AGGREGATION-DIFFUSION PHENOMENA: FROM MICROSCOPIC MODELS TO FREE BOUNDARY PROBLEMS

INWON KIM, ANTOINE MELLET, AND JEREMY SHEUNG-HIM WU

ABSTRACT. This paper reviews (and expands) some recent results on the modeling of aggregation-diffusion phenomena at various scales, focusing on the emergence of collective dynamics as a result of the competition between attractive and repulsive phenomena - especially (but not exclusively) in the context of attractive chemotaxis phenomena.

At microscopic scales, particles (or other agents) are represented by spheres of radius $\delta > 0$ and we discuss both soft-sphere models (with a pressure term penalizing the overlap of the particles) and hard-sphere models (in which overlap is prohibited). The first case leads to so-called "blob models" which have received some attention recently as a tool to approximate non-linear diffusion by particle systems. The hard-sphere model is similar to a classical model for congested crowd motion. We will review well-posedness results for these models and discuss their relationship to classical continuum description of aggregation-diffusion phenomena in the limit $\delta \to 0$: the classical nonlinear drift diffusion equation and its incompressible counterpart.

In the second part of the paper, we discuss recent results on the emergence and evolution of sharp interfaces when a large population of particles is considered at appropriate space and time scales: At some intermediate time scale, phase separation occurs and a sharp interface appears which evolves according to a Stefan free boundary problem (and the density function eventually relaxes to a characteristic function - metastable steady state for the original problem). At a larger time scale the attractive forces lead to surface tension phenomena and the evolution of the sharp interface can be described by a Hele-Shaw free boundary problem with surface tension. At that same time scale, we will also discuss the emergence of contact angle conditions for problems set in bounded domains.

CONTENTS

1. Introduction

1.1. **Aggregation-diffusion phenomena.** The study of emergent collective dynamics in large systems of interacting agents has applications in many areas of applied sciences and has been the source of many interesting mathematical problems for decades. Classical examples in mathematical biology include the collective dynamics of micro-organisms such as cells or amoeba via chemotaxis [83, 118] or that of larger living organisms (insect swarms, the motion of animal herds or human crowds) [134, 11, 102]. Throughout this paper, we will refer to these organisms as active particles and assume that they are experiencing two competing effects: repulsion, modeled by nonlinear diffusion (or even density constraint) and attraction, modeled by nonlocal interactions. The simple mathematical model at the center of this paper is the following aggregation-diffusion equation

$$(1.1) \qquad \partial_t \rho + \operatorname{div}\left(\rho \nabla (G * \rho)\right) = \operatorname{div}\left(\rho \nabla f'(\rho)\right) \qquad \text{in } [0, \infty) \times \mathbb{R}^d, \quad \rho(t = 0) = \rho_{in},$$

which describes the evolution of the density of particles $\rho(t, x)$ and provides a continuum description of the particle systems at some macroscopic scale (see [24] and references therein).

The non-linearity $\rho \mapsto f(\rho)$ describes the repulsive force which models a natural volume exclusion principle (or incompressibility of the particles) and will play a central role throughout this paper. The standard linear diffusion corresponds to $f(\rho) = \rho \log \rho$, but we will consider pressure laws which penalize values of the density above a critical threshold ρ^* and lead to degenerate diffusion for small values of ρ. Taking $\rho^* = 1$ (to simplify the notations), we will focus on the following power-law:

$$(1.2) \qquad\qquad f_m(\rho) := \frac{\rho^m}{m - 1}, \qquad \forall \rho \geq 0, \qquad\qquad m > 1,$$

and its limit $m \to \infty$ (which enforces the congestion constraint $\rho \leq 1$):

$$(1.3) \qquad\qquad f_\infty(\rho) := \begin{cases} 0 & \text{when } \rho \in [0, 1], \\ \infty & \text{when } \rho > 1, \end{cases}$$

(in that case (1.1) should be written as (1.10)). More general convex functions can be considered as well.

The interaction kernel G in (1.1) describes long range interactions. We will focus our analysis on kernels that are radially symmetric, **non-negative** and **integrable**:

$$G(x) = G(|x|) \geq 0, \qquad G \in L^1(\mathbb{R}^d).$$

In many applications, the interaction kernels will be radially decreasing[1] at least for large $|x|$ though we do not need such an assumption. One can think of the potential $\phi = G * \rho$ as a chemical sensing function and the drift term in (1.1) models the motion of the agents toward higher values of that function. An important example of an attractive kernel is G given by the equation

$$(1.4) \qquad\qquad \sigma G - \eta \Delta G = \delta_{x=0} \qquad \text{in } \mathbb{R}^d, \qquad \lim_{|x| \to \infty} G(x) = 0,$$

which satisfies $G(x) \sim |x|^{2-d}$ when $|x| \to 0$ and $\int_{\mathbb{R}^d} G(x)\, dx = \sigma^{-1}$. When f is given by (1.2) and G by (1.4), Equation (1.1) is the classical (parabolic-elliptic) Patlak-Keller-Segel model (PKS) for chemotaxis

---

[1] With the sign convention used in (1.1), radially decreasing kernels model an attractive force between particles. While the opposite convention is used in some of the references mentioned in this paper, this convention allows us to work with a non-negative kernel.

[83, 118] which describes the collective motion of cells that are attracted by a self-emitted chemical substance (we refer to Hillen and Painter [71] for a review on various modeling aspects):

$$(1.5) \qquad \begin{cases} \partial_t \rho + \operatorname{div}\left(\rho \nabla \phi\right) = \Delta \rho^m, \\ \sigma \phi - \eta \Delta \phi = \rho. \end{cases}$$

While we will not exclusively focus on this model, all the results presented here will hold in particular for (1.5) when $m > 2$.

Purpose of this paper: Aggregation-diffusion equations such as (1.1) have been an active area of mathematical research for the past few decades. Well-posedness, blow-ups, long time behavior, steady states and energy minimizers have all been studied intensely and we will not attempt to present a complete picture of the existing theory (see [7, 8, 21, 23, 24, 46, 91, 132] and references therein). Instead, our focus will be on two recent trends in the analysis of these phenomena at both smaller and larger scales than the continuum model (1.1).

In the first part of the paper, we will describe a microscopic model for aggregation phenomena that includes a repulsive term taking into account the finite size of the particles, modeled by balls of radius $\delta > 0$. When f is given by (1.2), this model is similar to the blob-method introduced for approximating nonlinear diffusion equations [19, 30, 31, 36, 37, 95]. Since this model allows for the overlap of the balls, we will refer to it as a soft-sphere model. We will also consider a hard-sphere microscopic model, which includes a strict non-overlapping condition, and which has been used to model congested crowd motion [58, 85, 100, 101, 102, 103]. At the macroscopic scale, this hard-sphere model formally corresponds to (1.1) with f given by (1.3) and is related to Hele-Shaw free boundary problems. The purpose of this first part of the paper is thus to review these microscopic models (their derivation and their properties) and to discuss the connection with the continuum model (1.1).

In the second part of this paper, we will consider these same phenomena at a larger scale: When $m \in (2, \infty]$, we will show that equation (1.5) leads to phase separation (or phase segregation) and to the emergence of sharp interfaces. Phase separation, or the ability of a given population to organize itself into segregated groups with sharp edges, is an important feature of biological aggregation. Examples include animals moving as a homogeneous patch [117, 96] (e.g. insect swarms), pattern formation in cell tissues [3, 28], or the formation of membraneless organelles in eukaryotic cells[2] [16, 72, 107]. The fact that attractive nonlocal interactions can explain this phenomena has been previously pointed out [134, 57] and we present here a rigorous approach based on the recent work [86, 104] to derive these phenomena from (1.1) in regimes corresponding to a very large population of particles observed from far away over large time scales (or, equivalently, small attraction range). Characterizing the dynamics of these interfaces, and thus describing the collective behavior of the particles after aggregation, is a fascinating and challenging problem and we will present some recent results in this direction: At some appropriate time scale, the evolution of this interface is described by free boundary problems (Stefan and Hele-Shaw), which include surface tension effects as well as contact angle conditions and which are reminiscent of classical models for the evolution of interfaces in fluid dynamics.

In the rest of this introduction, we recall some classical facts about aggregation-diffusion equations (1.1) and then present all the models (microscopic, macroscopic and geometric) that will be studied in the paper as well as their relationships. Figure 1.3 summarizes that presentation.

1.2. **Finite time blow-up VS global well-posedness.** Before presenting these various models, we recall that since the attractive potential G leads to concentration, the well-posedness of (1.1) globally in time is not obvious. Preventing blow-ups (the formation of Dirac masses) in the mathematical model

---

[2]Membraneless organelles are mesoscopic structures found in the nucleus and cytoplasm of cells, characterized by higher protein density and weaker molecular motion than the surrounding medium, but not separated from it by a membrane. They play a crucial role in the cell's operations by allowing for increased rates of biochemical reactions.

is relatively natural from a modeling point of view since one would expect that the particles (if they represent living organisms) have some built-in incompressibility which prevents overlap and thus limits the maximum density. For the classical PKS model (1.5) with $\sigma \geq 0$, it is known that the standard **linear diffusion** ($m = 1$) is not enough to prevent finite time blow-up, but for $m > 2 - \frac{2}{d}$, solutions remain bounded in L^∞ uniformly in time (see [21, 132, 8, 7]) and global well-posedness holds for this equation (when $m = 2 - \frac{2}{d}$, well-posedness depends on the initial mass). On the other hand, when $\sigma > 0$ (so that the interaction kernel G is integrable), the power $m_c = 2$ is known to be critical for the existence of stationary states and energy minimizers for (1.5). In particular, while stationary states exist for all $\sigma > 0$ when $m > 2$ (see [6]) they exist if and only if $\sigma < \sigma_0$ when $m = 2$ (see [18, 81]). We will see in the second half of the paper that the condition $m > 2$ is also the natural condition for solutions of (1.5) to experience phase separation phenomena.

1.3. **Micro, macro, and geometric models.**
We now present the various models that will be discussed in this paper. In order to keep this introduction relatively short, we did not include references below. Relevant references for each of these models will be discussed in the corresponding sections of the paper. Figure 1.3 (at the end of this section) summarizes the relationship between the various models.

Microscopic "agent based" models (soft-sphere and hard-sphere) At the microscopic level with $N \gg 1$ particles, each particle is represented by a ball $B_\delta(x_i(t))$ with radius $\delta > 0$, whose center $x_i(t)$ evolves according to a first order ODE for $i = 1, \ldots, N$. In the soft-sphere model, the overlap of each ball $B_\delta(x_i(t))$ is allowed, but penalized, resulting in a repulsive force depending on the local density. This leads to the following microscopic model (see Section 3):

$$\dot{x}_i(t) = \frac{1}{N} \sum_{j=1}^N \nabla \widetilde{G}_\delta(x_j(t) - x_i(t)) - \nabla \int K_\delta(x_i(t) - y) f' \left(\frac{1}{N} \sum_{j=1}^N K_\delta(y - x^j(t)) \right) dy, \quad \forall i = 1, \ldots, N,$$

where $K_\delta = \frac{1}{|B_\delta|} \chi_{B_\delta(0)}$ (or another approximation of unity supported in $B_\delta(0)$) and $\widetilde{G}_\delta = K_\delta * G * K_\delta$. The corresponding empirical distribution $\rho(t, x) = \frac{1}{N} \sum_{i=1}^N \delta(x - x_i(t))$ solves the non-local PDE

$$(1.6) \qquad \partial_t \rho + \operatorname{div}(\rho \nabla (\widetilde{G}_\delta * \rho)) = \operatorname{div}(\rho \nabla K_\delta * f'(K_\delta * \rho)) \qquad \text{in } [0, \infty) \times \mathbb{R}^d$$

with initial condition

$$(1.7) \qquad \rho(0, x) = \frac{1}{N} \sum_{i=1}^N \delta(x - x_i(0)).$$

Similar equations and corresponding systems of ODEs have been introduced and used in particular to develop numerical methods for nonlinear diffusion models. In Section 3, we will recall several results concerning the well-posedness of (1.6) for general initial data (not necessarily empirical distributions) and review convergence results as $N \to \infty$ (mean-field limit - see Theorem 3.1) and $\delta \to 0$ (Theorem 3.3).

The microscopic hard-sphere model is a similar model in which overlap is not permitted: The constraint

$$|x_i(t) - x_j(t)| \geq 2\delta, \qquad \forall i, j = 1, \ldots, N, \qquad i \neq j,$$

is strictly enforced, leading to a well-posed system of ODE with constraints (see Section 4.1). A corresponding PDE can be obtained formally by taking the limit $m \to \infty$ in (1.6), leading to the following equation (see Section 4):

$$(1.8) \qquad \partial_t \rho + \operatorname{div}(\rho \nabla (\widetilde{G}_\delta * \rho)) = \operatorname{div}(\rho \nabla K_\delta * p), \quad p \in \partial f_\infty(K_\delta * \rho) \qquad \text{in } [0, \infty) \times \mathbb{R}^d,$$

which includes a pressure function $p(t,x)$ playing the role of a Lagrange multiplier for the non-local constraint $K_\delta * \rho \leq 1$. The condition $p \in \partial f_\infty(K_\delta * \rho)$[3] is a shorthand notation for the three conditions

$$K_\delta * \rho(t,x) \leq 1, \qquad p(t,x) \geq 0 \qquad p(t,x)(1 - K_\delta * \rho(t,x)) = 0 \qquad \text{a.e. } x \in \mathbb{R}^d, \ t > 0$$

(note that (1.6) can be written in a similar form with $p = \partial f(K_\delta * \rho)$). However, while (1.8) is indeed related to a well-posed microscopic model when the initial condition is an empirical distribution (1.7), we will see that its well-posedness for general initial conditions is far from clear (see Section 4.5).

Macroscopic models (soft-sphere and hard-sphere). In the limit $\delta \to 0$, Equation (1.6) leads to the classical aggregation-diffusion equation which we recall here:

$$(1.9) \qquad \partial_t \rho + \text{div}\,(\rho \nabla (G * \rho)) = \text{div}\,(\rho \nabla f'(\rho)) \qquad \text{in } [0,\infty) \times \mathbb{R}^d,$$

while the hard-sphere model (1.8) yields:

$$(1.10) \qquad \partial_t \rho + \text{div}\,(\rho \nabla (G * \rho)) = \text{div}\,(\rho \nabla p), \qquad p \in \partial f_\infty(\rho) \qquad \text{in } [0,\infty) \times \mathbb{R}^d.$$

A similar equation was introduced as a model for congested crowd motion and some variations of it were derived in many frameworks, in particular as mechanical models for tumor growth. Interpretations and properties of this model are recalled in Sections 4.2-4.4. It is closely related to the classical Hele-Shaw free boundary problem with active potential (see Section 5).

Geometric models: Phase separation and free boundary problems. The second half of this paper is devoted to phase separation, sharp interface limits and the derivation of free boundary problems from (1.9) when f is given by (1.2) with $m > 2$ or from (1.10). The *sharp interface limit* corresponds to the regime in which the size of the support of ρ is very large compared to the interfacial region, and we will derive mathematical models for the evolution of this interface. The results of Sections 6 - 9 can be summarized as follows: When the total mass $\int \rho\,dx =: \varepsilon^{-d}$ is very large ($\varepsilon \ll 1$), we observe the development of an interface separating regions of low and high densities, which happens at time scale of order ε^{-2}. After that time, the evolution of the interface can be described by a *Stefan free boundary problem*. In that regime, the motion of the interface is driven by the evolution of the density in the bulk. This bulk density eventually relaxes toward a constant (stable) value θ. At that point, the motion of the interface stops, at least at this time scale. But such a state is only metastable and the slower evolution of the interface, at time scale ε^{-3}, can be described by a one-phase Hele-Shaw free boundary problem with surface tension.

To justify this analysis, we consider a large number of particles, $\int \rho_{in}(x)\,dx = \varepsilon^{-d} \gg 1$. Rescaling the variables $x \mapsto \varepsilon x$ and $t \mapsto \varepsilon^2 t$ leads to the equation

$$(1.11) \qquad \partial_t \rho + \text{div}\,(\rho \nabla (G_\varepsilon * \rho)) = \text{div}\,(\rho \nabla f'(\rho)) \qquad \text{in } [0,\infty) \times \mathbb{R}^d,$$

with $G_\varepsilon(x) := \varepsilon^{-d} G(\varepsilon^{-1} x)$ and $\int \rho_{in}(x)\,dx = 1$ (we can also consider the corresponding hard-sphere model). Recalling that $G \in L^1(\mathbb{R}^d)$ and denoting

$$\int_{\mathbb{R}^d} G(x)\,dx = \frac{1}{\sigma} < \infty,$$

---

[3]∂f_∞ denotes the subdifferential of the convex function f_∞ and is sometimes called the *Hele-Shaw graph*. It is a multi-valued function given by

$$\partial f_\infty(\rho) := \begin{cases} 0 & \text{if } \rho < 1 \\ [0,\infty] & \text{if } \rho = 1 \\ \infty & \text{if } \rho > 1 \end{cases}$$
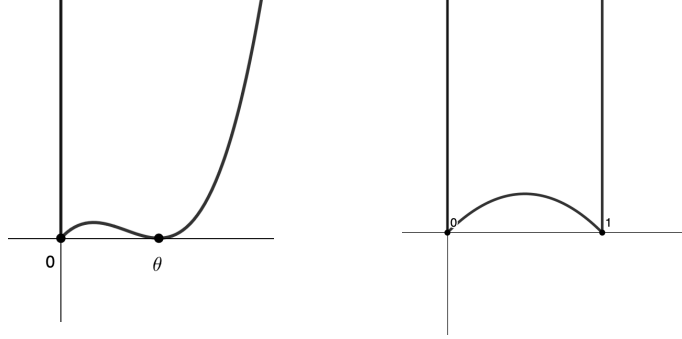
FIGURE 1. The double-well potential $h_m(\rho)$ when $m = 3$ (left) and $h_\infty(\rho)$ (right)

(which is satisfied for (1.4)), we rewrite (1.11) as follows:

$$
(1.12) \qquad \partial_t \rho + \mathrm{div}\left(\rho \nabla \int G_\varepsilon(x - y)(\rho(y) - \rho(x))\, dy\right) = \mathrm{div}\left(\rho \nabla h'(\rho)\right) \qquad \text{in } [0, \infty) \times \mathbb{R}^d,
$$

where

$$
(1.13) \qquad h(\rho) := f(\rho) - \frac{1}{2\sigma}\rho^2 + a\rho
$$

for some constant $a \in \mathbb{R}$ - the addition of the term $a\rho$ in the function h does not affect the equation but will be important for the energy. When f is given by (1.2) with $m > 2$, the function $h''(\rho) = f''(\rho) - \frac{1}{\sigma}$ is negative for small ρ and positive for large ρ. This is a fundamental property which implies that the limiting equation $\partial_t \rho = \mathrm{div}\left(\rho \nabla h'(\rho)\right)$ is an ill-posed forward-backward diffusion equation and leads to phase separation. In fact, the function h is a **double well-potential**: When f is given by (1.2) with $m > 2$, the function h in (1.13) becomes

$$
(1.14) \qquad h_m(\rho) = \frac{\rho}{m-1}\left(\rho^{m-1} - (m-1)\theta_m^{m-2}\rho + (m-2)\theta_m^{m-1}\right), \qquad \theta_m = \left(\frac{1}{2\sigma}\right)^{\frac{1}{m-2}}
$$

where we took $a = \frac{m-2}{m-1}\theta_m^{m-1}$. This function is a double-well potential with a (singular) well at $\rho = 0$ and a (smooth) well at $\rho = \theta_m$ (see Figure 1). When $f = f_\infty$, we take $a = \frac{1}{2\sigma}$ in (1.13), and find

$$
(1.15) \qquad h_\infty(\rho) := \begin{cases} \frac{1}{2\sigma}\rho(1 - \rho) & \text{if } 0 \leq \rho \leq 1, \\ \infty & \text{otherwise,} \end{cases}
$$

which is a double-well (or double obstacle) potential, with wells at $\rho = 0$ and $\rho = 1$ (see Figure 1).

We do not need to restrict ourselves to f given by (1.2) or (1.3) for the results below. All we need is that the function $\rho \mapsto h(\rho)$ is a double-well potential with stable region $\{0\} \cup (\theta, \infty)$ for some $\theta > 0$. Under this condition, (1.12) is a nonlocal approximation of a Cahn-Hilliard equation whose sharp interface limit $\varepsilon \to 0$ is a classical problem. We will show (in terms of Γ-convergence of the energy - see Theorem 7.2) that the limit $\varepsilon \to 0$ of (1.11) or (1.12) leads to a generalized Stefan problem:

$$
(1.16) \qquad \partial_t \rho = \mathrm{div}\left(\rho \nabla h^{**\prime}(\rho)\right)
$$

where h^{**} denotes the convex hull of h. When the initial density does not take values in the unstable phase, that is if $|\{\rho_{in}(x) \in (0, \theta)\}| = 0$, then (1.16) is a weak formulation of

$$(1.17) \qquad \begin{cases} \rho \geq \theta, \quad \partial_t \rho = \operatorname{div}\left(\rho \nabla h^{**\prime}(\rho)\right) & \text{in } E(t) \\ \rho = 0 & \text{in } \mathbb{R}^d \setminus E(t) \\ V = -\nabla h^{**\prime}(\rho) \cdot n & \text{on } \partial E(t). \end{cases}$$

where V denotes the normal velocity of $\partial E(t)$. The evolution of ρ in the set $E(t)$ (described by a nonlinear diffusion equation) is thus responsible for the motion of the interface $\partial E(t)$ and will eventually lead to constant density, that is $\rho(t) \to \theta \chi_E$ for some (constant) set E.

If $\rho_{in}(x) = \theta \chi_{E_{in}}$, then the solution of (1.16) is constant in time which indicates that the corresponding solution of (1.11) evolves very slowly when $\varepsilon \ll 1$ (characteristic functions are metastable for (1.11)). In order to characterize the evolution of such solutions over larger time scale (of order ε^{-3} compared to the microscopic time scale), we consider (in Section 8) the equation

$$(1.18) \qquad \varepsilon \partial_t \rho^\varepsilon + \operatorname{div}\left(\rho^\varepsilon \nabla(G_\varepsilon * \rho^\varepsilon)\right) = \operatorname{div}\left(\rho^\varepsilon \nabla f'(\rho^\varepsilon)\right) \qquad \text{in } [0, \infty) \times \mathbb{R}^d,$$

with initial data $\rho_{in}(x) = \theta \chi_{E_{in}}$ for some set E_{in} with finite perimeter. When G is the Newtonian kernel (1.4), the evolution of ρ^ε, solution of (1.18), is then asymptotically described by the Hele-Shaw free boundary problem with surface tension (see Theorem 8.3):

$$(1.19) \qquad \begin{cases} \Delta p = 0 & \text{in } E(t), \\ p = \gamma \kappa & \text{on } \partial E(t), \\ V = -\nabla p \cdot \nu & \text{on } \partial E(t), \end{cases}$$

where $\kappa(t, x)$ denotes the mean-curvature of $\partial E(t)$ (with the convention that $\kappa \geq 0$ when E is convex) and V denotes the normal velocity of the interface $\partial E(t)$. The constant γ depends on m and σ (see (8.3)).

This shows that while the Stefan problem (1.16) describes the phase separation phenomena that takes place for the solutions of (1.11) when $\varepsilon \ll 1$, the resulting clusters (regions when $\rho = \theta$) will continue to evolve over a much larger time scale. This evolution, due to to the attractive interactions between the particles, is akin to the evolution of an interface separating two immiscible fluids under the effect of surface tension phenomenon. This is not completely surprising: Surface tension phenomena can be explained by stating that the bond between two molecules of the same fluid (cohesion) is stronger than that between two molecules of different fluids (adhesion). The same can be said about our particles, which are happier when surrounded by other particles than by vacuum.

Another important phenomenon that often comes together with surface tension is the notion of **contact angle**, which appears when the free surface come in contact with a fixed boundary. In most of the paper we will look at problems set in the whole space $\mathbb{R}^d$, but in Section 9 we look at problems set in a bounded domain and show that contact angle conditions are a natural byproduct of the sharp interface limit discussed above. More precisely, we will consider the initial boundary value problem

$$(1.20) \qquad \begin{cases} \varepsilon \partial_t \rho + \operatorname{div}\left(\rho \nabla \phi_\varepsilon\right) = \operatorname{div}\left(\rho \nabla f'(\rho)\right) & \text{in } [0, \infty) \times \Omega, \\ \rho \nabla[-f'(\rho) + \phi_\varepsilon] \cdot n = 0 & \text{on } [0, \infty) \times \partial \Omega, \\ \rho(x, 0) = \rho_{in} & \text{in } \Omega. \end{cases}$$

In the context of chemotaxis, ϕ_ε will be the solution of an elliptic boundary value problem in Ω (see (9.1)). In that case we will derive a contact angle condition which depends on the boundary conditions for ϕ_ε (see (9.4)). In some experimental settings, it can also be interesting to keep $\phi_\varepsilon = G_\varepsilon * \rho$ (where ρ is extended by 0 to $\mathbb{R}^d$ to make sense of the convolution), modeling interactions that only depends on the distance between the particles. In that case, Ω acts as an obstacle and the limiting free boundary

problem is (1.19) supplemented with contact angle condition $\alpha = \pi$ (i.e. the contact is tangential). But general contact angle conditions can be recovered as well by taking into account the interactions of our particles with the fixed boundary (Section 9.2).

Longer time scale. As a final note, we observe that stationary solutions for (1.9) and global minimizers of the corresponding energy have been the subject of intense research in the past two decades. We will not attempt to give a full review of these results here, but we recall that in many settings it is known that stationary solutions must be radially decreasing. For example when G is the Newtonian attractive kernel and f is given by (1.2) with $m > 1$, the unique stationary state in two dimensions is radially symmetric and decreasing (see for instance [23, 46]). In contrast, both (1.16) and (1.19) have stationary solutions that are not radially decreasing, but consist of clusters separated by vacuum. This suggests that when $\varepsilon \ll 1$, the solutions of (1.11) and (1.18) may first approach some metastable equilibrium, close to some steady states of the limiting free boundary problems, but eventually these clusters will coalesce and converge as $t \to \infty$ to a radially symmetric stationary state (which consists of a unique cluster). Such dynamics are indeed observed numerically (see for instance [24]).

