SORTING PHENOMENA IN A MATHEMATICAL MODEL FOR TWO MUTUALLY ATTRACTING/REPELLING SPECIES

MARTIN BURGER, MARCO DI FRANCESCO, SIMONE FAGIOLI,
AND ANGELA STEVENS

Abstract. Macroscopic models for systems involving diffusion, short-range repulsion, and long-range attraction have been studied extensively in the last decades. In this paper we extend the analysis to a system for two species interacting with each other according to different inner- and intra-species attractions. Under suitable conditions on this self- and crosswise attraction an interesting effect can be observed, namely phase separation into neighbouring regions, each of which contains only one of the species. We prove that the intersection of the support of the stationary solutions of the continuum model for the two species has zero Lebesgue measure, while the support of the sum of the two densities is a connected interval.

Preliminary results indicate the existence of phase separation, i.e. spatial sorting of the different species. A detailed analysis is given in one spatial dimension. The existence and shape of segregated stationary solutions is shown via the Krein-Rutman theorem. Moreover, for small repulsion/nonlinear diffusion, also uniqueness of these stationary states is proved.

1. Introduction

The interplay between (nonlinear) diffusion and nonlocal attractive/repulsive interactions arises in a variety of contexts in the natural-, life-, and social sciences. As a non exhaustive list of examples, let us mention granular media physics [7, 19], astrophysics [13, 81], semiconductors [41, 81], chemotaxis [55, 72, 49, 2, 70, 42], ecology, animal swarming and aggregation [62, 66, 2, 18, 63, 77], alignment [71, 43], and opinion formation [76, 1]. One way of deriving such models from *first order* microscopic systems of (stochastic) interacting particles $x_1, \dots, x_n$ is the following. Each particle x_i is driven by nonlinear forces due to short range repulsion, see [63] - respectively undergoes an independent Brownian motion. Further, it moves towards higher concentrations of particles of its own kind, respectively those of an external signal, [73]. At the macroscopic and continuum level, this set of rules results in a nonlocal partial differential equation

$$\partial_t \rho = \operatorname{div}\left[\rho \nabla \left(a(\rho) - W * \rho\right)\right], \tag{1}$$

respectively, a chemotaxis-type system

$$\partial_t \rho = \operatorname{div}\left[\mu(\rho, v)\nabla\rho - \chi(\rho, v)\rho\nabla v\right] \quad , \quad \tau\partial_t v = \eta\Delta v + k(\rho, v)\ . \tag{2}$$

For $\mu(\rho, v) = \rho a'(\rho)$, $\chi(\rho, v) = \chi_0$, $\tau = 0$, and $k(\rho, v) = \rho - \beta v$, thus $v = (\beta I - \Delta)^{-1}\rho$, system (2) is equivalent to (1), with W being the Newtonian- or Bessel-potential ($\beta = 0$ or $\beta > 0$). Especially when $a(\rho) = \log(\rho)$, i.e. $\mu(\rho, v) = 1$ in this case, then a prototype model for nonlocal aggregation phenomena results, namely the simplified, classical parabolic-elliptic Keller-Segel model [55], or in physics a model for gravitational self-interacting clusters [81].

In (1), $\rho = \rho(x, t)$ denotes the density of particles, $a : [0, \infty) \to [0, +\infty)$ is a C^1 monotone increasing function with $a'(0) = 0$, and $W(x) = \tilde{W}(|x|)$ is a C^1 potential

with $\tilde{W}'(r) < 0$ for all $r > 0$. The PDE (1) can be interpreted as the gradient flow of the functional

$$\mathcal{F}[\rho] = \int A(\rho)dx - \frac{1}{2}\int \rho W * \rho\, dx \;\text{ where } A(\rho) = \int_0^\rho a(\xi)d\xi\,, \tag{3}$$

w.r.t. the d_2 Wasserstein distance arising in optimal transport theory. Compare also the different Lyapunov functionals and energies used for (2) in [81, 42, 70], e.g. $A(\rho) = \rho\log\rho$. Roughly speaking, this means that the velocity $\tilde{v} = -\nabla\left(a(\rho) - W*\rho\right)$ in the continuity equation $\rho_t + \operatorname{div}(\rho\tilde{v}) = 0$ can be interpreted as the sub-differential of $\mathcal{F}$ in the metric space $\mathcal{P}_2$ of probability measures with finite second moment and metric d_2, see [3] for more details.

Models of type (1), (2) with *two or more species* being involved are considered in the context of chemotaxis [40, 48, 82, 83], opinion formation [39], pedestrian dynamics [4], and population biology [30, 36]. A reasonable generalization of (3) for two species is

$$\mathcal{E}[\rho_1,\rho_2] = \int f(\rho_1,\rho_2)dx - \frac{1}{2}\int \rho_1 S_1*\rho_1 dx - \frac{1}{2}\int \rho_2 S_2*\rho_2 dx - \int \rho_1 K*\rho_2 dx, \tag{4}$$

where $f : [0,+\infty)\times[0,+\infty) \to \mathbb{R}$, $f \in C^1$, and S_1, S_2, K are C^1 and radially decreasing potentials like W above. The first term of $\mathcal{E}$ typically represents local repulsion, while the nonlocal terms model attractive forces. We call S_1 and S_2 *self-interaction* potentials and K *cross-interaction* potential. For $(\mathcal{P}_2(\mathbb{R}^d), d_2)\times(\mathcal{P}_2(\mathbb{R}^d), d_2)$ with the natural product topology, the formal gradient flow of $E[\rho_1,\rho_2]$ w.r.t. this metric structure is given by

$$\begin{cases} \partial_t\rho_1 = \operatorname{div}\big[\rho_1\nabla\big(f_{\rho_1}(\rho_1,\rho_2) - S_1*\rho_1 - K*\rho_2\big)\big] \\ \partial_t\rho_2 = \operatorname{div}\big[\rho_2\nabla\big(f_{\rho_2}(\rho_1,\rho_2) - S_2*\rho_2 - K*\rho_1\big)\big]\,. \end{cases} \tag{5}$$

