\int \left(\Phi - \frac{1}{2}\rho c\right)\mathrm{d}\mathbf{x} = \frac{\mathrm{d}}{\mathrm{d}t}\left[\int \left(\Phi - \frac{1}{2}c^2 - \frac{1}{2}|\nabla c|^2\right)\mathrm{d}\mathbf{x}\right]$$

$$= \int \left[H\partial_t\rho - c\left(\partial_t c - \Delta(\partial_t c)\right)\right] \mathrm{d}\mathbf{x} = \int (H-c)\partial_t\rho \,\mathrm{d}\mathbf{x}$$
$$= -\int \chi_0 q(\rho)\rho |\nabla(H-c)|^2 \,\mathrm{d}\mathbf{x} \le 0 \ .$$

□

Remark 2.4. Proposition 2.3 can be generalized to the equation

$$\partial_t\rho + \nabla\cdot(\chi_0 q(\rho)\rho(-H'(\rho)\nabla\rho + \nabla c)) = h(\rho),$$

where a proliferation term $h(\rho)$ is included satisfying $h(\rho) \ge 0$ and $h(\rho)H(\rho) \le 0$. In fact, with a similar argument as in the proof of Proposition 2.3, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) = \int (H-c)\partial_t\rho \,\mathrm{d}\mathbf{x} = -\int \chi_0 q(\rho)\rho|\nabla(H-c)|^2 \,\mathrm{d}\mathbf{x} + \int (H-c)h \,\mathrm{d}\mathbf{x} \le 0.$$

Specifically, when $q(\rho) = 1-\rho/\bar\rho$ and $h(\rho) = r_0\rho\left(1-\rho/\rho_{\max}\right)_+$, it can be checked that $\rho_{\max} + \rho_{\max}/\bar\rho \le 1$ is a sufficient condition for $h(\rho)H(\rho) \le 0$.

3. **Micro-macro decomposition.** The transport term $F_\varepsilon(\rho)$ and the turning operator T_ε introduced in (5) are highly nonlinear. To facilitate the design and analysis of an asymptotic preserving scheme based on micro-macro decomposition, a linearization of these terms is needed. From now on, we consider a slightly modified version of the kinetic model, where F_ε and T_ε are truncated to the first order in the expansions (9) and (10). To be precise, we solve the following equation

$$\varepsilon^2\partial_t f + \varepsilon\mathbf{v}\cdot\nabla(\tilde F_\varepsilon(\rho)f) = q(\rho)\left(-f + \rho\tilde T_\varepsilon(\mathbf{v},\rho,\nabla c)\right) + \varepsilon^2 r_0 f\left(1-\rho/\rho_{\max}\right)_+ \ , \quad (21)$$

where

$$\tilde F_\varepsilon = q(\rho(t,\mathbf{x})) + \varepsilon q'(\rho(t,\mathbf{x}))\mathbf{v}\cdot\nabla\rho(t,\mathbf{x}) \ ,$$
$$\tilde T_\varepsilon = \psi_0(\mathbf{v}) + \varepsilon\left(\psi_1(\mathbf{v},\nabla c) + \frac{q'(\rho)}{q(\rho)}(\mathbf{v}\cdot\nabla\rho)\psi_0(\mathbf{v})\right) \ .$$

With a similar argument as in the proof of Theorem 2.1, we can show that the approximate generalized kinetic model (21) converges to the same macroscopic model (7) in the diffusive limit $\varepsilon \to 0$.

To design an asymptotic preserving scheme which automatically preserves the macroscopic limit, a micro-macro formulation needs to be derived. We decompose the solution $f(t,\mathbf{x},\mathbf{v})$ as

$$f(t,\mathbf{x},\mathbf{v}) = \rho(t,\mathbf{x})\psi_0(\mathbf{v}) + \varepsilon g(t,\mathbf{x},\mathbf{v}) \ , \quad (22)$$

where $\langle g\rangle = 0$. Equations for ρ and g need to be derived then.

To get the equation for ρ, we substitute the micro-macro decomposition (22) into the modified model (21) and integrate over $\mathbf{v}$. Noticing that

$$\mathbf{v}\cdot\nabla(\tilde F_\varepsilon(\rho)f) = (\mathbf{v}\psi_0)\cdot\nabla(\rho q(\rho)) + \varepsilon\mathbf{v}\cdot\nabla\left[q(\rho)g + (\mathbf{v}\cdot\nabla\rho)\rho q'(\rho)\psi_0\right]$$
$$= (\mathbf{v}\psi_0)\cdot\nabla(\rho q(\rho)) + \varepsilon\mathbf{v}\cdot\nabla(q(\rho)g) + \varepsilon\nabla\cdot(\rho q'(\rho)(\mathbf{v}\otimes\mathbf{v})\psi_0\nabla\rho)$$

and $\langle\mathbf{v}\psi_0\rangle = 0$, $\langle\tilde T_\varepsilon\rangle = 1$, $\langle f\rangle = \rho$, we can finally derive that

$$\partial_t\rho + \langle\mathbf{v}\cdot\nabla(q(\rho)g)\rangle + \nabla\cdot(q'(\rho)\rho D_0\nabla\rho) = r_0\rho\left(1-\rho/\rho_{\max}\right)_+ \ . \quad (23)$$

To get the equation for g, we use the projection technique. For simplicity of notations, we introduce the identity operator I and the projection operator Π as

$$I(f(t,\mathbf{v},\mathbf{x})) = f(t,\mathbf{v},\mathbf{x}), \quad \Pi f(t,\mathbf{v},\mathbf{x}) = \langle f(t,\mathbf{v},\mathbf{x})\rangle\psi_0(\mathbf{v}) \ .$$

It is easy to check that,

$$
\begin{aligned}
&(I-\Pi)f = \varepsilon(I-\Pi)g = \varepsilon g \ ,\\
&(I-\Pi)(\mathbf{v}\cdot\nabla(\tilde{F}_\varepsilon f)) = (\mathbf{v}\psi_0)\cdot\nabla(q(\rho)\rho) + \varepsilon(I-\Pi)\Big[\mathbf{v}\cdot\nabla(q(\rho)g)\\
&\qquad\qquad + \nabla\cdot(\rho q'(\rho)(\mathbf{v}\otimes\mathbf{v})\psi_0\nabla\rho)\Big] \ ,\\
&(I-\Pi)\left(q(\rho)(-f+\rho\tilde{T}_\varepsilon)\right) = \varepsilon\left(\rho q(\rho)\psi_1 + \rho q'(\rho)(\mathbf{v}\cdot\nabla\rho)\psi_0 - q(\rho)g\right) \ ,\\
&(I-\Pi)\left(r_0 f\left(1-\rho/\rho_{\max}\right)\right) = \varepsilon r_0 g\left(1-\rho/\rho_{\max}\right)_+ \ .
\end{aligned}
$$

Then taking the operator $I-\Pi$ on equation (21), we get

$$
\begin{aligned}
&\partial_t g + \frac{1}{\varepsilon}(I-\Pi)\left[\mathbf{v}\cdot\nabla(q(\rho)g) + \nabla\cdot(\rho q'(\rho)(\mathbf{v}\otimes\mathbf{v})\nabla\rho\psi_0)\right]\\
=&\frac{1}{\varepsilon^2}\left[-q(\rho)(\mathbf{v}\cdot\nabla\rho)\psi_0 + \rho q(\rho)\psi_1 - q(\rho)g\right] + r_0 g\left(1-\rho/\rho_{\max}\right)_+ \ . \qquad (24)
\end{aligned}
$$

As a summary, by decomposing f as (22), the following micro-macro formulation of the system is derived

$$
\begin{cases}
\partial_t\rho + \langle\mathbf{v}\cdot\nabla(q(\rho)g)\rangle + D_0\nabla\cdot(\rho q'(\rho)\nabla\rho) = r_0\rho\left(1-\rho/\rho_{\max}\right)_+ \ ,\\
\partial_t g + \frac{1}{\varepsilon}(I-\Pi)K_\varepsilon = \frac{1}{\varepsilon^2}(S_\varepsilon - q(\rho)g) + r_0 g\left(1-\rho/\rho_{\max}\right)_+ \ ,\\
\Delta c + \rho - c = 0 \ ,
\end{cases} \qquad (25)
$$

where $D_0 = \langle(\mathbf{v}\otimes\mathbf{v})\psi_0\rangle$, and

$$
\begin{aligned}
K_\varepsilon &= \mathbf{v}\cdot\nabla(q(\rho)g) + \nabla\cdot(\rho\psi_0(\mathbf{v}\otimes\mathbf{v})\nabla q(\rho)) \ ,\\
S_\varepsilon &= -q(\rho)(\mathbf{v}\cdot\nabla\rho)\psi_0 + \rho q(\rho)\psi_1 \ .
\end{aligned}
$$

With a sufficiently large domain, we expect f as well as c will almost reach a steady state at the boundary.

Here we formally show that the micro-macro formulation derived recovers the macroscopic limit as $\varepsilon\to 0$. In fact, the leading order term in (24) shows that

$$
q(\rho)g = S_\varepsilon = -q(\rho)(\mathbf{v}\cdot\nabla\rho)\psi_0 + \rho q(\rho)\psi_1
$$

in the limit $\varepsilon\to 0$. Therefore,

$$
\langle\mathbf{v}\cdot\nabla(q(\rho)g)\rangle = \nabla\cdot\left[-D_0 q(\rho)\nabla\rho + \langle\mathbf{v}\psi_1\rangle\rho q(\rho)\right] \ .
$$

Substituting it into (23), we recover

$$
\begin{cases}
\partial_t\rho - \nabla\cdot\left[D_0(q(\rho)-\rho q'(\rho))\nabla\rho - \langle\mathbf{v}\psi_1\rangle\rho q(\rho)\right] = r_0\rho\left(1-\rho/\rho_{\max}\right)_+ \ ,\\
\Delta c + \rho - c = 0 \ .
\end{cases} \qquad (26)
$$

In Section 2.3 (see Remark 2.4), we show that equation (26) admits an energy functional that decreases in time providing we have the condition $\rho_{\max}(1+\frac{1}{\tilde{\rho}})\le 1$.

4. **An asymptotic preserving finite difference scheme.** To facilitate a more straightforward analysis of the asymptotic preserving property based on the micro-macro decomposition, we adopt the finite difference discretization. This choice enables a fully explicit formulation and simplifies the theoretical verification of the asymptotic preserving property. To describe the fully discretized scheme, we consider the 1D case for simplicity, i.e. $x, v \in [x_{\min}, x_{\max}]\times[v_{\min}, v_{\max}]$ with periodic boundary conditions in the x-direction and zero boundary conditions in the v-direction. The generalization to the multidimensional case with tensor product

grids is straightforward and is included in Section A. We use a uniformly distributed mesh with

$$t_n = n\Delta t\ , \quad x_j = j\Delta x\ , \quad v_k = k\Delta v\ ,$$

where $n \geq 0$, $j = 0, 1, \cdots, N_x - 1$, $k = 0, 1, \cdots, N_v$, $N_x = (x_{\max} - x_{\min})/\Delta x$ and $N_v = (v_{\max} - v_{\min})/\Delta v$. For the unknown functions $\rho(t, x)$ and $g(t, x, v)$, we compute their approximations ρ_j^n and $g_{j+\frac{1}{2},k}^n$ with

$$\rho_j^n \approx \rho(t_n, x_j)\ , \quad \text{and} \quad g_{j+\frac{1}{2},k}^n \approx g(t_n, x_{j+\frac{1}{2}}, v_k)\ .$$

Note that for the convenience of numerical computation, the approximation of the density function $\rho(t, x)$ is computed on grid points x_j, while the perturbation function $g(t, x, v)$ is computed on half grid points $x_{j+\frac{1}{2}}$. Approximations of the density function $\rho(t, x)$ at half grid points can then be efficiently computed by interpolation. To be more precise, $\rho(t_n, x_{j+\frac{1}{2}}) \approx \bar{\rho}_{j+\frac{1}{2}}^n := (\rho_j^n + \rho_{j+1}^n)/2$.