2. Energies and gradient flows.

Formally, all the models described in this paper have a gradient flow structure on the space of probability measures with finite second moment
$$\mathcal{P}_2(\mathbb{R}^d) := \left\{ \rho \in \mathcal{P}(\mathbb{R}^d) \,;\, \int_{\mathbb{R}^d} |x|^2 d\rho(x) < \infty \right\},$$

endowed with the 2-Wasserstein metric which we denote by d_W. This structure was first formalized for the Fokker-Planck equation in the seminal work of Jordan, Kinderlehrer and Otto [80] and extended to other PDE in various work, see for instance [27, 5, 22]. Every equation presented in the introduction can be written in the form

$$(2.1) \qquad \partial_t \rho + \operatorname{div}(\rho v) = 0, \qquad v = -\nabla \frac{\delta \mathcal{E}}{\delta \rho}(\rho) \text{ in } (0, \infty) \times \mathbb{R}^d,$$

for some energy functional $\mathcal{E} : \mathcal{P}_2 \to \mathbb{R} \cup \{+\infty\}$, which in turn provides a natural Lyapunov functional for the evolution equation: Weak solutions of (2.1) satisfy the energy dissipation inequality

$$\mathcal{E}[\rho(t)] + \int_0^t \int_{\mathbb{R}^d} |v|^2 d\rho \, ds \le \mathcal{E}[\rho(0)].$$

We briefly present below the energy functionals associated to our models. Throughout the paper we will rely on these functionals and their asymptotic behavior to develop heuristic arguments and to justify singular limits of the associated PDEs. For most of this paper, we will work with measures ρ which are absolutely continuous with respect to Lebesgue measure. This is denoted by $d\rho \ll dx$ and we will use ρ to denote both the measure and its density with respect to Lebesgue measure.

The (macroscopic soft-sphere) aggregation-diffusion equation (1.9) is classically associated with the energy
(2.2)
$$\mathcal{E}_f[\rho] := \mathcal{F}_f[\rho] - \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} G(x-y) \, d\rho(x) \, d\rho(y), \quad \mathcal{F}_f[\rho] := \begin{cases} \int_{\mathbb{R}^d} f(\rho(x)) \, dx, & \text{if } \rho \in \mathcal{P}_2(\mathbb{R}^d), \ d\rho \ll dx, \\ +\infty, & \text{otherwise.} \end{cases}$$

We use the notation $\mathcal{E}_m$ (resp. $\mathcal{E}_\infty$) and $\mathcal{F}_m$ (resp. $\mathcal{F}_\infty$) when f is given by (1.2) (resp. (1.3)). In particular, for $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ such that $d\rho \ll dx$, we have

$$\mathcal{F}_m[\rho] = \int_{\mathbb{R}^d} \frac{\rho(x)^m}{m-1} \, dx, \qquad \mathcal{F}_\infty[\rho] = \begin{cases} 0 & \text{if } \rho \le 1 \text{ a.e.} \\ +\infty & \text{otherwise,} \end{cases}$$

Soft Sphere models   **Hard Sphere (incompressible) models**

$\dot{x}_i(t) = \nabla\widetilde{G}_\delta(x_i(t)) - \nabla K_\delta * .f'_m(K_\delta * \rho_N)(x_i(t))$ $\dot{x}_i(t) = \nabla\widetilde{G}_\delta(x_i(t)), \quad |x_i(t) - x_j(t)| \geq 2\delta$

$\rho_t + \nabla \cdot (\rho\nabla\widetilde{G}_\delta * \rho) = \nabla \cdot (\rho\nabla K_\delta * f'_m(K_\delta * \rho))$
with ρ as empirical measure (e.m.).

$\rho_t + \nabla \cdot (\rho\nabla\widetilde{G}_\delta * \rho) = \nabla \cdot (\rho\nabla K_\delta * p)$
$\rho * K_\delta \leq 1, \quad p(1 - \rho * K_\delta) = 0, \quad \rho$ as e.m.

$N \to \infty, \delta > 0$?   ? $N \to \infty$

$\rho_t + \nabla \cdot (\rho\nabla\widetilde{G}_\delta * \rho) = \nabla \cdot (\rho\nabla K_\delta * f'_m(K_\delta * \rho))$
with ρ as general measure (g.m.).

$m \to \infty$?

$\rho_t + \nabla \cdot (\rho\nabla\widetilde{G}_\delta * \rho) = \nabla \cdot (\rho\nabla K_\delta * p)$
$\rho * K_\delta \leq 1, \quad p(1 - \rho * K_\delta) = 0, \quad \rho$ as g.m.

$\delta \to 0$ $N \to \infty$, $\delta \to 0$, $\delta^d N = 1$? $N \to \infty$

$\rho_t + \nabla \cdot (\rho\nabla G * \rho) = \nabla \cdot (\rho\nabla f'_m(\rho))$ $\rho_t + \nabla \cdot (\rho\nabla G * \rho) = \nabla \cdot (\rho\nabla p),$
$\rho \leq 1, \quad p(1 - \rho) = 0.$

$m \to \infty$

Rescaling
$x \sim \varepsilon^{-1}, \ t \sim \varepsilon^{-2}$
$\varepsilon \to 0$

$x \sim \varepsilon^{-1}, \ t \sim \varepsilon^{-3},$
$\varepsilon \to 0$

Stefan FBP
$\rho_t = \nabla \cdot (\rho\nabla h^{**}(\rho))$

Hele-Shaw FBP with active potential

$x \sim \varepsilon^{-1}, \ t \sim \varepsilon^{-3}$
$\varepsilon \to 0$

Hele-Shaw FBP with Surface tension, $\rho = \theta\chi_E$
$\begin{cases} \Delta p = 0 \text{ in } E(t), \\ p = \gamma\kappa \text{ on } \partial E(t) \\ V = -\nabla p \cdot n \text{ on } \partial E(t). \end{cases}$

Microscopic (agent based) models

Macroscopic PDE models
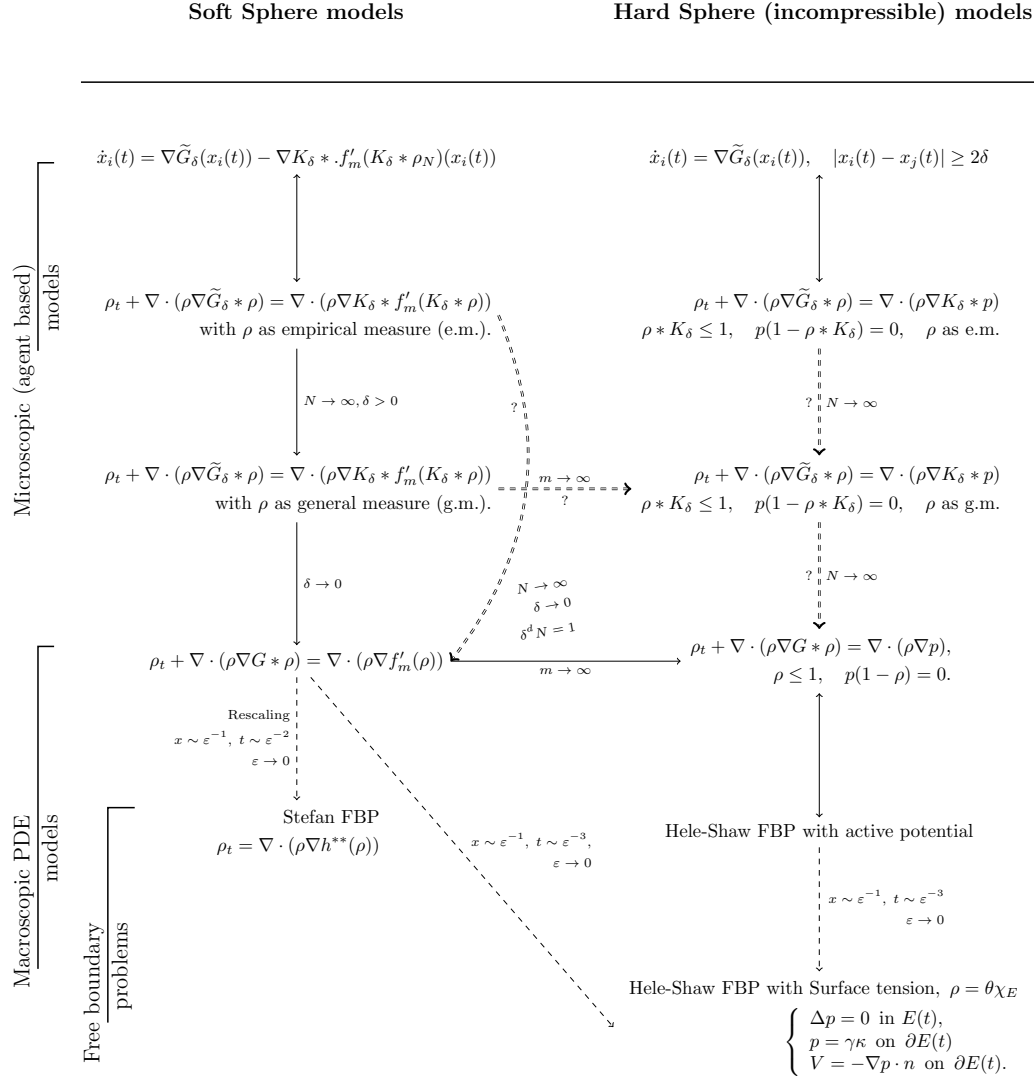
Free boundary problems

FIGURE 2. The various models discussed in this paper and their relationships. The dashed arrows indicate conditional convergence results. The double dashed arrows indicate open problems

and the energy $\mathscr{E}_\infty$, given by

$$(2.3) \qquad \mathscr{E}_\infty[\rho] = \mathscr{F}_\infty[\rho] - \frac{1}{2}\iint G(x - y)\, d\rho(x)\, d\rho(y),$$

is associated with the hard-sphere model (1.10).

Similarly, the microscopic model (1.6) is a gradient flow for the regularized energy

$$(2.4) \qquad \mathscr{E}_\delta[\rho] := \mathscr{F}_f[\rho * K_\delta] - \frac{1}{2}\iint_{\mathbb{R}^d \times \mathbb{R}^d} \widetilde{G}_\delta(x - y)\, d\rho(x)\, d\rho(y).$$

Notice that $\mathscr{E}_\delta[\rho] = \mathscr{E}_f[K_\delta * \rho]$. In particular, thanks to the convolution with K_δ, the measure ρ does not have to be absolutely continuous with respect to the Lebesgue measure for $\mathscr{E}_\delta$ to be finite. These energies play an important role in our proofs, and the Γ-convergence of $\mathscr{E}_\delta$ to $\mathscr{E}_f$ as $\delta \to 0$ and that of $\mathscr{E}_m$ to $\mathscr{E}_\infty$ as $m \to \infty$ are good indicators of the relationship between the corresponding PDEs.

Such Γ-convergence results play an even more central role in the second part of the paper and the derivation of the free boundary problems (1.16) and (1.19). First, we note that the rescaled equation (1.11) is a gradient flow for the rescaled energy

$$(2.5) \qquad \mathscr{F}_f[\rho] - \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} G_\varepsilon(x - y)\, d\rho(x)\, d\rho(y), \qquad G_\varepsilon(x) = \varepsilon^{-d} G(\varepsilon^{-1} x).$$

Up to a constant, this energy can also be written as

$$(2.6) \qquad \mathscr{J}_\varepsilon[\rho] = \begin{cases} \displaystyle\int_{\mathbb{R}^d} h(\rho)\, dx + \frac{1}{4} \iint_{\mathbb{R}^d \times \mathbb{R}^d} G_\varepsilon(x - y)[\rho(x) - \rho(y)]^2\, dx\, dy & \text{if } \rho \in \mathcal{P}_2(\mathbb{R}^d),\ d\rho \ll dx, \\ \infty & \text{otherwise,} \end{cases}$$

where h is the double well-potential defined by (1.13) (with wells at $\rho = 0$ and $\rho = \theta$). This functional is a non local approximation of the Allen-Cahn energy

$$(2.7) \qquad \int_{\mathbb{R}^d} h(\rho(x))\, dx + \varepsilon^2 \beta \int_{\mathbb{R}^d} \frac{1}{2}|\nabla \rho|^2\, dx,$$

commonly used in problems that exhibit phase separation. We will prove in particular that as long as $\int_{\mathbb{R}^d}(1 + |x|^2)G(x)\, dx < \infty$, $\mathscr{J}_\varepsilon$ Γ-converges to (see Proposition 7.1):

$$(2.8) \qquad \mathscr{J}^*[\rho] = \int_{\mathbb{R}^d} h^{**}(\rho)\, dx,$$

whose corresponding gradient flow is the generalized Stefan free boundary problem (1.16). Characteristic functions $\theta\chi_E$ are stationary solutions of (1.16) and $\mathscr{J}_\varepsilon(\theta\chi_E) \to \mathscr{J}^*(\theta\chi_E) = 0$ as $\varepsilon \to 0$. But the rescaled functional $\mathscr{G}_\varepsilon[\rho] := \varepsilon^{-1} \mathscr{J}_\varepsilon[\rho]$ Γ-converges to the perimeter functional (see Theorem 8.1)

$$\mathscr{G}_0(\rho) = \begin{cases} \gamma\theta P(E) & \text{if } \rho \in \mathcal{P}_2(\mathbb{R}^d), \quad \rho = \theta\chi_E \in \mathrm{BV}(\mathbb{R}^d), \\ \infty & \text{otherwise,} \end{cases}$$

for some constants γ depending on σ and θ where the perimeter $P(E)$ of the set $E \subset \mathbb{R}^d$ is defined by

$$(2.9) \qquad P(E) = \sup\left\{ \int_E \operatorname{div} g\, dx\,;\ g \in C_0^1(\mathbb{R}^d; \mathbb{R}^d),\quad |g(x)| \le 1,\ \forall x \in \mathbb{R}^d \right\}.$$

Such results, for general kernels G and smooth double-well potentials were first derived in [1]. This convergence is also proved by a different approach in [104] when G is given by (1.4) and f is given by (1.2) with $m > 2$ and in [106, 86] when $f = f_\infty$. In these papers, the role of boundary conditions (when the problem is set in a bounded domain) plays a crucial role (see Theorem 9.1) and led to the derivation of contact angle conditions.

The final observation is that the Wasserstein gradient flow for the perimeter functional is the Hele-Shaw free boundary problem with surface tension (1.19) (see [114]). The Γ-convergence of $\mathscr{G}_\varepsilon$ to $\mathscr{G}_0$ is thus consistent with the convergence of (1.11) to (1.19). However, the Γ-convergence of the energy alone does not guarantee the convergence of the corresponding gradient flows. Additional sufficient conditions for such a convergence were established by Sandier-Serfaty [129] and used, for example, to justify the convergence of (1.5) to (1.10) in the limit $m \to \infty$ in [39]. These additional conditions can be difficult to establish in other frameworks and we will not review or extend such results in this paper. For the derivation of the free boundary problems (1.16) and (1.19) we will take a different approach and state conditional convergence results (see Theorems 7.2 and 8.3), which imply the convergence of the solutions

under an assumption on the convergence of the energy (see (7.3)). Such conditional convergence results have a long history with free boundary problems involving mean-curvature (see [114, 94, 55, 78]), and proving rigorous convergence results without this assumption remains an important and challenging problem.

Part 1. **Microscopic/macroscopic soft-sphere and hard-sphere models**

3. Soft-sphere models: The blob method

The main goal of this section is to derive the microscopic model (1.6) from particle dynamics and review some of the recent literature on this and related models. The convolution with the kernel K_δ is the key feature of this model. It accounts for the finite size ($\delta > 0$) of the particles, as opposed to treating them as points. We will see that (1.6) has solutions in the form of empirical distributions (sums of Dirac masses centered at points $x_i(t)$), a fact that allows us to connect (1.6) to the corresponding microscopic model describing the motion of individual particles via a system of coupled ODEs.

From a mathematical perspective, (1.6) is the gradient flow of a λ-convex energy functional (under appropriate assumptions on K_δ and f). This classical framework yields the well-posedness of (1.6) in the set of probability measures. Finally, we will discuss the connection between (1.6) and the classical aggregation-diffusion equation (1.9) in the limit $\delta \to 0$.

3.1. **Transport equation with interaction potential.** Transport equations without any diffusion and with sufficiently smooth interaction potential $\widetilde{G} : \mathbb{R}^d \to \mathbb{R}$ (satisfying $\nabla \widetilde{G}(0) = 0$) are naturally associated to deterministic particle systems via the empirical distribution: Denoting by $\{x_i(t)\}_{i=1,\ldots,N}$ the positions of N particles, solutions to the following system of ODEs

$$(3.1) \qquad \dot{x}_i(t) = \frac{1}{N} \sum_{j=1}^{N} \nabla \widetilde{G}(x_j(t) - x_i(t)) \qquad \forall i = 1, \ldots, N,$$

the corresponding empirical distribution

$$\rho_N(t, x) = \frac{1}{N} \sum_{j=1}^{N} \delta(x - x_j(t))$$

is the solution (in the sense of distribution) of the continuity equation

$$(3.2) \qquad \partial_t \rho + \operatorname{div}(\rho \nabla \widetilde{G} * \rho) = 0$$

with initial data $\rho_N(0, x) = \frac{1}{N} \sum_{j=1}^{N} \delta(x - x_j(0))$ (see for instance [22, 29]).

Equation (3.2) appears in many settings, particularly in the mathematical modeling of collective behaviors. We refer to [22] and the many references therein for a detailed discussion of the associated well-posedness theory. We recall that (3.2) is a gradient flow for the energy

$$-\frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \widetilde{G}(x - y) d\rho(x) \, d\rho(y)$$

defined on $\mathcal{P}_2(\mathbb{R}^d)$ (this was formalized in [27] following ideas that were first introduced in [115]). In that context, a key assumption for the well-posedness of (3.2) is the λ-convexity of the kernel $-\widetilde{G}$ for some $\lambda \leq 0$ (see [22]). Under this assumption, we also get the following stability estimate

$$(3.3) \qquad d_W(\rho^1(t), \rho^2(t)) \leq d_W(\rho^1(0), \rho^2(0)) e^{-\lambda t}$$

for any two weak measure solutions $\rho^1(t)$ and $\rho^2(t)$ of (3.2) in $C([0, \infty); \mathcal{P}(\mathbb{R}^d))$. Stability estimates have a long history and a similar inequality was first proved by Dobrushin [49] (see also Golse [65]) with the 1-Wasserstein distance. Deriving such estimates (which can lead to useful numerical approximations)

when the interaction potential $\widetilde{G}$ is less regular and possibly in different topologies is still an active area of research. Inequality (3.3) implies in particular that if $\rho_N(0)$ converges to ρ_{in} (with respect to d_W), the measure $\rho_N(t, x)$ converges to the corresponding solution of (3.2). We will say that (3.2) is the mean-field limit of the ODE system (3.1).

The mean field limit (but not (3.3)) can also be justified for some more singular potentials (see for instance [29] when $\widetilde{G}(x) \sim |x|^b$ with $b > 2 - d$), but not, to our knowledge, with the Newtonian kernel as in the PKS model (in dimension 1, some results have been obtained by Bonashi-Carrillo-Di Francesco-Peletier [14] for the attractive Newtonian kernel). Other classical references for mean-field limit of related model include Sznitman [133] and Golse [65] as well as the recent reviews with both theoretical and numerical perspectives [32, 33].

3.2. **Linear/Nonlinear repulsion.** When diffusion (linear or nonlinear) is added to (3.2), solutions starting with empirical measures are instantly regularized by the diffusion. This makes the approximation of the corresponding PDE by a deterministic ODE system challenging (see [44, 128]). A classical way to deal with this issue is to consider instead a system of stochastic ODEs (adding a Brownian motion to (3.1)) as is done with the discretization of the linear Fokker-Planck equation by Langevin dynamics. Some nonlinear diffusion equations have also been handled via such methods [112, 110, 122, 59]. We also refer to some recent work treating singular potentials via the so-called *relative-entropy method* [76, 77, 17] and *modulated energy method* [126] (the result of [17] applies to the Patlak-Keller-Segel model in dimension 2).

We will now describe a different approach, based on the so-called *blob method*, which is a deterministic approach in which the (nonlinear) diffusion is approximated by a nonlocal interaction term which preserves the empirical distribution (see [36, 30, 95, 37, 19, 31]). This approach was also used by Motsch and Peurichard in [111] to derive a model for tumor growth.

3.3. **Nonlocal nonlinear repulsion.** The *blob method* amounts to taking into account the finite size $\delta > 0$ of the particles (at the microscopic level) and some volume exclusion principle. A key features of the microscopic models we have in mind is a repulsive effect with finite radius: Particles that are within a certain distance of each others feel the effect of a strong repulsion, which prevents the density of particles from becoming too large.

Given a compactly supported function $\bar{K}_\delta : \mathbb{R}^d \to [0, \infty)$, we introduce the counting function

$$\mu_N(x) = \sum_{j=1}^{N} \bar{K}_\delta(x - x_j(t)).$$

When $\bar{K}_\delta = \chi_{B_\delta(0)}$ (an important example we will come back to regularly), $\mu_N(x)$ counts how many particles are located in $B_\delta(x)$. If we think of the particles as *hard-spheres* of radius δ, it is natural to impose the constraint $\mu_N \le 1$. We will discuss this setting, and its many challenges, in the next section, but for now we will consider a *soft-sphere* model which instead of this hard constraint includes a repulsion force that drives the particles away from crowded regions with a strength that depends on μ_N.

Before introducing the model, we note that we can interpret the quantity $\int \bar{K}_\delta(x) \, dx = |B_\delta| = \omega_d \delta^d$ as the volume occupied by one particle, so the total volume occupied by all particles is equal to $m_0 := N \int \bar{K}_\delta(x) \, dx$. Introducing the normalized kernel $K_\delta = \frac{1}{\int \bar{K}_\delta(x) \, dx} \bar{K}_\delta$, we can write

$$\mu_N(x) = N\bar{K}_\delta * \rho_N(x) = m_0 K_\delta * \rho_N(x).$$

In order to simplify the notations, we will take $m_0 = 1$ below so that the counting function μ_N can be written as

$$\mu_N(x) = K_\delta * \rho_N(x), \qquad \int K_\delta(x) \, dx = 1, \quad \text{Supp } K_\delta = B_\delta.$$

Note that we can choose $K_\delta(x) = \frac{1}{\omega_d \delta^d} K(\frac{x}{\delta})$ with $K(x) = \chi_{B_1}(x)$, but when convenient we will replace this kernel with a smooth radially symmetric decreasing function supported in B_1. In what follows N and δ will be treated as independent parameters, but we should remember that this interpretation of the microscopic model imposes the constraint $N \sim \delta^{-d}$.

The soft-sphere model is obtained by adding a repulsive term to (3.1) as follows:

$$(3.4) \qquad \dot{x}_i(t) = \frac{1}{N} \sum_{j=1}^{N} \nabla \widetilde{G}(x_j(t) - x_i(t)) - \int K_\delta(x_i(t) - y) \nabla f'(\mu_N)(y) \, dy, \qquad i = 1, \ldots N,$$

for a convex function f (for instance given by (1.2)). Equation (3.4) can also be written as

$$(3.5) \qquad \dot{x}_i(t) = \nabla \widetilde{G} * \rho_N(x_i(t)) - \nabla K_\delta * f'(K_\delta * \rho_N(x))(x_i(t)), \qquad i = 1, \ldots N.$$

The particular structure of this repulsive term, with the double convolution, is convenient from a mathematical view point (it leads to an interpretation of the corresponding PDE as a gradient flow). From a modeling view point, it can be interpreted as follows (when $K(x) = \chi_{B_1}(x)$): Since a particle is represented by the ball $B_\delta(x_i(t))$, the repulsion force exerted on its center $x_i(t)$ is the average over $B_\delta(x_i(t))$ of the gradient of the pressure $f'(\mu_N(x))$ – which depends on the counting function μ_N.

System (3.5) is similar to the deterministic system used for example in [36, 30, 37, 31] to approximate non-linear diffusion equation via what has been called the blob method. One advantage of this approach is the fact that given a solution $\{x_i(t)\}_{i=1,\ldots,N}$ of (3.4), the empirical distribution $\rho_N(t, x)$ is a solution (in the sense of distribution) of the nonlinear transport equation

$$(3.6) \qquad \partial_t \rho + \operatorname{div}(\rho \nabla \widetilde{G} * \rho) = \operatorname{div}(\rho \nabla K_\delta * f'(K_\delta * \rho)).$$

However, it has been observed (see for example [111]) that the long-time behavior of (3.6) exhibits discrepancies with that of the microscopic dynamics (3.5): When $\widetilde{G}$ is zero and K_δ is compactly supported, numerical simulations show that solutions of (3.6) with absolutely continuous densities spread and converge (locally) to zero (while preserving the mass) as $t \to \infty$ (which is consistent with the behavior of solutions of the porous media equation) whereas, for a finite number of particles N, the solutions to (3.5) converge to a finite non-zero value (in particular, the dynamics stops when all the points $x_i(t)$ are at distance greater than 2δ from one another). So, while we have a one-to-one correspondence between the solutions of (3.5) and empirical measures solutions of (3.6), these empirical measures are *unstable* solutions of (3.6) (in the sense that perturbations of the Dirac masses diffuse in space under the dynamics of (3.6), hence departing further from their original Dirac structure). This discrepancy between the microscopic (ODE) and macroscopic (PDE) dynamics is not unique to this model and we shall see this again for the hard-sphere (or hard congestion) models in the next section [4].

Note that (3.6) is reminiscent of the transport equation with Brinkman's law, which can be written as

$$(3.7) \qquad \partial_t \rho = \operatorname{div}(\rho \nabla \widetilde{G} * \rho) + \operatorname{div}(\rho \nabla K_\delta * f'(\rho)), \qquad K_\delta - \delta \Delta K_\delta = \delta_{x=0}.$$

This model has a similar behavior as (3.6) when $f'(\rho) = \rho$, but when $f(\rho)$ is given by (1.2) with $m \gg 1$, (3.7) imposes a height constraint on ρ while (3.6) imposes a constraint on $K_\delta * \rho$. Both the dynamics and the mathematical theory are quite different in that case.