The functionals we investigate in this paper are of the form (4), with

$$f(\rho_1,\rho_2) = \epsilon(\rho_1+\rho_2)^2/2, \quad \epsilon > 0,$$

leading to

$$\begin{cases} \partial_t\rho_1 = \operatorname{div}\big[\epsilon\rho_1\nabla(\rho_1+\rho_2) - \rho_1\nabla S_1*\rho_1 - \rho_1\nabla K*\rho_2\big] \\ \partial_t\rho_2 = \operatorname{div}\big[\epsilon\rho_2\nabla(\rho_1+\rho_2) - \rho_2\nabla S_2*\rho_2 - \rho_2\nabla K*\rho_1\big] \end{cases}. \tag{6}$$

Besides modelling local repulsion and global attraction between different types of particles (cf. the introduction and references in [22]) our particular motivation is to consider cell sorting due to differential attraction and the resulting pattern formation. Similar cells or species of equal size (consequently the terms $(\rho_1+\rho_2)$ in (6)), but with different reactions to attraction forces undergo a reorganization process, where cells with stronger self-attraction finally sort into the center of the total cell population and those with weaker self-attraction to the outside. Differential attraction can either be by an external (chemical) signal or directly between the species. Such phenomena are observed in developmental processes. In the first case one may assume that S_1, S_2, and K are multiples of the same kernel, modelling indirectly a chemo-attractant; like the fundamental solution of the elliptic equation for the chemo-attractant in the prototype Keller-Segel system, which can be rewritten into the single prototype aggregation equation. We are interested in suitable conditions for the self- and the cross-attraction, which result in the above mentioned sorting phenomenon. In [40] for differential attraction, i.e. different chemotactic sensitivities, of two species towards higher concentrations of one chemo-attractant, it was proved that if the solution for the more strongly attracted species blows up, than also the second one blows up at the same time. The amount of mass,

which concentrates in the joint blowup is different though, and controlled by the system parameters. This last, formal result, hints towards a possible separation of the main amount of masses of the two species. The more strongly attracted species accumulates more mass in the blowup than the other one. Cell sorting due to differential adhesion/attraction was also discussed in [80, 58], where stochastic particle models and numerical studies of continuum models were considered. Based on experimental results, see e.g. [61], sorting due to differential chemotaxis and chemical spiral waves was modeled and simulated in [78, 68]. In [54] a rigorous analysis was given.

We are interested in steady states of (6) and minimizers of (4), hence let us first discuss some relevant results in the single species setting (2) and (1). Varying χ and k in (2) plays a similar role as varying W in (1) for the pattern formation properties of the respective solution. Though a one-to-one connection between all variants of aggregation equations and chemotaxis systems as in the case of the Newtonian- and Besselpotential for $(-W)$ and the prototype Keller-Segel system has not been established, both types of models and their dynamics are strongly intertwined.

A wealth of mathematical results on existence of global solutions, blowup phenomena and pattern formation exists for these models, see e.g. the summaries in [46, 47, 69]. The longtime behavior of the prototype models is especially interesting in two dimensions as a biological model for the peculiar chemotactic self-organisation of *Dictyostelium discoideum* [49], and for self-graviational collapse and star formation in astrophysics in dimension three [13]. In both cases, see also [66], blowup phenomena in the respective dimension, and point support of stationary Dirac-type solutions are of crucial relevance for the respective application. Non-trivial stationary solutions and their qualitative features have been analyzed in [72] for the general system (2) with Neumann boundary conditions, for $\chi(\rho, v) = \chi_0 \mu(\rho, v)\rho\phi'(v)$, and $\tau = 1$. Especially the one dimensional steady states and their stability are well understood. These results should be compared with aggregation equation analoga. With $\rho = C\exp(\chi_0\phi(v))$ the steady state analysis in [72] reduces to analyzing the elliptic equation $\eta\Delta v + k(\psi(v, \lambda), v) = 0$, where ψ is the flux of $r'(s) = \chi(r, s)/\mu(r, s) = \chi_0\rho\phi'(v)$.

In [2] density dependent diffusion-drift equations for aggregation of the form

$$\partial_t \rho = \partial_x \left[\mu(\rho)\partial_x \rho - \gamma(\cdot, \rho)\right], \tag{7}$$

are analyzed with e.g. $\mu(\rho) = m\rho^{m-1}$ and $\gamma(t, x, \rho)$ being proportional to ρ, respectively depending on a functional of the density distribution $\rho(t, \cdot)$. Thus integral terms are included, as discussed in [62, 66], and (7) can be compared with (1). The connections between chemotaxis systems and (7) are given, and a Lyapunov function is constructed, i.e. a functional decreasing along solutions as time increases. This provides a general theorem on global existence of solutions. Asymptotic convergence to steady states is proved, which can be non-trivial, for instance plateau-like.

Free energies in such models and in chemotaxis are based on an entropic term, e.g. $A(\rho) = \rho\log\rho - \rho + 1$ (or just $A(\rho) = \rho\log\rho$) in (3). In [81] not only global minimizers of (3) are studied, but the whole solution set. Here ρ is a steady solution of the governing PDE - the associated gradient flow - if and only if it is a critical point of the variational problem with constraint $\int \rho = M$. Due to the logarithmic term in the entropy we have $a(\rho) \sim \log\rho$ and inverting this relation to an *entropy variable* $\rho \sim \exp a(\rho)$ one can obtain simplified systems for stationary states. With this exponential transformation and the Moser-Trudinger inequality as given in [65, 64] the critical mass 8π for graviational collapse is deduced. Existence, uniqueness, stability and symmetry breaking of stationary solutions are proved via

the precise connections between the free energy and the respective Vlasov-Fokker-Planck (chemotaxis-like) equation.