For simplicity of notations, we further introduce the standard finite difference operators δ_t^+ and δ_x, which are numerical approximations of ∂_t and ∂_x, respectively, and defined as

$$\delta_t^+ \rho_j^n = \frac{\rho_j^{n+1} - \rho_j^n}{\Delta t}\ , \ \delta_x \rho_{j+\frac{1}{2}}^n = \frac{\rho_{j+1}^n - \rho_j^n}{\Delta x}\ , \ \delta_x g_{j,k}^n = \frac{g_{j+\frac{1}{2},k}^n - g_{j-\frac{1}{2},k}^n}{\Delta x}\ .$$

The composite of two operators δ_x, which is denoted as δ_x^2, is then defined to be

$$\delta_x^2 \rho_j^n = \frac{\delta_x \rho_{j+\frac{1}{2}}^n - \delta_x \rho_{j-\frac{1}{2}}^n}{\Delta x} = \frac{\rho_{j+1}^n - 2\rho_j^n + \rho_{j-1}^n}{(\Delta x)^2}\ ,$$

which is the numerical approximation of ∂_x^2. The standard finite difference operators can be applied to a multiplication of two functions. As an example, we define

$$\delta_x (q(\bar{\rho}_*^n) g_{*,k}^n)_j = \frac{q(\bar{\rho}_{j+\frac{1}{2}}^n) g_{j+\frac{1}{2},k}^n - q(\bar{\rho}_{j-\frac{1}{2}}^n) g_{j-\frac{1}{2},k}^n}{\Delta x}\ ,$$

where we use * to denote the positions where the sub-index j is substituted. Another important notation to be introduced is $\langle \cdot \rangle_h$, which is defined as

$$\langle \eta_{j,k}^n \rangle_h := \Delta v \sum_k \eta_{j,k}^n\ ,$$

where $\eta_{j,k}^n \approx \eta(t_n, x_j, v_k)$ for some general function $\eta(t, x, v)$. Obviously, $\langle \eta_{j,k}^n \rangle_h$ is the finite difference approximation of $\langle \eta(t_n, x_j, v) \rangle := \int_V \eta(t_n, x_j, v)\, \mathrm{d}v$. Then $D_0 := \langle v^2 \psi_0(v) \rangle$ can be approximated by

$$D_h := \langle v_k^2 \psi_0(v_k) \rangle_h\ . \tag{27}$$

Finally, to better approximate $q(\rho)\rho$ at $x = x_{j+\frac{1}{2}}$ and $t = t_n$, we introduce the notation $\Phi_{j+\frac{1}{2}}^{n_1,n}$, which is defined as

$$\Phi_{j+\frac{1}{2}}^{n_1,n} = \begin{cases} \rho_j^{n_1} q(\rho_{j+1}^n)\ , & \text{if } \delta_x c_{j+\frac{1}{2}}^n \geq 0\ , \\ \rho_{j+1}^{n_1} q(\rho_j^n)\ , & \text{if } \delta_x c_{j+\frac{1}{2}}^n < 0\ , \end{cases} \tag{28}$$

where $n_1 = n$ or $n+1$. As shown in Ref. [1], $\Phi_{j+\frac{1}{2}}^{n_1,n}$ approximates $q(\rho)\rho$ at $t = t_n$ and $x = x_{j+\frac{1}{2}}$ in an upwind manner and thus helps improve the numerical stability when discretizing the macroscopic model (7). To guarantee that the numerical scheme designed for the micro-macro formulation (25) shares the same stability property in the limit $\varepsilon \to 0$, we keep using $\Phi_{j+\frac{1}{2}}^{n_1,n}$ to approximate $q(\rho)\rho$ appearing in (25).

With the notations defined, the system (25) can be discretized as

$$\begin{cases} \delta_t^+ \rho_j^n + \langle v_k \delta_x (q(\bar\rho_*^n) g_{*,k}^{n+1})_j \rangle_h + D_h \delta_x (\bar\rho_*^n q'(\bar\rho_*^n) \delta_x \rho_*^{n+1})_j = r_0 \rho_j^n \left(1 - \frac{\rho_j^n}{\rho_{\max}}\right)_+ , \\ \delta_t^+ g_{j+\frac12,k}^n + \frac{(I-\Pi_h) K_{j+\frac12,k}^n}{\varepsilon} = \frac{S_{j+\frac12,k}^{n,n+1} - q(\bar\rho_{j+\frac12}^n) g_{j+\frac12,k}^{n+1}}{\varepsilon^2} + r_0 g_{j+\frac12,k}^n \left(1 - \frac{\rho_{j+\frac12}^n}{\rho_{\max}}\right)_+ , \\ \delta_x^2 c_j^{n+1} + \rho_j^{n+1} - c_j^{n+1} = 0 , \end{cases} \tag{29}$$

where Π_h is the discrete projection operator defined as $\Pi_h \eta_{j,k}^n = \langle \eta_{j,k}^n \rangle_h \psi_0(v_k)$ for some general function $\eta(t,x,v)$, and

$$K_{j+\frac12,k}^n = v_k^+ \delta_x (q(\bar\rho_*^n) g_{*,k}^n)_j - v_k^- \delta_x (q(\bar\rho_*^n) g_{*,k}^n)_{j+1} + v_k^2 \psi_0(v_k) \delta_x (\bar\rho_*^n q'(\bar\rho_*^n) \delta_x \rho_*^n)_{j+\frac12} ,$$
$$S_{j+\frac12,k}^{n,n+1} = -v_k \psi_0(v_k) q(\bar\rho_{j+\frac12}^n) \delta_x \rho_{j+\frac12}^{n+1} + \psi_1(v_k, \delta_x c_{j+\frac12}^n) \Phi_{j+\frac12}^{n+1,n} ,$$

where $v^+ = \max\{v, 0\}$ and $v^- = \max\{-v, 0\}$.

Following the idea in Ref. [25], the scheme (29) can be solved efficiently. Instead of solving the system (29) directly, where all densities ρ_j^{n+1} and perturbations $g_{j+\frac12,k}^{n+1}$ are coupled so that a large linear system needs to be inverted, we introduce $\tilde g_{j+\frac12,k}^{n+1}$, which satisfies

$$\tilde\delta_t^+ g_{j+\frac12,k}^n + \frac{(I-\Pi_h) K_{j+\frac12,k}^n}{\varepsilon} = \frac{\tilde S_{j+\frac12,k}^n - q(\bar\rho_{j+\frac12}^n) \tilde g_{j+\frac12,k}^{n+1}}{\varepsilon^2} + r_0 g_{j+\frac12,k}^n \left(1 - \frac{\rho_{j+\frac12}^n}{\rho_{\max}}\right)_+ , \tag{30}$$

where

$$\tilde\delta_t^+ g_{j+\frac12,k}^n = \frac{\tilde g_{j+\frac12,k}^{n+1} - g_{j+\frac12,k}^n}{\Delta t} ,$$
$$\tilde S_{j+\frac12,k}^n = -v_k \psi_0(v_k) q(\bar\rho_{j+\frac12}^n) \delta_x \rho_{j+\frac12}^n + \psi_1(v_k, \delta_x c_{j+\frac12}^n) \Phi_{j+\frac12}^{n,n} .$$

By reformulating (30), it is easy to see that all the unknowns $\tilde g_{j+\frac12,k}^{n+1}$ can be solved explicitly. Furthermore, the unknowns $\tilde g_{j+\frac12,k}^{n+1}$ and $g_{j+\frac12,k}^{n+1}$ are closely related by

$$g_{j+\frac12,k}^{n+1} = \tilde g_{j+\frac12,k}^{n+1} + \frac{1}{\varepsilon^2} \left(\frac{1}{\Delta t} + \frac{q(\bar\rho_{j+\frac12}^n)}{\varepsilon^2} \right)^{-1} (S_{j+\frac12,k}^{n,n+1} - \tilde S_{j+\frac12,k}^n), \tag{31}$$

which can be easily derived by comparing (30) and the second equation in (29). Then by substituting (31) into the first equation in (29), a system that contains only the unknowns ρ_j^{n+1} is derived. Specifically, we have

$$\frac{\rho_j^{n+1}}{\Delta t} - \delta_x (a_*^n \delta_x \rho_*^{n+1})_j + \delta_x (b_*^n \Phi_*^{n+1,n})_j + D_h \delta_x (\rho_*^n q'(\rho_*^n) \delta_x \rho_*^{n+1})_j = r_j^n , \tag{32}$$

where the coefficients $a_{j+\frac12}^n$, $b_{j+\frac12}^n$ and residuals r_j^n can be explicitly computed via

$$a_{j+\frac12}^n = \frac{q(\bar\rho_{j+\frac12}^n)\Delta t}{\varepsilon^2 + q(\bar\rho_{j+\frac12}^n)\Delta t} D_h q(\bar\rho_{j+\frac12}^n) ,$$
$$b_{j+\frac12}^n = \frac{q(\bar\rho_{j+\frac12}^n)\Delta t}{\varepsilon^2 + q(\bar\rho_{j+\frac12}^n)\Delta t} \langle v_k \psi_1(v_k, \delta_x c_{j+\frac12}^n) \rangle_h ,$$

$$r_j^n = \frac{\rho_j^n}{\Delta t} - \langle v_k \delta_x (q_*^n \tilde{g}_{*,k}^{n+1})_j \rangle_h + r_0 \rho_j^n \left(1 - \frac{\rho_j^n}{\rho_{\max}}\right)_+ - \delta_x (a_*^n \delta_x \rho_*^n)_j + \delta_x (b_*^n \Phi_*^{n,n})_j \ .$$

To solve all the unknowns ρ_j^{n+1} from the system (32), only a tridiagonal matrix needs to be inverted. And then the unknowns $g_{j+\frac{1}{2},k}^{n+1}$ can be solved explicitly via (31). In this way, we efficiently update the system (29) from $t = t_n$ to $t = t_{n+1}$.