---

[4]Motsch and Peurichard [111] proposed a stabilizing method which consists in assuming that $f(\rho) = 0$ for $\rho \leq \rho^*$. This is to ensure that no repulsion occurs when the particle density ρ falls below some threshold value ρ_*.

3.4. **A nonlocal, nonlinear Keller-Segel model.** In the parabolic-elliptic approximation of the Patlak-Keller-Segel model, the attractive potential $\phi = G * \rho$ solves $\sigma\phi - \eta\Delta\phi = \rho$ and represents the concentration of a self-emitted chemo-attractant substance. At the microscopic scale, we should also take into account the finite size of the particles B_δ (here, we take $K = \chi_{B_1}$ again): First, the production of the chemo-attractant occurs over the whole ball $B_\delta(x_i)$ instead of being concentrated at the center x_i, leading to

$$(3.8) \qquad \sigma\phi_\delta - \eta\Delta\phi_\delta = K_\delta * \rho_N.$$

Next, the particles sensors are also located over the whole ball $B_\delta(x_i)$, rather than only in the center x_i, so that the velocity is given by $\nabla K_\delta * \phi_\delta(x_i(t))$:

$$(3.9) \qquad \dot{x}_i(t) = \nabla K_\delta * \phi_\delta(x_i(t)) - \nabla K_\delta * f'(K_\delta * \rho_N(x))(x_i(t)).$$

This system is exactly (3.4) if we replace $\widetilde{G}$ with $\widetilde{G}_\delta := K_\delta * \widetilde{G} * K_\delta$. The good news is that while G is too singular to apply the existing well-posedness theory, the regularized kernel $\widetilde{G}_\delta$ is much nicer. In particular[5], we have $\widetilde{G}_\delta \in W^{2,\infty}$ as soon as $K_\delta \in L^2$.

Combining the results mentioned above (for instance [22] and [31]), we deduce that when $\{x_i(t)\}_{i=1,\dots,N}$ solves (3.8)-(3.9), the empirical distribution $\rho_N(t, x)$ solves (1.6) which is the gradient flow in $\mathcal{P}_2(\mathbb{R}^d)$ for the energy functional (2.4). This gradient flow is studied in [22] when $f = 0$ (in particular $\widetilde{G}_\delta$ satisfies the assumptions of [22] when $K_\delta \in L^2$) and in [31] when $\widetilde{G}_\delta = 0$ and f satisfies some assumptions satisfied in particular when $f(\rho) = \frac{1}{m-1}\rho^m$ (but with additional assumption on K_δ). We also refer to earlier work in a similar spirit by Carrillo, Craig, and Patacchini [30]. The special case $m = 2$ has been given detailed attention by Burger and Esposito [19] and Craig, Elamvazhuthi, Haberland and Turanova [37] (with applications in cross-diffusion and sampling, respectively). Adapting the method developed in these papers (in particular [22, 30, 37]), we get:

Theorem 3.1. *Fix $\delta > 0$, assume that $K \in C^2(\mathbb{R}^d)$, $K \in W^{2,\infty}(\mathbb{R}^d)$, and take $f = f_m$ for some $m > 1$. For all $\rho_{in} \in \mathcal{P}_2(\mathbb{R}^d)$ with $\mathscr{E}_\delta[\rho_{in}] < \infty$ (where E_δ is defined by (2.4)), there exists a unique gradient flow solution $\rho^\delta(t)$ in $AC_{loc}([0,\infty); \mathcal{P}_2(\mathbb{R}^d))$ of (1.6). This solution is characterized by:*

$$(3.10) \qquad \begin{cases} \partial_t\rho^\delta + \operatorname{div}(\rho^\delta\nabla\widetilde{G}_\delta * \rho^\delta) = \operatorname{div}(\rho^\delta\nabla K_\delta * f'_m(K_\delta * \rho^\delta)) & in \ \mathcal{D}'((0,\infty) \times \mathbb{R}^d) \\ \lim_{t\to 0^+} d_W(\rho^\delta(t), \rho_{in}) = 0. \end{cases}$$

Furthermore, given two such solutions $\rho^1(t)$ and $\rho^2(t)$, the following stability estimate holds

$$(3.11) \qquad d_W(\rho^1(t), \rho^2(t)) \leq d_W(\rho^1(0), \rho^2(0))e^{-\lambda_0 t}$$

with

$$(3.12) \qquad \lambda_0 := -C(\|K\|_{W^{2,\infty}(\mathbb{R}^d)})\delta^{-d(m-1)-2} + C(\|K\|_{L^2})\delta^{-d}.$$

Inequality (3.11) implies the mean-field limit for fixed δ: if the empirical measure converges to a density ρ_{in}, in the sense that $d_W(\rho_N(t = 0), \rho_{in}) \to 0$, then the empirical distribution $\rho_N(t, x)$ converges weakly to the solution ρ of (3.10). The constant λ_0 given by (3.12) is a (lower bound for the) modulus of convexity for the energy functional $\mathscr{E}_\delta$ (defined by (2.4)) with respect to the underlying optimal transport geodesic structure on $\mathcal{P}_2$. Importantly, we note that when $\delta \to 0$, we have $\lambda_0 \to -\infty$ as $\delta \to 0$. On the other hand, it is known that the limiting energy $\mathscr{E}$ is convex. This discrepancy suggests that this (3.12) is far from optimal.

---

[5]Indeed, the function $u = G * K_\delta$ solves $\sigma u - \eta\Delta u = K_\delta \in L^2$ and is thus in H^2. We deduce $\|D^2\widetilde{G}_\delta\|_{L^\infty} = \|K_\delta * D^2 u\|_{L^\infty} \leq \|K_\delta\|_{L^2}\|D^2 u\|_{L^2} \leq C\|K_\delta\|_{L^2}^2$.

Sketch of the proof of Theorem 3.1. Our assumptions on K_δ and the properties of $\widetilde{G}_\delta$ ensures that $\mathscr{E}_\delta$: $\mathcal{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ is proper, lower-semicontinuous and λ-convex with modulus of convexity given by (3.12) (see [22, 30]). The differentiability of $\mathscr{E}_\delta$ can be proved using the same arguments as in [22] (for the G_δ term) and [30] (for the f_m-term) and leads to

$$\frac{\delta \mathscr{E}_\delta}{\delta \rho}[\rho] = -\widetilde{G}_\delta * \rho + f'_m(K_\delta * \rho) * K_\delta.$$

Existence and uniqueness of the gradient flow ρ^δ, as well as the fact that the gradient flow is a curve of maximal slope, follows from [5, Theorem 11.2.1]. Furthermore, ρ^δ solves the continuity equation

$$\partial_t \rho^\delta + \operatorname{div}(\rho^\delta v^\delta) = 0, \quad \text{in } \mathcal{D}'(\mathbb{R}^d \times (0, \infty)),$$

with velocity

$$v^\delta(t) = \nabla \widetilde{G}_\delta * \rho^\delta(t) - \nabla K_\delta * f'_m(K_\delta * \rho^\delta(t)) \qquad \text{a.e. } t > 0,$$

and satisfies the energy inequality

$$(3.13) \qquad \mathscr{E}_\delta[\rho^\delta(t)] + \int_0^t \int_{\mathbb{R}^d} |v^\delta|^2 d\rho^\delta(s)\, ds \le \mathscr{E}_\delta[\rho_{\text{in}}].$$

We note that the presence of the convolution and the assumption on K_δ implies that v^δ is uniformly bounded (and C^1), so ρ^δ solves (3.10). Conversely, this also implies (see [37, Proposition 3.12]) that any solution of (3.10) is in fact in $AC_{loc}([0, \infty); \mathcal{P}_2(\mathbb{R}^d))$ and is a gradient flow of $\mathscr{E}_\delta$.

Finally we get (3.11) as a consequence of the usual stability estimate for λ-gradient flows [5, Theorem 11.2.1] (see also [37, Theorem 1.4]). $\qquad \square$

For future references, we also recall that ρ^δ satisfies the propagation of the second moment:

$$(3.14) \qquad \int_{\mathbb{R}^d} |x|^2 \rho^\delta(t, x)\, dx \le \left(\int_{\mathbb{R}^d} |x|^2 \rho_{\text{in}}(x)\, dx + \mathcal{E}_\delta[\rho_{\text{in}}] \right) e^t,$$

which follows from (3.13) since we have

$$\frac{d}{dt} \int_\Omega |x|^2 \rho^\delta(t, x)\, dx = 2 \int_\Omega x \cdot v^\delta \rho^\delta\, dx \le 2 \left(\int_\Omega |x|^2 \rho^\delta\, dx \right)^{1/2} \left(\int_\Omega |v^\delta|^2 \rho^\delta\, dx \right)^{1/2}.$$

3.5. **Convergence to the macroscopic model: The limit $\delta \to 0$.** We now turn to the question of the convergence of the nonlocal model (1.6) to the classical aggregation-diffusion equation (1.9) when $\delta \to 0$. We note that such a convergence cannot hold unless (1.9) has a solution, and it is well-known that aggregation diffusion equations can experience finite time blow-up. When f is given by (1.2) and G is the chemotaxis potential (1.4), then global in time (bounded) solutions exist in the subcritical regime $m > m_c := 2 - \frac{2}{d}$ (more generally, if $G(x) \sim \frac{1}{|x|^k}$, then the critical power is given by $m_c := 1 + \frac{k}{d}$, see [24]). In fact, we have:

Theorem 3.2. *Let $f = \frac{1}{m-1}\rho^m$ with $m > 2 - \frac{2}{d}$ and G be given by (1.4). Given $\rho_{in} \in \mathcal{P}_2(\mathbb{R}^d)$ such that $\|\rho_{in}\|_{L^\infty(\mathbb{R}^d)} < \infty$ and $\operatorname{Supp}\rho_{in} \in B_R$, equation (1.9) has a unique global solution $\rho(t, x)$ with initial data $\rho_{in}(x)$.*

Furthermore, for all $T > 0$ there exist $C(T)$ and $R(T)$ such that

$$\|\rho\|_{L^\infty((0,T) \times \mathbb{R}^d)} \le C(T), \qquad \operatorname{Supp}\rho(t) \subset B_{R(T)} \quad \forall t \in [0, T].$$

We refer for example to [132, 91, 21, 8] for the existence of a global bounded solution. The propagation of the L^∞ norm and compact support can also be proved as in [68, Appendix A and B] and uniqueness is proved in [8].

While this result only requires $m > 2 - \frac{2}{d}$, we will now make the stronger assumption $m > 2$ to simplify the analysis (the assumption $m > 2$ will be crucial in the second part of this paper but the result below should hold for all $m > 2 - \frac{2}{d}$). Indeed, since $\int_{\mathbb{R}^d} G(x)\,dx = \frac{1}{\sigma}$, we can rewrite the energy (2.2) as

$$
\begin{aligned}
\mathscr{E}_m[\rho] &= \int_{\mathbb{R}^d} f_m(\rho)\,dx - \frac{1}{2}\iint_{\mathbb{R}^d \times \mathbb{R}^d} G(x-y)\,\rho(x)\,\rho(y)\,dx\,dy \\
&= \int_{\mathbb{R}^d} f_m(\rho) - \frac{1}{2\sigma}\rho^2\,dx + \frac{1}{2}\iint_{\mathbb{R}^d \times \mathbb{R}^d} G(x-y)\,\left[\rho(x)^2 - \rho(x)\,\rho(y)\right]dx\,dy \\
&= \int_{\mathbb{R}^d} f_m(\rho) - \frac{1}{2\sigma}\rho^2\,dx + \frac{1}{4}\iint G(x-y)[\rho(x) - \rho(y)]^2\,dx\,dy.
\end{aligned}
$$

When $m > 2$, the function $\rho \mapsto f_m(\rho) - \frac{1}{2\sigma}\rho^2$ is bounded below. Since the equation preserves $\int \rho(t,x)\,dx$ over time, we can add a term $a\rho$ to make it non-negative (see (1.13)). This leads to the function $\rho \mapsto h_m(\rho)$ defined by (1.14), which satisfies $h_m(\rho) \geq 0$ for all $\rho \geq 0$. We can then work with the energy functional

$$
\overline{\mathscr{E}_m}[\rho] := \begin{cases} \displaystyle\int h_m(\rho(x))\,dx + \frac{1}{4}\iint G(x-y)[\rho(x) - \rho(y)]^2\,dx\,dy & \text{if } \rho \in \mathcal{P}_2(\mathbb{R}^d),\ d\rho(x) \ll dx, \\ \infty & \text{otherwise,} \end{cases}
$$

which differs from $\mathscr{E}_m$ by a constant. Similarly, we can rewrite the energy (2.4), up to a constant, as

$$
\overline{\mathscr{E}_\delta}[\rho] := \int h_m(K_\delta * \rho(x))\,dx + \frac{1}{4}\iint G(x-y)[K_\delta * \rho(x) - K_\delta * \rho(y)]^2\,dx\,dy, \qquad \rho \in \mathcal{P}_2(\mathbb{R}^d).
$$

In particular, the energy inequality (3.13) holds with $\mathscr{E}_\delta$ replaced by $\overline{\mathscr{E}_\delta}$.

We can now prove:

Theorem 3.3. *Assume that G is given by (1.4) and that $f(\rho) = \frac{1}{m-1}\rho^m$ with $m > 2$ and let ρ_{in} be as in Theorem 3.2. Let $\rho^\delta(t,x)$ be the solution of (1.6) given by Theorem 3.1 and $\rho(t,x)$ be the solution of (1.9) given by Theorem 3.2. Then, the entire sequence $\rho^\delta(t)$ converges narrowly to $\rho(t)$ (uniformly in $[0,T]$).*

We provide a proof of this result in Appendix A by adapting the arguments developed in [31] to prove a similar results when $G = 0$.

Theorem 3.3 does not address the convergence of the microscopic model (3.5) when $\delta \to 0$ since the condition $\overline{\mathscr{E}_m}[\rho_{in}] < \infty$ excludes empirical measures. As explained in [31], this result can however be combined with (3.11): If we approximate the measure ρ_{in} by a sequence of empirical measures such that $d_W(\rho_{N,\delta}(t=0), \rho_{in}) \leq \frac{C}{N}$ and $N = o\left(e^{-\frac{1}{\delta^{2+d(m-1)}}}\right)$, then using (3.11) and the convergence of ρ^δ to ρ, we can show that the empirical distribution $\rho_{N,\delta}(t,x)$ converges to $\rho(t,x)$ solution of (1.9) with initial condition ρ_{in}. This restriction on N is very far from the scaling $N \sim \delta^{-d}$ necessary to preserve the total mass in our microscopic interpretation of the blob model. Numerical evidence (see [37, 36]) suggests that a good level of approximation holds for much smaller number of particles but justifying this limit is a challenging open problem.

To summarize, equation (3.6) preserve the empirical measure and thus describe precisely the evolution of a system of particles evolving according to the microscopic model (3.5). When $\delta \to 0$, this equation (1.6) is an approximation of (1.9) (which does not preserve the Dirac structure of ρ_N in its evolution).

To the best of our knowledge, this approach to approximating nonlocal diffusion equations by particle methods originates from the works of Mas-Gallic [99] for kinetic equations and Oelschläger [113] for the quadratic porous medium equation. We also mention [95, 44, 59, 25] in this direction. The convolution

by K_δ alters the diffusion in the sense that particles remain particles: $\rho_N(t) = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i(t)}$ is a solution to the regularised equation where each particle $x_i(t)$ evolves according to the system of characteristics.

3.6. **Other microscopic models with repulsion.** A similar regularization approach can be used, at least formally, to justify other macroscopic models. For example, Karper, Lindgren and Tadmor [82] followed a similar approach and proposed the following microscopic model for chemotaxis:

$$\dot{x}_i(t) = \nabla\phi(x_i(t)) - \frac{\beta(\mu_N(x_i))}{\mu_N(x_i(t))}\nabla\mu_N(x_i(t)),$$

with $\beta(\mu) := \frac{\mu}{1-\mu}$. The corresponding empirical distribution then solves

$$(3.15) \qquad \partial_t\rho_N + \operatorname{div}(\rho_N\nabla\phi) = \operatorname{div}\left(\frac{\rho_N}{K_\delta * \rho_N}\beta(K_\delta * \rho_N)\nabla(K_\delta * \rho_N)\right)$$

and in the limit $N \to \infty$ and $\delta \to 0$, we formally obtain

$$(3.16) \qquad \partial_t\rho + \operatorname{div}(\rho\nabla\phi) = \operatorname{div}\left(\frac{\rho}{1-\rho}\nabla\rho\right)$$

which corresponds to (1.9) with $f(\rho) = (1-\rho)\ln(1-\rho)$. The use of such singular pressure is a classical way to enforce a congestion constraint $\rho \le 1$ (see for instance [116]) and a model similar to (3.16) was derived in [98] from a stochastic particle system with volume effect, by enforcing the fact that two particles (which are assumed to be rigid squares in [98]) cannot overlap. Justifying the limit from (3.15) to (3.16) rigorously, in the spirit of Theorem 3.3, is, to the best of our knowledge an interesting open problem.

Other microscopic models have been proposed to describe short range repulsion. For example, Fischer, Kanzler and Schmeiser [60] consider a one-dimensional model of particles $x_0(t) < x_1(t) < \cdots < x_N(t)$ in which each agent x_i only interacts with its two immediate neighbors x_{i-1} and x_{i+1} via a repulsive force which depends only on their distance. They derive a nonlinear diffusion model depending on the repulsive force (in [84] we use a similar approach to study the hard sphere model in dimension 1). The extension of this model to higher dimensions is highly non-trivial and is very much related to similar difficulties faced in the so-called 'upwind' scheme for hyperbolic equations [47, 56] for continuity equations with non-linear mobilities. Other topological models, in which interactions depend on the "rank" of a neighbor rather than its distance, have been studied for example by Blanchet and Degond [12, 13]. Finally, let us point out that there are many models of pedestrian flows that also include the congestion constraint (see for instance [74, 93, 35]).

4. HARD-SPHERE (INCOMPRESSIBLE) MODELS

We now turn our attention to the hard-sphere (or incompressible) model. At the microscopic scale, the particles, still represented by balls of radius $\delta > 0$, are prevented from overlapping (congestion constraint). At the macroscopic scale, this will lead to the assumption that the population density cannot increase above a prescribed value.

On the one hand, we will see that the relationship between microscopic and macroscopic models is a lot more complex in this case than in the soft sphere-case of the previous section. On the other hand, we will show that the macroscopic hard-sphere model can easily be derived as the singular limit $m \to \infty$ of the (macroscopic) soft sphere models. In the first part of the discussion, we assume that the "desired" velocity field is given a priori, but we also discuss the well-posedness properties of this macroscopic model when this velocity field is the gradient of the attractive chemotaxis potential.

4.1. **The microscopic hard-sphere model.** The microscopic model we now introduce has been extensively studied in the context of congested crowd motion, in particular by Maury, Santambrogio et al. (see [101, 102, 103, 58]). In this context the domain Ω and the boundary condition (existence of "exits") play a very important role; this aspect will not be discussed in this paper.

Unlike the soft sphere model, in which the velocity is modified by a repulsive force, we now impose the hard-sphere constraint

$$(4.1) \qquad D_{ij}(\mathbf{x}(t)) := |x_i(t) - x_j(t)| \geq 2\delta \qquad \text{for all } i \neq j, \text{ for all } t \geq 0.$$

In terms of the empirical measure $\rho_N(t,x) = \frac{1}{N} \sum_{j=1}^{N} \delta(x - x_j(t))$, and using the notation of the previous section with $K = \chi_{B_1}$, we can also write this constraint as

$$(4.2) \qquad \mu_N = K_\delta * \rho_N \leq 1.$$

As a consequence, the actual velocity of a particle will be the projection of the desired velocity onto the set of admissible velocity, namely velocities that preserve the constraint (4.1).

Following [102], we assume that the particle i has a desired velocity v_i which might depend on the positions of all the particles. Below, we use the notation $\mathbf{x} = (x_1, \ldots, x_N) \in (\mathbb{R}^d)^N$ to denote the position of the N particles, and $\mathbf{v}(\mathbf{x}) = (v_1(\mathbf{x}), \ldots, v_N(\mathbf{x}))$ for their desired velocity. The constrained microscopic model can then be written (see [102]):

$$(4.3) \qquad \dot{\mathbf{x}}(t) = \mathbf{u}(\mathbf{x}), \quad \mathbf{u} = P_{C(\mathbf{x})}\mathbf{v},$$

where $P_{C(\mathbf{x})}\mathbf{v}$ denotes the projection of $\mathbf{v}$ onto the set $C(\mathbf{x})$ of admissible velocities. If we define $D_{ij}(\mathbf{x}) := |x_i - x_j|$, then admissible velocities must be such that $D_{ij}(\mathbf{x})$ cannot decrease if the i and j particles are already touching. This leads to

$$C(\mathbf{x}) := \{ \mathbf{u} \in (\mathbb{R}^d)^N \ : \ \nabla D_{ij}(\mathbf{x}) \cdot \mathbf{u} \geq 0 \text{ for all } i,j \text{ such that } D_{ij}(\mathbf{x}) = 2\delta \}.$$

Alternatively, using the fact that $\frac{d}{dt}|x_i(t) - x_j(t)|^2 = 2(x_i(t) - x_j(t)) \cdot (u_i(\mathbf{x}(t)) - u_j(\mathbf{x}(t)))$, we can write

$$(4.4) \qquad C(\mathbf{x}) = \{ \mathbf{u} \in (\mathbb{R}^d)^N \ : \ (u_i - u_j) \cdot (x_i - x_j) \geq 0 \text{ for all } i,j \text{ such that } |x_i - x_j| = 2\delta \}.$$

Following [103, 102], one can show that the projection $\mathbf{u} = P_{C(\mathbf{x})}\mathbf{v}$ in $(\mathbb{R}^d)^N$ can be characterized as follows: Using the notation $i \sim j$ to indicate that the particle i is in contact with the particle j and $e_{ij} = \frac{x_j - x_i}{|x_j - x_i|}$, we have

$$(4.5) \qquad u_i = v_i - \sum_{j \sim i} p_{ij} e_{ij},$$

where the pressure $p_{ij} = p_{ji}$ satisfies

$$p_{ij} \geq 0, \quad u_i \cdot e_{ij} + u_j \cdot e_{ji} \leq 0, \qquad \forall i \sim j$$

(this last condition expresses the non-overlapping constraint) and

$$u_i \cdot e_{ij} + u_j \cdot e_{ji} = 0 \text{ if } p_{ij} \neq 0.$$

This condition expresses the fact that if $p_{ij} \neq 0$, then the particles will stay in contact, i.e. the pressure prevents overlapping but does not push the particles away from each other.

The existence of a unique solution to (4.3) (for $\mathbf{v}$ Lipschitz and bounded) is proved in [103] (see also [101, Theorem 3.2]). Except in the case of $d = 1$, this does not fall into standard ODE theory because the set of admissible particle configurations

$$Q = \{ \mathbf{x} = (x_1, x_2, \ldots, x_N) \in (\mathbb{R}^d)^N \ : \ |x_i - x_j| \geq 2\delta, \quad \forall i \neq j \}$$

is not convex in general. But it satisfies the weaker property of *unifom prox-regularity* which is sufficient for well-posedness.

As a final comment, we recall that in this microscopic model, it is possible to consider that each particle x_i has its own desired velocity v_i (see [102]). This distinction is not possible in the macroscopic model, which does not distinguish particles from each other. Thus in consideration of passing to the macroscopic model it makes sense to assume that the velocity of the ith particles $v_i(\mathbf{x})$ is equal to $\nabla\phi(x_i)$ for some potential $\phi : \mathbb{R}^d \to \mathbb{R}$. Given the corresponding projection $\mathbf{u}$, we denote by $\mathbb{P}_C(\nabla\phi)$ a vector field $\mathbb{R}^d \to \mathbb{R}^d$ such that $\mathbb{P}_C(\nabla\phi)(x_i) = u_i(\mathbf{x})$ (only the values of this vector field at the points x_i are uniquely defined). We can then write that the empirical distribution satisfies

$$(4.6) \qquad \partial_t \rho_N + \operatorname{div}\left(\rho_N \mathbb{P}_C(\nabla\phi)\right) = 0, \qquad \mu_N(t) = K_\delta * \rho_N \leq 1.$$

4.2. **The macroscopic hard-sphere model.** Formally, the same principle can be applied to the macroscopic system: Here, we consider the formulation (4.6) of (4.3) and we write that the density $\rho(t, x)$ solves the continuity equation

$$\partial_t \rho + \operatorname{div}\left(\rho\, \mathbb{P}_{C_\rho}(\nabla\phi)\right) = 0, \qquad \rho \leq 1.$$

The set of admissible velocities, which preserves the local constraint $\rho \leq 1$, can now be defined by duality as

$$C_\rho = \left\{ U \in L^2(\mathbb{R}^d; \mathbb{R}^d)\,;\, \int_{\mathbb{R}^d} U \cdot \nabla q \leq 0, \quad \forall q \in H^1_\rho(\mathbb{R}^d) \right\},$$

where

$$H^1_\rho(\mathbb{R}^d) = \left\{ q \in H^1(\mathbb{R}^d)\,;\, q \geq 0 \text{ a.e. in } \mathbb{R}^d \text{ and } q(x)(1 - \rho(x)) = 0 \text{ a.e.} \right\}$$

(we can also write $H^1_\rho(\mathbb{R}^d) = \left\{ q \in H^1(\mathbb{R}^d)\,;\, q(x) \in \partial f_\infty(\rho(x)) \text{ a.e.} \right\}$, but the condition $q(x)(1 - \rho(x)) = 0$ is a common way of writing the condition $q(x) = 0$ a.e. in $\{\rho < 1\}$)

Formally at least, this projection operator guarantees that the density satisfies the incompressibility constraint $\rho \leq 1$ since it implies $\operatorname{div} U \geq 0$ in (the interior of) the set $\{\rho = 1\}$. In order to develop a well posedness theory for this model, we first show that it can be rewritten without the projection operator via the introduction of a pressure term (as is commonly done in incompressible fluid dynamic - although in our case the constraint is $\rho \leq 1$ and not $\rho = 1$).