In [42] a Lyapunov functional - c.f. the energy functional in (3) - for the prototype chemotaxis-model in two dimensions is provided, and extended to more general equations in [70]. Also here the connection to the exponential transformation and the elliptic equation given in [65] is notified. Geometric criteria for not necessarily trivial stationary states are derived. Now any metric ds^2 on a two-dimensional sphere determines a Gauss curvature function K satisfying the Gauss-Bonnet formula, $\int\int_{S^2} K d\tilde{\mu} = 4\pi$. Here $d\tilde{\mu}$ is the volume element of S^2. To characterize all K belonging to metrices ds^2 and relating to the standard metric ds_0^2 via $ds^2 = p\, ds_0^2$, with a positive function p on the sphere, one has to determine $p = p(K)$ uniquely. In [65] it is proved that the transformation $p = \exp(v)$ reduces this question to solving

$$\Delta v + Ke^{2v} - 1 = 0,$$

on the sphere, which is done by a variational approach. This is a specific form of the elliptic problem analyzed in [72], where the more general transformation $\rho = C\exp(\chi_0\phi(v))$ was used for the steady states analysis of generalized chemotaxis systems. With $k(\psi(v,\lambda), v) = Ke^{2v} - 1$, i.e. $\psi(v,\lambda) = Ke^{2v}$, respectively $\chi_0\phi = 2v$ in [72], the above elliptic PDE results. Qualitative features of simplified chemotaxis systems with non-linear diffusion have been discussed in [74, 75]. See also further references therein.

Existence and uniqueness for (1) are discussed e.g. in [28], the last part of the book [3], and earlier via the related works on chemotaxis, c.f. [49, 45, 46, 47, 74, 75]. In [8] also singular potentials W are considered, e. g. Coulomb potential as in electrodynamics, $W(x) = |x|$, $W(x) = |x|^\alpha$, $\alpha > 2 - d$. We also refer to [17, 14, 15] for further discussion.

For $a(\rho) = \log\rho$ in (3), without any nonlocal effects, but with the addition of a confinement external potential V, in [50] weak solutions to the linear Fokker-Planck equation in a gradient flow setting were derived by constructing the finite time-step minimizing movement

$$\rho^n \mapsto \rho^{n+1} = \mathrm{argmin}_\rho \left(F[\rho] + \frac{1}{2\Delta t} d_2^2(\rho^n, \rho) \right),$$

and taking the limit $\Delta t \to 0$. This idea was applied to (1) in [28], and to a more general metric framework in [3].

The large time asymptotics of (1) depend on the competing effects of $\int A(\rho)$, which promotes particle spreading, such that the density ρ stabilises at a constant state (which is zero if we consider $\mathbb{R}^d$), whereas the nonlocal term drives the particles towards aggregation. A Dirac delta results, which is located at the (preserved) center of mass of the system. To prove existence of a global minimum of $\mathcal{F}$ is therefore a challenging problem. In [5] this issue was tackled in detail by using Lions' concentrated compactness technique.

The limiting case $a = 0$ was analyzed in [33, 24], see also the references therein. The case $a(\rho) = \rho$ was derived as a nonlocal *repulsive* effect under the action of a potential V_ϵ which converges weakly to a Dirac delta as $\epsilon \to 0$. For a formal argument see [63], see also [67] for a rigorous derivation via interacting particle systems with short-repulsion (the repulsion range shrinking to zero as the number of particles goes to infinity). The results in [22] and [5] provide a clear picture of the behaviour of the corresponding functional

$$\mathcal{G}[\rho] = \frac{1}{2}\int \rho(\epsilon\rho + W * \rho)dx,$$

with a given diffusion constant $\epsilon > 0$, and a $W^{1,1}$ attractive potential $W \leq 0$. If $\epsilon \geq \|W\|_{L^1}$, then $\mathcal{G}$ is uniformly convex, and hence zero is its global minimizer. This suggests that in this case $\rho \to 0$ for large times. On the other hand, if $\epsilon < \|W\|_{L^1}$, a nontrivial global minimizer exists in the class $\int \rho\, dx = 1$. This suggests existence of nontrivial steady state for large times, but the solution to the Cauchy problem could still decay to zero if a large enough variance is present initially. This problem is still open in the case of slow diffusion, see [6]. In one dimension a more refined analysis can be given. In [22] it was proved that a unique steady state with given mass and center of mass exists for

$$\rho_t = [\rho(\epsilon\rho + W * \rho)_x]_x,$$

provided that $\epsilon < \|W\|_{L^1}$ and W is radially increasing, negative, and supported on the whole of $\mathbb{R}$. Such a steady state is also the unique global minimizer for $\mathcal{G}$ for fixed mass and fixed center of mass, it is compactly supported, symmetric, and in $W^{1,\infty}$, with a shape similar to a Barenblatt profile for the porous medium equation, see [79]. Such steady state is locally stable for large times, see the recent [38]. The result in [22] was partly extended to more general nonlinear diffusions in [23], where the existence of a unique diffusion constant for a given support of the steady state is proven. Both results rely on the Krein-Rutman theorem in order to characterise the steady states as eigenvectors of a certain nonlocal operator. Further generalizations and completions of some open questions in those papers were recently given in [53, 52]. One major issue solved in [22] is to prove that a one-to-one correspondence exists between the diffusion constant (eigenvalue) and the support of the steady state. The results in [22] improved a previous result in [20] about the existence and uniqueness of nontrivial steady states for small diffusion constant $\epsilon \ll 1$ performed via an implicit function theorem argument.

Now let us come back to the functional (4). The first term of $\mathcal{E}$ typically represents local repulsion. Apart from nonlocal cross-diffusion via the kernels S_i also f, when containing mixed terms in ρ_1 and ρ_2, introduces cross diffusion. Indeed, (5) can be formally rewritten as

$$\partial_t U = \operatorname{div}\left(D(U)\nabla U\right) - \operatorname{div}\left(\begin{pmatrix}\rho_1 & 0\\ 0 & \rho_2\end{pmatrix}\nabla\begin{pmatrix}S_1 * \rho_1 + K * \rho_2\\ S_2 * \rho_2 + K * \rho_1\end{pmatrix}\right), \tag{8}$$

with $U = (\rho_1, \rho_2)$,

$$D(U) = \begin{pmatrix}\rho_1 f_{\rho_1,\rho_1} & \rho_1 f_{\rho_1,\rho_2}\\ \rho_2 f_{\rho_1,\rho_2} & \rho_2 f_{\rho_2,\rho_2}\end{pmatrix}.$$

If $D(U)$ is symmetric and semi-positive definite, then the theory in [51] applies, see also the recent [34] in the context of Wasserstein gradient flows. However, in our situation $D(U)$ is never symmetric. This makes existence proofs for solutions of (5) in full generality a delicate problem. In most cases, $D(U)$ does not even have a semi-positive definite symmetric part.