4.1. **Asymptotic preserving property.** Here, we formally check the asymptotic preserving property of the scheme by taking $\varepsilon \to 0$ in the system (29). We only need to show that the scheme for the kinetic model (21) converges to a scheme for the corresponding macroscopic model (7), or the first equation in (26). By checking the order of ε of each term in the equation for the perturbation function g, i.e. the second equation in (29), it is easy to see that, as $\varepsilon \to 0$, we should have $S_{j+\frac{1}{2},k}^{n,n+1} = q(\bar{\rho}_{j+\frac{1}{2}}^n) g_{j+\frac{1}{2},k}^{n+1}$, namely

$$q(\bar{\rho}_{j+\frac{1}{2}}^n) g_{j+\frac{1}{2},k}^{n+1} = -v_k \psi_0(v_k) q(\bar{\rho}_{j+\frac{1}{2}}^n) \delta_x \rho_{j+\frac{1}{2}}^{n+1} + \psi_1(v_k, \delta_x c_{j+\frac{1}{2}}^n) \Phi_{j+\frac{1}{2}}^{n+1,n} \ . \tag{34}$$

Combining (27) and (34), a direct computation shows that

$$\langle v_k \delta_x (q(\bar{\rho}_*^n) g_{*,k}^{n+1})_j \rangle_h = \delta_x (-D_h q(\bar{\rho}_*^n) \delta_x \rho_*^{n+1} + \langle v_k \psi_1(v_k, \delta_x c_*^n) \rangle_h \Phi_*^{n+1,n})_j \ . \tag{35}$$

Finally, by substituting (35) into the first equation in (29), we get

$$\begin{cases} \delta_t^+ \rho_j^n - \delta_x \left[D_h d(\bar{\rho}_*^n) \delta_x \rho_*^{n+1} - \langle v_k \psi_1(v_k, \delta_x c_*^n) \rangle_h \Phi_*^{n+1,n} \right]_j = r_0 \rho_j^n \left(1 - \frac{\rho_j^n}{\rho_{\max}}\right)_+ , \\ \delta_x^2 c_j^{n+1} + \rho_j^{n+1} - c_j^{n+1} = 0 \ , \end{cases} \tag{36}$$

where $d(\bar{\rho}_*^n) = q(\bar{\rho}_*^n) - \bar{\rho}_*^n q'(\bar{\rho}_*^n)$, which is indeed the finite difference system for solving (26). In this way, we verified the asymptotic preserving property of our scheme (29).

4.2. **Bound-preserving property.** Though the scheme (32) might not be positive preserving for a general fixed $\Delta t > 0$ and $\varepsilon > 0$, the following proposition shows that its limit (36) as $\varepsilon \to 0^+$ is positive preserving if $q(\rho) = 1 - (\rho/\bar{\rho})^\gamma$, where $\gamma \geq 1$ and $\bar{\rho} \geq \rho_{\max}$. The above choice of the squeezing probability function is commonly used for semi-elastic entities as described in the Introduction. A direct computation shows that, with $q(\rho) = 1 - (\rho/\bar{\rho})^\gamma$, the following is always non-negative,

$$d(\rho) = q(\rho) - \rho q'(\rho) = 1 + (\gamma - 1) \left(\frac{\rho}{\bar{\rho}}\right)^\gamma \geq 0 \ .$$

Moreover, the limiting scheme preserves the lower and upper bounds of the solution under the CFL condition

$$\Delta t \leq \Delta x \ \min \Big(\frac{\bar{\rho} - \rho_{\max}}{2\bar{\rho} \max_j |\eta_{j+1/2}| + \rho_{\max} r_0 \Delta x}, \frac{\Gamma}{2 \max_j |\eta_{j+1/2}|} \Big), \tag{37}$$

where $\Gamma = \inf_{s \in [\rho_{\max}, \bar{\rho}]} \frac{\bar{\rho} - s}{\bar{\rho} q(s)}$ and $\eta_{j+1/2} = \langle v_k \psi_1(v_k, \delta_x c_{j+\frac{1}{2}}^n) \rangle_h$. This CFL condition ensures that if $0 \leq \rho_j^0 \leq \bar{\rho}$ for all i then $0 \leq \rho_j^n \leq \bar{\rho}$ for all j and n.

Proposition 4.1. *With a general non-negative function $d(\rho) := q(\rho) - \rho q'(\rho)$, and under the CFL condition (37), if $0 \leq \rho_0^n \leq \bar{\rho}$ for all j, then, we have $0 \leq \rho_j^n \leq \bar{\rho}$ for all j in (36).*

Proof. Positive preservation

We prove by induction. Assuming that $\rho_j^n \geq 0$ for all j, we aim to show that $\rho_j^{n+1} \geq 0$ holds true for all j. Noticing that $\eta^n_{j+\frac{1}{2}}\Phi^{n+1,n}_{j+\frac{1}{2}} = (\eta^n_{j+\frac{1}{2}})_+ q(\rho^n_{j+1})\rho_j^{n+1} - (\eta^n_{j+\frac{1}{2}})_- q(\rho^n_j)\rho^{n+1}_{j+1}$ via (28), the numerical scheme (36) can be reformulated in the matrix form

$$M^n \boldsymbol{\rho}^{n+1} = \boldsymbol{r}^n \ , \tag{38}$$

where $M^n = (m^n_{i,j})$ is a tridiagonal matrix and $\boldsymbol{r}^n = (r^n_j)$ is a vector with

$$m^n_{j,j} = 1 + \Delta t \left[D_h \frac{d(\bar\rho^n_{j+\frac{1}{2}}) + d(\bar\rho^n_{j-\frac{1}{2}})}{(\Delta x)^2} + \frac{(\eta^n_{j+\frac{1}{2}})_+ q(\rho^n_{j+1}) + (\eta^n_{j-\frac{1}{2}})_- q(\rho^n_{j-1})}{\Delta x} \right] \geq 0 \ ,$$

$$m^n_{j,j+1} = -\frac{\Delta t}{(\Delta x)^2} D_h d(\bar\rho^n_{j+\frac{1}{2}}) - \frac{\Delta t}{\Delta x}(\eta^n_{j+\frac{1}{2}})_- q(\rho^n_j) \leq 0 \ ,$$

$$m^n_{j,j-1} = -\frac{\Delta t}{(\Delta x)^2} D_h d(\bar\rho^n_{j-\frac{1}{2}}) - \frac{\Delta t}{\Delta x}(\eta^n_{j-\frac{1}{2}})_+ q(\rho^n_j) \leq 0 \ ,$$

$$r^n_j = \rho^n_j + \Delta t\, r_0 \rho^n_j \left(1 - \frac{\rho^n_j}{\rho_{\max}}\right)_+ \geq 0 \ ,$$

where $(\eta)_+ = \max\{\eta, 0\} \geq 0$ and $(\eta)_- = \max\{-\eta, 0\} \geq 0$. Noticing that

$$m^n_{j,j} + m^n_{j-1,j} + m^n_{j+1,j} = 1 \ ,$$

the matrix M^n is strictly diagonal dominant in columns with all diagonal elements positive and off-diagonal elements non-positive. As a result, the matrix M^n is an M-matrix and thus inverse positive, i.e. all elements of its inverse $(M^n)^{-1}$ are non-negative. As a result, we must have $\boldsymbol{\rho}^{n+1} = (M^n)^{-1}\boldsymbol{r}^n \geq 0$.

Upper-bound. The upper-bound preservation is proved by induction as well. We assume that $\rho^n_j \leq \bar\rho$ and aim to show that $\rho^{n+1}_j \leq \bar\rho$ for all j. From (38), we write:

$$M^n(\mathbf{1}\bar\rho - \boldsymbol{\rho}^{n+1}) = M^n \mathbf{1}\bar\rho - \boldsymbol{r}^n \ ,$$

where $\mathbf{1} = (1, 1 \ldots, 1)$ is the unit vector. Using the fact that M^n is a M-matrix, we now aim to show that the vector $V^n = M^n\mathbf{1}\bar\rho - \boldsymbol{r}^n \geq 0$, ensuring that $(\mathbf{1}\bar\rho - \boldsymbol{\rho}^{n+1}) = (M^n)^{-1}\big(M^n\mathbf{1}\bar\rho - \boldsymbol{r}^n\big) \geq 0$, i.e $\rho^{n+1}_j \leq \bar\rho$ for all j.

Using Eq. (38), we compute

$$\begin{aligned} V^n_j &= \bar\rho(m^n_{j,j} + m^n_{j,j-1} + m^n_{j,j+1}) - r^n_j = \bar\rho + \frac{\bar\rho\Delta t}{\Delta x}[(\eta^n_{j+\frac{1}{2}})_+ q(\rho^n_{j+1}) + (\eta^n_{j-\frac{1}{2}})_- q(\rho^n_{j-1}) \\ &\quad - (\eta^n_{j+\frac{1}{2}})_- q(\rho^n_j) - (\eta^n_{j-\frac{1}{2}})_+ q(\rho^n_j)] - \rho^n_j(1 + \Delta t r_0(1 - \frac{\rho^n_j}{\rho_{\max}})_+) \\ &\geq \bar\rho - \frac{\bar\rho\Delta t}{\Delta x}[(\eta^n_{j+\frac{1}{2}})_- q(\rho^n_j) + (\eta^n_{j-\frac{1}{2}})_+ q(\rho^n_j)] - \rho^n_j(1 + \Delta t r_0(1 - \frac{\rho^n_j}{\rho_{\max}})_+) \\ &\geq \bar\rho - 2\frac{\bar\rho\Delta t}{\Delta x}\max_j |\eta^n_{j+\frac{1}{2}}| q(\rho^n_j) - \rho^n_j(1 + \Delta t r_0(1 - \frac{\rho^n_j}{\rho_{\max}})_+). \end{aligned}$$

For $\rho^n_j = \bar\rho$, we directly obtain $V^n_j \geq 0$. For $\rho^n_j < \rho_{\max}$, we have

$$\rho^n_j(1 + \Delta t r_0(1 - \frac{\rho^n_j}{\rho_{\max}})_+) < \rho_{\max}(1 + \Delta t r_0)$$

and therefore

$$V^n_j \geq \bar\rho - 2\frac{\bar\rho\Delta t}{\Delta x}\max_j |\eta^n_{j+\frac{1}{2}}| - \rho_{\max}(1 + \Delta t r_0),$$

arriving at the condition

$$\Delta t \le \frac{(\bar\rho - \rho_{\max})\Delta x}{2\bar\rho \max_j |\eta^n_{j+1/2}| + \rho_{\max} r_0 \Delta x}. \tag{39}$$

Finally, for $\rho_{\max} < \rho^n_j < \bar\rho$, we have

$$\rho^n_j(1 + \Delta t r_0(1 - \frac{\rho^n_j}{\rho_{\max}})_+) = \rho^n_j$$

and therefore

$$V^n_j \ge \bar\rho - 2\frac{\bar\rho \Delta t}{\Delta x} \max_j |\eta^n_{j+\frac12}| q(\rho^n_j) - \rho^n_j.$$

This case leads to requiring

$$\Delta t \le \frac{\Delta x}{2\max_j |\eta_{j+1/2}|} \min_{\rho_{\max} < \rho^n_j < \bar\rho} \frac{\bar\rho - \rho^n_j}{\bar\rho q(\rho^n_j)}. \tag{40}$$

Combining (39) and (40), we arrive at the CFL condition (37). □

The CFL condition (37) derived is not sharp. In practice, violating this condition does not necessarily lead to computational breakdown. Nevertheless, it provides useful guidance on the choice of time steps for ensuring stability.

5. **Numerical experiments.** In this section we present several numerical examples. In particular, we numerically verify the convergence of the kinetic model proposed in (21) (numerically solved with (29) from the micro-macro decomposition (25)), which we denote as $\rho^\varepsilon_{\text{kinetic}}$, to the volume-exclusion Keller-Segel model (26) (numerically solved with (36)), denoted as ρ_{macro}, as $\varepsilon \to 0$ in one and two dimensions.