To derive this formulation, we note that the projection $U := \mathbb{P}_{C_\rho}(\nabla\phi)$ is characterized by the following inequality:

$$(4.7) \qquad \int_{\mathbb{R}^d} (U - \nabla\phi) \cdot (U - Q)\, dx \leq 0, \qquad \forall Q \in C_\rho.$$

This projection gives rise to a Lagrange multiplier in the form of a pressure term: One can show[6] that there must exist $p \in H^1_\rho(\mathbb{R}^d)$ such that

$$U = \nabla\phi - \nabla p.$$

The velocity $U = \mathbb{P}_{C_\rho}(\nabla\phi)$ is in C_ρ and thus satisfies $\int_{\mathbb{R}^d} U \cdot \nabla q\, dx \leq 0$ for all $q \in H^1_\rho(\Omega)$ and the pressure satisfies the orthogonality condition (which follows from (4.7) by taking $Q = 0$ and $Q = 2U$): $\int_{\mathbb{R}^d} U \cdot \nabla p\, dx = 0$. We deduce that $-\int_{\mathbb{R}^d} U \cdot \nabla(p - q)\, dx \leq 0$ for all $q \in H^1_\rho$, which implies that p solves the variational inequality

$$(4.8) \qquad \begin{cases} p \in H^1_\rho \\ \displaystyle\int_{\mathbb{R}^d} [\nabla p - \nabla\phi] \cdot \nabla(p - q) \leq 0, \qquad \forall q \in H^1_\rho. \end{cases}$$

---

[6] This follows from Moreau's decomposition theorem since the set $\mathcal{K}_\rho = \{\nabla p\,;\, p \in H^1_\rho(\mathbb{R}^d)\}$ is a closed convex cone in $L^2(\mathbb{R}^d; \mathbb{R}^d)$ and its polar cone is exactly $\mathcal{K}^\circ_\rho = C_\rho$. It follows that all $\nabla\phi \in L^2(\mathbb{R}^d; \mathbb{R}^d)$ can be written in a unique way as $\nabla\phi = \mathbb{P}_{C_\rho}(\nabla\phi) + \mathbb{P}_{\mathcal{K}_\rho}(\nabla\phi)$.

This variational inequality is in fact an obstacle problem (because of the constraint $p \geq 0$). We can summarize this discussion with the following proposition (see [85, Appendix B]):

Proposition 4.1. *Assume $\phi \in H^1$ and let $p(x)$ be the (unique) solution of the variational inequality (4.8). Then $\mathbb{P}_{C_\rho}(\nabla\phi) = \nabla\phi - \nabla p$.*

The macroscopic hard-sphere model can thus be written as:

$$\partial_t\rho + \operatorname{div}(\rho\nabla\phi) = \operatorname{div}(\rho\nabla p), \qquad p \in \partial f_\infty(\rho). \tag{4.9}$$

4.3. **Well-posedness of the PKS hard-sphere model.** In this section, we discuss the well-posedness of the model (4.9) when the velocity field V is the attractive chemotaxis gradient from (1.10). We can also write this as:

$$\partial_t\rho + \operatorname{div}(\rho\nabla G * \rho) = \operatorname{div}(\rho\nabla p), \qquad p \in \partial f_\infty(\rho). \tag{4.10}$$

A solution of (4.10) can be constructed either as a gradient flow or as the singular limit of the nonlinear diffusion model.

The gradient flow approach was first developed in the context of congested crowd motion in [101]. In that paper, the velocity field was given by $V = -\nabla\phi$ for a continuous and λ-convex potential ϕ. A similar approach was later carried out with the chemotaxis potential in [85]: Equation (4.10) is the 2-Wasserstein gradient flow of the energy $\mathscr{E}_\infty$ defined by (2.3) and existence of a solution can be proved via a discrete time approximation based on a classical JKO minimization scheme: For a given initial data ρ_{in} such that $\mathscr{E}_\infty[\rho_{in}] < \infty$ (where $\mathscr{E}_\infty$ is defined by (2.3)), we fix a time step $\tau > 0$ and define the sequence ρ^n by:

$$\rho^0 = \rho_{in}, \qquad \rho^n \in \operatorname{argmin}\left\{\frac{1}{2\tau}W_2^2(\rho, \rho^{n-1}) + \mathscr{E}_\infty[\rho]\right\} \qquad \forall n \geq 1. \tag{4.11}$$

The piecewise constant function defined by $\rho^\tau(t) := \rho^{n+1}$ for $t \in [n\tau, (n+1)\tau)$ can then be proved to converge (as $\tau \to 0$) to a solution of (4.10) which gives the following theorem (see [85]):

Theorem 4.2. *Let G be given by (1.4). Given $\rho_{in} \in \mathcal{P}_2(\mathbb{R}^d)$ such that $\rho_{in} \leq 1$ and $\operatorname{Supp}\rho_{in} \in B_R$. Then Equation (1.10) (or (4.10)) has a unique global weak solution $\rho(t, x)$ with initial data $\rho_{in}(x)$. This solution is such that $\rho \in L^1(0, \infty; L^\infty(\Omega)) \cap C^{1/2}(0, \infty; \mathcal{P}(\Omega))$, $p \in L^2(0, \infty; H^1(\Omega))$ with*

$$0 \leq \rho \leq 1, \quad p \geq 0, \quad p(1 - \rho) = 0 \quad a.e. \ in \ [0, \infty) \times \mathbb{R}^d$$

(note that the conditions $\rho \in C^{1/2}([0, \infty); \mathcal{P}_2(\mathbb{R}^d))$ and $p \in L^2(0, \infty; H^1(\mathbb{R}^d))$ are enough to make sense of the product $p(1 - \rho) = 0$). Furthermore, there exists $R(t)$ such that

$$\operatorname{Supp}\rho(t) \subset B_{R(T)} \quad \forall t \in [0, T].$$

When the drift term $\nabla\phi$ is fixed, uniqueness with density constraint is proved in [48] assuming some monotonicity for $\nabla\phi$, and in [75] in the spirit of DiPerna-Lions theory for transport equations, assuming some sobolev regularity for ϕ. Uniqueness for a related model (without the drift term) was also proved in [120] using a duality method that turns out to be very adaptable and was used in [85] and [68] to prove uniqueness for (4.10).

Theorem 4.2 (existence and uniqueness) also follows from [68], where a solution is constructed by passing to the limit $m \to \infty$ in the soft sphere model (1.9). Note that the incompressible limit $m \to \infty$ for (1.9) is a classical problem that goes back to the 80's: the "stiff (or incompressible) limit" of the porous media equation. We can understand this limit by noticing that (1.9) has the same form as (4.10) but with $p = f'_m(\rho)$. We can also make sense of this limit at the level of the energy. Indeed, equation (1.9) is the 2-Wasserstein gradient flow for the energy $\mathscr{E}_m$ defined by (2.2) which Γ-converges to $\mathscr{E}_\infty[\rho]$. The Γ-convergence of the energy does not imply the convergence of the corresponding gradient flow, but it is a good indication that such a convergence might take place.

There are numerous references for this limit $m \to \infty$ in various frameworks: We refer to [135, 63, 52] for the classical porous media equation $\partial_t \rho = \mathrm{div}\,(\rho \nabla f'_m(\rho))$ and to [120, 41, 90, 105, 87, 89, 42] for more recent results in the context of tumor growth modeling, which includes a non-negative source term $\partial_t \rho = \mathrm{div}\,(\rho \nabla f'_m(\rho)) + \rho F(f'_m(\rho))$ (more recently [67], these results were extended to the case of the porous media with an absorption term, which leads to interesting behavior). For the equation with a fixed drift term, the limit was first investigated with a fixed potential $\partial_t \rho + \mathrm{div}\,(\rho \nabla \phi) = \mathrm{div}\,(\rho \nabla f'_m(\rho))$ which is an important pedestrian model for congested crowd motion (see [38, 2, 88, 40]). Finally, the limit was established for the attractive chemotaxis problem (1.4) with pure Newtonian attraction $\sigma = 0$ (the analysis can be extended to the case $\sigma > 0$, which does not affect the singularity of the kernel at 0): This was first done in [38] for particular solutions (patch solutions) using variational methods and the viscosity solution approach. The convergence of the gradient flow of $\mathscr{E}_m$ for general initial conditions was then proved in [39], and finally, the convergence of weak solutions was proved in [68] by a more direct approach (which can be generalized to include source terms for example, see [79]).

The key feature of this limit is the fact that $p_m(\rho) = f'_m(\rho) \to \partial f_\infty(\rho)$ when $m \to \infty$. There are other natural pressure terms that lead to a similar result. For example, one can replace the pressure $p_m(\rho)$ with a singular pressure such as

$$(4.12) \qquad p_\alpha(\rho) = \begin{cases} \alpha \frac{\rho}{1-\rho} & \text{if } \rho < 1, \\ \infty & \text{if } \rho \geq 1, \end{cases}$$

or

$$(4.13) \qquad p_\alpha(\rho) = \begin{cases} -\alpha \log(1-\rho) & \text{if } \rho < 1, \\ \infty & \text{if } \rho \geq 1. \end{cases}$$

Such pressure terms lead to solutions satisfying $\rho < 1$ (for $\alpha > 0$), but in the limit $\alpha \to 0$, we find $p_\alpha(\rho) \to \partial f_\infty(\rho)$. The corresponding limits were made rigorous in [69, 43, 40] for (4.12) (with a fixed potential ϕ and with convergence rates) and in [82] for (4.13) (with the attractive chemotaxis potential with $\sigma = 0$).

4.4. **The projected velocity.** In the approach discussed above, solutions of (4.10) are constructed by imposing the constraint $\rho \leq 1$, rather than by using the projection $\mathbb{P}_{C_\rho}(\nabla \phi)$. However, it is possible to show that the effective velocity in (4.10), $\nabla \phi - \nabla p$, is indeed the expected projection. Following [101], one can show that

$$\int_{\mathbb{R}^d} \nabla q(x) \cdot (\nabla p(t_0, x) - \nabla \phi(t_0, x))\, dx = 0 \qquad \forall q \in H^1_{\rho(t_0)}, \text{ a.e. } t_0 > 0.$$

Since $p \in H^1_{\rho(t_0)}$, this implies that (4.8) holds for a.e. $t_0 > 0$. In fact it implies the stronger statement

$$\int_\Omega [\nabla p - \nabla \phi] \cdot \nabla(p - q) = 0 \qquad \forall q \in H^1_\rho.$$

By taking $q(x) = p(t_0, x)\varphi(t_0, x))$, we deduce in particular

$$\int_{\mathbb{R}^d} \nabla(p(t_0, x)\varphi(t_0, x)) \cdot (\nabla p(t_0, x) - \nabla \phi(t_0, x))\, dx = 0 \qquad \text{a.e. } t_0 > 0,$$

which implies the so-called *complementarity condition*

$$(4.14) \qquad p(\Delta p - \Delta \phi) = 0, \text{ in } \mathcal{D}'((0, T) \times \mathbb{R}^d), \quad p(t) \in H^1_{\rho(t)}.$$

It is also possible to derive these properties as part of the construction process. Indeed, both the solution of the nonlinear diffusion equation and the discrete time approximation given by the JKO scheme satisfy approximation of (4.14). Intuitively, passing to the limit in these approximations require the strong convergence of the pressure in $H^1(\Omega)$. This is done in [120] in the time-monotone model for

tumor growth. The case of the more complex model with a term representing nutrients (which is no longer monotone), was then treated in [41] by generalizing the so-called Aronson-Bénilan estimates to get some strong convergence on the gradient of the pressure.

This derivation was also performed in [67] by showing directly that for all $t > 0$, the pressure $p(t)$ is the unique solution of the variational inequality (4.8). This approach has two advantages: This formulation of the complementarity equation can be derived without requiring any strong convergence on ∇p, and unlike (4.14), it identifies the pressure for all time (rather than a.e.). In problems without monotonicity in time, the support of the pressure can indeed be smaller than the set $\{\rho = 1\}$, an important phenomenon which is clearly identified with this obstacle problem formulation. This obstacle problem (4.8) was derived for the limit of the JKO solution in [85] as well.

4.5. **Relation between the microscopic and macroscopic models.** Since both the microscopic particles model (4.3) and macroscopic model (4.9) have solutions, it is natural to ask whether these solutions converge to one another when $\delta \to 0$ (for the purpose of this discussion, we focus on the case of a fixed velocity field $V(x) = \nabla\phi(x)$). However, unlike the soft sphere model discussed in Section 3, the relation between the microscopic and macroscopic models is far from clear in the hard-sphere case. We recall that in the soft sphere case, this relationship was addressed in two steps: First there was a one-to-one correspondence between the solutions of the ODE system (3.9) and the empirical measures solutions of the PDE (1.6). Second, solutions of (1.6) (for absolutely continuous initial data) were proved to converge to solutions of (1.9) as $\delta \to 0$. We will show below that the first step still holds in the hard-sphere models: Solutions of the system of ODE with hard-sphere constraint (4.1) are associated to empirical measures solutions of (1.8). But the second step in the limiting process ($\delta \to 0$) is delicate at best. In fact, it is not obvious that the equation (1.8) is well-posed outside of the particular case of empirical measures associated to (3.9). Without a well-posedness theory, a convergence result (the equivalent of Theorem 3.3) seems out of reach.

Beyond the technical difficulties, we should also point out that the dynamics of the microscopic and macroscopic models are quite different. For example, when the particle i is in contact with several other particles, which is expected in congested configurations, the microscopic condition requires control on the velocity v_i in every direction $x_i - x_j$ for which $D_{i,j} = 0$. By contrast, the macroscopic constraint is a unique scalar constraint on the divergence of the velocity field. The dynamic of the microscopic model is thus much richer than that of the macroscopic model. Simple examples (see [100]) show that slightly different arrangements of the particles at the microscopic level (corresponding to the same macroscopic distribution) can lead to completely different behavior.

We now show that the nonlocal PDE (1.8) corresponds to the microscopic hard-sphere model when $K_\delta = \frac{1}{\omega_d \delta^d}\chi_{B_\delta}$. We already noted (see (4.2)) that the non overlapping constraint can be written as $\mu_N = K_\delta * \rho_N \leq 1$. Next, we describe the set of admissible velocity: if $\partial_t \rho_N + \operatorname{div}(\rho_N U) = 0$, and $p \geq 0$ is a measure supported on the set $\{\mu_N = 1\}$, then the condition "$\partial_t \mu_N \leq 0$ when $\mu_N = 1$" implies

$$\frac{d}{dt}\int K_\delta * \rho_N\, p\, dx = \int \rho_N\, U \cdot \nabla K_\delta * p\, dx \leq 0.$$

We can thus define the set of admissible velocities as

$$(4.15) \quad \widetilde{C}_\delta(\rho_N) = \left\{ U : \mathbb{R}^d \to \mathbb{R}^d;\ \int \rho_N\, U \cdot \nabla K_\delta * p\, dx \leq 0,\ \text{for all } p \geq 0,\quad p(1 - K_\delta * \rho_N) = 0 \right\}.$$

Importantly, this description at the level of empirical measures is consistent with (4.4): if $U(x_i) = u_i(\mathbf{x})$, then the condition $\mathbf{u} \in C_\delta(\mathbf{x})$ from (4.4) is equivalent to $U \in \widetilde{C}_\delta(\rho_N)$. For example, if the two particles i and j are touching (up to a translation, we can take $x_i = -\delta e$ and $x_j = +\delta e$), then taking $p(x) = (\delta - |x \cdot e|)_+ \mathcal{H}^1|_{\mathbb{R}e}$ in (4.15) will lead to the condition in (4.4).
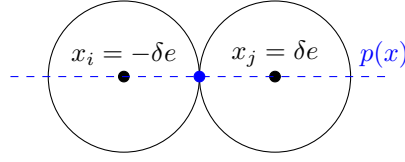
FIGURE 3. Schematic of two particles with radius δ in contact along the axis e. The pressure $p(x) = (\delta - |x \cdot e|)_+ \mathcal{H}^1|_{\mathbb{R}e}$ can be highly singular since it only needs to activate along the axis e.

The simple illustration in Figure 3 sheds some light on another discrepancy between the microscopic model (4.3) and the macroscopic model (4.9) (already discussed in [102, 101]): the set of feasible velocities C_δ depends not only on the local density but also on the orientation of particles. In the macroscopic model (4.9), the feasible velocity only depends on the local density.

The set $\widetilde{C}_\delta(\rho_N)$ is the polar cone of

$$\widetilde{\mathcal{K}}_\delta(\rho_N) = \{\nabla K_\delta * p \, ; \, p \geq 0, \quad p(1 - K_\delta * \rho_N) = 0\},$$

in the Hilbert space $L^2(\mathbb{R}^d, d\rho_N)^d$. In keeping with Moreau's decomposition theorem, we expect to be able to write

$$U = \mathbb{P}_{\widetilde{C}_\delta(\rho_N)}(\nabla \phi) = \nabla \phi - \nabla K_\delta * p,$$

for some pressure $p \geq 0$, satisfying $p(1 - \mu_N) = 0$. This is in fact correct when ρ_N is an empirical distribution: This decomposition is equivalent to the (finite dimensional) projection of $\mathbf{v}$ given by (4.5). The existence of a solution to the microscopic model (see Section 4.1) thus implies the existence of a solution (in the form of the associated empirical measure) of

(4.16) $$\partial_t \rho + \operatorname{div}(\rho \nabla \phi) = \operatorname{div}(\rho \nabla K_\delta * p), \qquad p \in \partial f_\infty(K_\delta * \rho),$$

when $\rho_{in} = \frac{1}{N}\sum_{i=1}^{N}\delta(x - x_i(0))$ (where we recall that $p \in \partial f_\infty(K_\delta * \rho)$ is equivalent to $K_\delta * \rho \leq 1$, $p \geq 0$, $p(1 - K_\delta * \rho) = 0$).

It is now natural to try to mirror the analysis described in Section 3 with this model for general density, namely to prove a mean field limit, as in Theorem 3.1 (the constant λ_0 given by (3.12) depends on m, so the extension to the case $m = \infty$ is not obvious) and the limit $\delta \to 0$. There is however a serious obstruction: Except for the particular case of empirical measures (for which the existence of a solution follows from the existence of a solution for the microscopic model), the existence of solution for (4.16) is far from clear. A discrete time approximation can be constructed via a JKO scheme with the energy

$$- \int_\Omega \phi(x)\rho(x)\,dx + \mathcal{F}_\infty[K_\delta * \rho],$$

but proving that the resulting velocity is indeed the projection of $\nabla \phi$ onto $\widetilde{C}_\delta(\rho)$ is challenging. Even then, it should also be noted that for general density ρ, it is not clear that the cone $\widetilde{\mathcal{K}}_\delta(\rho)$ is closed and thus Moreau's decomposition theorem does not apply. Similar difficulties arise when trying to obtain the existence of a solution of (1.8) by passing to the limit $m \to \infty$ in the soft sphere model (1.6). This limit has been studied for the related Brinkman law model (3.7) [121, 89], but this case is somewhat easier since the condition $\rho \leq 1$ (as opposed to $\rho * K \leq 1$) is being enforced.

5. Hard-sphere models and Hele-Shaw Free boundary problems

In some settings, the hard-sphere model (1.10) might lead to the formation and propagation of sharp edges. In fact the connection between the hard-sphere model (1.10) and Hele-Shaw free boundary problem (with active potential, but without surface tension) is well documented: If we assume that $\rho(t,x) = \chi_{E(t)}$ and denote $\phi(t,x) = G * \chi_{E(t)}(x)$, then (1.10) implies (formally at least):

$$(5.1) \qquad \Delta(p - \phi) = 0 \quad \text{in } E(t), \qquad V = -\nabla(p - \phi) \cdot \nu \quad \text{on } \partial E(t),$$

where V denotes the normal velocity of the interface $\partial E(t)$ and ν is the outward unit normal vector along $\partial E(t)$, with the condition

$$(5.2) \qquad p(t,x) = 0 \quad \text{on } \partial E(t).$$

Equations (5.1)-(5.2) (which one can write with the effective pressure $q = p - \phi$) is reminiscent of a Hele-Shaw free boundary problem (without surface tension) or one-phase Muskat problem.

However, the assumption $\rho(t,x) = \chi_{E(t)}$ is not typically satisfied by all solutions of (1.10). Since the equation only imposes the constraint $\rho(t,x) \leq 1$, there are no obvious reasons why ρ should be a characteristic function. But it can sometimes be proved that phase separation is propagated by the equation: If $\rho_{in} = \chi_{E_{in}}$, then $\rho(t) = \chi_{E(t)}$ for all $t > 0$. Such a result was proved in [105] and [87] for a different model, involving a non-negative source term but no drift. The crucial ingredient in these papers was the monotonicity of ρ with respect to t, and it was proved that $\rho(t) = \chi_{E(t)}$ a.e. $t > 0$ for some open set $E(t)$ with finite perimeter. For the aggregation equation, a similar result was proved in [38, Theorem 1.1 (b)] when G is the purely Newtonian attraction kernel using viscosity solution approach. Similar results were proved in [2] when ϕ is a fixed potential satisfying $-\Delta\phi \geq 0$ (see also [88] where the assumption is on the source term). This assumption, which implies some monotonicity of ρ along characteristic curves, is essential in these works. When G is given by (1.4), the condition $-\Delta\phi \geq 0$ is no longer satisfied but the fact that this potential is always attractive suggests that there is still some intrinsic monotonicity in the model. Indeed, the propagation of phase separation was proved in [85] for the hard-sphere PKS model (that is (1.10) with G given by (1.4)). The proof is very different from those mentioned above and the result is somewhat weaker: it is proved that for appropriate initial data and for a.e. $t > 0$ there is a set $E(t)$ such that $\rho(t) = 1$ a.e. in $E(t)$ and $\rho(t) = 0$ a.e. in $\mathbb{R}^d \setminus \overline{E(t)}$. The condition $\rho = \chi_{E(t)}$ would follow if one could show that $|\partial E(t)| = 0$, a property that appears difficult to get.

Making rigorous the connection between (1.10) and the common notions of weak solutions for the Hele-Shaw free boundary problem (e.g. viscosity solutions) is a delicate issue. Besides the question of phase separation, there is the issue of giving a meaning to (5.1) when the set $E(t)$ is not smooth. As discussed in Section 4.4, the first condition in (5.1) can be derived rigorously (it is the complementarity condition). The fact that the Hele-Shaw velocity law is satisfied in some viscosity solutions sense was proved in [38, Theorem 1.1] for the Newtonian kernel.

Part 2. Sharp interface limits and free boundary problems

In the second part of this paper, we review and extend recent results showing that when $m > 2$ and at appropriate scales, the competition between local repulsion and non-local attraction as described by equation (1.9) leads to phase separation and to the formation of an interface separating regions of high and low (or 0) density, interface whose evolution can be described by models of geometric nature. As noted in the introduction, phase separation plays a critical role in many biological processes and it has been shown that it can be explained by the simple attraction-repulsion mechanisms modeled by equation (1.1). We refer for instance to [134, 11] (for biological aggregations such as insect swarms) and to [57] (where phase separation is discussed for a continuum model of cell-cell adhesion). From a mathematical

view point, the main idea behind this phenomena is the close relationship between the aggregation-diffusion equation (1.1) and a degenerate Cahn-Hilliard equation. In this part of the paper, we describe an approach, based on the recent work [86, 104], to rigorously investigate phase separation phenomena and derive effective free boundary problems by considering appropriate singular limits of (1.1).

6. A non-local Cahn-Hilliard equation

Our goal is to derive effective models for aggregation-diffusion phenomena when the total mass $\int \rho(t,x)\,dx$ is very large and the phenomena are observed from far away (at an appropriate time scale): We introduce $\varepsilon \ll 1$ such that $\int_{\mathbb{R}^d} \rho(t,x)\,dx = \varepsilon^{-d}$ and rescale the space and time variables as follows:

$$x = \varepsilon \bar{x}, \qquad t = \varepsilon^2 \bar{t}.$$

The function $\rho^{\varepsilon}(t,x) := \rho(t/\varepsilon^2, x/\varepsilon)$ now satisfies $\int_{\mathbb{R}^d} \rho^{\varepsilon}(t,x)\,dx = 1$ and solves (1.11) which we recall here for the reader's convenience (note in particular that $G_{\varepsilon} * \rho^{\varepsilon}(x) = G * \rho(x/\varepsilon)$):

$$\partial_t \rho + \operatorname{div}\left(\rho \nabla (G_{\varepsilon} * \rho)\right) = \operatorname{div}\left(\rho \nabla f'(\rho)\right), \qquad G_{\varepsilon}(x) = \varepsilon^{-d} G(\varepsilon^{-1} x)$$

(alternatively, this amounts to studying (1.9) when the range of the nonlocal attraction is small compared to typical macroscopic length). Throughout this section, we will assume that f satisfies

(6.1) $$f(\rho) = f_m(\rho), \quad m > 2, \quad \text{or} \quad f(\rho) = f_{\infty}(\rho).$$

This assumption ensures that the function $h(\rho)$ defined by (1.13) is a double-well potential (see (1.14) and (1.15)), which is all we really need for the results presented here to hold.