A comprehensive existence and uniqueness theory for (5) with $f \equiv 0$ was given in [35], where convergence of the JKO scheme leads to existence of weak measure valued solutions for 'mildly singular' potentials. A suitable notion of displacement convexity for systems provides a uniqueness result. An implicit-explicit variation of the JKO scheme yields an existence result in [35], also for nonlocal terms which are not of gradient flow type, i.e. for non symmetric cross-interaction terms. Without the nonlocal interaction part, an existence theory with a direct application of the JKO scheme is given in [57] for a fully coupled system of two second-order parabolic degenerate equations arising as a thin film approximation to the Muskat problem. The peculiar form of the cross-diffusion in this case allows for an energy functional with good coercivity properties. In [57] the regularity needed in order

to identify the suitable Euler-Lagrange equation in the variational scheme was obtained. Although not strictly related to our model, let us also mention the hybrid variational scheme generalizing the JKO scheme. This has been introduced in [16] for the parabolic-parabolic Keller-Segel model in $\mathbb{R}^2$ working in the product space $\mathscr{P}_2(\mathbb{R}^2) \times L^2(\mathbb{R}^2)$. Using a modified Wasserstein distance between vector-valued densities on $\mathbb{R}$, in [84] the variational structure of systems with degenerate diffusion and nonlinear reaction terms was investigated. A more general approach has been discussed in [59], where a convexity concept for reaction-diffusion systems was developed, that allows to analyze some important examples.

In system (6) the symmetric part of $D(U)$ is not semi-positive definite for all $\rho_1, \rho_2 \geq 0$ in general. A brute-force JKO approach would imply a good dissipation estimate only for $w = \rho_1 + \rho_2$, but not for both ρ_1 and ρ_2. Thus a regularising effect for the sum w is still possible. On the other hand, such a degenerate dissipation estimate does not prevent the formation of discontinuities for ρ_1 and ρ_2 separately. In order to partly validate this hypothesis, we focus on stationary states in one space dimension, and show that *discontinuous stationary patterns* may arise. More precisely, we prove that ρ_1 and ρ_2 can separate completely in the stationary state, i.e. they feature a jump discontinuity at exactly the same point, with the sum $w = \rho_1 + \rho_2$ remaining smooth at that point. We call such a configuration a *fully segregated steady state*.

Segregation in multi-species systems with nonlinear diffusion terms like in (6) and possible reaction terms has been widely investigated in the literature. We refer to [11, 10, 12, 9] and references therein. It is well known that segregated initial data produce a unique segregated solution. The problem of existence of a mixed solution is partly open. Difficulties in our case occur due to the presence of the nonlocal terms, which typically do not allow for a local comparison principle and may produce aggregation phenomena. The emergence vs. non-emergence of segregated steady states in two species systems with nonlocal attraction has recently been treated in [31], where the repulsive effect of the nonlinear diffusion has been replaced by an upper bound for $\rho_1 + \rho_2$. Then the minimisation problem for $\mathcal{E}$ with $f = 0$ and $0 \leq \rho_1 + \rho_2 \leq 1$ with all interaction potentials being multiples of a given function K is analyzed. Sharp conditions on these multiplying factors are provided, which results in complete segregation of the two species. In [60] rearrangements for three already separated domains has been considered in a different context.

More specifically on system (6) (almost parallel to our result), the recent [27] shows how to construct explicit segregated and non-segregated steady states with power-law interaction potentials (with possible confinement effects) and produces numerical evidence that segregation may not occur for large ϵ. At the same time, [56] shows that initially segregated initial data yield segregated solutions for all times for a system with the same (cross-)diffusion term of (6) and external potentials (under suitable assumptions on the latter). We also mention at this stage the result in [25], in which an existence theory for a reaction diffusion system with the same (cross-)diffusion of (6) has been proven via JKO scheme with BV initial conditions, without any initial separation assumption.

Let us know briefly summarise our results. For the one-dimensional case, we first investigate conditions for existence or non-existence of non trivial stationary solutions. For small diffusion coefficient ϵ we show existence of segregated stationary states via the implicit function theorem. We extend the strategies in [21, 22] to the multi species case. We also relate stationary solutions to energy minimizers in a rigorous way and provide some results characterizing their structure. In particular we verify that for stationary solutions and energy minimizers, the sum w is

supported on a connected interval and we give a rigorous result on the segregation in the case of dominant self-attraction. In a case of weak self-attraction we can characterized the support of the minimizers by the same arguments as in [31].

For the interaction kernels in (6) we assume unless further noted:

(A1) $S_1, S_2, K \in C^2(\mathbb{R})$,
(A2) S_1, S_2, K are radially symmetric and decreasing w.r.t. the radial variable,
(A3) S_1, S_2, K are nonnegative and have finite mass on $\mathbb{R}$.

The paper is organized as follows. In Section 2 we give some preliminary results on phase separation for (6). In Section 3 we analyze the relation of stationary solutions with critical points of the associated energy functional and provide sufficient conditions on ϵ and on the interaction kernels yielding non existence of non-trivial steady states. Section 4 is devoted to the existence analysis of stationary solutions involving phase-separation via the Krein-Rutman based approach given in [22]. Finally, in Section 5 existence and uniqueness results for stationary solutions are proved in case of small repulsion, corresponding to small nonlinear diffusion in (6). We complement our results with some numerical simulations in Section 6 showing segregated behavior as well as mixing and diffusion-dominated behaviour for large diffusions.

2. Segregation due to differential aggregation

First, we provide some preliminary results in arbitrary dimensions about pattern formation for model (6), and more specifically on the emergence or non-existence of segregation. Consider the canonical model

$$\begin{cases} \partial_t \rho_1 = \operatorname{div}\big(\epsilon \rho_1 \nabla(\rho_1+\rho_2) - \rho_1 \nabla V_1\big), \\ \partial_t \rho_2 = \operatorname{div}\big(\epsilon \rho_2 \nabla(\rho_1+\rho_2) - \rho_2 \nabla V_2\big), \end{cases} \tag{9}$$

where V_1 and V_2 are two given smooth external potentials. We can rewrite (6) in this form, with V_1 and V_2 being determined by convolutions of ρ_1 and ρ_2 with aggregation kernels, c.f. (11).