5.1. **Energy dissipation and convergence tests in 1D.** In Section 2.3 we proved that, under some assumptions [5, 6, 1], the volume-exclusion Keller-Segel model (26) is energy dissipative, where the energy is defined by the functional

$$\mathcal{E}(t) = \int \Phi(\rho)\,\mathrm{d}x - \frac12 \int \rho c\,\mathrm{d}x\ , \tag{41}$$

with Φ satisfing Eq. (19). Unfortunately, it is difficult to prove the energy dissipation property theoretically for the numerical scheme. Instead, we will show it numerically for both the macro model and the kinetic model. Via numerical integration, we can accurately approximate $\Phi(\rho)$. The energy $\mathcal{E}(t)$ in (41) can then be numerically approximated via quadrature rules. For clarity, we will denote by $\mathcal{E}_\varepsilon(t)$ the value of the functional (41) computed on the solution of the kinetic system $\rho^\varepsilon_{\text{kinetic}}$ for a given ε at a given time t. The convergence of density profiles as $\varepsilon \to 0$ will be numerically tested as well. We will compare $\rho_{\text{macro}}(t,x)$ and $\rho^\varepsilon_{\text{kinetic}}(t,x)$ at specific time points and show the convergence rate by checking $\frac{\|\rho_{\text{macro}} - \rho^\varepsilon_{\text{kinetic}}\|_2}{\|\rho_{\text{macro}}\|_2}$ in the limit $\varepsilon \to 0$, where $||\cdot||_2$ is the L_2 norm.

For simplicity, we consider the 1D problem within the domain $(x,v) \in (-20,20)^2$. We use a uniform mesh with $\Delta x = 0.1, \Delta v = 0.2$. The periodic boundary condition is applied in the x-direction and the zero boundary condition is applied in the v-direction. In the simulations, we choose $r_0 = 0.1$, $\rho_{\max} = 0.5$, $\bar\rho = 1$ and

$$\psi_0(v) = \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}\ , \quad \psi_1(v, \partial_x c) = \frac{v}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} \partial_x c\ . \tag{42}$$

It is easy to check that the choices of $\psi_0(v)$ and $\psi_1(v)$ satisfy the Hypothesis H1 and Hypothesis H2.

Although the numerical scheme (29) works well in the diffusion limit $\varepsilon \to 0$, it is conditionally stable, and numerical experiments show that the restrictions on time step Δt might be quite strong for large ε values. In Figure 1, we evaluate the stability and computational efficiency of the scheme for different time steps and ε values. As demonstrated, maintaining stability for large values of ε requires significantly small time steps, which in turn greatly increases computational costs. Consequently, in subsequent 1D simulations, we limit ε to a maximum of 0.2 to avoid overly restrictive time step constraints. The instability observed for large values of ε is mainly due to the failure of the AP scheme to ensure the boundedness of the kinetic solution. Although it has been proven that the scheme ensures boundedness for the macroscopic limit model with sufficiently small time steps, no such theoretical result exists for the full kinetic model. This highlights a limitation of the current scheme in the kinetic regime with $\varepsilon = O(1)$, and addressing this issue will be an important direction for future work.

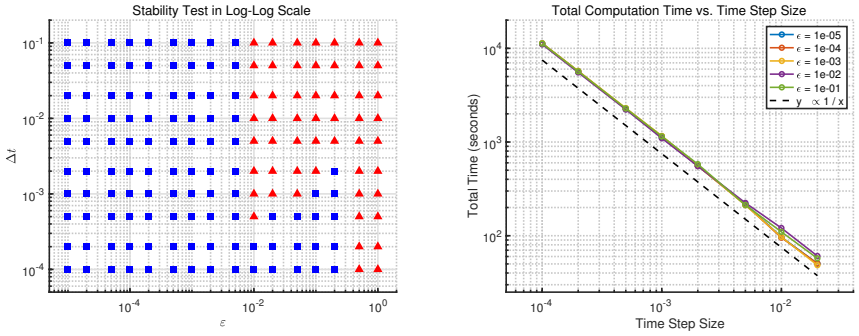


FIGURE 1. Stability and efficiency of the AP scheme for the kinetic model. The left panel evaluates stability by examining the boundedness of numerical solutions at $T = 20$ for various values of ε and Δt. Solutions confined within the range $[0, 1]$ are marked with blue squares, whereas those extending beyond this range are denoted by red triangles. The right panel illustrates the efficiency, demonstrating that the total computational time is roughly inversely proportional to the time step size for all ε values.

By choosing $\varepsilon = 0.2, 0.1, 0.05$ and a relatively small time step $\Delta t = 10^{-4}$, along with the initial data

$$\rho_{0_{\text{macro}}}(x) = \rho^{\varepsilon}_{0_{\text{kinetic}}}(x) = 0.5 + u(x)\ , \quad c_0(x) = c_0^{\varepsilon} = 0.5\ , \quad g_0^{\varepsilon}(x, v) = 0\ ,$$

where $u(x)$ is a uniformly distributed random function ranging in $(-0.1, 0.1)$, we compute the solution until $t = 40$. In Figure 2 we start with a comparison between $\rho^{\varepsilon}_{\text{kinetic}}$, for different values of ε ($\varepsilon = 0.2$, purple curves, $\varepsilon = 0.1$, yellow curves and $\varepsilon = 0.05$, red curves) and ρ_{macro} (blue curves), for different simulation times $t = 5,\ 20,\ 32,\ 40$ (from upper left to bottom right panels, respectively). When the aggregates are forming ($t = 5$) or merging together ($t = 32$), the discrepancy between the kinetic and the macroscopic solutions are larger, specially for large values of ε (purple line). As time progresses ($t = 40$) this difference becomes

smaller and we observe a very good agreement between the solutions of the kinetic and the macroscopic models for small values of the scaling parameter ε.

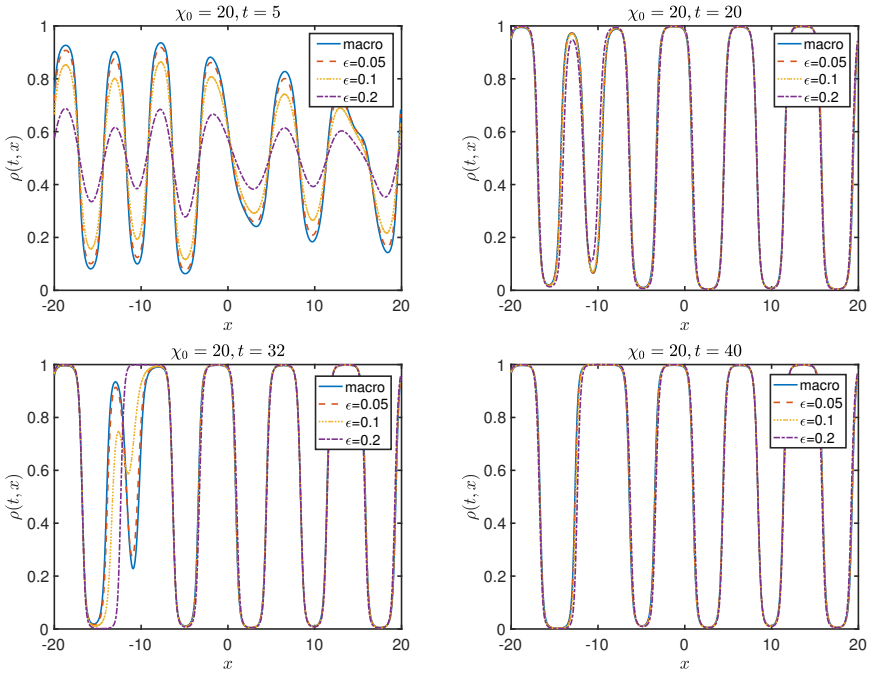


FIGURE 2. Comparison of ρ_{macro} and $\rho^{\varepsilon}_{\text{kinetic}}$ for $\chi_0 = 20$, and different values of ε.

In Figure 3 we show the evolution of the energy quantities $\mathcal{E}(t)$ given by (41) (blue curve) and $\mathcal{E}_\varepsilon(t)$ as functions of time, for different values of ε: $\varepsilon = 0.2$ (purple curve), $\varepsilon = 0.1$ (yellow curve) and $\varepsilon = 0.05$ (red curve), up to $t = 50$. This figure shows that the energies of the kinetic and macroscopic models are in very good agreement. Consistent with the analysis of Section 2.3, we recover that the energy is decreasing in time for the macroscopic model. It is noteworthy that the energy defined by (41) also decreases along the solutions of the kinetic model for $\epsilon > 0$. These numerical results indicate that the kinetic model may possess an energy dissipation structure, the formal proof of which is left for future work. The inset figures show the evolution in time of the macroscopic density (continuous blue line) and the kinetic density for different values of ε (lines with the same style as for the energy). It is clear from these figures that the larger discrepancies between the kinetic and macro energies are indeed related with changes in the density profiles, for example when two aggregates merge together (see the inset plots at $t = 5$ and $t = 32$). Even in this critical case of aggregation formation we observe that the kinetic solution for $\varepsilon = 0.05$ agrees with the macroscopic solution.

A similar behaviour is observed in Figure 4 (left) where we plot the relative L_2-error between the kinetic and the macroscopic solutions as a function of time and for different values of ε ($\varepsilon = 0.2$, black curve, $\varepsilon = 0.1$, blue curve and $\varepsilon = 0.05$, red curve). In agreement with the behaviour observed in Figure 3 this error is larger at times $t = 5$ and $t = 32$, approximately, which corresponds to times where aggregates are merging.

In Figure 4 (right) we show the rate of convergence of the relative L_2-error between the kinetic, $\rho^{\varepsilon}_{\text{kinetic}}$, and macroscopic, ρ_{macro}, solutions for different values of ε, at different times. We observe that the error between both solutions decreases as ε decreases, and the convergence order is around 1.5 in ℓ^2 norm. Altogether, these first results suggest that the macroscopic and kinetic models are in good agreement for small values of ε, and that the kinetic model converges towards the macroscopic model as $\varepsilon \to 0$ in the 1D case. In Figure 4 (left) we see formations of aggregates at $t = 5$ and around $t = 32$, in agreement with Figure 3. For the different values of ε we observe that the formation of the aggregate at $t = 32$ occurs at different times, i.e., the kinetic model seems to converge faster to the "aggregated-state" compared to the macroscopic dynamics. These changes in speed could be a result of the diffusion scaling, where the macroscopic model is obtained in a regime where there are many velocity jumps but small net displacements in one order of time (see [3, 15, 34, 35] for thorough interpretations of the diffusion scaling). In the next section, we take a step further and analyze the evolution of the pattern sizes in time as function of the chemotactic sensitivity χ_0.