Most of the rigorous results in this direction have only been proved when G is the chemotaxis potential (1.4) or some other very particular kernel (such as the heat kernel). Extending the theory (in particular Theorem 8.3) to general interaction potentials G describing local attraction is an interesting and important open problem. For the formal discussion below, it is enough to assume that

$$G(x) = G(|x|) \geq 0, \qquad \int G(x)\,dx = \sigma^{-1} \in (0, \infty) \quad \text{and} \quad \int_{\mathbb{R}^d} |x|^2 G(x)\,dx < \infty.$$

A crucial observation mentioned in the introduction is the fact that (1.11) can be written as a non-local Cahn-Hilliard equation (with a singular double-well potential). Indeed, since $G_{\varepsilon} * \rho \to \sigma^{-1} \rho$ when $\varepsilon \to 0$, it makes sense to rewrite (1.11) as

(6.2) $$\partial_t \rho = \operatorname{div}\left(\rho \nabla [f'(\rho) - \sigma^{-1} \rho]\right) - \operatorname{div}\left(\rho \nabla [G_{\varepsilon} * \rho - \sigma^{-1} \rho]\right).$$

Introducing the symmetric nonlocal operator

$$B_{\varepsilon}[\rho](x) := (G_{\varepsilon} * \rho - \sigma^{-1} \rho)(x) = \int_{\mathbb{R}^d} G_{\varepsilon}(x-y)[\rho(y) - \rho(x)]\,dy,$$

and using the function h defined by (1.13), we can write (6.2) as

(6.3) $$\partial_t \rho + \operatorname{div}\left(\rho v\right) = 0, \qquad v = -\nabla h'(\rho) + \nabla B_{\varepsilon}[\rho].$$

This equation is the gradient flow for the energy $\mathscr{J}_{\varepsilon}$ defined in (2.6) (which can be obtained from the energy (2.5) by a similar computation) with respect to the 2-Wasserstein distance. Equation (6.3) is closely related to a model derived in [61] by Giacomin and Lebowitz as a macroscopic model for phase separation in particle systems with long range interactions. This equation can be seen as a non-local approximation of the degenerate Cahn-Hilliard equation [20, 50]: When ρ is a smooth function, we have (using the symmetry of G):

$$B_{\varepsilon}[\rho] = \varepsilon^2 \beta \Delta \rho + \mathcal{O}(\varepsilon^4), \qquad \beta := \int_{\mathbb{R}^d} (z \cdot e)^2 G(z)\,dz$$

(for any $e \in \mathbb{S}^{d-1}$) and so (6.3) is close to

$$(6.4) \qquad \partial_t \rho + \operatorname{div}(\rho v) = 0, \qquad v = -\nabla h'(\rho) + \varepsilon^2 \beta \nabla \Delta \rho.$$

This connection between aggregation-diffusion equations and degenerate Cahn-Hilliard equations is discussed and used for instance in [11, 24, 57, 86, 104]. When h is a double-well potential (as is the case when (6.1) holds), it is well-known that (6.4) leads to phase-separation and can be approximated – at appropriate time scales – by Stefan or Hele-Shaw free boundary problems. The sharp interface limit $\varepsilon \to 0$ was first investigated via formal expansion method by Pego [119] (for the non-degenerate equation) and by Glasner [64] (for the degenerate equation). Rigorous justification of these formal asymptotic were established by Alikakos, Bates and Chen [4, 34] for the non-degenerate equation. Formal results were also obtained for the nonlocal equation (6.3) by Giacomin and Lebowitz in [62]. In this part, we present some rigorous results for such limits for the aggregation diffusion equation (6.3).

As a final comment, we note that the convergence of the rescaled version of (6.3)

$$\partial_t \rho + \operatorname{div}(\rho v) = 0, \qquad v = -\nabla h'(\rho) + \varepsilon^{-2} \nabla B_\varepsilon[\rho],$$

to the degenerate Cahn-Hilliard equation,

$$\partial_t \rho + \operatorname{div}(\rho v) = 0, \qquad v = -\nabla h'(\rho) + \beta \Delta \rho,$$

was recently proved in [50] on the torus and when G_ε is a compactly supported approximation of unity (see also [26] for a similar result for systems of Cahn-Hilliard equations).

7. Phase separation: Stefan free boundary problems

Going back to (6.3), we now investigate the asymptotic behavior of the solutions as $\varepsilon \to 0$. Since $B_\varepsilon[\rho]$ vanishes as ε vanishes, we expect the asymptotic dynamics to be close to that of the forward-backward diffusion equation

$$(7.1) \qquad \partial_t \rho = \operatorname{div}(\rho \nabla h'(\rho)),$$

(recall that under (6.1), $\rho \mapsto h(\rho)$ is not convex). Such equations are commonly associated with the modeling of phase separation mechanisms, but are ill-posed from a mathematical point of view (see [73]). One would expect the limits of solutions of (6.3) to inherit some additional properties, but characterizing these limits is quite challenging, even for the classical Cahn-Hilliard equation. The simplest setting is when the initial data does not take values in the mushy region $(0, \theta)$: such initial data are usually referred to as well-prepared. In that case, the limit $\varepsilon \to 0$ of (6.3) leads to the equation

$$(7.2) \qquad \partial_t \rho = \operatorname{div}(\rho \nabla h^{**\prime}(\rho)),$$

where h^{**} denotes the convex hull of h. When $h = h_m$ in (1.14), we have

$$h_m^{**}(\rho) = \begin{cases} 0 & \text{for } \rho \in [0, \theta_m]; \\ h_m(\rho) & \text{for } \rho \geq \theta_m. \end{cases}$$

With this choice of h (7.2) corresponds to a generalized Stefan free boundary problem, where ρ is the enthalpy variable. This equation propagates phase separation in the sense that the mushy region $\{\rho(t, x) \in (0, \theta)\}$ is non-increasing with respect to t (see for example [66]). So when $|\{\rho(x, 0) \in (0, \theta)\}| = 0$, this property remains true for all $t > 0$, and there exists a set $E(t)$ such that

$$\rho > \theta \text{ in } E(t), \qquad \rho = 0 \text{ in } \mathbb{R}^d \setminus E(t).$$

The evolution of the sharp interface $\partial E(t)$ is then described by the Stefan problem (7.2) which is a weak formulation of (1.17).

When the initial data is not well-prepared, further characterization of the limits for solutions of (6.3) is required to understand how phase separation arises. This has been studied for the non-degenerate

Cahn-Hilliard equation. First we note (see Figure 1) that there exists $\underline{\theta} \in (0, \theta)$ such that $h'' > 0$ in $(\underline{\theta}, \infty)$ and $h'' < 0$ in $(0, \underline{\theta})$. When ρ_{in} does not take values in the unstable region $(0, \underline{\theta})$, Plotnikov [123] characterized the limits in dimension 1 and showed that the limiting problem may exhibit hysteresis phenomena. In general the problem is very unstable and numerical computations show that oscillations appear in the region where $\rho_{in}(x) \in (0, \underline{\theta})$ (see [51, 10]). Passing to the limit $\varepsilon \to 0$ requires a weak formulation of (7.1) involving Young's measures (see for example Plotnikov [124] and Slemrod [130]). We will not review these results here, but we stress the fact that the dynamics of this phase separation process in the set $\{\rho_{in} \in (0, \underline{\theta})\}$ is very rich, even in dimension 1 (see [10]). We are not aware of any similar analysis for the degenerate Cahn-Hilliard equation or for its non-local approximation (6.3).

Even with well-prepared initial data, justifying the convergence to the Stefan problem (7.2) is delicate. A similar limit has been studied for the degenerate Cahn-Hilliard equation (6.4). The convergence of (6.4) to (7.2), studied formally in [119], was established rigorously only in dimension 1 by Delgadino in [45], for well-prepared initial data. Delgadino relied on the approach developed by Sandier and Serfaty [129] to prove the convergence of gradient flows using the fact that the Allen-Cahn energy (2.7) Γ-converges to $\int_{\mathbb{R}^d} h^{**}(\rho(x)) \, dx$. Convergence to a Stefan free boundary problem was also proved by a similar approach for the 1-dimensional non-degenerate Cahn-Hilliard equation in [9]. But as far as we know no such results are known for our non-local equation (6.3) except for some formal results for a nonlocal equation similar to (6.3) in [62] (via formal expansion and matched asymptotic analysis). Extending the rigorous analysis of [45] to this equation is an interesting open problem. We note that such a result appears plausible at the level of the energy: indeed a corresponding Γ-convergence result can be obtained for the energy $\mathscr{J}_\varepsilon$ given by (2.6).

Proposition 7.1. *Assume that $G \in L^1(\mathbb{R}^d)$ and that $\int_{\mathbb{R}^d} |z|^2 G(z) \, dz < \infty$. Then $\mathscr{J}_\varepsilon$ (defined by (2.6)) Γ-converges to $\mathscr{J}^*$ (defined by (2.8)). More precisely, we have*

(i) *For any sequence $\rho^\varepsilon \in \mathcal{P}_2(\mathbb{R}^d)$ that converges weakly to $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, we have*

$$\liminf_{\varepsilon \to 0} \mathscr{J}_\varepsilon(\rho^\varepsilon) \geq \mathscr{J}^*(\rho).$$

(ii) *For any $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, there exists a sequence $\rho^\varepsilon \in \mathcal{P}_2(\mathbb{R}^d)$ such that ρ_ε converges weakly to ρ and*

$$\limsup_{\varepsilon \to 0} \mathscr{J}_\varepsilon(\rho^\varepsilon) \leq \mathscr{J}^*(\rho).$$

We provide a short proof of this result in Appendix B.

While we will not attempt to extend the analysis of [45] to our nonlocal equation (6.3), we can prove the following conditional result:

Theorem 7.2. *Assume that (6.1) holds and that G is given by (1.4). Let ρ_{in}^ε be a sequence of initial data in $\mathcal{P}_2(\mathbb{R}^d)$ satisfying $\mathscr{J}_\varepsilon(\rho_{in}^\varepsilon) \leq C$ and let $\rho^\varepsilon(t, x)$ be the corresponding solution of (1.11) up to time $T > 0$. There exists a subsequence ε_n such that $\rho^{\varepsilon_n}(t)$ converges narrowly (locally uniformly in t) to $\rho \in AC([0, T]; \mathcal{P}_2(\mathbb{R}^d))$. Furthermore if we have*

$$(7.3) \qquad \lim \int_0^T \mathscr{J}_{\varepsilon_n}(\rho^{\varepsilon_n}(t)) \, dt \to \int_0^T \mathscr{J}^*(\rho(t)) \, dt,$$

then ρ solves (7.2) on $[0, T] \times \mathbb{R}^d$ and ρ^{ε_n} converges strongly (in L^2) to ρ in $\{h^{**}(\rho) > 0\}$.

As mentioned in the introduction, the convergence assumption (7.3) offers a classical approach to (conditional) convergence results for free boundary problems (see also the next section). We recall that the Γ-convergence of $\mathscr{J}_\varepsilon$ implies the liminf inequality, so (7.3) says that there is no loss of energy in the limit. Theorem 7.2 appears to be new and we provide a proof in Appendix B. While we do not explicitly require the initial data to be well-prepared in this result, assumption (7.3) is in fact likely stronger: In [45] it is proved (in dimension 1 and for the Cahn-Hilliard equation) that condition (7.3) is satisfied when the convergence of the energy holds for the initial data, that is, $\mathscr{J}_\varepsilon(\rho_{in}^\varepsilon) \to \mathscr{J}^*(\rho_{in})$.

8. Surface tension phenomena: Hele-Shaw flows

When $t \to \infty$, solutions of the Stefan problem (7.2) converge to functions of the form $\theta \chi_{E_\infty}$, which are steady states of the Stefan problem (for any set E_∞). Going back to our aggregation-diffusion equation (1.11) (or (6.3)), this means that solutions experience phase separation. However, these simple characteristic functions are not steady states of (1.11) and will thus continue to evolve at a much slower time scale. In order to characterize this evolution, we now take an initial condition of the form

$$\rho_{in} = \theta \chi_{E_{in}}$$

in (6.3) and rescale the time variable $t \mapsto \varepsilon^{-1} t$, leading to equation (1.18), which we rewrite as

$$(8.1) \qquad\qquad \varepsilon \partial_t \rho = \operatorname{div}(\rho \nabla h'(\rho)) - \operatorname{div}(\rho \nabla B_\varepsilon[\rho]).$$

In this section, we describe how the limit $\varepsilon \to 0$ gives rise to surface tension (mean-curvature) effects. In the next section 9, we will show that when the equation is set on a bounded domain this limit also yields contact angle conditions. This is not surprising, since for the corresponding rescaled Cahn-Hilliard equation,

$$\varepsilon \partial_t \rho = \operatorname{div}(\rho \nabla [h'(\rho) - \varepsilon^2 \Delta \rho]),$$

it was formally shown by Glasner [64] (using formal expansion and matched asymptotics) that the limit $\varepsilon \to 0$ leads to a one-phase Hele-Shaw free boundary problem with surface tension. We can also make sense of this limit by looking at the associated energy functional: Since $h(\rho_{in}) = 0$, (2.6) reduces to

$$\mathscr{J}_\varepsilon[\rho_{in}] = \frac{\theta^2}{4} \iint_{\mathbb{R}^d \times \mathbb{R}^d} G_\varepsilon(x-y)[\chi_{E_{in}}(x) - \chi_{E_{in}}(y)]^2 \, dx \, dy$$

$$= \frac{\theta^2}{4} \iint_{\mathbb{R}^d \times \mathbb{R}^d} G_\varepsilon(x-y)|\chi_{E_{in}}(x) - \chi_{E_{in}}(y)| \, dx \, dy$$

$$= \varepsilon \gamma_0 P(E_{in}) + o(1)$$

for some $\gamma_0 > 0$ (see for instance [15, 106]) where $P(E)$ denotes the perimeter of E, defined by (2.9). We thus introduce the rescaled energy $\mathscr{G}_\varepsilon[\rho] := \varepsilon^{-1} \mathscr{J}_\varepsilon[\rho]$, that is

$$(8.2) \quad \mathscr{G}_\varepsilon[\rho] = \begin{cases} \dfrac{1}{\varepsilon} \displaystyle\int_{\mathbb{R}^d} h(\rho) \, dx + \dfrac{1}{4\varepsilon} \iint_{\mathbb{R}^d \times \mathbb{R}^d} G_\varepsilon(x-y)[\rho(x) - \rho(y)]^2 \, dx \, dy & \text{if } \rho \in \mathcal{P}_2(\mathbb{R}^d), \ d\rho \ll dx, \\ \infty & \text{otherwise}, \end{cases}$$

which is a nonlocal approximation of the rescaled Allen-Cahn functional

$$\frac{1}{\varepsilon} \int_{\mathbb{R}^d} h(\rho) \, dx + \frac{\varepsilon}{2} \int_{\mathbb{R}^d} |\nabla \rho|^2 \, dx,$$

whose Γ-convergence toward the perimeter functional is a classical result of Modica-Mortola [109] (see also Modica [108] and Sternberg [131]). A corresponding result for $\mathscr{G}_\varepsilon$ was proved for rather general kernels G when h is a smooth double-well functional in [1]. We recall that in our framework h is not a smooth potential at 0 (since $h(\rho) = -\infty$ when $\rho < 0$ and $h(\rho) = +\infty$ when $\rho > 1$ in the hard-sphere case $m = \infty$). Nevertheless, the proof can be adapted to this case. When G is given by (1.4), we have the following result (see [106] for the case $m = +\infty$ and [104] for the case $m \in (2, \infty)$):

Theorem 8.1. *Assume that G is given by* (1.4) *and that f satisfies* (6.1). *Then the energy functional $\mathscr{G}_\varepsilon$ defined by* (8.2) *Γ-converges to $\mathscr{G}_0$ defined by*

$$\mathscr{G}_0(\rho) = \begin{cases} \gamma \displaystyle\int_{\mathbb{R}^d} |\nabla \rho| & \text{if } \rho \in \mathcal{P}_2(\mathbb{R}^d), \quad \rho \in \mathrm{BV}(\mathbb{R}^d; \{0, \theta\}), \\ \infty & \text{otherwise}, \end{cases}$$

for some constant γ defined by (8.3) *below. More precisely, we have*

(i) **liminf property**: *For all sequences $\rho^\varepsilon \in L^1(\Omega)$ such that $\rho^\varepsilon \to \rho$ in $L^1(\Omega)$, we have*

$$\liminf_{\varepsilon \to 0} \mathscr{G}_\varepsilon(\rho^\varepsilon) \geq \mathscr{G}_0(\rho).$$

(ii) **limsup property**: *For all $\rho \in L^1(\Omega)$, there exists a sequence $\rho^\varepsilon \in L^1(\Omega)$ such that $\rho^\varepsilon \to \rho$ in $L^1(\Omega)$ and*

$$\limsup_{\varepsilon \to 0} \mathscr{G}_\varepsilon(\rho^\varepsilon) \leq \mathscr{G}_0(\rho).$$

The definition of the constant γ appearing in the definition of $\mathscr{G}_0$ requires the introduction of the function

$$g(s) := \inf_{\rho \geq 0} \left[h(\rho) + \frac{1}{2\sigma}(\rho - \sigma s)^2 \right].$$

We note that $g(s) \geq 0$ for all $s \geq 0$ and $g(0) = g(\theta/\sigma) = 0$ (in fact g is a double-well potential). The constant γ is then given by [7]

$$(8.3) \qquad \gamma := \frac{1}{\theta} \int_0^{\theta/\sigma} \sqrt{2g(s)}\, ds.$$

A key tool in the proof of this theorem in [106, 104] is the following alternative formula for $\mathscr{G}_\varepsilon$ (with $\phi_\varepsilon = G_\varepsilon * \rho$):

$$(8.4) \qquad \mathscr{G}_\varepsilon(\rho) = \frac{1}{\varepsilon}\int_{\mathbb{R}^d} h(\rho) + \frac{1}{2\sigma}(\rho - \sigma\phi^\varepsilon)^2\, dx + \varepsilon \int_{\mathbb{R}^d} \frac{1}{2}|\nabla\phi_\varepsilon|^2\, dx$$

$$(8.5) \qquad \geq \frac{1}{\varepsilon}\int_{\mathbb{R}^d} g(\phi_\varepsilon)dx + \varepsilon \int_{\mathbb{R}^d} \frac{1}{2}|\nabla\phi_\varepsilon|^2\, dx.$$

Inequality (8.5) shows a stronger connection between $\mathscr{G}_\varepsilon$ and the Modica-Mortola functional with double-well potential g (which explains the role of the function g in determining the constant γ in $\mathscr{G}_0$). Formula (8.4) is the key to proving another important property of this energy: The convergence of the energy implies the convergence of the corresponding first variation. Such results have a long history: Reshetnyak [125] proved that if χ_{E_ε} converges to χ_E strongly in L^1, then the convergence of the perimeter $P(E_\varepsilon) \to P(E)$ implies the convergence of its first variation

$$\delta P(E_\varepsilon)[\xi] := \int (\operatorname{div}\xi - \nu_\varepsilon \otimes \nu_\varepsilon : D\xi)\, |D\chi_{E_\varepsilon}| \longrightarrow \int (\operatorname{div}\xi - \nu \otimes \nu : D\xi)|D\chi_E| \qquad \forall \xi \in C_c^1(\mathbb{R}^d; \mathbb{R}^d),$$

where ν is the density $\frac{D\chi_E}{|D\chi_E|}$ (which exists by Radon-Nikodym's differentiation theorem) and can be interpreted as the normal vector to ∂E. We recall that the first variation of P is the mean-curvature. Indeed, for a smooth interface ∂E we have

$$(8.6) \qquad \int (\operatorname{div}\xi - \nu \otimes \nu : D\xi)|D\chi_E| = \int_{\partial E} \operatorname{div}\xi - \nu \otimes \nu : D\xi\, d\mathcal{H}^{d-1} = \int_{\partial E} \kappa\, \xi \cdot \nu\, d\mathcal{H}^{d-1}$$

where κ denotes the mean curvature of ∂E. A similar result was proved by Luckhaus and Modica [97] for the Ginzburg-Landau functional

$$\int_\Omega \frac{1}{\varepsilon}(1 - \rho^2)^2 + \varepsilon|\nabla\rho|^2\, dx.$$

But as far as we know, corresponding results for the nonlocal functional $\mathscr{G}_\varepsilon$ are only known for particular choices of the kernel G: Laux and Otto [94] proved it when G is the Gaussian and used it in the study of the thresholding scheme for multi-phase mean-curvature flow (see also [55, 54, 53]). Jacobs, Kim and Mészáros [78] proved a similar result and used it to derive the Muskat problem. It should be noted however that in these references, the convergence is only proved when $\mathscr{G}_\varepsilon$ is restricted to characteristic

---

[7]When $h = h_\infty$, we have $\theta = 1$ and $g(s) = \frac{\sigma}{2}\min\{s^2, (s - \sigma^{-1})^2\}$ which leads to $\gamma = \frac{1}{4\sigma^{3/2}}$.

functions. In [86], we proved the corresponding result when G is given by (1.4) in the hard-sphere case $m = +\infty$, and then in [104] for the same kernel in the soft-sphere case $m \in (2, \infty)$ without restriction to characteristic functions. We state here the corresponding result in our framework (see [104] - see also [86] for the case $m = \infty$):

Proposition 8.2. *Assume that G is given by (1.4) and that f satisfies (6.1). Given any sequence $\rho^\varepsilon \in L^1$ which strongly converges to ρ in $L^1(\mathbb{R}^d)$, if*

$$\lim_{\varepsilon \to 0} \mathscr{G}_\varepsilon(\rho^\varepsilon) = \mathscr{G}_0(\rho) < \infty,$$

then for all $\xi \in C_c^1(\mathbb{R}^d, \mathbb{R}^d)$ we have:

$$\lim_{\varepsilon \to 0} -\varepsilon^{-1} \int_{\mathbb{R}^d} (h'(\rho^\varepsilon) - B_\varepsilon[\rho^\varepsilon]) \nabla \rho^\varepsilon \cdot \xi \, dx = \gamma \int_{\mathbb{R}^d} [\operatorname{div} \xi - \nu \otimes \nu : D\xi] \, |\nabla \rho|,$$

where $\nu = \frac{\nabla \rho}{|\nabla \rho|}$.

This proposition is the cornerstone of the proof of the following theorem which makes clear the connection between the rescaled aggregation-diffusion equation (8.1) and the Hele-Shaw free boundary problem with surface tension (1.19), where the latter is the gradient flow of the perimeter functional $\mathscr{G}_0$ with respect to the 2-Wasserstein distance.

Theorem 8.3. *Given a sequence of initial data $\rho_{in}^{\varepsilon_n}$ such that $\mathscr{G}_{\varepsilon_n}(\rho_{in}^{\varepsilon_n}) \leq M$ and $\rho_{in}^{\varepsilon_n} \to \rho_{in} = \theta \chi_{E_{in}} \in BV(\mathbb{R}^d; \{0, \theta\})$, let $\rho^{\varepsilon_n}(t, x)$ be the weak solution of (1.18) with initial data $\rho_{in}^{\varepsilon_n}$. Then the following hold:*
(i) Along a subsequence, the density $\rho^{\varepsilon_n}(t, x)$ converges strongly in $L^\infty(0, T; L^1(\mathbb{R}^d))$ to

$$\rho(t, x) = \theta \chi_{E(t)} \in L^\infty(0, \infty; BV(\mathbb{R}^d; \{0, \theta\}))$$

and there exists a velocity function $v(t, x)$ such that

$$(8.7) \qquad \begin{cases} \partial_t \rho + \operatorname{div}(\rho v) = 0, \\ \rho(x, 0) = \theta \chi_{E_{in}}. \end{cases}$$

(ii) If the following energy convergence assumption holds:

$$\lim_{n \to \infty} \int_0^T \mathscr{G}_{\varepsilon_n}(\rho^{\varepsilon_n}(t)) \, dt = \int_0^T \mathscr{G}_0(\rho(t)) \, dt,$$

then there exists $p \in L^2(0, T; (C^s(\mathbb{R}^d))^*)$ (for any $s > 0$) such that

$$(8.8) \qquad \int_0^\infty \int_{\mathbb{R}^d} (\rho v \cdot \xi - \theta p \operatorname{div} \xi) \, dx \, dt = -\gamma \int_0^\infty \int_{\mathbb{R}^d} [\operatorname{div} \xi - \nu \otimes \nu : D\xi] \, |\nabla \rho| \, dt$$

for any vector field $\xi \in C_c^\infty((0, \infty) \times \mathbb{R}^d; \mathbb{R}^d)$.[8] Together, equations (8.7) and (8.8) say that $E(t)$ is a weak solution of the Hele-Shaw free boundary problem with surface tension (1.19) with initial condition E_{in}.

The notion of weak solution of (1.19) that we recover with this theorem is similar to the definition considered for example in [78, 86, 92]: The continuity equation (8.7) encodes the incompressibility condition $\operatorname{div} v = 0$ in $E(t)$ and the free boundary condition $V = -v \cdot \nu$. By taking test functions ξ supported in either $E(t)$ or $\mathbb{R}^d \setminus E(t)$, we see that (8.8) implies $\nabla p = 0$ outside of $E(t)$ and $v = -\nabla p$ in $E(t)$. For general test functions ξ, and taking into account the right hand side of (8.8) we further get the surface tension condition $[p] = \gamma \kappa$ on $\partial E(t)$ as a consequence of (8.6). In the radial setting, (8.6) implies that p equals the constant curvature $\partial E(t)$ in a strong sense, thus yielding discontinuity of the pressure across the interface.

---

[8] The integral $\int_{\mathbb{R}^d} p \operatorname{div} \xi \, dx$ above should be understood as the duality bracket $\langle q, \operatorname{div} \xi \rangle_{(C^s(\Omega))^*, C^s(\Omega)}$.

Proof. This result is proved in [86] when $m = \infty$ and [104] when $m \in (2, \infty)$ when the equation is set in a bounded domain. The proof can be extended to $\mathbb{R}^d$ with the use of the estimate on the second moment $\int_{\mathbb{R}^d} |x|^2 \rho \, dx$. Indeed, the only required adjustment is in the proof of [104, Proposition 4.2], where the argument only yields the strong convergence of ρ^{ε_n} to ρ in $L^\infty(0, T; L^1_{loc}(\mathbb{R}^d))$. But we can write

$$\limsup \int_{\mathbb{R}^d} |\rho^{\varepsilon_n}(t) - \rho(t)| \, dx \leq \limsup \int_{B_R} |\rho^{\varepsilon_n}(t) - \rho(t)| \, dx + \limsup \int_{\mathbb{R}^d \setminus B_R} |\rho^{\varepsilon_n}(t) - \rho(t)| \, dx$$

$$\leq 0 + \frac{1}{R^2} \int_{\mathbb{R}^d} |x|^2 (\rho^{\varepsilon_n}(t) + \rho(t)) \, dx$$

and the strong convergence in $L^\infty(0, T; L^1(\mathbb{R}^d))$ follows by taking $R \to \infty$. $\square$

More general nonlinearity f and interaction kernel G: Theorem 8.3 holds for general convex nonlinearities $f(\rho)$, provided the function $h(\rho)$ is a double-well potential. For technical reasons, the proof also requires that $f(\rho) \geq \left(\frac{1}{2\sigma} + \eta\right) \rho^2$ for large ρ. In particular, adding a small linear diffusion term to the equation, which adds the term $\nu\rho \log \rho$ to the function $h(\rho)$, will displace the well at $\rho = 0$ to some (small) positive value. In that case we expect the limit $\varepsilon \to 0$ to lead to a two-phase free boundary problem (Mullins-Sekerka).