Proposition 2.1. *Let $V_i \in L^1_{loc}(\mathbb{R}^d) \cap W^{1,\infty}(\mathbb{R}^d)$ be given external potentials. If $(\rho_1^\infty, \rho_2^\infty)$ is a C^1 (weak) stationary solution of (9), then we have*

$$supp(\rho_1^\infty) \cap supp(\rho_2^\infty) \subseteq \{\nabla V_1 = \nabla V_2\}. \tag{10}$$

Proof. Let $x \in \operatorname{supp}(\rho_1^\infty) \cap \operatorname{supp}(\rho_2^\infty)$. Then the weak formulation of (9) implies

$$\epsilon \nabla(\rho_1^\infty + \rho_2^\infty)(x) = \nabla V_1(x) = \nabla V_2(x),$$

from which the assertion follows. □

An immediate consequence of Proposition 2.1 is the following result, which deals with the special case of all the interaction kernels being multiples of the fundamental solution of the Laplace equation.

Proposition 2.2. *Let K be the fundamental solution of the Laplace equation in $\mathbb{R}^d$, and $S_1 = \sigma_1 K$, $S_2 = \sigma_2 K$ with $\sigma_1 \le 1 \le \sigma_2$ and $\sigma_1 \neq \sigma_2$. Then, every C^1 stationary solution (ρ_1, ρ_2) of (6) is fully segregated, i. e. $supp(\rho_1) \cap supp(\rho_2)$ has empty interior.*

Proof.

$$\text{Let } V_1 = S_1 * \rho_1 + K * \rho_2, \qquad V_2 = S_2 * \rho_2 + K * \rho_1. \tag{11}$$

Assume that there is a non-empty open set $\mathcal{O} \subset \operatorname{supp}(\rho_1) \cap \operatorname{supp}(\rho_2)$. Then Proposition 2.1 implies $\mathcal{O} \subset \{\nabla V_1 = \nabla V_2\}$. Hence $V_1 - V_2 = c = const.$ on $\mathcal{O}$, i.e. $(S_1 - K) * \rho_1 - (S_2 - K) * \rho_2 = c$. With the assumptions on the kernels, we can apply the Laplace operator and obtain $(\sigma_1 - 1)\rho_1(x) + (1 - \sigma_2)\rho_2(x) = 0$, for

all $x \in \mathcal{O}$. The assumptions on σ_1 and σ_2 imply $\rho_1(x) = \rho_2(x) = 0$, which is a contradiction. □

Proposition 2.2 shows segregation of steady states in a particular case. However, this effect occurs for a much wider class of aggregation kernels S_1, S_2, K, as we will argue and partially prove below.

First, take a closer look at the dynamics of segregation by using a transformation of variables similar to the one in [12], where strong reaction terms induced the segregation though. Let

$$w := \rho_1 + \rho_2, \qquad \zeta := \frac{\rho_1 - \rho_2}{w}, \tag{12}$$

where the relative difference ζ is only considered on the support of the total density w. Adding both equations in (9) and using the notion of (11) yields

$$\begin{aligned} \partial_t w &= \operatorname{div}\Big(\epsilon w \nabla w - \rho_1 \nabla V_1 - \rho_2 \nabla V_2\Big), \\ &= \operatorname{div}\left(\epsilon w \nabla w - w\frac{1+\zeta}{2}\nabla V_1 - w\frac{1-\zeta}{2}\nabla V_2\right). \end{aligned}$$

Thus the dynamics of w are governed by a porous medium equation with additional convective terms. The dynamics of ζ are obtained by subtracting the equations for ρ_1 and ρ_2 and then inserting the equation for w, which yields

$$w\partial_t \zeta = \left(\epsilon w \nabla w - w\frac{1+\zeta}{2}\nabla V_1 - w\frac{1-\zeta}{2}\nabla V_2\right)\cdot\nabla\zeta - \frac{1-\zeta^2}{2}\operatorname{div}\big(w\nabla(V_1 - V_2)\big).$$

As in [12] the evolution of ζ is governed by a first-order equation, which gives particular insight into the dynamics of the system. The first of the two terms on the right-hand side of the above equation is along the flux of w, which determines the spatial position and shape of the solution. The crucial part for the segregation dynamics is the second term, a reaction term w.r.t. ζ and two fixed points $\zeta = \pm 1$, corresponding to segregation. Depending on the sign of $\operatorname{div}(w\nabla(V_1 - V_2))$ one of them is stable. Thus there is always some dynamics towards a segregated state driven by the differences in the attraction forces. Instead of pursuing a time-dependent analysis inspired by the above considerations, we restrict our search for segregated states to the analysis of stationary solutions, and leave the stability vs. instability analysis of segregated patterns to future work.

3. Steady states vs energy minimization

We now explore the relation between one dimensional steady states of (6), namely solutions (ρ_1, ρ_2) of

$$\begin{cases} 0 = \big(\epsilon\rho_1(\rho_1+\rho_2)_x - \rho_1 S_1' * \rho_1 - \rho_1 K' * \rho_2\big)_x \\ 0 = \big(\epsilon\rho_2(\rho_1+\rho_2)_x - \rho_2 S_2' * \rho_2 - \rho_2 K' * \rho_1\big)_x, \end{cases} \tag{13}$$

and the minimisers of the energy functional

$$\mathcal{F}[\rho_1,\rho_2] = \frac{\epsilon}{2}\int(\rho_1+\rho_2)^2 dx - \frac{1}{2}\int \rho_1 S_1 * \rho_1 dx - \frac{1}{2}\int \rho_2 S_2 * \rho_2 dx - \int \rho_1 K * \rho_2 dx.$$

Note, that $\rho_1, \rho_2 \in L^2(\mathbb{R})$ is a natural setting for the minimization of $\mathcal{F}$, since assumption (A3) ensures that the nonlocal terms in $\mathcal{F}$ are finite. More precisely, we will look for minimizers within the set

$$\mathcal{M} = \left\{(\rho_1,\rho_2) \in (L^2(\mathbb{R})\cap L^1(\mathbb{R}))^2 \;\middle|\; \rho_i \geq 0,\ \int_{\mathbb{R}} \rho_i\, dx = m_i\,,\ i = 1,2\right\}, \tag{14}$$

for given masses m_1, m_2. The functional setting for the weak solutions of (13) should involve some space derivative, since the cross-diffusion term in (13) is not of

type $\Delta F(\rho_1, \rho_2)$ with some nonlinear vector field F. Therefore one cannot integrate by parts twice in the distributional formulation. For a weak formulation involving spatial derivatives of ρ_1 and ρ_2 and allowing for discontinuities of each species at the same time, we choose $BV(\mathbb{R})$ as the functional setting for the steady states, whose sum $w = \rho_1 + \rho_2$ is Lipschitz continuous. Such a choice is partly supported by the results in [25], in which the preservation of the BV regularity is proven for a similar system.