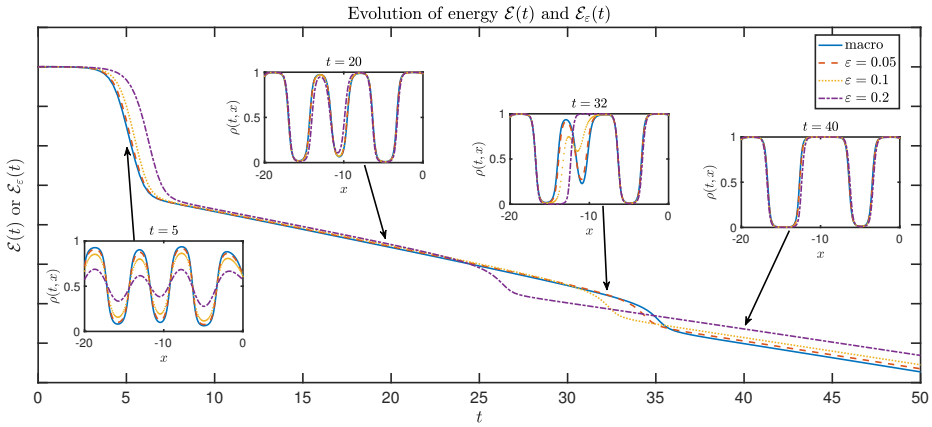


FIGURE 3. Evolution of $\mathcal{E}(t)$ and $\mathcal{E}_\varepsilon(t)$ along with the comparison between the kinetic solutions $\rho^{\varepsilon}_{\text{kinetic}}$ and the macroscopic solutions $\rho_{\text{macro}}(t,x)$ at $t = 5,\ 20,\ 32,\ 40$ for $\chi_0 = 20$.

5.2. **Pattern formation from a perturbed 1D initial data.** With a strong chemotaxis effect, cells will aggregate to form patterns in regions where the chemoattractant is highly concentrated. For the volume-exclusion Keller-Segel model (26), the relation between the aggregate size from an initial data perturbed from some homogeneous steady state and the strength of chemotaxis effect χ_0 was proven in [2] via linear stability analysis.

In this section, we numerically verify this relation for both the kinetic (21) and the macroscopic model (26). Again, we only consider here the 1D case with periodic boundary conditions in space. More specifically, we consider the domain $(x, v) \in (-20, 20)^2$ with a uniform mesh $\Delta x = 0.1$, $\Delta v = 0.2$. We choose the time step $\Delta t = 10^{-3}$ and starting from a randomly perturbed initial data

$$\rho_{0_{\text{macro}}}(t,x) = \rho^{\varepsilon}_{0_{\text{kinetic}}}(t,x) = 0.5 + u(x)\ ,$$

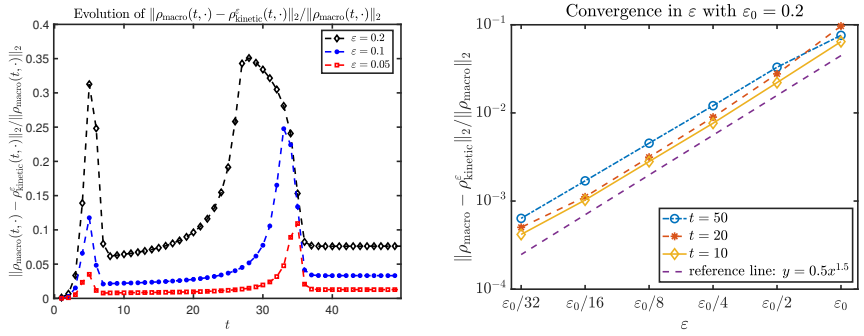


FIGURE 4. Left: Evolution of the relative L_2-error $\frac{\|\rho_{\text{macro}}(t,\cdot)-\rho_{\text{kinetic}}(t,\cdot)\|_2}{\|\rho_{\text{macro}}(t,\cdot)\|}$ over time with $\chi_0 = 20$. Right: Convergence of the relative L_2-error in ε at $t = 10,\ 20,\ 50$. The numerical setting is the same as in Figure 2.

we let the simulation run until $t = 20$. To avoid effects due to the randomness of the initial data, we will compute the pattern size for 10 solutions, each evolved from some random initial data, and simply average.

To numerically compute the pattern sizes, we consider the Fourier transform of the density function $\rho(t,x)$ (macro and kinetic) and extract the frequency that corresponds to the maximal Fourier mode. Specifically, we consider

$$k_{\max} = \text{argmax}_\lambda(|\hat{\rho}(\lambda)|)\ ,$$

where $\hat{\rho}(\lambda) = \mathcal{F}(\rho)(t,x)$ is the Fourier transform of the density function $\rho(t,x)$. Then, $1/k_{\max}$ can be used to describe the pattern size. With the numerical experimental settings in Section 5.1, here we compare the numerical results with the following formula, which was proposed in [2],

$$k_{\max}^{\text{analytical}} = \text{argmax}_k \left\{ -\frac{k^4 - \left(\frac{\chi_0}{4} - 1.1\right)k^2 + 0.1}{k^2+1} \right\} = \sqrt{\frac{\sqrt{\chi_0}}{2} - 1}.$$

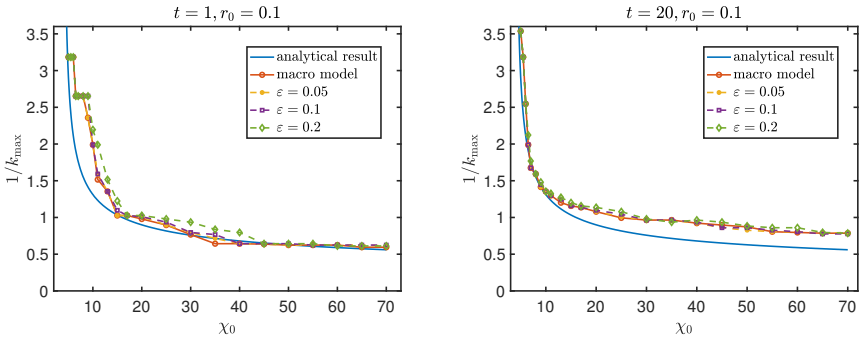


FIGURE 5. Comparison of $\frac{1}{k_{\max}}$ between the analytical, the numerical results of the macro and the kinetic model. The numerical result is computed by averaging 10 solutions, each evolved from a random initial data around 0.5.

In Figure 5 we show the overall pattern sizes $1/k_{\max}$ as a function of the chemotaxis sensitivity χ_0 at times $t = 1$ (left panel) and $t = 20$ (right panel). For each time, we plot the analytical prediction of $\frac{1}{k_{\max}}$ (blue line, see Ref. [2]), the numerical result for the macroscopic model (red line) and the results for the kinetic model with various values of ε ($\varepsilon = 0.05$ in yellow, $\varepsilon = 0.1$ in purple and $\varepsilon = 0.2$ in green lines). As one can observe, we obtain a very good agreement between the predicted pattern sizes and the ones computed numerically for both the macroscopic and kinetic models. With a strong chemotaxis sensitivity, cells aggregate more quickly to form patterns and, therefore, we can observe the agreement within a short time. It will take longer time to observe the agreement for cases with a relatively weak chemotaxis sensitivity. It explains why in Figure 5 the good agreement between the predicted pattern sizes and the numerical ones can be observed at $t = 1$ when χ_0 is large while the same phenomenon for small χ_0 can only be observed when $t = 20$. As predicted by the stability analysis performed in [2], the pattern sizes decrease as the chemotactic sensitivity χ_0 increases, and we recover the critical value $\chi_0^* \approx 6.9$ below which there are no patterns, i.e for which the perturbations are damped and the solution comes back to a homogeneous distribution.

5.3. **2D numerical examples.** The numerical schemes for both the kinetic model (29) and the macroscopic model (36) can be generalized to multi-dimensional problems, where the tensor-product grid is adopted (see Section A for a detailed description of the 2D numerical scheme for the kinetic model).

In this section we perform 2D simulations for both the kinetic model (21) and the volume-exclusion Keller-Segel model (26). We consider the computation domain $\Omega_{\mathbf{x}} = \{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : -20 \le x_1, x_2 \le 20\}$ with a uniform mesh $\Delta x_1 = \Delta x_2 = 0.1$ and periodic boundary conditions. For the kinetic model (21), we need to further define the domain of the velocity $\Omega_{\mathbf{v}} = \{\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2 : -10 \le v_1, v_2 \le 10\}$ with a uniform mesh $\Delta v_1 = \Delta v_2 = 0.2$ and zero boundary conditions. We choose $r_0 = 0.1$, $\rho_{\max} = 0.5$, $\bar{\rho} = 1$ and

$$\psi_0(\mathbf{v}) = \frac{1}{2\pi} e^{-\frac{|\mathbf{v}|^2}{2}}, \quad \psi_1(\mathbf{v}\ , \nabla_{\mathbf{x}} c) = \frac{\mathbf{v}}{2\pi} e^{-\frac{|\mathbf{v}|^2}{2}} \cdot \nabla_{\mathbf{x}} c\ .$$

As in the 1D case, we can check that the choices of $\psi_0(\mathbf{v})$ and $\psi_1(\mathbf{v})$ satisfy the Hypotheses H1 and H2.

We fix $\Delta t = 10^{-3}$ and choose the initial data to be

$$\rho_{0_{\text{macro}}}(\mathbf{x}) = \rho^{\varepsilon}_{0_{\text{kinetic}}}(\mathbf{x}) = \rho_0 + u(x)\ , \quad c_0(\mathbf{x}) = c_0^{\varepsilon}(\mathbf{x}) = 0.5\ ,$$

where $u(x)$ is a randomly chosen uniformly distributed function ranging in $(-0.1, 0.1)$. In Figure 6 we show the numerical results at $t = 5$ (first row), $t = 20$ (second row) and $t = 50$ (third row) for the macroscopic model with different chemotaxis sensitivities $\chi_0 = 6$ (first column), $\chi_0 = 20$ (second column) and $\chi_0 = 50$ (third column). As one can observe, starting from an initial data perturbed around the homogeneous value 0.5, we obtain the formation of labyrinthic patterns for a chemotactic sensitivity $\chi_0 > 6$ (middle and right columns), while the solution dampens to the homogeneous state for $\chi_0 = 6$ (left column), in agreement with the predictions of the stability analysis performed in [2] and the results in Figure 5. Moreover, we observe that larger values of the chemotactic sensitivity χ_0 lead to sharper layers near the boundary of the patterns (compare middle and right columns) as expected.

The convergence of the numerical patterns towards $\varepsilon \to 0$ is illustrated in Fig. 7. In this test, we choose $\varepsilon = 10^{-1}, 10^{-2}, 10^{-3}$ and consider chemotactic sensitivities

$\chi_0 = 20$ and $\chi_0 = 50$, along with two different initial densities: $\rho_0 = 0.5$ and $\rho_0 = 0.1$. As observed in [34], different initial conditions lead to qualitatively different pattern types for both kinetic and macroscopic models: labyrinthine patterns for $\rho_0 = 0.5$, and spot-like patterns for $\rho_0 = 0.1$. Moreover, we note that increasing χ_0 leads to smaller and sharper patterns in the kinetic model as well.

In order to quantify the differences between the kinetic and macroscopic 2D models, we show in Figure 8 (left) the evolution in time of the relative L^2-error between the macroscopic and kinetic models for different values of ε: $\varepsilon = 10^{-2}$ (blue curve), $\varepsilon = 10^{-4}$ (red curve), $\varepsilon = 10^{-6}$ (yellow curve), $\varepsilon = 10^{-8}$ (purple curve), and Figure 8 (right) shows this relative error as function of ε for different time points: $t = 1$ (blue curve), $t = 3$ (red curve), $t = 5$ (yellow curve) and $t = 10$ (purple curve). We observe that the relative error between both models decreases as ε decreases, and the reference line $y = x$ (green curve) shows that the rate of convergence of the kinetic model towards the macroscopic one is roughly $\mathcal{O}(\varepsilon)$.