While it is not difficult to extend Theorem 8.3 to general nonlinearities, the proofs provided in [86, 104] are strongly dependent on the particular form of the interaction kernel G. In fact the proofs make use of the elliptic equation (1.4). One could extend the proof to slightly more general equations, for example by replacing the Laplace operator by a general elliptic operator (which would introduce both space and direction dependence in the free boundary condition in (1.19)), but treating the case of general interaction kernels seems to require new ideas. As noted earlier, the Γ-convergence of the energy (Theorem 8.1) is known to hold for a large class of such kernel, and the main obstacle is rather to prove Proposition 8.2 in a more general framework.

9. Bounded domains and contact angle conditions

We have so far only considered problems set in the whole space. But in many experimental settings, the particles are confined to a bounded domain $\Omega \subset \mathbb{R}^d$. The equation for ρ is then set on Ω and supplemented with no-flux boundary conditions on $\partial\Omega$. But the domain can also have a non-trivial influence on the interaction kernel.

In some settings, it might make sense to stick to interactions that depend only on the distance between the particles (mathematically, this is done by extending ρ by zero outside the domain and keeping the convolution $G_\varepsilon * \rho$ in the equation). But in the context of chemotaxis (1.5), the potential $\phi = \phi_\varepsilon = G_\varepsilon * \rho$ represents the concentration of a chemoattractant chemical which is now diffusing in the same bounded domain Ω so that this convolution should be replaced with the solution of some elliptic boundary value problem.

In the first part of this section, we present results that have been proved in this later setting. We will then discuss the case of interactions that depend only on the distance between the particles.

9.1. **Chemotaxis potential in bounded domains.** In the context of chemotaxis (1.5), we replace the convolution $G_\varepsilon * \rho$ by the solution ϕ_ε of the following elliptic boundary value problem:

$$(9.1) \qquad \begin{cases} \sigma\phi_\varepsilon - \varepsilon^2 \Delta\phi_\varepsilon = \rho & \text{in } \Omega, \\ a\phi_\varepsilon + \varepsilon b \nabla\phi_\varepsilon \cdot n = 0 & \text{on } \partial\Omega. \end{cases}$$

We are taking general Robin type boundary conditions at the boundary $\partial\Omega$ for further generality, even though Neumann conditions ($a = 0$) are probably the most natural type of conditions. When $a > 0$ (9.1)

takes into account the possible leakage (or destruction) of the chemical at the boundary. We can take a and b to be constant or functions of $x \in \partial\Omega$ and we will always assume that $a \geq 0$, $b \geq 0$ and $a + b > 0$. The solution of (9.1) can be written as $\phi_\varepsilon(x) = \int_\Omega G_\varepsilon(x, y)\rho(y)\, dy$ with $G_\varepsilon(x, y) \geq 0$ a symmetric kernel and we set

$$B_\varepsilon[\rho] := \phi_\varepsilon - \sigma^{-1}\rho,$$

so that (8.1) corresponds to

(9.2)
$$\begin{cases} \varepsilon\partial_t\rho = \operatorname{div}\left(\rho\nabla h'(\rho)\right) - \operatorname{div}\left(\rho\nabla B_\varepsilon[\rho]\right) & \text{in } (0, T) \times \Omega; \\ \rho\nabla(h'(\rho) - B_\varepsilon[\rho]) \cdot n = 0 & \text{on } (0, T) \times \partial\Omega; \\ \rho(x, 0) = \rho_{in} & \text{in } \Omega. \end{cases}$$

However, we do not have $\int_\Omega G_\varepsilon(x, y)\, dy = \sigma^{-1}$ so we cannot write $B_\varepsilon[\rho]$ as $\int_\Omega G_\varepsilon(x, y)[\rho(y) - \rho(x)]\, dy$. We can still write $\int_\Omega G_\varepsilon(x, y)\, dy = \sigma^{-1} - \tau_\varepsilon(x)$ where $\tau_\varepsilon \geq 0$ solves

$$\begin{cases} \sigma\tau_\varepsilon - \varepsilon^2\Delta\tau_\varepsilon = 0 & \text{in } \Omega, \\ a\tau_\varepsilon + \varepsilon b\nabla\tau_\varepsilon \cdot n = \frac{a}{\sigma} & \text{on } \partial\Omega. \end{cases}$$

We thus have

$$B_\varepsilon[\rho](x) = \int_\Omega G_\varepsilon(x, y)[\rho(y) - \rho(x)]\, dy - \tau_\varepsilon(x)\rho(x)$$

and the corresponding energy functional can be written as

$$\begin{aligned} \mathscr{G}_\varepsilon[\rho] &= \frac{1}{\varepsilon}\int_\Omega h(\rho)\, dx - \frac{1}{2\varepsilon}\int_\Omega \rho B_\varepsilon[\rho]\, dx \\ &= \frac{1}{\varepsilon}\int_\Omega h(\rho)\, dx + \frac{1}{4\varepsilon}\iint_{\Omega\times\Omega} G_\varepsilon(x, y)[\rho(x) - \rho(y)]^2\, dx\, dy + \frac{1}{2\varepsilon}\int_\Omega \tau_\varepsilon(x)|\rho(x)|^2\, dx. \end{aligned}$$

The last term in this energy accounts for the boundary conditions on ϕ_ε. In the case of Neumann boundary condition $a = 0$, we have $\tau_\varepsilon = 0$, and this new term vanishes. But in general this new term will contribute to the limiting energy. Indeed, when $\partial\Omega$ is regular enough, we can show[9] that $\frac{1}{\varepsilon}\tau_\varepsilon$ converges as a distribution to $\frac{1}{\sigma^{3/2}}\frac{a}{a+b\sqrt{\sigma}}d\mathcal{H}^{n-1}|_{\partial\Omega}$. This suggests that if E is a subset of Ω with C^1 boundary, then

$$\mathscr{G}_\varepsilon[\theta\chi_E] \to \gamma\theta P(E, \Omega) + \frac{1}{2\sigma^{3/2}}\frac{a\theta}{a+b\sqrt{\sigma}}\mathcal{H}^{n-1}(\partial\Omega \cap E),$$

where the local perimeter $P(E, \Omega)$ is defined by

$$P(E, \Omega) := \sup\left\{\int_E \operatorname{div} g\, dx\, ;\, g \in (C_0^1(\Omega))^d, \quad |g(x)| \leq 1\, \forall x \in \Omega\right\} = \int_\Omega |\nabla\chi_E|.$$

This limit was proved rigorously in [106] (see also [86, Appendix C]) for the hard-sphere case $f = f_\infty$. In that case $\theta = 1$ and $\gamma = \frac{1}{4\sigma^{3/2}}$, so the formula for the limit becomes

(9.3)
$$\frac{1}{4\sigma^{3/2}}\left(P(E, \Omega) + \frac{2a}{a+b\sqrt{\sigma}}\mathcal{H}^{n-1}(\partial\Omega \cap E)\right).$$

Here we make a couple of remarks: First, functionals like (9.3), in which the area of the free surface $\partial E \cap \Omega$ and the contact area $\partial\Omega \cap E$ are weighted differently, appear naturally in various applications, such as the capillary energy for the modeling of sessile liquid drop in contact with a solid support. It is well known that the minimization of such energies leads to contact angle conditions. Second, we note

---

[9]To see this, notice that when Ω is the half space $\{x_d > 0\}$, we have $\tau_\varepsilon(x) = \frac{1}{\sigma}\frac{a}{a+b\sqrt{\sigma}}e^{-\frac{\sqrt{\sigma}}{\varepsilon}x_d}$ and so $\varepsilon^{-1}\tau_\varepsilon \to \frac{1}{\sigma^{3/2}}\frac{a}{a+b\sqrt{\sigma}}\delta(x_d)$.

that when $\frac{2a}{a+b\sqrt{\sigma}} > 1$, the energy (9.3) is not lower semi-continuous and can therefore not be the Γ-limit of $\mathscr{G}_\varepsilon$. In fact, the following result was proved in [106, 86] when $f = f_\infty$:

Theorem 9.1. *Assume that Ω is a bounded open set with $C^{1,\alpha}$ boundary and take $f = f_\infty$ and G given by* (1.4). *Assume further that a and b are non-negative constants such that $a + b > 0$. The functional $\mathscr{G}_\varepsilon$ Γ-converges, when $\varepsilon \to 0$ to*

$$\mathscr{G}_0(\rho) := \begin{cases} \dfrac{1}{4\sigma^{3/2}} \left[\displaystyle\int_\Omega |\nabla \rho| + \int_{\partial\Omega} \min\left(1, \dfrac{2a}{a+\sqrt{\sigma}b}\right) \rho \, d\mathcal{H}^{n-1}(x) \right] & \text{if } \rho \in \mathrm{BV}(\Omega; \{0,1\}), \\ \infty & \text{otherwise.} \end{cases}$$

The result also holds when $a(x)$, $b(x)$ are bounded, Lipschitz non-negative functions defined on $\partial\Omega$ such that $\alpha(x) + \beta(x) \geq \eta > 0$ and that $\mathcal{H}^{n-2}\left(\partial\left\{\frac{2a}{a+\sqrt{\sigma}b} > 1\right\}\right) < \infty$ (this technical condition ensures that the function $x \mapsto \frac{2a(x)}{a(x)+\sqrt{\sigma}b(x)}$ does not oscillate too much around the value 1). A crucial tool for the proof of this result is the following alternative formula for $\mathscr{G}_\varepsilon$:

$$\mathscr{G}_\varepsilon(\rho) = \frac{1}{2\varepsilon} \int_\Omega h_\infty(\rho) \, dx + \frac{1}{2\sigma\varepsilon} \int_\Omega (\rho - \sigma\phi_\varepsilon)^2 \, dx + \frac{\varepsilon}{2} \int_\Omega |\nabla\phi_\varepsilon|^2 \, dx$$
$$+ \frac{\varepsilon^2}{2} \int_{\partial\Omega} \frac{b}{a+b} |\nabla\phi_\varepsilon \cdot n|^2 \, d\mathcal{H}^{n-1}(x) + \frac{1}{2} \int_{\partial\Omega} \frac{a}{a+b} |\phi_\varepsilon|^2 \, d\mathcal{H}^{n-1}(x),$$

where $\phi_\varepsilon(x) = \int_\Omega G_\varepsilon(x, y)\rho(y) \, dy$ solves (9.1).

Theorem 9.1 was also proved in [104] for the soft-sphere case $f = f_m$ with $m > 2$ when $a = 0$ (Neumann boundary condition). The approach of [106] and [104] can be combined to get the result when $f = f_m$ and for general coefficients a and b (although the formula for the coefficient in the boundary integral is not straightforward).

When $a = 0$ then $\mathscr{G}_0(\chi_E)$ is proportional to the local perimeter $P(E, \Omega)$ and thus only measures the area of $\partial E \cap \Omega$, while when $b = 0$ (Dirichlet boundary conditions), then $\mathscr{G}_0$ is proportional to the full perimeter $P(E)$. When $\frac{2a}{a+\sqrt{\sigma}b} \in (0, 1)$, the areas of $\partial E \cap \Omega$ and $E \cap \partial\Omega$ are weighted differently.

As mentioned above, the boundary term in $\mathscr{G}_0$ translates into a contact angle condition in the limiting Hele-Shaw free boundary problem: It was proved in [85] that the limit $\varepsilon \to 0$ of (9.1)-(9.2) leads to the same Hele-Shaw free boundary problem (1.19) in Ω, supplemented with Neumann boundary condition

$$\nabla p \cdot n = 0 \qquad \text{on } \partial\Omega$$

and the condition (see Figure 4)

$$(9.4) \qquad\qquad \cos(\alpha) := -\min\left(1, \frac{2a}{a+\sqrt{\sigma}b}\right) \quad \text{on } \partial E(t) \cap \partial\Omega,$$

where α denotes the angle between the free surface ∂E and the fixed boundary $\partial\Omega$ (see Figure 4).
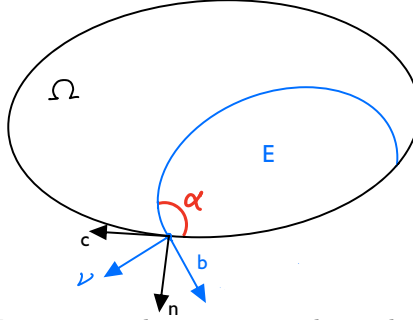
FIGURE 4. The contact angle condition

In particular, for Neuman condition $a = 0$ (no absorption of the chemical on $\partial\Omega$) we get $\alpha = \frac{\pi}{2}$ so the contact must be orthogonal, while for Dirichlet condition $b = 0$ (and whenever the absorption rate satisfies $a/b \geq \sqrt{\sigma}$), the contact must be tangential ($\alpha = \pi$). Let us note that a similar contact angle condition was derived recently in [70] for the geometric mean-curvature flow.

The precise result in this case, which is proved by establishing the equivalent of Proposition 8.2 with the boundary term (see [86, Proposition 5.2]), is as follows:

Theorem 9.2. [86, Theorem 1.6] *Under the assumptions of Theorem 9.1, and given a sequence ε_n and initial data $\rho_{in}^{\varepsilon_n}$ such that $\mathscr{G}_{\varepsilon_n}(\rho_{in}^{\varepsilon_n}) \leq M$ and $\rho_{in}^{\varepsilon_n} \to \rho_{in} = \chi_{E_{in}} \in BV(\Omega)$ as $\varepsilon_n \to 0$, let $\rho^{\varepsilon_n}(t,x)$ be the weak solution of (9.2).*

Then the conclusion of Theorem 8.3 holds, with (8.8) replaced by

(9.5)
$$\int_0^\infty \int_\Omega \rho v \cdot \xi - \theta p \operatorname{div} \xi(t)\, dx\, dt = -\gamma \int_0^\infty \int_\Omega [\operatorname{div} \xi - \nu \otimes \nu : D\xi]\, |\nabla\rho|\, dt$$
$$- \gamma \min\left(1, \frac{2a}{a + \sqrt{\sigma}b}\right) \int_0^\infty \int_{\partial\Omega} [\operatorname{div} \xi - n \otimes n : D\xi]\, \rho\, d\mathcal{H}^{n-1}(x)\, dt,$$

for any vector field $\xi \in C_c^\infty((0,\infty) \times \overline{\Omega}; \mathbb{R}^d)$ such that $\xi \cdot n = 0$ on $\partial\Omega$.

The fact that (9.5) implies the contact angle condition (9.4) can be seen by using the classical formula (for a smooth interface Σ):

$$\int_\Sigma \operatorname{div} \xi - \nu \otimes \nu : D\xi = \int_\Sigma \kappa\, \xi \cdot \nu + \int_\Gamma b \cdot \xi,$$

where ν is the normal vector to Σ, κ denotes the mean curvature of Σ and b is the conormal vector along $\Gamma = \partial\Sigma$. Indeed, formally at least, the right hand side of (9.5) is (using the fact that $\xi \cdot n = 0$ on $\partial\Omega$):

$$-\gamma\left[\int_\Sigma \kappa\, \xi \cdot \nu + \int_\Gamma \vec{b} \cdot \xi\right] + \gamma \min\left(1, \frac{2a}{a + \sqrt{\sigma}b}\right)\left[\int_{\partial\Omega \cap \Omega_s} \kappa\, \xi \cdot n + \int_\Gamma \vec{c} \cdot \xi\right]$$
$$= \gamma\left[-\int_\Sigma \kappa\, \xi \cdot \nu + \int_\Gamma \min\left(1, \frac{2a}{a + \sqrt{\sigma}b}\right) \vec{c} \cdot \xi - \vec{b} \cdot \xi\right],$$

where $\vec{b}$ and $\vec{c}$ are unit conormal vectors along $\Gamma = \partial\Sigma \cap \partial\Omega$: $\vec{b}$ is tangent to Σ while $\vec{c}$ is tangent to $\partial\Omega$ (see Figure 4). When $\xi = \vec{c}$ on $\partial\Omega$, we deduce $\cos\alpha = \vec{b} \cdot \vec{c} = \min\left(1, \frac{2a}{a + \sqrt{\sigma}b}\right)$.

9.2. **Interactions dependent only on the distance in a bounded domain.** While the framework discussed above makes sense in the context of chemical interactions, there are other settings in which interactions depend on the distance between particles, even in the presence of fixed boundaries (or obstacles). This is often the case of social interactions which are facilitated by sight [11]. Another example is the modeling of cell-cell adhesion [28, 57] by an attractive force between cells that are within a given distance from each others and the formation of membraneless organelles in eukaryotic cells [72, 107] (in that case, the bounded domain is the cell). In that case, we have

$$\phi_\varepsilon(x) = \int_\Omega G_\varepsilon(x-y)\rho(y)\,dy = \int_{\mathbb{R}^d} G_\varepsilon(x-y)\bar\rho(y)\,dy,$$

where $\bar\rho$ is the extension of ρ to $\mathbb{R}^d$ by 0. For simplicity, we focus on the hard-sphere case $f = f_\infty$, although the soft-sphere case could be treated similarly. The energy functional then reads

$$\varepsilon^{-1}\int_\Omega h_\infty(\rho(x)) - \frac{1}{2}\rho(x)\phi_\varepsilon(x)\,dx = \varepsilon^{-1}\int_{\mathbb{R}^d} h_\infty(\bar\rho(x)) - \frac{1}{2}\bar\rho(x)G_\varepsilon * \bar\rho(x)\,dx,$$

which is the same energy as studied in Section 8 applied to $\bar\rho$ instead of ρ. Its Γ-limit is proportional to the (full) perimeter functional $P(E)$. The limit $\varepsilon \to 0$ of the aggregation-diffusion equation (1.20) then leads to the Hele-Shaw free boundary problem (1.19) supplemented with the contact angle condition

$$\cos\alpha = -1$$

that is $\alpha = \pi$: The boundary of the domain $\partial\Omega$ acts like an hydrophobic surface and any contact of the free surface with the fixed boundary will be tangential.

Other contact angle conditions may be derived by taking into account the interactions of the particles with the boundary of the domain (in the modeling of cell-cell adhesion, the fixed boundary may represent a different species with a different inter-species adhesion strength). To illustrate this effect, we introduce the potential $\tau_\varepsilon(x) = G_\varepsilon * \chi_{\mathbb{R}^d \setminus \Omega}$ and consider the equation

$$(9.6) \quad \begin{cases} \varepsilon\partial_t\rho + \operatorname{div}(\rho\nabla(G_\varepsilon * \rho)) + \eta\operatorname{div}(\rho\nabla\tau_\varepsilon) = \operatorname{div}(\rho\nabla p), & p \in \partial f_\infty(\rho) & \text{in } [0,\infty) \times \Omega, \\ \rho\nabla[G_\varepsilon * \rho + \eta\tau_\varepsilon - p] \cdot n = 0 & & \text{on } [0,\infty) \times \partial\Omega, \\ \rho(x,0) = \rho_{in}^\varepsilon(x) & & \text{in } \Omega. \end{cases}$$

The new drift term appearing in this equation models the attraction of the particles toward $\mathbb{R}^d \setminus \Omega$ (or repulsion if $\eta < 0$). It describes the interactions of our particles with (different) particles located outside of Ω. These interactions are assumed to be of similar nature but possibly different strength (another classical example is that of a fluid resting on a solid surface: The fluid molecules interact with each others as well as with the solid).

The respective strength of the interactions – the coefficient η – determines the contact angle condition. This can be seen by considering the energy associated to (9.6), which reads:

$$\mathscr{G}_{\eta,\varepsilon}(\rho) = \varepsilon^{-1}\int_\Omega f_\infty(\rho) + \frac{1}{2\sigma}\rho - \frac{1}{2}\rho\phi_\varepsilon - \eta\rho\tau_\varepsilon\,dx.$$

Using the fact that $\tau_\varepsilon(x) = \frac{1}{\sigma} - \int_\Omega G_\varepsilon(x-y)\,dy$, we can write

$$\mathscr{G}_{\eta,\varepsilon}(\rho) = \varepsilon^{-1}\int_\Omega f_\infty(\rho) + \frac{1}{2}\phi_\varepsilon(1-\rho) - \eta\rho\tau_\varepsilon + \frac{1}{2}\left(\frac{1}{\sigma}\rho - \phi_\varepsilon\right)dx$$

$$= \varepsilon^{-1}\int_\Omega f_\infty(\rho) + \frac{1}{2}\phi_\varepsilon(1-\rho)\,dx + \left(\frac{1}{2} - \eta\right)\varepsilon^{-1}\int_\Omega \rho\tau_\varepsilon\,dx.$$

This energy has two terms: The first term was studied for example in [78] where it was shown (for a different kernel, but the proof extends to (1.4) using the arguments of [86]) that it Γ-converges to the

local perimeter

$$\frac{1}{4\sigma^{3/2}} P(E, \Omega).$$

The second term yield a contribution from the boundary. When G is the Newtonian attractive kernel (1.4), then τ_ε solves

$$(9.7) \qquad\qquad \tau_\varepsilon - \varepsilon^2 \Delta \tau_\varepsilon = \chi_{\mathbb{R}^d \setminus \Omega}.$$

When Ω is the half space $\{x \in \mathbb{R}^d \,;\, x_d > 0\}$, we can solve (9.7) explicitly and find $\tau_\varepsilon(x) = \frac{1}{2\sigma} e^{-\frac{\sqrt{\sigma} x_d}{\varepsilon}}$ for $x_d \geq 0$ so that $\frac{1}{\varepsilon} \tau_\varepsilon(x) \chi_{x_d > 0}$ converges to $\frac{1}{2\sigma^{3/2}} \delta(x_d)$ as $\varepsilon \to 0$ and

$$\varepsilon^{-1} \int_\Omega \rho \tau_\varepsilon \, dx \to \frac{1}{2\sigma^{3/2}} \int_{\partial\Omega} \rho \, d\mathcal{H}^{n-1}(x).$$

This convergence can be justified for general domains Ω with smooth boundary. In fact, we can show:

Proposition 9.3. *Assume that Ω is a bounded open set with $C^{1,\alpha}$ boundary and take $f = f_\infty$ and G given by (1.4). For all $\eta \in [0, 1/2]$, the energy functional $\mathscr{G}_{\eta,\varepsilon}$ Γ-converges to*

$$(9.8) \qquad\qquad \frac{1}{4\sigma^{3/2}} P(E, \Omega) + \frac{1}{2\sigma^{3/2}} \left(\frac{1}{2} - \eta \right) H^{n-1}(\partial\Omega \cap E).$$

Note that the result holds for $\eta \in [0, 1)$ but the proof is more delicate than the proof given below when $\eta > 1/2$. This Proposition suggests that a result similar to Theorem 9.2 holds for (9.6): Under some assumptions on the convergence of the energy, the limit $\varepsilon \to 0$ in (9.6) leads to the Hele-Shaw free boundary problem (with surface tension) (1.19), supplemented with the contact angle condition

$$\cos \alpha = 2\eta - 1$$

at the triple junction. We can thus have contact angles between π and 0 when η takes values between 0 and 1. When $\eta \geq 1$, the attraction towards the boundary dominates the dynamics and the particles will spread over the whole boundary to maximize the contact area. A rigorous justification of this result would require proving the convergence of the first variation of the energy (the equivalent of Proposition 8.2).

Idea of the proof of Proposition 9.3. We note that

$$\int_\Omega \rho \tau_\varepsilon \, dx = \int_\Omega \rho * G_\varepsilon \chi_{\mathbb{R}^d \setminus \Omega} \, dx = \int_\Omega \phi_\varepsilon \chi_{\mathbb{R}^d \setminus \Omega} \, dx = \int_\Omega \frac{1}{\sigma} \rho - \phi_\varepsilon \, dx$$

and we write

$$\mathscr{G}_{\eta,\varepsilon}(\rho) = \varepsilon^{-1}(1 - 2\eta) \int_\Omega f_\infty(\rho) + \frac{1}{2\sigma}\rho - \frac{1}{2}\rho\phi_\varepsilon \, dx + \varepsilon^{-1} 2\eta \int_\Omega f_\infty(\rho) + \frac{1}{2\sigma}\rho - \frac{1}{2}\rho\phi_\varepsilon - \frac{1}{2}\rho\tau_\varepsilon \, dx$$

$$= \varepsilon^{-1}(1 - 2\eta) \int_{\mathbb{R}^d} h_\infty(\bar\rho) + \frac{1}{2}\bar\rho(1 - \phi_\varepsilon) \, dx + \varepsilon^{-1} 2\eta \int_\Omega f_\infty(\rho) + \frac{1}{2}\phi_\varepsilon(1 - \rho) \, dx$$

where every term is non-negative. The first term is (up to a multiplicative constant) the energy $\mathscr{G}_\varepsilon$ and thus Γ-converges to $\frac{1}{4\sigma^{3/2}}(1 - 2\eta)P(E)$ (when $\eta \leq 1/2$). The second term can be written as

$$\varepsilon^{-1} 2\eta \int_\Omega f_\infty(\rho) + \frac{1}{2}\tau_\varepsilon(x)\rho(x)(1 - \rho(x)) \, dx + \frac{1}{4} \iint_{\Omega \times \Omega} G(x - y)[\rho(y) - \rho(x)]^2 \, dx \, dy.$$

where the first integral is the double well-potential and the second integral is similar to the energy studied, for example, in [1]. This energy Γ-converges to $2\eta \frac{1}{4\sigma^{3/2}} P(E, \Omega)$. The limiting energy is thus

$$\frac{1}{4\sigma^{3/2}}(1 - 2\eta)P(E) + 2\eta \frac{1}{4\sigma^{3/2}} P(E, \Omega),$$

which is equivalent to (9.8). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

10. Conclusion

In this paper, we presented a road map to investigate phase separation phenomena in biological (and other) systems. We start from an agent-based microscopic description of attractive behavior limited by a volume exclusion principle, accounting for the natural incompressibility of the agents. The key takeaway is that phase separation is natural when repulsion is modeled by nonlinear diffusion with $m > 2$ (soft-sphere models) or with density constraint (hard-sphere models), and when the long-range attractive kernel is integrable. We note that the restriction $m > 2$, which is required for the function $h(\rho)$ defined by (1.13) to be a double-well potential, is also consistent with known results on the existence of stationary states and energy minimizers for (1.5) when $\sigma > 0$ (see [6, 18, 81]) – the existence of such stationary states seems necessary for our analysis which concerns the behavior of solutions of (1.5) over large time scales.