Definition 3.1. *The pair* $(\rho_1, \rho_2) \in \mathcal{M} \cap BV(\mathbb{R})^2$ *is a* weak solution *to* (13) *if* $w = \rho_1 + \rho_2 \in \mathrm{Lip}(\mathbb{R})$, *and*

$$0 = \int \rho_1 \Big(\epsilon(\rho_1 + \rho_2)_x - S_1' * \rho_1 - K' * \rho_2 \Big) U_x \, dx,$$
$$0 = \int \rho_2 \Big(\epsilon(\rho_1 + \rho_2)_x - S_2' * \rho_2 - K' * \rho_1 \Big) V_x \, dx,$$

for arbitrary $U, V \in C_c^1(\mathbb{R})$.

Proposition 3.1. *Let* $(\rho_1, \rho_2) \in \mathcal{M} \cap BV(\mathbb{R})^2$ *be a minimizer of*

$$\mathcal{F}[\rho_1, \rho_2] = \frac{\epsilon}{2} \int_{\mathbb{R}} (\rho_1 + \rho_2)^2 dx - \frac{1}{2} \int_{\mathbb{R}} \rho_1 S_1 * \rho_2 dx - \frac{1}{2} \int_{\mathbb{R}} \rho_2 S_2 * \rho_2 dx - \int_{\mathbb{R}} \rho_1 K * \rho_2 dx,$$

such that $w = \rho_1 + \rho_2 \in \mathrm{Lip}(\mathbb{R})$. *Then* (ρ_1, ρ_2) *is a weak solutions of* (13) *according to Definition 3.1. Moreover, every weak solution* (ρ_1, ρ_2) *of* (13) *according to Definition 3.1 is a critical point of* $\mathcal{F}$.

Proof. Let $(\rho_1, \rho_2) \in \mathcal{M} \cap BV(\mathbb{R})^2$ be a minimizer of $\mathcal{F}$. We calculate the first Gateaux derivative of $\mathcal{F}$

$$\frac{d}{d\rho_1} \mathcal{F}[\rho_1, \rho_2](\mu) = \lim_{\delta \to 0} \frac{1}{\delta} \Big(\mathcal{F}[\rho_1 + \delta\mu, \rho_2] - \mathcal{F}[\rho_1, \rho_2] \Big),$$

along an arbitrary direction $\mu \in L^2(\mathbb{R})$, such that $(\rho_1 + \delta\mu, \rho_2) \in \mathcal{M}$. We have

$$\begin{aligned}
\mathcal{F}[\rho_1 + \delta\mu, \rho_2] - \mathcal{F}[\rho_1, \rho_2] &= \frac{\epsilon}{2} \int_{\mathbb{R}} (\rho_1 + \delta\mu + \rho_2)^2 dx - \frac{\epsilon}{2} \int_{\mathbb{R}} (\rho_1 + \rho_2)^2 dx \\
&\quad - \frac{1}{2} \int_{\mathbb{R}} (\rho_1 + \delta\mu) S_1 * (\rho_1 + \delta\mu) dx + \frac{1}{2} \int_{\mathbb{R}} \rho_1 S_1 * \rho_1 dx \\
&\quad - \int_{\mathbb{R}} (\rho_1 + \delta\mu) K * \rho_2 dx + \int_{\mathbb{R}} \rho_1 K * \rho_2 dx \\
&= \frac{\epsilon}{2} \int_{\mathbb{R}} (\delta\mu)^2 + 2\delta\mu(\rho_1 + \rho_2) dx - \frac{1}{2} \int_{\mathbb{R}} 2\delta\mu S_1 * \rho_1 dx \\
&\quad - \frac{1}{2} \int_{\mathbb{R}} \delta^2 \mu S_1 * \mu dx - \int_{\mathbb{R}} \delta\mu K * \rho_2 dx.
\end{aligned}$$

Dividing by δ and taking the limit $\delta \to 0$ results in

$$\frac{d}{d\rho_1} \mathcal{F}[\rho_1, \rho_2](\mu) = \int_{\mathbb{R}} \mu \Big(\epsilon(\rho_1 + \rho_2) - S_1 * \rho_1 - K * \rho_2 \Big) dx.$$

Let $U \in C_c^1(\mathbb{R})$ be an arbitrary vector field and let $\mu_\gamma = \partial_x(\rho_{1,\gamma} U_{\gamma,x})$, where the subscript γ denotes convolution with a standard (compactly supported) C^∞-mollifier. Then

$$\frac{d}{d\rho_1} \mathcal{F}[\rho_1, \rho_2] = - \int_{\mathbb{R}} \rho_{1,\gamma} \partial_x \Big(\epsilon(\rho_1 + \rho_2) - S_1 * \rho_1 - K * \rho_2 \Big) \cdot U_{\gamma,x} \, dx \,,$$

is well defined, since w is Lipschitz continuous. For $\gamma \searrow 0$ one obtains the first equation of (13) in weak form. The second equation follows similarly.