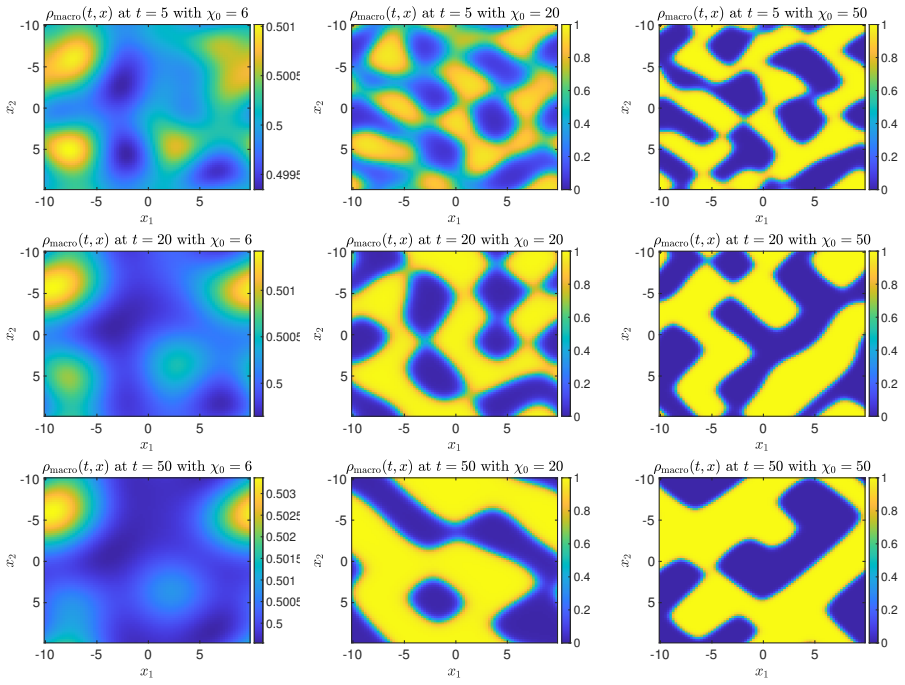


FIGURE 6. Plots of density $\rho_{\text{macro}}(t, \mathbf{x})$ at $t = 5,\ 20,\ 50$ (from top to bottom) with $\chi_0 = 6,\ 20,\ 50$ (from left to right), respectively.

6. **Conclusions.** In this paper we have derived a model for chemotaxis incorporating a density dependence in the chemotactic sensitivity function that takes into account the finite size of the cells and volume limitations. We showed, with formal arguments, that the macroscopic chemotactic system can be seen as the diffusion limit of a kinetic 'velocity-jump' model, provided that both the transport term and the turning operator are density dependent. This derivation provides a more direct interpretation of the diffusion tensor and chemotactic sensitivity in terms of more fundamental characteristics of the motion.

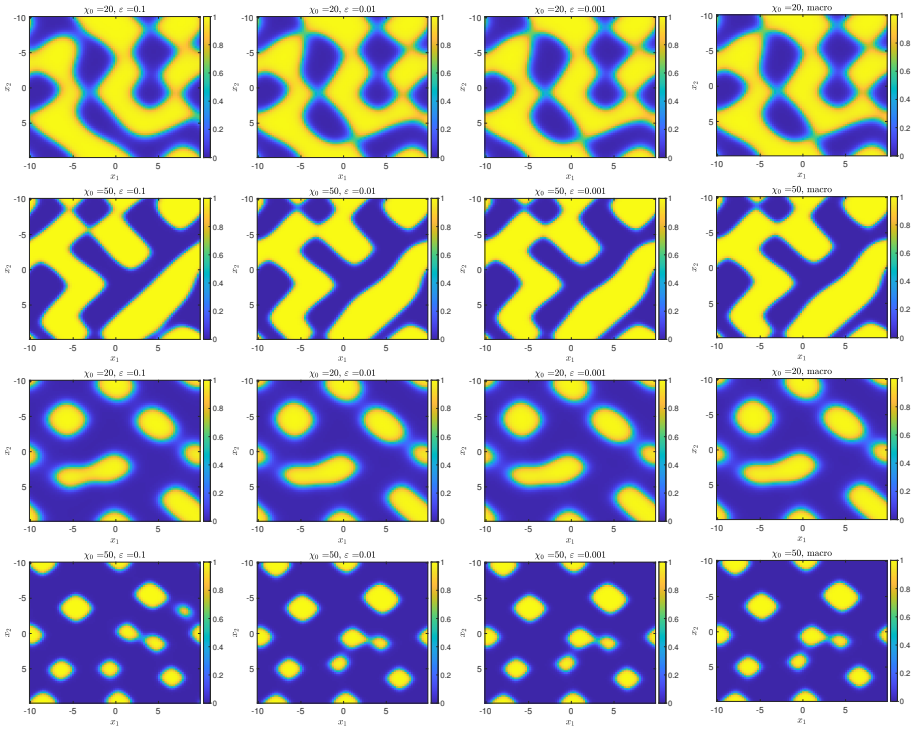


FIGURE 7. Comparison of $\rho^{\varepsilon}_{\text{kinetic}}(t,\mathbf{x})$ for $\varepsilon = 10^{-1}, 10^{-2}, 10^{-3}$ and $\rho_{\text{macro}}(t,\mathbf{x})$ at $T = 20$. In the numerical tests, we set $\chi_0 = 20$ for the first and third rows, and $\chi_0 = 50$ for the second and fourth rows. Additionally, the initial values are approximately 0.5 for the first and second rows, and around 0.1 for the third and fourth rows.

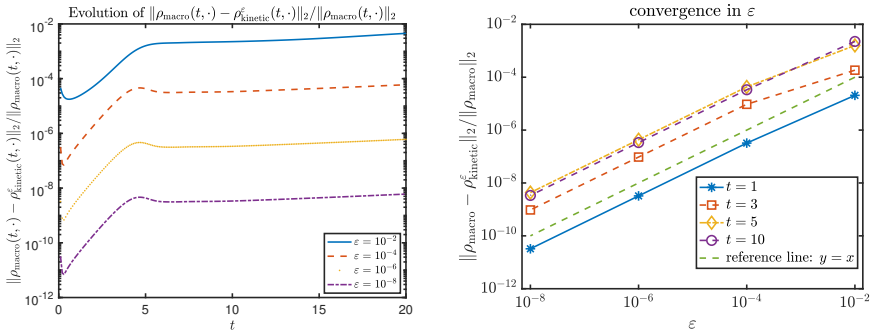


FIGURE 8. Left: Evolution of the relative L_2-error. Right: Convergence of the relative L_2-error in ε at $t = 1,\ 3,\ 5,\ 10$ with $\chi_0 = 20$ and $\Delta t = 10^{-2}$.

We further studied this macroscopic limit numerically using an asymptotic preserving finite difference scheme based on a micro-macro decomposition of the unknown in the sense of [25], a projection technique to obtain a coupled system of two evolution equations for the microscopic and macroscopic components, and a

suitable semi-implicit time discretization. The scheme was successfully extended to account for nonlinear terms by implicit-explicit discretization in an upwind manner, allowing for accurate approximations in the case of strong chemosensitivity. This scheme enabled us to explore numerically the different behaviours observed by the kinetic and macroscopic models in 1D and 2D, and we showed that both models are in good agreement as the diffusion scaling parameter becomes smaller. Moreover, the numerical simulations of the kinetic model revealed the same pattern sizes as obtained with the macroscopic model and predicted theoretically, with very good precision as ε goes to zero in the kinetic setting.

From the modeling perspective, it would be natural to extend the derivation to consider different turning kernels, to take into account cell-cell adhesion or nonlocal movement, for instance. The idea to construct the scheme could be generalized to include these cases, but we stress the fact that the detailed discretization is problem-dependent. Moreover, the rigorous derivation of volume-filling chemotactic equations from stochastic processes of interacting populations could be considered by adapting ideas from [37] for instance.

Appendix A. **A finite difference scheme for the 2D kinetic model.** The finite difference scheme (29) can be generalized to multi-dimensional problems where a tensor product grid is applied. Here we consider the 2D kinetic model with the special choice $\psi_0(v_1, v_2)$ and $\psi_1(v_1, v_2) = \vec{\phi}(v_1, v_2) \cdot \nabla c$, where $\vec{\phi}(v_1, v_2) = (\phi_1(v_1, v_2), \phi_2(v_1, v_2))^T$, such that $DI_2 = \int (\mathbf{v} \otimes \mathbf{v})\psi_0 \, \mathrm{d}v_1 \, \mathrm{d}v_2$ and $\chi_0 I_2 = \int \mathbf{v} \otimes \vec{\phi} \, \mathrm{d}v_1 \, \mathrm{d}v_2$ with $\mathbf{v} = (v_1, v_2)^T$ and I_2 being the 2-dimentional identity matrix.

We introduce the notations $\rho^n_{j_1,j_2}$, $c^n_{j_1,j_2}$, $g^{(1),n}_{j_1+\frac{1}{2},j_2,k_1,k_2}$ and $g^{(2),n}_{j_1,j_2+\frac{1}{2},k_1,k_2}$ to be the numerical approximations of $\rho(t_n, x_{j_1}, x_{j_2})$, $c(t_n, x_{j_1}, x_{j_2})$, $g(t_n, x_{j_1+\frac{1}{2}}, x_{j_2}, v_{k_1}, v_{k_2})$ and $g(t_n, x_{j_1}, x_{j_2+\frac{1}{2}}, v_{k_1}, v_{k_2})$, respectively. The approximations of $\rho(t, \mathbf{x})$ at half grid points such as $(x_{j_1}, x_{j_2+\frac{1}{2}})$ can be then easily approximated by the average $\bar{\rho}_{j_1,j_2+\frac{1}{2}} := (\rho_{j_1,j_2} + \rho_{j_1,j_2+1})/2$. It is worth noticing that we used different notations for approximating $g(t_n, x_{j_1+\frac{1}{2}}, x_{j_2}, *, *)$ and $g(t_n, x_{j_1}, x_{j_2+\frac{1}{2}}, *, *)$ since different upwind discretizations will be used depending on whether the half grid is in x_1-direction or x_2-direction. An illustration of the grids in $\mathbf{x}$-space for computing $\rho(t, \mathbf{x})$ and $g(t, \mathbf{x}, \mathbf{v})$ in 1D and 2D can be found in Figure 9.

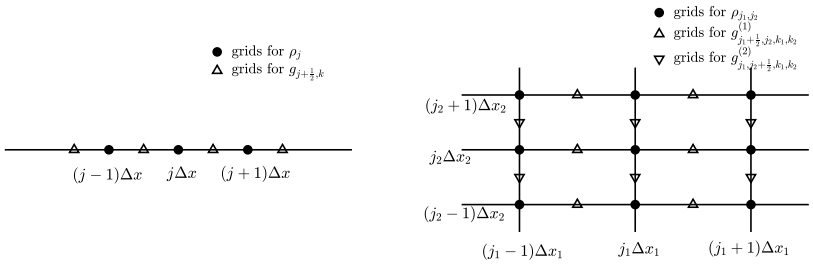


FIGURE 9. Illustration of the grids in $\mathbf{x}$-space for computing $\rho(t, \mathbf{x})$ and $g(t, \mathbf{x}, \mathbf{v})$ in 1D and 2D.