Based on recent works, we proposed a rigorous mathematical approach to derive effective models of geometric nature which describe the evolution of the sharp interface over appropriate time scales after aggregation occurs. In particular, the results presented here show that surface tension and contact angle dynamics, classically associated with fluid interfaces, are natural consequences of volume-limited aggregation. This is in agreement with many real-life observations: for instance, the liquid-like behavior of membraneless organelles in eukaryotic cells has been well documented [16, 72].

This review leads to several important open questions. Among them is the well-posedness of the hard-sphere microscopic model (1.8) for general initial conditions and its convergence as $\delta \to 0$ to the macroscopic model (1.10). Solutions to (4.16) can be formally obtained in the limit $m \to \infty$ in the soft-sphere model (1.6) or via a JKO type discrete scheme, but neither approaches has been made rigorous.

Concerning the derivation of the free boundary problems, we emphasize once more that all stated rigorous results are conditional, which require an assumption on the energy convergence. Obtaining non-conditional convergence results remains out of reach, not only in the context of the problems discussed in this review, but for related free boundary problems as well. Last but not least, these conditional results have been proved only for the particular kernel G, solution of (1.4). On the other hand, the Γ-convergence of the energies $\mathscr{J}_\varepsilon$ and $\mathscr{G}_\varepsilon$ has been proved for very general G. Thus we may conjecture that these conditional results can hold for a large class of non-negative integrable kernels. Further potential generalizations of our analysis also include the derivation of two-phase free boundary problems, either by considering pressure function f for which the two wells of the function h are both positive (leading to a Mullins-Sekerka free boundary problem), or for multi-species models such as those proposed in [57].

Appendix A. From micro to macro: convergence $\delta \to 0$

Proof of Theorem 3.3. The continuity equation and (3.13) imply that ρ^δ is bounded in $C^{1/2}([0,T]; \mathcal{P}_2(\mathbb{R}^d))$. It follows from a refined version of Ascoli-Arzelà theorem, as in [31, Proposition 4.1], that there exists a subsequence such that $\rho^\delta(t)$ narrowly converge to some function $\bar\rho(t)$ (uniformly in t).

Since K_δ is supported in B_δ, we have $K_\delta * \varphi \to \varphi$ uniformly for any uniformly continuous function φ. Using the bound on the second moment (3.14), we can then show that $K_\delta * \rho^\delta(t)$ converges narrowly to $\bar\rho(t)$.

We also note that (3.13) implies that $\rho^\delta * K_\delta$ is bounded in $L^\infty(0,T; L^m(\mathbb{R}^d))$ and thus also in $L^\infty(0,T; L^2(\mathbb{R}^d))$. Furthermore, using the fact that $\phi^\delta = G * \rho^\delta$ solves $\sigma\phi^\delta - \eta\Delta\phi^\delta = \rho^\delta$, we see that ϕ^δ and $K_\delta * \phi^\delta$ are bounded in $L^\infty(0,T; H^1(\mathbb{R}^d))$.

However, this convergence is not enough to pass to the limit in the equation and the main idea in [31] is to rewrite the limiting equation (1.9) (when f is given by (1.2)) as

$$\text{(A.1)} \qquad \partial_t \rho = -\operatorname{div}(\rho \nabla G * \rho) + 2\nabla \cdot (\rho^{\frac{m}{2}} \nabla \rho^{\frac{m}{2}})$$

and to adjust, in a similar way, the weak formulation of the regularized equation (1.6). Indeed, given a test function $\psi \in C_c^\infty([0,T] \times \mathbb{R}^d)$, a straightforward computation with integration by parts yields

$$\int \psi(T,x)\, d\rho^\delta(T,x) - \int \psi(0,x)\, d\rho_{\text{in}}(x) - \int_0^T \int \partial_t \psi\, d\rho^\delta(t,x)\, dt$$

$$= \int_0^T \int \nabla(K_\delta * K_\delta * \phi^\delta) \cdot \nabla\psi\, d\rho^\delta(t,x)\, dt - \frac{m}{m-1}\int_0^T \int \nabla K_\delta * (K_\delta * \rho^\delta)^{m-1} \cdot \nabla\psi\, d\rho^\delta(t,x)\, dt$$

$$= \int_0^T \int \nabla(K_\delta * \phi^\delta) \cdot [K_\delta * (\rho^\delta \nabla\psi)] - m\int_0^T \int (K_\delta * \rho^\delta)^{m-2}\nabla(K_\delta * \rho^\delta) \cdot [K_\delta * (\rho^\delta \nabla\psi)]$$

$$(A.2) \quad = \int_0^T \int (K_\delta * \rho^\delta)\nabla(K_\delta * \phi^\delta) \cdot \nabla\psi - 2\int_0^T \int (K_\delta * \rho^\delta)^{\frac{m}{2}}\nabla(K_\delta * \rho^\delta)^{\frac{m}{2}} \cdot \nabla\psi$$

$$(A.3) \quad + \int_0^T \int \nabla(K_\delta * \phi^\delta) \cdot E^\delta - 2\int_0^T \int (K_\delta * \rho^\delta)^{\frac{m}{2}-1}\nabla(K_\delta * \rho^\delta)^{\frac{m}{2}} \cdot E^\delta.$$

The error term in (A.3) comes from the commutation between convolution and multiplication, i.e. $E^\delta(x) := [K_\delta * (\rho^\delta \nabla\psi)](x) - (K_\delta * \rho^\delta)\nabla\psi(x)$. We will see that this error term vanishes in the limit. Passing to the limit in the second term of (A.2) (which has the same structure as (A.1)) requires the following convergences:

$$(K_\delta * \rho^\delta)^{\frac{m}{2}} \to \rho^{\frac{m}{2}}, \quad \text{(strong) in } L^2_{t,x}, \quad \text{and} \quad \nabla(K_\delta * \rho^\delta)^{\frac{m}{2}} \rightharpoonup \nabla\rho^{\frac{m}{2}}, \quad \text{(weak) in } L^2_{t,x}.$$

This weak-strong convergence pair is proved by considering the evolution of $\int \rho^\delta \log \rho^\delta$. Indeed, since $\mathscr{E}[\rho_{\text{in}}] < +\infty$ and $\rho \log \rho \lesssim \rho^m$ for $m > 1$, we see that $\int \rho_{\text{in}} \log \rho_{\text{in}} < +\infty$. Formally (these steps can be made rigorous by using the piecewise constant interpolation given by the JKO scheme as in [31]), by differentiating $\int \rho^\delta \log \rho^\delta$ in time we get:

$$\int \rho^\delta(t,x) \log \rho^\delta(t,x) - \int \rho_{\text{in}} \log \rho_{\text{in}} = -\frac{2}{m}\int_0^t \int |\nabla(K_\delta * \rho^\delta)^{\frac{m}{2}}|^2 + \int_0^t \int \nabla(K_\delta * \phi^\delta) \cdot \nabla K_\delta * \rho^\delta$$

where

$$\int_0^t \int \nabla(K_\delta * \phi^\delta) \cdot \nabla K_\delta * \rho^\delta = -\int_0^t \int \Delta(K_\delta * \phi^\delta) K_\delta * \rho^\delta = -\frac{\sigma}{\eta}\int_0^t \int (K_\delta * \phi^\delta) K_\delta * \rho^\delta + \frac{1}{\eta}\int_0^t \int |K_\delta * \rho^\delta|^2$$

and so

$$\int \rho^\delta(t,x) \log \rho^\delta(t,x) + \frac{2}{m}\int_0^t \int |\nabla(K_\delta * \rho^\delta)^{\frac{m}{2}}|^2 + \frac{\sigma}{\eta}\int_0^t \int (K_\delta * \phi^\delta) K_\delta * \rho^\delta = \int \rho_{\text{in}} \log \rho_{\text{in}} + \frac{1}{\eta}\int_0^t \int |K_\delta * \rho^\delta|^2.$$

We recall (see [31, Remark 4.2]) that the entropy is bounded below by the second moment as follows:

$$\int \rho^\delta(t,x) \log \rho^\delta(t,x) \geq -d\log(2\pi) - \frac{1}{2}\int |x|^2 \rho^\delta$$

and since $m > 2$, we have

$$\int_0^t \int |K_\delta * \rho^\delta|^2 \leq C \int_0^t \int K_\delta * \rho^\delta + |K_\delta * \rho^\delta|^m\, dx$$

$$\leq C(1 + \mathscr{E}_\delta[\rho_{\text{in}}])\, t.$$

This implies that the sequence $(K_\delta * \rho^\delta)^{\frac{m}{2}}$ is bounded in $L^2(0,T;H^1(\mathbb{R}^d))$ uniformly in $\delta > 0$. Proceeding as in [31, Proposition 4.3] (using a refined version of the Aubin-Lions Lemma [127, Theorem 2]) we can then show that up to another subsequence we have

$$K_\delta * \rho^\delta \to \rho \quad \text{in } L^m([0,T] \times \mathbb{R}^d), \quad \nabla(K_\delta * \rho^\delta)^{\frac{m}{2}} \rightharpoonup \nabla\rho^{\frac{m}{2}} \text{ in } L^2([0,T] \times \mathbb{R}^d).$$

We already mentioned that $K_\delta * \phi^\delta$ was bounded in $L^\infty(0, T; H^1(\mathbb{R}^d))$, which implies

$$\nabla(K_\delta * \phi^\delta) \rightharpoonup \nabla\phi \text{ in } L^2([0, T] \times \mathbb{R}^d).$$

Finally, (see [31, Lemma 5.1]) we have that $E^\delta(t, x)$ converges to zero in $L^m([0, T] \times \mathbb{R}^d)$. These convergences are enough to show that (A.3) vanishes as $\delta \to 0$ and to pass to the limit in (A.2). $\qquad\square$

APPENDIX B. CONVERGENCE RESULTS FOR THE STEFAN PROBLEM

Proof of Proposition 7.1. We have $\mathscr{J}_\varepsilon[\rho^\varepsilon] \geq \mathscr{J}^*[\rho^\varepsilon]$ and the functional $\mathscr{J}^*$ is convex and thus lower semicontinuous with respect to the weak topology. The liminf property follows.

To prove the limsup property, we will simply use the fact that $\mathscr{J}_\varepsilon$ can be bounded by some Allen-Cahn functional for which the Γ convergence has been proved elsewhere. Indeed, we can write

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} G_\varepsilon(x - y)[\rho(x) - \rho(y)]^2 \, dx \, dy = \iint_{\mathbb{R}^d \times \mathbb{R}^d} G_\varepsilon(z)[\rho(y + z) - \rho(y)]^2 \, dz \, dy$$

where

$$|\rho(y + z) - \rho(y)|^2 = \left| \int_0^1 \nabla\rho(y + tz) \cdot z \, dt \right|^2$$

$$\leq \int_0^1 |\nabla\rho(y + tz)|^2 \, dt \, |z|^2.$$

We deduce

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} G_\varepsilon(x - y)[\rho(x) - \rho(y)]^2 \, dx \, dy \leq \int_0^1 \iint_{\mathbb{R}^d \times \mathbb{R}^d} |z|^2 G_\varepsilon(z) |\nabla\rho(y + tz)|^2 \, dz \, dy \, dt$$

$$\leq \int_0^1 \int_{\mathbb{R}^d} |z|^2 G_\varepsilon(z) \int_{\mathbb{R}^d} |\nabla\rho(y)|^2 \, dy \, dz \, dt$$

$$\leq \varepsilon^2 \int_{\mathbb{R}^d} |z|^2 G(z) \, dz \int_{\mathbb{R}^d} |\nabla\rho(y)|^2 \, dy.$$

We thus have

$$\mathscr{J}_\varepsilon[\rho^\varepsilon] \leq \int_{\mathbb{R}^d} h(\rho^\varepsilon) \, dx + \varepsilon^2 D \int_{\mathbb{R}^d} |\nabla\rho^\varepsilon(x)|^2 \, dx$$

for some constant D. The functional in the right hand side is an Allen-Cahn functional for which the Γ convergence is proved, for example in [9]. Any recovery sequence for this functional will be a recovery sequence for $\mathscr{J}_\varepsilon$. $\qquad\square$

Proof of Theorem 7.2. The first part of the proof is classical (See for instance [104, Proposition 4.2] we recall the main steps here: We note that ρ^ε solves (6.3) with

$$\mathscr{J}_\varepsilon[\rho^\varepsilon(t)] + \int_0^t \int_{\mathbb{R}^d} \rho^\varepsilon |v^\varepsilon|^2 \, dx \, dt \leq \mathscr{J}_\varepsilon[\rho^\varepsilon_{in}], \quad \sup_{t \in [0,T]} \int_{\mathbb{R}^d} |x|^2 \rho^\varepsilon(t) \, dx \leq C(T).$$

We have in particular (see [5])

$$d_W(\rho^\varepsilon(t), \rho^\varepsilon(s)) \leq \left(\int_0^T \int_{\mathbb{R}^d} \rho^\varepsilon |v^\varepsilon|^2 \, dx \, dt \right)^{1/2} (t - s)^{1/2}$$

so ρ^ε is bounded in $C^{1/2}([0, T]; \mathcal{P}_2(\mathbb{R}^d))$. Together with the bound on the second moment, this implies that there exists a subsequence ρ^{ε_n} such that $\rho^{\varepsilon_n}(t)$ narrowly converges to $\rho(t)$ locally uniformly in t (see for example [19, Proposition 4.1]). In addition, the fact that $m > 2$ together with the bound on the energy implies that ρ^ε is bounded in $L^\infty(0, \infty; L^2(\mathbb{R}^d))$ and so $j^\varepsilon = \rho^\varepsilon v^\varepsilon$ is bounded in $L^2(0, \infty; L^{4/3}(\mathbb{R}^d))$. Up to another subsequence, we can thus assume that j^ε converges to j (weakly) and a classical argument

implies that $j = \rho v$ for some $v \in L^2(d\rho(t))$ (see for example [86, Proposition 4.1]). We can pass to the limit in the continuity equation to show that ρ, v satisfy the continuity equation

$$\partial_t \rho + \operatorname{div}(\rho v) = 0.$$

To complete the proof, we will use the assumption of convergence of the energy (7.3) to pass to the limit in the momentum equation. First, since $h(\rho^\varepsilon) \geq h^{**}(\rho^\varepsilon) \geq 0$ and

$$\liminf_{\varepsilon \to 0} \int_0^T \int_{\mathbb{R}^d} h^{**}(\rho^\varepsilon) \, dx \, dt \geq \int_0^T \int_{\mathbb{R}^d} h^{**}(\rho) \, dx \, dt$$

(by the convexity of h^{**}) and since $\frac{1}{4} \iint_{\mathbb{R}^d \times \mathbb{R}^d} G_\varepsilon(x-y)[\rho(x) - \rho(y)]^2 \, dx \, dy \geq 0$, we see that (7.3) implies that

$$(B.1) \qquad \int_0^T \int_{\mathbb{R}^d} h(\rho^{\varepsilon_n}) \varphi \, dx \, dt \to \int_0^T \int_{\mathbb{R}^d} h^{**}(\rho)\varphi \, dx \, dt \text{ for any } \varphi \in L^\infty,$$

$$(B.2) \qquad h(\rho^{\varepsilon_n}) - h^{**}(\rho^{\varepsilon_n}) \to 0 \quad \text{in } L^1((0,T) \times \mathbb{R}^d),$$

$$(B.3) \qquad \int_0^T \iint_{\mathbb{R}^d \times \mathbb{R}^d} G_{\varepsilon_n}(x-y)[\rho^{\varepsilon_n}(x) - \rho^{\varepsilon_n}(y)]^2 \, dx \, dy \, dt \to 0.$$

We now write the weak formulation of the momentum equation, using the fact that $v^\varepsilon = -\nabla h'(\rho^\varepsilon) + \nabla B_\varepsilon(\rho^\varepsilon)$ (see (6.3)):

$$\int_0^T \int_{\mathbb{R}^d} \rho^{\varepsilon_n} v^{\varepsilon_n} \cdot \xi \, dx \, dt = \int_0^T \int_{\mathbb{R}^d} \rho^{\varepsilon_n} \nabla[-h'(\rho^{\varepsilon_n}) + B_\varepsilon[\rho^{\varepsilon_n}]] \cdot \xi \, dx \, dt$$

$$(B.4) \qquad = \int_0^T \int_{\mathbb{R}^d} [\rho^{\varepsilon_n} h'(\rho^{\varepsilon_n}) - h(\rho^{\varepsilon_n})] \operatorname{div} \xi \, dx \, dt - \int_0^T \int_{\mathbb{R}^d} \rho^{\varepsilon_n} \nabla B_{\varepsilon_n}[\rho^{\varepsilon_n}] \cdot \xi \, dx \, dt.$$

We will complete the proof of the theorem by passing to the limit in (B.4) thanks to the following Lemma:

Lemma B.1. *Assume* (B.1), (B.2), *and* (B.3). *Then:*
For all $\xi \in C_c^\infty((0,\infty) \times \mathbb{R}^d; \mathbb{R}^d)$, there holds

$$(B.5) \qquad \lim_{n \to \infty} \int_0^T \int_{\mathbb{R}^d} \rho^{\varepsilon_n} \nabla B_{\varepsilon_n}[\rho^{\varepsilon_n}] \cdot \xi \, dx \, dt = 0.$$

For all $\xi \in C_c^\infty((0,\infty) \times \mathbb{R}^d; \mathbb{R}^d)$, there exists a subsequence ε_n' such that

$$(B.6) \qquad \lim_{n \to \infty} \int_0^T \int_{\mathbb{R}^d} \left[\rho^{\varepsilon_n'} h'(\rho^{\varepsilon_n'}) - h(\rho^{\varepsilon_n'}) \right] \operatorname{div} \xi \, dx \, dt = \int_0^T \int_{\mathbb{R}^d} \left[\rho h^{**'}(\rho) - h^{**}(\rho) \right] \operatorname{div} \xi \, dx \, dt.$$

This lemma implies (for a given $\xi \in C_c^\infty((0,\infty) \times \mathbb{R}^d; \mathbb{R}^d)$) and with the corresponding subsequence such that (B.6) holds)

$$\int_0^T \int_{\mathbb{R}^d} \rho v \cdot \xi \, dx \, dt = \lim_{n \to \infty} \int_0^T \int_{\mathbb{R}^d} \rho^{\varepsilon_n'} v^{\varepsilon_n'} \cdot \xi \, dx \, dt = \int_0^T \int_{\mathbb{R}^d} \left[\rho h^{**'}(\rho) - h^{**}(\rho) \right] \operatorname{div} \xi \, dx \, dt$$

and since this holds for all $\xi \in C_c^\infty((0,\infty) \times \mathbb{R}^d; \mathbb{R}^d)$, the theorem follows. $\qquad \square$

It remains now to prove the above lemma.

Proof of Lemma B.1. In order to prove (B.5), we use the same computation used in [106, 86]: Recalling that $\phi^\varepsilon = G_\varepsilon * \rho$ solves $\sigma \phi^\varepsilon - \varepsilon^2 \Delta \phi^\varepsilon = \rho$, we can write:

$$(\text{B.7}) \quad \frac{1}{2} \int_0^T \iint_{\mathbb{R}^d \times \mathbb{R}^d} G_\varepsilon(x-y)[\rho(x) - \rho(y)]^2 \, dx \, dy$$

$$= \int_0^T \iint_{\mathbb{R}^d \times \mathbb{R}^d} \rho \left[\frac{1}{\sigma} \rho - \phi_\varepsilon \right] dx \, dy$$

$$= \sigma \int_0^T \iint_{\mathbb{R}^d \times \mathbb{R}^d} \left[\frac{1}{\sigma} \rho - \phi_\varepsilon \right]^2 dx \, dy + \sigma \int_0^T \iint_{\mathbb{R}^d \times \mathbb{R}^d} \phi^\varepsilon \left[\frac{1}{\sigma} \rho - \phi_\varepsilon \right] dx \, dy$$

$$= \sigma \int_0^T \iint_{\mathbb{R}^d \times \mathbb{R}^d} \left[\frac{1}{\sigma} \rho - \phi_\varepsilon \right]^2 dx \, dy + \varepsilon^2 \int_0^T \iint_{\mathbb{R}^d \times \mathbb{R}^d} |\nabla \phi_\varepsilon|^2 \, dx \, dy$$

and

$$\int_0^T \int_{\mathbb{R}^d} \rho^\varepsilon \nabla B_\varepsilon[\rho^\varepsilon] \cdot \xi \, dx \, dt = \int_0^T \int_{\mathbb{R}^d} \rho^\varepsilon \nabla[\phi_\varepsilon - \sigma^{-1}\rho^\varepsilon] \cdot \xi \, dx \, dt$$

$$= \int_0^T \int_{\mathbb{R}^d} [\rho^\varepsilon - \sigma\phi_\varepsilon] \nabla[\phi_\varepsilon - \sigma^{-1}\rho^\varepsilon] \cdot \xi \, dx \, dt + \int_0^T \int_{\mathbb{R}^d} -\sigma\phi_\varepsilon \nabla[\phi_\varepsilon - \sigma^{-1}\rho^\varepsilon] \cdot \xi \, dx \, dt$$

$$= \frac{\sigma}{2} \int_0^T \int_{\mathbb{R}^d} [\phi_\varepsilon - \sigma^{-1}\rho^\varepsilon]^2 \mathrm{div}\, \xi \, dx \, dt + \varepsilon^2 \int_0^T \int_{\mathbb{R}^d} \phi_\varepsilon \nabla \Delta \phi_\varepsilon \cdot \xi \, dx \, dt$$

$$= \frac{\sigma}{2} \int_0^T \int_{\mathbb{R}^d} [\phi_\varepsilon - \sigma^{-1}\rho^\varepsilon]^2 \mathrm{div}\, \xi \, dx \, dt$$

$$+ \varepsilon^2 \int_0^T \int_{\mathbb{R}^d} \frac{|\nabla \phi_\varepsilon|^2}{2} \mathrm{div}\, \xi + D\xi : \nabla \phi_\varepsilon \otimes \nabla \phi_\varepsilon + \phi_\varepsilon \nabla \phi_\varepsilon \cdot \Delta \xi \, dx \, dt,$$

which together imply

$$\left| \int_0^T \int_{\mathbb{R}^d} \rho^{\varepsilon_n} \nabla B_{\varepsilon_n}[\rho^{\varepsilon_n}] \cdot \xi \, dx \, dt \right| \leq \|\mathrm{div}\, \xi\|_{L^\infty} \int_0^T \int_{\mathbb{R}^d} [\phi_\varepsilon - \sigma^{-1}\rho^\varepsilon]^2 \, dx \, dt$$

$$+ C(1 + \|D\xi\|_{L^\infty})\varepsilon^2 \int_0^T \int_{\mathbb{R}^d} |\nabla \phi_\varepsilon|^2 \, dx \, dt + C\varepsilon^2 \|\Delta \xi\|_{L^2}$$

$$\leq C(\xi) \left(\int_0^T \iint_{\mathbb{R}^d \times \mathbb{R}^d} G_{\varepsilon_n}(x-y)[\rho^{\varepsilon_n}(x) - \rho^{\varepsilon_n}(y)]^2 \, dx \, dy + \varepsilon^2 \right).$$

Thus (B.3) yields (B.5).

We now turn to the proof of (B.6). Given $\xi \in C_c^\infty((0,\infty) \times \mathbb{R}^d; \mathbb{R}^d)$, we fix $Q = (0,T) \times B_R$ such that $\mathrm{Supp}\, \xi \in Q$ and we introduce the measurable sets

$$E = \{h^{**}(\rho) = 0\} \cap Q, \qquad F = \{h^{**}(\rho) > 0\} \cap Q.$$

First we claim that there exists a subsequence ε_n' such that

$$(\text{B.8}) \qquad \lim_{n \to \infty} \iint_E \left[\rho^{\varepsilon_n'} h'(\rho^{\varepsilon_n'}) - h(\rho^{\varepsilon_n'}) \right] \mathrm{div}\, \xi \, dx \, dt \to 0.$$

Indeed, using (B.1), we get $h(\rho^{\varepsilon_n'}) \to 0$ a.e. in E (for some subsequence). Furthermore, for any $\delta > 0$, there exists C_δ such that $|\rho h'(\rho) - h(\rho)| \leq \delta + C_\delta h(\rho)$ (this is because the function in the left hand side

vanishes when $\rho = 0$ or θ which are the only zeroes of h and the two functions have the same growth at ∞). We deduce

$$\limsup_{n \to \infty} \iint_E \left| \rho^{\varepsilon'_n} h'(\rho^{\varepsilon'_n}) - h(\rho^{\varepsilon'_n}) \right| |\text{div } \xi| \, dx \, dt \le \|\text{div } \xi\|_{L^\infty} \limsup_{n \to \infty} \iint_E \delta + C_\delta h(\rho^{\varepsilon'_n}) \, dx \, dt \le \|\text{div } \xi\|_{L^\infty} |Q| \delta$$

and (B.8) follows by taking $\delta \to 0$.