In order to prove the last assertion in the statement, we work by contradiction. Let ρ_1 and ρ_2 be as assumed and let a direction $\mu \in L^2$, such that $(\rho_1+\delta\mu, \rho_2) \in \mathcal{M}$ for which, e.g.

$$\frac{d}{d\rho_1}\mathcal{F}[\rho_1,\rho_2](\mu) = \int_{\mathbb{R}} \mu\Big(\epsilon(\rho_1+\rho_2) - S_1 * \rho_1 - K*\rho_2\Big)dx \neq 0.$$

By a density argument, one finds a function $U \in C_c^2(\mathbb{R})$ such that

$$\frac{d}{d\rho_1}\mathcal{F}[\rho_1,\rho_2]((\rho_1 U_x)_x) \neq 0, \text{ which implies}$$

$$0 \neq \int \rho_1 U_x \Big(\epsilon(\rho_1+\rho_2) - S_1*\rho_1 - K*\rho_2\Big)_x dx,$$

which is well defined since $\rho_1+\rho_2$ is Lipschitz. Thus (ρ_1,ρ_2) cannot be a steady state. □

Next, introduce a technical result that will be useful to investigate sufficient conditions for a steady state to be a local minimizer of $\mathcal{F}$ in the next subsection.

Lemma 3.1. *The second Gateaux derivatives for $\mathcal{F}$ on (ρ_1,ρ_2)*

$$\frac{d^2\mathcal{F}}{d\rho_i d\rho_j}(\rho_1,\rho_2)[\mu,\nu] = \lim_{\delta\to 0}\frac{1}{\delta}\left(\frac{d\mathcal{F}}{d\rho_i}(\rho_1+\delta_{1,j}\mu, \rho_2+\delta_{2,j}\nu) - \frac{d\mathcal{F}}{d\rho_i}(\rho_1,\rho_2)\right),$$ [1]

are given by

$$\begin{aligned} H[\mu,\nu] &= \begin{pmatrix} \frac{d^2\mathcal{F}}{d\rho_1^2}(\rho_1,\rho_2)[\mu,\nu] & \frac{d^2\mathcal{F}}{d\rho_1 d\rho_2}(\rho_1,\rho_2)[\mu,\nu] \\ \frac{d^2\mathcal{F}}{d\rho_1 d\rho_2}(\rho_1,\rho_2)[\mu,\nu] & \frac{d^2\mathcal{F}}{d\rho_2^2}(\rho_1,\rho_2)[\mu,\nu] \end{pmatrix} \\ &= \begin{pmatrix} \int_{\mathbb{R}}(\epsilon\mu^2 - \mu S_1*\mu)dx & \int_{\mathbb{R}}(\epsilon\mu\nu - \mu K*\nu)dx \\ \int_{\mathbb{R}}(\epsilon\mu\nu - \mu K*\nu)dx & \int_{\mathbb{R}}(\epsilon\nu^2 - \nu S_2*\nu)dx \end{pmatrix}, \end{aligned}$$

where $\mu,\nu \in L^2(\mathbb{R})^2$ are arbitrary and such that $(\rho_1+\delta\mu,\rho_2)$ and $(\rho_1,\rho_2+\delta\nu)$ are in $\mathcal{M}$.

Proof. Computing the upper left element of $H[\mu,\nu]$ gives

$$\begin{aligned} &\frac{d}{d\rho_1}\mathcal{F}[\rho_1+\delta\mu,\rho_2] - \frac{d}{d\rho_1}\mathcal{F}[\rho_1,\rho_2] \\ &\quad= \int_{\mathbb{R}} \mu\Big(\epsilon(\rho_1+\delta\mu+\rho_2) - S_1*(\rho_1+\delta\mu) - K*\rho_2\Big)dx \\ &\qquad - \int_{\mathbb{R}} \mu\Big(\epsilon(\rho_1+\rho_2) - S_1*\rho_1 - K*\rho_2\Big)dx \\ &\quad= \delta\int_{\mathbb{R}} \epsilon\mu^2 - \mu S_1*\mu\, dx. \end{aligned}$$

Therefore, we obtain

$$\frac{d^2}{d\rho_1^2}\mathcal{F}[\rho_1,\rho_2](\mu) = \int_{\mathbb{R}}(\epsilon\mu^2 - \mu S_1*\mu)dx.$$

The other entries of $H[\mu,\nu]$ can be computed similarly. □

[1] $\delta_{i,j}$ denotes the Kronecker symbol.

3.1. **Non-Existence of Steady States.** We now establish a simple necessary condition for the existence of non trivial steady states, based on the idea that $\mathcal{F}$ is strictly convex when the diffusion part is dominant. Thus zero is the unique global minimizer.

Lemma 3.2. *Assume that the Fourier transforms of the interaction kernels satisfy*

$$(i)\quad \epsilon > \max\{\hat{S}_1(\xi), \hat{S}_2(\xi)\},\ \ \textit{and}\ \ (ii)\quad (\epsilon - \hat{S}_1(\xi))(\epsilon - \hat{S}_2(\xi)) > (\epsilon - \hat{K}(\xi))^2, \tag{15}$$

for all $\xi \in \mathbb{R}$. Then there does not exist any global minimiser $(\rho_1, \rho_2) \in \mathcal{M}$.

Proof. Thanks to assumption (A3), there exists $C > 0$ such that

$$\mathcal{F}[\rho_1, \rho_2] \leq C\left(\|\rho_1\|_{L^2}^2 + \|\rho_2\|_{L^2}^2\right). \tag{16}$$

Applying the Fourier transform we get

$$\begin{aligned}
\mathcal{F}[\rho_1, \rho_2] &= \frac{1}{2}\int \left(\epsilon - \hat{S}_1(\xi)\right)\hat{\rho}_1^2(\xi)d\xi + \frac{1}{2}\int \left(\epsilon - \hat{S}_2(\xi)\right)\hat{\rho}_2^2(\xi)d\xi \\
&\quad + \int \left(\epsilon - \hat{K}(\xi)\right)\hat{\rho}_1(\xi)\hat{\rho}_2(\xi)d\xi \\
&= \int (\hat{\rho}_1(\xi), \hat{\rho}_2(\xi))^T \cdot A(\xi) \cdot (\hat{\rho}_1(\xi), \hat{\rho}_2(\xi))d\xi, \\
&\text{with } A(\xi) = \begin{pmatrix} \frac{1}{2}(\epsilon - \hat{S}_1(\xi)) & \frac{1}{2}(\epsilon - \hat{K}(\xi)) \\ \frac{1}{2}(\epsilon - \hat{K}(\xi)) & \frac{1}{2}(\epsilon - \hat{S}_2(\xi)) \end{pmatrix}.
\end{aligned}$$