With the notations defined, the 2D kinetic model (21) can be discretized as

$$
\begin{cases}
\delta_t^+ \rho^n_{j_1,j_2} + \sum_{k_1,k_2} \left[v_{k_1} \delta_{x_1} (q(\bar\rho^n_{*,j_2}) g^{(1),n+1}_{*,j_2,k_1,k_2})_{j_1} + v_{k_2} \delta_{x_2} (q(\bar\rho^n_{j_1,*}) g^{(2),n+1}_{j_1,*,k_1,k_2})_{j_2} \right] \Delta v_1 \Delta v_2 \\
\quad + D_h \left[\delta_{x_1} (\bar\rho^n_{*,j_2} q'(\bar\rho^n_{*,j_2}) \delta_{x_1} \rho^{n+1}_{*,j_2})_{j_1} + \delta_{x_2} (\bar\rho^n_{j_1,*} q'(\bar\rho^n_{j_1,*}) \delta_{x_2} \rho^{n+1}_{j_1,*})_{j_2} \right] = r_0 \rho^n_{j_1,j_2} \left(1 - \frac{\rho^n_{j_1,j_2}}{\rho_{\max}} \right)_+ , \\
\delta_t^+ g^{(1),n}_{j_1+\frac12,j_2,k_1,k_2} + \frac{1}{\varepsilon} (I - \Pi_h) K^{(1),n}_{j_1+\frac12,j_2,k_1,k_2} \\
\quad = \frac{1}{\varepsilon^2} (S^{(1),n,n+1}_{j_1+\frac12,j_2,k_1,k_2} - q(\bar\rho^n_{j_1+\frac12,j_2}) g^{(1),n+1}_{j_1+\frac12,j_2,k_1,k_2}) + r_0 g^{(1),n}_{j_1+\frac12,j_2,k_1,k_2} \left(1 - \frac{\rho^n_{j_1+\frac12,j_2}}{\rho_{\max}} \right)_+ , \\
\delta_t^+ g^{(2),n}_{j_1,j_2+\frac12,k_1,k_2} + \frac{1}{\varepsilon} (I - \Pi_h) K^{(2),n}_{j_1,j_2+\frac12,k_1,k_2} \\
\quad = \frac{1}{\varepsilon^2} (S^{(2),n,n+1}_{j_1,j_2+\frac12,k_1,k_2} - q(\bar\rho^n_{j_1,j_2+\frac12}) g^{(2),n+1}_{j_1,j_2+\frac12,k_1,k_2}) + r_0 g^{(2),n}_{j_1,j_2+\frac12,k_1,k_2} \left(1 - \frac{\rho^n_{j_1,j_2+\frac12}}{\rho_{\max}} \right)_+ , \\
(\delta_{x_1}^2 + \delta_{x_2}^2) c^{n+1}_{j_1,j_2} + \rho^{n+1}_{j_1,j_2} - c^{n+1}_{j_1,j_2} = 0,
\end{cases}
\tag{43}
$$

where $D_h = \sum_{k_1,k_2} v_{k_i}^2 \psi_0(v_{k_1}, v_{k_2}) \Delta v_1 \Delta v_2$, $i = 1$ or 2, and Π_h is the discrete projection operator defined as

$$
\Pi_h \eta^n_{j_1,j_2,k_1,k_2} = \sum_{k_1,k_2} \eta^n_{j_1,j_2,k_1,k_2} \psi_0(v_{k_1}, v_{k_2}) \Delta v_1 \Delta v_2
$$

for some general function $\eta(t, \mathbf{x}, \mathbf{v})$ and

$$
\begin{aligned}
K^{(1),n}_{j_1+\frac12,j_2,k_1,k_2} &= v^+_{k_1} \delta_{x_1} (q(\bar\rho^n_{*,j_2}) g^{(1),n}_{*,j_2,k_1,k_2})_{j_1} - v^-_{k_1} \delta_{x_1} (q(\bar\rho^n_{*,j_2}) g^{(1),n}_{*,j_2,k_1,k_2})_{j_1+1} \\
&\quad + v^+_{k_2} \delta_{x_2} (q(\bar\rho^n_{j_1+\frac12,*}) g^{(1),n}_{j_1+\frac12,*,k_1,k_2})_{j_2-\frac12} - v^-_{k_2} \delta_{x_2} (q(\bar\rho^n_{j_1+\frac12,*}) g^{(1),n}_{j_1+\frac12,*,k_1,k_2})_{j_2+\frac12} \\
&\quad + \psi_0(v_{k_1}, v_{k_2}) \delta_{x_1} \left\{ \rho^n_{*,j_2} q'(\rho^n_{*,j_2}) \left[v^2_{k_1} \delta_{x_1} \bar\rho^n_{*,j_2} + v_{k_1} v_{k_2} (\delta_{x_2} \bar\rho^n_{*,**})_{j_2} \right] \right\}_{j_1+\frac12} , \\
&\quad + \psi_0(v_{k_1}, v_{k_2}) \delta_{x_2} \left\{ \bar\rho^n_{j_1+\frac12,*} q'(\bar\rho^n_{j_1+\frac12,*}) \left[v_{k_1} v_{k_2} (\delta_{x_1} \bar\rho^n_{**,*})_{j_1+\frac12} + v^2_{k_2} \delta_{x_2} \bar\rho^n_{j_1+\frac12,*} \right] \right\}_{j_2} , \\
K^{(2),n}_{j_1,j_2+\frac12,k_1,k_2} &= v^+_{k_1} \delta_{x_1} (q(\bar\rho^n_{*,j_2+\frac12}) g^{(1),n}_{*,j_2+\frac12,k_1,k_2})_{j_1-\frac12} - v^-_{k_1} \delta_{x_1} (q(\bar\rho^n_{*,j_2+\frac12}) g^{(1),n}_{*,j_2+\frac12,k_1,k_2})_{j_1+\frac12} \\
&\quad + v^+_{k_2} \delta_{x_2} (q(\bar\rho^n_{j_1,*}) g^{(1),n}_{j_1,*,k_1,k_2})_{j_2} - v^-_{k_2} \delta_{x_2} (q(\bar\rho^n_{j_1,*}) g^{(1),n}_{j_1,*,k_1,k_2})_{j_2+1} \\
&\quad + \psi_0(v_{k_1}, v_{k_2}) \delta_{x_1} \left\{ \bar\rho^n_{*,j_2+\frac12} q'(\bar\rho^n_{*,j_2+\frac12}) \left[v^2_{k_1} \delta_{x_1} \bar\rho^n_{*,j_2+\frac12} + v_{k_1} v_{k_2} (\delta_{x_2} \bar\rho^n_{*,**})_{j_2+\frac12} \right] \right\}_{j_1} , \\
&\quad + \psi_0(v_{k_1}, v_{k_2}) \delta_{x_2} \left\{ \rho^n_{j_1,*} q'(\rho^n_{j_1,*}) \left[v_{k_1} v_{k_2} (\delta_{x_1} \bar\rho^n_{**,*})_{j_1} + v^2_{k_2} \delta_{x_2} \bar\rho^n_{j_1,*} \right] \right\}_{j_2+\frac12} , \\
S^{(1),n,n+1}_{j_1+\frac12,j_2,k_1,k_2} &= -\psi_0(v_{k_1}, v_{k_2}) q(\bar\rho^n_{j_1+\frac12,j_2}) \left[v_{k_1} \delta_{x_1} \rho^{n+1}_{j_1+\frac12,j_2} + v_{k_2} \delta_{x_2} \bar\rho^{n+1}_{j_1+\frac12,j_2} \right] \\
&\quad + \phi_1(v_{k_1}, v_{k_2}) \delta_{x_1} c^n_{j_1+\frac12,j_2} \Phi^{(1),n+1,n}_{j_1+\frac12,j_2} + \phi_2(v_{k_1}, v_{k_2}) \delta_{x_2} c^n_{j_1+\frac12,j_2} \Phi^{(2),n+1,n}_{j_1+\frac12,j_2} , \\
S^{(2),n,n+1}_{j_1,j_2+\frac12,k_1,k_2} &= -\psi_0(v_{k_1}, v_{k_2}) q(\bar\rho^n_{j_1,j_2+\frac12}) \left[v_{k_1} \delta_{x_1} \bar\rho^{n+1}_{j_1,j_2+\frac12} + v_{k_2} \delta_{x_2} \rho^{n+1}_{j_1,j_2+\frac12} \right] \\
&\quad + \phi_1(v_{k_1}, v_{k_2}) \delta_{x_1} c^n_{j_1,j_2+\frac12} \Phi^{(1),n+1,n}_{j_1,j_2+\frac12} + \phi_2(v_{k_1}, v_{k_2}) \delta_{x_2} c^n_{j_1,j_2+\frac12} \Phi^{(2),n+1,n}_{j_1,j_2+\frac12} ,
\end{aligned}
$$

where, as in the 1D case, both $\Phi^{(1),n_1,n}$ and $\Phi^{(1),n_1,n}$ are upwind approximations of $q(\rho)\rho$ at $t = t_n$ and defined as

$$
\Phi^{(1),n_1,n}_{j_1,j_2} = \begin{cases} \rho^{n_1}_{j_1-\frac12,j_2} q(\rho^n_{j_1+\frac12,j_2}) , & \text{if } \delta_{x_1} c^n_{j_1,j_2} \geq 0 , \\ \rho^{n_1}_{j_1+\frac12,j_2} q(\rho^n_{j_1-\frac12,j_2}) , & \text{if } \delta_{x_1} c^n_{j_1,j_2} < 0 , \end{cases}
$$

$$
\Phi^{(2),n_1,n}_{j_1,j_2} = \begin{cases} \rho^{n_1}_{j_1,j_2-\frac12} q(\rho^n_{j_1,j_2+\frac12}) , & \text{if } \delta_{x_2} c^n_{j_1,j_2} \geq 0 , \\ \rho^{n_1}_{j_1,j_2+\frac12} q(\rho^n_{j_1,j_2-\frac12}) , & \text{if } \delta_{x_2} c^n_{j_1,j_2} < 0 , \end{cases}
$$

As for the 1D case, we can formally prove the asymptotic preserving property of the 2D scheme (43) in a similar way. In fact, when $\varepsilon \to 0$, we expect that

$$
S^{(1),n,n+1}_{j_1+\frac12,j_2,k_1,k_2} = q(\bar\rho^n_{j_1+\frac12,j_2}) g^{(1),n+1}_{j_1+\frac12,j_2,k_1,k_2}, \quad S^{(2),n,n+1}_{j_1,j_2+\frac12,k_1,k_2} = q(\bar\rho^n_{j_1,j_2+\frac12}) g^{(2),n+1}_{j_1,j_2+\frac12,k_1,k_2},
$$

from where we have

$$q(\bar\rho^n_{j_1+\frac12,j_2})g^{(1),n+1}_{j_1+\frac12,j_2,k_1,k_2} = -\psi_0(v_{k_1},v_{k_2})q(\bar\rho^n_{j_1+\frac12,j_2})\left[v_{k_1}\delta_{x_1}\rho^{n+1}_{j_1+\frac12,j_2} + v_{k_2}\delta_{x_2}\bar\rho^{n+1}_{j_1+\frac12,j_2}\right]$$
$$+ \phi_1(v_{k_1},v_{k_2})\delta_{x_1}c^n_{j_1+\frac12,j_2}\Phi^{(1),n+1,n}_{j_1+\frac12,j_2} + \phi_2(v_{k_1},v_{k_2})\delta_{x_2}c^n_{j_1+\frac12,j_2}\Phi^{(2),n+1,n}_{j_1+\frac12,j_2},$$
$$q(\bar\rho^n_{j_1,j_2+\frac12})g^{(2),n+1}_{j_1,j_2+\frac12,k_1,k_2} = -\psi_0(v_{k_1},v_{k_2})q(\bar\rho^n_{j_1,j_2+\frac12})\left[v_{k_1}\delta_{x_1}\bar\rho^{n+1}_{j_1,j_2+\frac12} + v_{k_2}\delta_{x_2}\rho^{n+1}_{j_1,j_2+\frac12}\right]$$
$$+ \phi_1(v_{k_1},v_{k_2})\delta_{x_1}c^n_{j_1,j_2+\frac12}\Phi^{(1),n+1,n}_{j_1,j_2+\frac12} + \phi_2(v_{k_1},v_{k_2})\delta_{x_2}c^n_{j_1,j_2+\frac12}\Phi^{(2),n+1,n}_{j_1,j_2+\frac12}.$$