Next, we will prove that the strict convexity if h^{**} in $\{h^{**}(s) > 0\}$ implies the strong convergence of ρ^ε to ρ in F. We define $w^\varepsilon = \max\{\rho^\varepsilon, \theta\}$, which only takes values in the set where $\rho \mapsto h^{**}(\rho)$ is strictly convex. Since w^{ε_n} and ρ^{ε_n} are bounded in $L^2(Q)$, they both have some weak limit w and ρ (up to some subsequence). And since $w^{\varepsilon_n} \ge \rho^{\varepsilon_n}$, we have $w \ge \rho$. Furthermore, we have

$$h^{**}(w^{\varepsilon_n}) = h^{**}(\rho^{\varepsilon_n}) \le h(\rho^{\varepsilon_n})$$

(the first equality holds since $h^{**}(\rho) = 0$ when $\rho < \theta$) and the convexity of h^{**} implies

$$\int_Q h^{**}(w) \le \liminf \int_Q h^{**}(w^{\varepsilon'_n}) \le \limsup \int_Q h^{**}(w^{\varepsilon'_n}) \le \limsup \int_Q h(\rho^{\varepsilon'_n}) = \int_Q h^{**}(\rho) \le \int_Q h^{**}(w)$$

where we used the assumption of convergence of the energy (7.3). We deduce that we have equality in all these inequalities. In particular

$$\lim \int_Q h^{**}(w^{\varepsilon'_n}) = \int_Q h^{**}(w)$$

and

$$\int_Q h^{**}(\rho) = \int_Q h^{**}(w).$$

The first equality and strict convexity of h^{**} in the set of values of $w^{\varepsilon'_n}(t, x)$ implies that $w^{\varepsilon'_n}$ converges strongly (and a.e.) to w. The second equality (together with the fact that $w \ge \rho$ and the strict monotonicity of h^{**} in (θ, ∞)) implies that $\rho = w$ in F. Finally, since $w^{\varepsilon'_n} \ge \rho^{\varepsilon'_n}$, we have

$$\int_F |w^{\varepsilon'_n} - \rho^{\varepsilon'_n}| = \int_F w^{\varepsilon'_n} - \rho^{\varepsilon'_n} \to \int_F w - \rho = 0$$

and so $\rho^{\varepsilon'_n}$ converges to $w = \rho$ in $L^1(F)$ and a.e. (up to yet another subsequence). Together with (B.8), this implies (B.6). $\qquad\square$

References

[1] Giovanni Alberti and Giovanni Bellettini. A non-local anisotropic model for phase transitions: asymptotic behaviour of rescaled energies. *European J. Appl. Math.*, 9(3):261–284, 1998.

[2] Damon Alexander, Inwon Kim, and Yao Yao. Quasi-static evolution and congested crowd transport. *Nonlinearity*, 27(4):823–858, 2014.

[3] Volkening Alexandria and Sandstede Björn. Modelling stripe formation in zebrafish: an agent-based approach. *J. R. Soc. Interface*, 12, 2015.

[4] Nicholas D. Alikakos, Peter W. Bates, and Xinfu Chen. Convergence of the Cahn-Hilliard equation to the Hele-Shaw model. *Arch. Rational Mech. Anal.*, 128(2):165–205, 1994.

[5] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[6] Jacob Bedrossian. Global minimizers for free energies of subcritical aggregation equations with degenerate diffusion. *Appl. Math. Lett.*, 24(11):1927–1932, 2011.

[7] Jacob Bedrossian and Nancy Rodríguez. Inhomogeneous Patlak-Keller-Segel models and aggregation equations with nonlinear diffusion in $\mathbb{R}^d$. *Discrete Contin. Dyn. Syst. Ser. B*, 19(5):1279–1309, 2014.

[8] Jacob Bedrossian, Nancy Rodríguez, and Andrea L Bertozzi. Local and global well-posedness for aggregation equations and Patlak–Keller–Segel models with degenerate diffusion. *Nonlinearity*, 24(6):1683, 2011.

[9] Giovanni Bellettini, Lorenzo Bertini, Mauro Mariani, and Matteo Novaga. Convergence of the one-dimensional Cahn-Hilliard equation. *SIAM J. Math. Anal.*, 44(5):3458–3480, 2012.

[10] Giovanni Bellettini, Giorgio Fusco, and Nicola Guglielmi. A concept of solution and numerical experiments for forward-backward diffusion equations. *Discrete Contin. Dyn. Syst.*, 16(4), 2006.

[11] Andrew J. Bernoff and Chad M. Topaz. Biological aggregation driven by social and environmental factors: a nonlocal model and its degenerate Cahn-Hilliard approximation. *SIAM J. Appl. Dyn. Syst.*, 15(3):1528–1562, 2016.

[12] Adrien Blanchet and Pierre Degond. Topological interactions in a Boltzmann-type framework. *J. Stat. Phys.*, 163(1):41–60, 2016.

[13] Adrien Blanchet and Pierre Degond. Kinetic models for topological nearest-neighbor interactions. *J. Stat. Phys.*, 169(5):929–950, 2017.

[14] Giovanni A. Bonaschi, José A. Carrillo, Marco Di Francesco, and Mark A. Peletier. Equivalence of gradient flows and entropy solutions for singular nonlocal interaction equations in 1D. *ESAIM Control Optim. Calc. Var.*, 21(2):414–441, 2015.

[15] Jean Bourgain, Haim Brezis, and Petru Mironescu. Another look at Sobolev spaces. In *Optimal control and partial differential equations*, pages 439–455. IOS, Amsterdam, 2001.

[16] C. P. et al. Brangwynne. Germline p granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, 324:1729–1732, 2009.

[17] Didier Bresch, Pierre-Emmanuel Jabin, and Zhenfu Wang. On mean-field limits and quantitative estimates with a large class of singular kernels: application to the Patlak-Keller-Segel model. *C. R. Math. Acad. Sci. Paris*, 357(9):708–720, 2019.

[18] Martin Burger, Marco di Francesco, and Marzena Franek. Stationary states of quadratic diffusion equations with long-range attraction. *Commun. Math. Sci.*, 11(3):709–738, 2013.

[19] Martin Burger and Antonio Esposito. Porous medium equation and cross-diffusion systems as limit of nonlocal interaction. *Nonlinear Anal.*, 235:Paper No. 113347, 30, 2023.

[20] John W. Cahn and John E. Hilliard. Free Energy of a Nonuniform System. I. Interfacial Free Energy. *The Journal of Chemical Physics*, 28:258–267, 1958.

[21] Vincent Calvez and José A. Carrillo. Volume effects in the Keller-Segel model: energy estimates preventing blow-up. *J. Math. Pures Appl. (9)*, 86(2):155–175, 2006.

[22] J. A. Carrillo, M. DiFrancesco, A. Figalli, T. Laurent, and D. Slepčev. Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. *Duke Math. J.*, 156(2):229–271, 2011.

[23] J. A. Carrillo, S. Hittmeir, B. Volzone, and Y. Yao. Nonlinear aggregation-diffusion equations: Radial symmetry and long time asymptotics. *Inventiones mathematicae*, 218:889–977, 2019.

[24] José A. Carrillo, Katy Craig, and Yao Yao. Aggregation-diffusion equations: dynamics, asymptotics, and singular limits. In *Active particles. Vol. 2. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 65–108. Birkhäuser/Springer, Cham, 2019.

[25] José A. Carrillo, Matias G. Delgadino, and Jeremy S. H. Wu. Convergence of a particle method for a regularized spatially homogeneous Landau equation. *Math. Models Methods Appl. Sci.*, 33(5):971–1008, 2023.

[26] José A. Carrillo, Charles Elbar, and Jakub Skrzeczkowski. Degenerate Cahn-Hilliard systems: From nonlocal to local. *arXiv preprint arXiv:2303.11929*, 2023.

[27] José A. Carrillo, Robert J. McCann, and Cédric Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.*, 179(2):217–263, 2006.

[28] Jose A. Carrillo, Hideki Murakawa, Makoto Sato, Hideru Togashi, and Olena Trush. A population dynamics model of cell-cell adhesion incorporating population pressure and density saturation. *Journal of Theoretical Biology*, 474:14–24, 2019.

[29] José Antonio Carrillo, Young-Pil Choi, and Maxime Hauray. The derivation of swarming models: mean-field limit and Wasserstein distances. In *Collective dynamics from bacteria to crowds*, volume 553 of *CISM Courses and Lect.*, pages 1–46. Springer, Vienna, 2014.

[30] José Antonio Carrillo, Katy Craig, and Francesco S. Patacchini. A blob method for diffusion. *Calc. Var. Partial Differential Equations*, 58(2):Paper No. 53, 53, 2019.

[31] José Antonio Carrillo, Antonio Esposito, and Jeremy Sheung-Him Wu. Nonlocal approximation of nonlinear diffusion equations. *Calculus of Variations and Partial Differential Equations*, 63(4):100, 2024.

[32] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: a review of models, methods and applications. I. Models and methods. *Kinet. Relat. Models*, 15(6):895–1015, 2022.

[33] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: a review of models, methods and applications. II. Applications. *Kinet. Relat. Models*, 15(6):1017–1173, 2022.

[34] Xinfu Chen. Global asymptotic limit of solutions of the Cahn-Hilliard equation. *J. Differential Geom.*, 44(2):262–311, 1996.

[35] Rinaldo M Colombo, Mauro Garavello, and Magali Lécureux-Mercier. A class of nonlocal models for pedestrian traffic. *Mathematical Models and Methods in Applied Sciences*, 22(04):1150023, 2012.

[36] Katy Craig and Andrea L. Bertozzi. A blob method for the aggregation equation. *Math. Comp.*, 85(300):1681–1717, 2016.

[37] Katy Craig, Karthik Elamvazhuthi, Matt Haberland, and Olga Turanova. A blob method for inhomogeneous diffusion with applications to multi-agent control and sampling. *Math. Comp.*, 92(344):2575–2654, 2023.

[38] Katy Craig, Inwon Kim, and Yao Yao. Congested aggregation via Newtonian interaction. *Archive for Rational Mechanics and Analysis*, 227(1):1–67, 2018.

[39] Katy Craig and Ihsan Topaloglu. Aggregation-diffusion to constrained interaction: minimizers & gradient flows in the slow diffusion limit. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 37(2):239–279, 2020.

[40] Noemi David, Tomasz Dębiec, and Benoît Perthame. Convergence rate for the incompressible limit of nonlinear diffusion-advection equations. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 40(3):511–529, 2023.

[41] Noemi David and Benoît Perthame. Free boundary limit of a tumor growth model with nutrient. *Journal de Mathématiques Pures et Appliquées*, 155:62–82, 2021.

[42] Noemi David and Markus Schmidtchen. On the incompressible limit for a tumour growth model incorporating convective effects. *Communications on Pure and Applied Mathematics*, 77(5):2613–2650, 2024.

[43] Pierre Degond, Sophie Hecht, and Nicolas Vauchelet. Incompressible limit of a continuum model of tissue growth for two cell populations. *Netw. Heterog. Media*, 15(1):57–85, 2020.

[44] Pierre Degond and Francisco-José Mustieles. A deterministic approximation of diffusion equations using particles. *SIAM J. Sci. Statist. Comput.*, 11(2):293–310, 1990.

[45] Matias G. Delgadino. Convergence of a one-dimensional Cahn-Hilliard equation with degenerate mobility. *SIAM J. Math. Anal.*, 50(4):4457–4482, 2018.

[46] Matias G. Delgadino, Xukai Yan, and Yao Yao. Uniqueness and nonuniqueness of steady states of aggregation-diffusion equations. *Comm. Pure Appl. Math.*, 75(1):3–59, 2022.

[47] M. Di Francesco and M. D. Rosini. Rigorous derivation of nonlinear scalar conservation laws from follow-the-leader type models via many particle limit. *Arch. Ration. Mech. Anal.*, 217(3):831–871, 2015.

[48] Simone Di Marino and Alpár Richárd Mészáros. Uniqueness issues for evolution equations with density constraints. *Math. Models Methods Appl. Sci.*, 26(9):1761–1783, 2016.

[49] R. L. Dobrušin. Vlasov equations. *Funktsional. Anal. i Prilozhen.*, 13(2):48–58, 96, 1979.

[50] Charles Elbar and Jakub Skrzeczkowski. Degenerate Cahn-Hilliard equation: From nonlocal to local. *Journal of Differential Equations*, 364:576–611, 2023.

[51] Charles M. Elliott. The Stefan problem with a nonmonotone constitutive relation. *IMA J. Appl. Math.*, 35(2):257–264, 1985. Special issue: IMA conference on crystal growth (Oxford, 1985).

[52] CM Elliott, MA Herrero, JR King, and JR Ockendon. The mesa problem: Diffusion patterns for $u_t = \Delta u^m$ as $m \to \infty$. *IMA journal of applied mathematics*, 37(2):147–154, 1986.

[53] Matt Elsey and Selim Esedoḡlu. Threshold dynamics for anisotropic surface energies. *Math. Comp.*, 87(312):1721–1756, 2018.

[54] Selim Esedoḡlu and Matt Jacobs. Convolution kernels and stability of threshold dynamics methods. *SIAM J. Numer. Anal.*, 55(5):2123–2150, 2017.

[55] Selim Esedoḡlu and Felix Otto. Threshold dynamics for networks with arbitrary surface tensions. *Comm. Pure Appl. Math.*, 68(5):808–864, 2015.

[56] Simone Fagioli and Oliver Tse. On gradient flow and entropy solutions for nonlocal transport equations with nonlinear mobility. *Nonlinear Anal.*, 221:Paper No. 112904, 35, 2022.

[57] Carles Falcó, Ruth E. Baker, and José A. Carrillo. A local continuum model of cell-cell adhesion. *SIAM Journal on Applied Mathematics*, pages S17–S42, 2023.

[58] Sylvain Faure and Bertrand Maury. Crowd motion from the granular standpoint. *Math. Models Methods Appl. Sci.*, 25(3):463–493, 2015.

[59] Alessio Figalli and Robert Philipowski. Convergence to the viscous porous medium equation and propagation of chaos. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:185–203, 2008.

[60] Michael Fischer, Laura Kanzler, and Christian Schmeiser. One-dimensional short-range nearest-neighbor interaction and its nonlinear diffusion limit. *SIAM Journal on Applied Mathematics*, 84(1):1–18, 2024.

[61] Giambattista Giacomin and Joel L. Lebowitz. Phase segregation dynamics in particle systems with long range interactions. I. Macroscopic limits. *J. Statist. Phys.*, 87(1-2):37–61, 1997.

[62] Giambattista Giacomin and Joel L. Lebowitz. Phase segregation dynamics in particle systems with long range interactions. II. Interface motion. *SIAM J. Appl. Math.*, 58(6):1707–1729, 1998.

[63] Omar Gil and F Quirós. Boundary layer formation in the transition from the porous media equation to a Hele-Shaw flow. *Annales de l'IHP Analyse non linéaire*, 20(1):13–36, 2003.

[64] Karl Glasner. A diffuse interface approach to Hele-Shaw flow. *Nonlinearity*, 16(1):49–66, 2003.

[65] François Golse. On the dynamics of large particle systems in the mean field limit. In *Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity*, volume 3 of *Lect. Notes Appl. Math. Mech.*, pages 1–144. Springer, [Cham], 2016.

[66] I. G. Götz and B. B. Zaltzman. Nonincrease of mushy region in a nonhomogeneous Stefan problem. *Quart. Appl. Math.*, 49(4):741–746, 1991.

[67] Nestor Guillen, Inwon Kim, and Antoine Mellet. A Hele-Shaw limit without monotonicity. *Archive for Rational Mechanics and Analysis*, pages 1–40, 2022.

[68] Qingyou He, Hai-Liang Li, and Benoît Perthame. Incompressible limits of Patlak-Keller-Segel model and its stationary state. *Acta Applicandae Mathematicae*, 188(1):11, 2023.

[69] Sophie Hecht and Nicolas Vauchelet. Incompressible limit of a mechanical model for tissue growth with non-overlapping constraint. *Commun. Math. Sci.*, 15(7):1913–1932, 2017.

[70] Sebastian Hensel and Tim Laux. BV solutions for mean curvature flow with constant contact angle: Allen-Cahn approximation and weak-strong uniqueness. *arXiv preprint arXiv: 2112.11150*, 2021.

[71] Thomas Hillen and Kevin J Painter. A user's guide to PDE models for chemotaxis. *Journal of mathematical biology*, 58(1):183–217, 2009.

[72] T. Hirose, K. Ninomiya, and S. et al. Nakagawa. A guide to membraneless organelles and their various roles in gene regulation. *Nat Rev Mol Cell Biol*, 24:288–304, 2023.

[73] Klaus Höllig. Existence of infinitely many solutions for a forward backward heat equation. *Trans. Amer. Math. Soc.*, 278(1):299–316, 1983.

[74] Roger L Hughes. A continuum theory for the flow of pedestrians. *Transportation Research Part B: Methodological*, 36(6):507–535, 2002.

[75] Noureddine Igbida. L^1-theory for Hele-Shaw flow with linear drift. *Math. Models Methods Appl. Sci.*, 33(7):1545–1576, 2023.

[76] Pierre-Emmanuel Jabin and Zhenfu Wang. Mean field limit for stochastic particle systems. *Active Particles, Volume 1 : Advances in Theory, Models, and Applications*, pages 379–402, 2017.

[77] Pierre-Emmanuel Jabin and Zhenfu Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.*, 214(1):523–591, 2018.

[78] Matt Jacobs, Inwon Kim, and Alpár R Mészáros. Weak solutions to the Muskat problem with surface tension via optimal transport. *Archive for Rational Mechanics and Analysis*, 239(1):389–430, 2021.

[79] Matt Jacobs, Inwon Kim, and Jiajun Tong. Darcy's law with a source term. *Archive for Rational Mechanics and Analysis*, 239(3):1349–1393, 2021.

[80] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

[81] Gunnar Kaib. Stationary states of an aggregation equation with degenerate diffusion and bounded attractive potential. *SIAM J. Math. Anal.*, 49(1):272–296, 2017.

[82] Trygve K Karper, Erik Lindgren, and Eitan Tadmor. A new model for chemotaxis with saturation effects and its incompressible limit. *Preprint*.

[83] Evelyn F Keller and Lee A Segel. Model for chemotaxis. *Journal of theoretical biology*, 30(2):225–234, 1971.

[84] Inwon Kim, Antoine Mellet, and Jeremy Sheung-Him Wu. Mean-field limit for congestion dynamics in dimension one. *Work in preparation*, 2023.

[85] Inwon Kim, Antoine Mellet, and Yijing Wu. A density-constrained model for chemotaxis. *Nonlinearity*, 36(2):1082–1119, 2023.

[86] Inwon Kim, Antoine Mellet, and Yijing Wu. Density-constrained chemotaxis and Hele-Shaw flow. *Transactions of the American Mathematical Society*, 377(01):395–429, 2024.

[87] Inwon Kim and Norbert Požár. Porous medium equation to Hele-Shaw flow with general initial density. *Transactions of the American Mathematical Society*, 370(2):873–909, 2018.

[88] Inwon Kim, Norbert Požár, and Brent Woodhouse. Singular limit of the porous medium equation with a drift. *Advances in Mathematics*, 349:682–732, 2019.

[89] Inwon Kim and Olga Turanova. Uniform convergence for the incompressible limit of a tumor growth model. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 35(5):1321–1354, 2018.

[90] Inwon C Kim, Benoît Perthame, and Panagiotis E Souganidis. Free boundary problems for tumor growth: a viscosity solutions approach. *Nonlinear Analysis*, 138:207–228, 2016.

[91] R. Kowalczyk. Preventing blow-up in a chemotaxis model. *J. Math. Anal. Appl.*, 305(2):566–588, 2005.

[92] Milan Kroemer and Tim Laux. The Hele-Shaw flow as the sharp interface limit of the Cahn-Hilliard equation with disparate mobilities. *Communications in Partial Differential Equations*, 47(12):2444–2486, 2022.

[93] Aimé Lachapelle and Marie-Therese Wolfram. On a mean field game approach modeling congestion and aversion in pedestrian crowds. *Transportation research part B: methodological*, 45(10):1572–1589, 2011.

[94] Tim Laux and Felix Otto. Convergence of the thresholding scheme for multi-phase mean-curvature flow. *Calculus of Variations and Partial Differential Equations*, 55(5):1–74, 2016.

[95] Pierre-Louis Lions and Sylvie Mas-Gallic. Une méthode particulaire déterministe pour des équations diffusives non linéaires. *C. R. Acad. Sci. Paris Sér. I Math.*, 332(4):369–376, 2001.

[96] Q.X. Liu, A. Doelman, and Rottschäfer V. et al. Phase separation explains a new class of self-organized spatial patterns in ecological systems. *Proc. Natl. Acad. Sci. U.S.A.*, 110, 2013.

[97] Stephan Luckhaus and Luciano Modica. The Gibbs-Thompson relation within the gradient theory of phase transitions. *Archive for Rational Mechanics and Analysis*, 107(1):71–83, 1989.

[98] Pavel M. Lushnikov, Nan Chen, and Mark Alber. Macroscopic dynamics of biological cells interacting via chemotaxis and direct contact. *Physical review E.*, 78, 2008.

[99] Sylvie Mas-Gallic. A deterministic particle method for the linearized Boltzmann equation. *Transport Theory Statist. Phys.*, 16(4-6):855–887, 1987.

[100] Bertrand Maury. Congested transport and microscopic and macroscopic scales. In *European Congress of Mathematics*, pages 427–442. Eur. Math. Soc., Zürich, 2018.

[101] Bertrand Maury, Aude Roudneff-Chupin, and Filippo Santambrogio. A macroscopic crowd motion model of gradient flow type. *Mathematical Models and Methods in Applied Sciences*, 20(10):1787–1821, 2010.

[102] Bertrand Maury, Aude Roudneff-Chupin, Filippo Santambrogio, and Juliette Venel. Handling congestion in crowd motion modeling. *Netw. Heterog. Media*, 6(3):485–519, 2011.

[103] Bertrand Maury and Juliette Venel. A discrete contact model for crowd motion. *ESAIM Math. Model. Numer. Anal.*, 45(1):145–168, 2011.

[104] Antoine Mellet. Hele-shaw flow as a singular limit of a keller-segel system with nonlinear diffusion. *arXiv preprint arXiv:2305.04742*, 2023.

[105] Antoine Mellet, Benoît Perthame, and Fernando Quiros. A Hele-Shaw problem for tumor growth. *Journal of Functional Analysis*, 273(10):3061–3093, 2017.

[106] Antoine Mellet and Yijing Wu. Γ-convergence of some nonlocal perimeters in bounded domains with general boundary conditions. *SIAM Journal on Mathematical Analysis*, 55(4):3226–3261, 2023.

[107] D.M. Mitrea and R.W. Kriwacki. Phase separation in biology; functional organization of a higher order. *Cell Commun. Signal*, 14, 2016.

[108] Luciano Modica. The gradient theory of phase transitions and the minimal interface criterion. *Arch. Rational Mech. Anal.*, 98(2):123–142, 1987.

[109] Luciano Modica and Stefano Mortola. Un esempio di Γ^--convergenza. *Boll. Un. Mat. Ital. B (5)*, 14(1):285–299, 1977.

[110] Daniela Morale, Vincenzo Capasso, and Karl Oelschläger. An interacting particle system modelling aggregation behavior: from individuals to populations. *J. Math. Biol.*, 50(1):49–66, 2005.

[111] Sebastien Motsch and Diane Peurichard. From short-range repulsion to Hele-Shaw problem in a model of tumor growth. *J. Math. Biol.*, 76(1-2):205–234, 2018.

[112] K. Oelschläger. A sequence of integro-differential equations approximating a viscous porous medium equation. *Z. Anal. Anwendungen*, 20(1):55–91, 2001.

[113] Karl Oelschläger. Large systems of interacting particles and the porous medium equation. *J. Differential Equations*, 88(2):294–346, 1990.

[114] Felix Otto. Dynamics of labyrinthine pattern formation in magnetic fluids: a mean-field theory. *Arch. Rational Mech. Anal.*, 141(1):63–103, 1998.

[115] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[116] Kevin J. Painter and Thomas Hillen. Volume-filling and quorum-sensing in models for chemosensitive movement. *Can. Appl. Math. Q.*, 10(4):501–543, 2002.

[117] J.K. Parrish and L. Edelstein-Keshet. Complexity, pattern, and evolutionary trade-offs in animal aggregation. *Science*, 284:99–101, 1999.

[118] Clifford S Patlak. Random walk with persistence and external bias. *The bulletin of mathematical biophysics*, 15(3):311–338, 1953.

[119] R. L. Pego. Front migration in the nonlinear Cahn-Hilliard equation. *Proc. Roy. Soc. London Ser. A*, 422(1863):261–278, 1989.

[120] Benoît Perthame, Fernando Quirós, and Juan Luis Vázquez. The Hele-Shaw asymptotics for mechanical models of tumor growth. *Arch. Ration. Mech. Anal.*, 212(1):93–127, 2014.

[121] Benoît Perthame and Nicolas Vauchelet. Incompressible limit of a mechanical model of tumour growth with viscosity. *Philos. Trans. Roy. Soc. A*, 373(2050):20140283, 16, 2015.

[122] Robert Philipowski. Interacting diffusions approximating the porous medium equation and propagation of chaos. *Stochastic Process. Appl.*, 117(4):526–538, 2007.

[123] P. I. Plotnikov. Forward-backward parabolic equations and hysteresis. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 233:183–209, 257–258, 1996.

[124] P. I. Plotnikov. Passage to the limit with respect to a small parameter in Cahn-Hilliard equations. *Sibirsk. Mat. Zh.*, 38(3):638–656, iii, 1997.

[125] Ju. G. Rešetnjak. The weak convergence of completely additive vector-valued set functions. *Sibirsk. Mat. .*, 9:1386–1394, 1968.

[126] Matthew Rosenzweig and Sylvia Serfaty. Global-in-time mean-field convergence for singular Riesz-type diffusive flows. *Ann. Appl. Probab.*, 33(2), 2023.

[127] Riccarda Rossi and Giuseppe Savaré. Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 2(2):395–431, 2003.

[128] Giovanni Russo. Deterministic diffusion of particles. *Comm. Pure Appl. Math.*, 43(6):697–733, 1990.

[129] Etienne Sandier and Sylvia Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.*, 57(12):1627–1672, 2004.

[130] M. Slemrod. Dynamics of measure valued solutions to a backward-forward heat equation. *J. Dynam. Differential Equations*, 3(1):1–28, 1991.

[131] Peter Sternberg. The effect of a singular perturbation on nonconvex variational problems. *Arch. Rational Mech. Anal.*, 101(3):209–260, 1988.

[132] Yoshie Sugiyama. Time global existence and asymptotic behavior of solutions to degenerate quasi-linear parabolic systems of chemotaxis. *Differential Integral Equations*, 20(2):133–180, 2007.

[133] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[134] Chad M. Topaz, Andrea L. Bertozzi, and Mark A. Lewis. A nonlocal continuum model for biological aggregation. *Bull. Math. Biol.*, 68(7):1601–1623, 2006.

[135] Juan Luis Vázquez. *The porous medium equation*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, Oxford, 2007. Mathematical theory.

Department of Mathematics, UCLA, Los Angeles, CA
E-mail address: `ikim@math.ucla.edu`

Department of Mathematics, University of Maryland, College Park, MD
E-mail address: `mellet@umd.edu`

Department of Mathematics, UCLA, Los Angeles, CA
E-mail address: `jeremywu@math.ucla.edu`