Substituting into the first equation in (43) and using the fact that

$$\sum_{k_1,k_2} v_{k_i}v_{k_j}\psi_0(v_{k_1},v_{k_2})\Delta v_1\Delta v_2 = D_h\delta_{i,j}, \quad \sum_{k_1,k_2} v_{k_i}\phi_j(v_{k_1},v_{k_2})\Delta v_1\Delta v_2 = \chi_0\delta_{i,j},$$

for $i,j=1,2$, we recover the finite difference scheme for the macro model

$$\delta_t^+\rho^n_{j_1,j_2} - D_h\delta_{x_1}\left(d(\bar\rho^n_{*,j_2})\delta_{x_1}\rho^{n+1}_{*,j_2}\right)_{j_1} - D_h\delta_{x_2}\left(d(\bar\rho^n_{j_1,*})\delta_{x_2}\rho^{n+1}_{j_1,*}\right)_{j_2}$$
$$+ \chi_0\delta_{x_1}\left[\Phi^{(1),n+1,n}_{*,j_2}\delta_{x_1}c^n_{*,j_2}\right]_{j_1} + \chi_0\delta_{x_2}\left[\Phi^{(1),n+1,n}_{j_1,*}\delta_{x_2}c^n_{j_1,*}\right]_{j_2} = r_0\rho^n_{j_1,j_2}\left(1-\frac{\rho^n_{j_1,j_2}}{\rho_{\max}}\right)_+,$$

where $d(\rho^n_{j_1,j_2}) = q(\rho^n_{j_1,j_2}) - \rho^n_{j_1,j_2}q'(\rho^n_{j_1,j_2})$.

Acknowledgments. The authors wish to thank L. Almeida and K. J. Painter for helpful discussions and guidance. This work was supported by Sorbonne Alliance University with an Emergence project MATHREGEN, grant number S29-05Z101, by Agence Nationale de la Recherche (ANR) under the project grant number ANR-22-CE45-0024-01, the Fondation Sciences Mathématiques de Paris (FSMP), the Advanced Grant Nonlocal-CPD (Nonlocal PDEs for Complex Particle Dynamics: Phase Transitions, Patterns and Synchronization) of the European Research Council Executive Agency (ERC), the project MoGlimaging, Plan Cancer THE Call, from INSERM, France, the R&D Program of Beijing Municipal Education Commission from China, grant number KM202310028016, and the National Natural Science Foundation of China, grant number 12201436.

REFERENCES

[1] L. Almeida, F. Bubba, B. Perthame and C. Pouchol, Energy and implicit discretization of the Fokker-Planck and Keller-Segel type equations, *Netw. Heterog. Media*, **14** (2019), 23-41.
[2] L. Almeida, G. Estrada-Rodriguez, L. Oliver, D. Peurichard, A. Poulain and F. Vallette, Treatment-induced shrinking of tumour aggregates: A nonlinear volume-filling chemotactic approach, *J. Math. Biol.*, **83** (2021), Paper No. 29, 41 pp.
[3] W. Alt, Biased random walk models for chemotaxis and related diffusion approximations, *J. Math. Biol.*, **9** (1980), 147-177.
[4] M. Brenner, P. Constantin, L. Kadanoff, A. Schenkel and S. Venkataramani, Diffusion, attraction and collapse, *Nonlinearity*, **12** (1999), 1071-1098.
[5] V. Calvez and J. A. Carrillo, Volume effects in the Keller-Segel model: Energy estimates preventing blow-up, *J. Math. Pure. Appl.*, **86** (2006), 155-175.
[6] J. A. Carrillo, A. Jüngel, P. A. Markowich, G. Toscani and A. Unterreiter, Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities, *Monatsh. Math.*, **133** (2001), 1-82.
[7] J. A. Carrillo and B. Yan, An asymptotic preserving scheme for the diffusive limit of kinetic systems for chemotaxis, *Multiscale Model. Sim.*, **11** (2013), 336-361.
[8] F. A. Chalub, P. A. Markowich, B. Perthame and C. Schmeiser, Kinetic models for chemotaxis and their drift-diffusion limits, *Nonlinear Differential Equation Models*, (2004), 123-141.
[9] A. Chertock and A. Kurganov, High-resolution positivity and asymptotic preserving numerical methods for chemotaxis and related models, *Active Particles*, **2** (2019), 109-148.

[10] Y. Dolak and C. Schmeiser, The Keller-Segel model with logistic sensitivity function and small diffusivity, *SIAM J. Appl. Math.*, **66** (2005), 286-308.
[11] C. Emako and M. Tang, Well-balanced and asymptotic preserving schemes for kinetic models, preprint, (2016), arXiv:1603.03171.
[12] F. Filbet and S. Jin, A class of asymptotic-preserving schemes for kinetic equations and related problems with stiff sources, *J. Comput. Phys.*, **229** (2010), 7625-7648.
[13] Y. Han, Z. Li, J. Tao and M. Ma, Pattern formation for a volume-filling chemotaxis model with logistic growth, *J. Math. Anal. Appl.*, **448** (2017), 885-907.
[14] M. Herrero, E. Medina and J. Velazquez, Self-similar blow-up for a reaction-diffusion system, *J. Comp. Appl. Math.*, **97** (1998), 99-119.
[15] T. Hillen and K. Painter, A user's guide to PDE models for chemotaxis, *J. Math. Biol.*, **58** (2009), 183-217.
[16] D. Horstmann, From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. i, *Jahresberichte DMV*, **105** (2003), 103-165.
[17] D. Horstmann, From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. ii, *Jahresberichte DMV*, **106** (2004), 51-69.
[18] M. Ibrahim and M. Saad, On the efficacy of a control volume finite element method for the capture of patterns for a volume-filling chemotaxis model, *Comput. Math. Appl.*, **68** (2014), 1032-1051.
[19] S. Jin, Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations, *SIAM J. Sci. Comput.*, **21** (1999), 441-454.
[20] S. Jin and B. Yan, A class of asymptotic-preserving schemes for the Fokker-Planck-Landau equation, *J. Comput. Phys.*, **230** (2011), 6420-6437.
[21] T. Jin, X. Xu and D. Hereld, Chemotaxis, chemokine receptors and human disease, *Cytokine*, **44** (2008), 1-8.
[22] U. B. Kaupp, 100 years of sperm chemotaxis, *J. Gen. Physiol.*, **406** (2012), 583-586.
[23] E. F. Keller and L. A. Segel, Initiation of slime mold aggregation viewed as an instability, *J. Theor. Biol.*, **26** (1970), 399-415.
[24] E. F. Keller and L. A. Segel, Model for chemotaxis, *J. Theor. Biol.*, **30** (1971), 225-234.
[25] M. Lemou and L. Mieussens, A new asymptotic preserving scheme based on micro-macro formulation for linear kinetic equations in the diffusion limit, *SIAM J. Sci. Comput.*, **31** (2008), 334-368.
[26] M. Ma, C. Ou and Z.-A. Wang, Stationary solutions of a volume-filling chemotaxis model with logistic growth and their stability, *SIAM J. Appl. Math.*, **72** (2012), 740-766.
[27] O. Meacock, A. Doostmohammadi, K. Foster, J. Yeomans and W. M. Durham, Bacteria solve the problem of crowding by moving slowly, *Nat. Phys.*, **17** (2021), 205-210.
[28] J. D. Murray, *Mathematical Biology. I. An Introduction*, Interdiscip. Appl. Math., 17. Mathematical Biology, Springer, 2002.
[29] J. D. Murray, *Mathematical Biology. II. Spatial Models and Biomedical Applications*, Interdiscip. Appl. Math., 18. Mathematical Biology, Springer, 2003.
[30] T. Nagai, Global existence of solutions to a parabolic system for chemotaxis in two space dimensions, *Proceedings of the Second World Congress of Nonlinear Analysts*, **30**, 5381-5388.
[31] H. G. Othmer and A. Stevens, Aggregation, blowup and collapse: The ABC's of taxis in reinforced random walks, *SIAM J. Appl. Math.*, (1997), 1044-1081.
[32] H. G. Othmer and T. Hillen, The diffusion limit of transport equations derived from velocity-jump processes, *SIAM J. Appl. Math.*, **61** (2000), 751-775.
[33] H. G. Othmer and T. Hillen, The diffusion limit of transport equations. II. Chemotaxis equations, *SIAM J. Appl. Math.*, **62** (2002), 1222-1250.
[34] K. J. Painter and T. Hillen, Volume-filling and quorum-sensing in models for chemosensitive movement, *Can. Appl. Math. Q.*, **10** (2002), 501-543.
[35] C. S. Patlak, Random walk with persistence and external bias, *B. Math. Biophys.*, **15** (1953), 311-338.
[36] E. Roussos, J. Condeelis and A. Patsialou, Chemotaxis in cancer, *Nat. Rev. Cancer*, **11** (2011), 573-587.
[37] A. Stevens, The derivation of chemotaxis equations as limit dynamics of moderately interacting stochastic many-particle systems, *SIAM J. Appl. Math.*, **61** (2000), 183-212.
[38] F. G. Vital-Lopez, A. Armaou, M. Hutnik and C. D. Maranas, Modeling the effect of chemotaxis on glioblastoma tumor progression, *AIChE J.*, **57** (2011), 778-792.

[39] L. Wang and B. Yan, An asymptotic-preserving scheme for the kinetic equation with anisotropic scattering: Heavy tail equilibrium and degenerate collision frequency, *SIAM J. Sci. Comput.*, **41** (2019), A422-A451.
[40] Z.-A. Wang, On chemotaxis models with cell population interactions, *Math. Model. Nat. Pheno.*, **5** (2010), 173-190.
[41] Z. Wang and T. Hillen, Classical solutions and pattern formation for a volume filling chemotaxis model, *Chaos*, **17** (2007), 037108, 13 pp.
[42] D. Wrzosek, Volume filling effect in modelling chemotaxis, *Math. Model. Nat. Pheno.*, **5** (2010), 123-147.
[43] P. Zheng, C. Mu and X. Hu, Boundedness and blow-up for a chemotaxis system with generalized volume-filling effect and logistic source, *Discret. Contin. Dyn. S.*, **35** (2015), 2299-2323.

Received October 2024; revised March 2025; early access May 2025.