Therefore, conditions (15) imply that the above quadratic form is positive definite, and hence $\mathcal{F}[\rho_1, \rho_2] > 0$ for all $\rho_1, \rho_2 \in L^2(\mathbb{R})$. Now, assume by contradiction that there exists a global minimiser (ρ_1, ρ_2) with the prescribed mass constraints. Then $\mathcal{F}[\rho_1, \rho_2] > 0$. We now rescale (ρ_1, ρ_2) by a parameter $\lambda > 0$ in such a way to preserve the total mass of both components, i.e.

$$\rho_{1,\lambda}(x) = \lambda^{-1}\rho_1(\lambda^{-1}x), \qquad \rho_{2,\lambda}(x) = \lambda^{-1}\rho_2(\lambda^{-1}x),$$

notice that $(\rho_{1,\lambda}, \rho_{2,\lambda}) \in \mathcal{M}$. We can use (16) as follows

$$0 \leq \mathcal{F}[\rho_{1,\lambda}, \rho_{2,\lambda}] \leq C\left(\|\rho_{1,\lambda}\|_{L^2}^2 + \|\rho_{2,\lambda}\|_{L^2}^2\right) = C\lambda^{-1}\left(\|\rho_1\|_{L^2}^2 + \|\rho_2\|_{L^2}^2\right).$$

The latter right-hand side converges to zero as $\lambda \to +\infty$. Therefore, for a large enough λ the value of the functional on $(\rho_{1,\lambda}, \rho_{2,\lambda})$ can be made smaller than $\mathcal{F}[\rho_1, \rho_2]$, thus contradicting the fact that (ρ_1, ρ_2) is a nontrivial global minimiser. □

We now establish a reasonable necessary condition for the existence of a non-trivial steady state.

Theorem 3.1. *Under conditions (15), there exists no nonzero stationary solution for (13).*

Proof. Let $(\rho_1, \rho_2) \in \mathcal{M}$. We use Lemma 3.1 to compute the second derivative $H[\mu, \nu]$ of $\mathcal{F}[\rho_1, \rho_2]$ in the direction (μ, ν). If

$$\int_{\mathbb{R}} (\epsilon\mu^2 - \mu S_1 * \mu)dx \int_{\mathbb{R}} (\epsilon\nu^2 - \nu S_2 * \nu)dx > \left(\int_{\mathbb{R}} (\epsilon\mu\nu - \mu K * \nu)dx\right)^2, \tag{17}$$

then $\det(H[\mu, \nu]) > 0$. Indeed, applying the Fourier transform on both sides of the inequality above, we obtain the equivalent condition

$$\int_{\mathbb{R}} (\epsilon - \hat{S}_1)\hat{\mu}^2 d\xi \int_{\mathbb{R}} (\epsilon - \hat{S}_2)\hat{\nu}^2 d\xi > \left(\int_{\mathbb{R}} \hat{\mu}(\epsilon\hat{\nu} - \hat{K}\hat{\nu})d\xi\right)^2.$$

Now from (15)(ii) we deduce

$$\left(\int(\epsilon-\hat{K}(\xi))\hat{\mu}(\xi)\hat{\nu}(\xi)d\xi\right)^2 < \left(\int|\epsilon-\hat{S}_1(\xi)|^{1/2}|\epsilon-\hat{S}_2(\xi)|^{1/2}|\hat{\mu}(\xi)||\hat{\nu}(\xi)|d\xi\right)^2$$
$$\leq \left(\int|\epsilon-\hat{S}_1(\xi)|\hat{\mu}^2(\xi)d\xi\right)\left(\int|\epsilon-\hat{S}_2(\xi)|\hat{\nu}^2(\xi)d\xi\right),$$

which implies $\det(H[\mu,\nu])>0$. From (15)(i) one deduces that the second derivative of $\mathcal{F}$ in (arbitrary) direction (μ,ν) is positive definite. Hence, $\mathcal{F}$ is strictly convex on $L^2(\mathbb{R})\times L^2(\mathbb{R})$. Now, assume $(\bar{\rho}_1,\bar{\rho}_2)$ is a non trivial steady state. Since $\mathcal{F}$ is strictly convex, then the critical point $(\bar{\rho}_1,\bar{\rho}_2)$ is also the unique global minimiser. This contradicts Lemma 3.2. □

Remark 3.1. Conditions (15) generalise the diffusion-dominated condition established in [22] for the one species case. In that case, the only interaction kernel S should satisfy $\|S\|_{L^1}>\epsilon$ in order to ensure the existence of a non trivial steady state, or equivalently the condition $\epsilon\geq\|S\|_{L^1}$ implies that no global minima exist with fixed positive mass. In the case with two species, the two conditions $\|S_i\|_{L^1}<\epsilon$ $i=1,2$, ensure that

$$\hat{S}_i(\xi)\leq\|\hat{S}_i\|_{L^\infty}\leq\|S_i\|_{L^1}<\epsilon\,,\qquad i=1,2,$$

and therefore the trace condition in (15) is trivially satisfied. This condition has a similar interpretation to the one in [22] for the one species case, i. e. the diffusion part is stronger than the (self) attraction parts, and so the spreading behaviour of particles dominates in the large time dynamics. However, (15)(ii) is a more specific feature of the two species case. In order to provide a heuristic interpretation of such condition, let us consider for simplicity Gaussian potentials of the form

$$S_1(x)=S_2(x)=S(x)=\frac{A}{\sigma\sqrt{\pi}}e^{-\frac{x^2}{2\sigma^2}}\,,\qquad K(x)=\frac{B}{\lambda\sqrt{\pi}}e^{-\frac{x^2}{2\lambda^2}},$$

for some positive constants A,B,σ,λ. Applying the Fourier transform we see that condition (i) is equivalent to

$$\epsilon>A.$$

Condition (15)(ii) is satisfied e.g. if

$$A<B<\epsilon\,,\qquad \lambda\ll\sigma, \tag{18}$$

i.e. if K has a larger mass than S and the variance of K is much smaller compared to that of S. Roughly speaking, this means that cross interaction should be relevant only at very small distances between particles of different species. This is not surprising. Consider for instance two separated patterns for ρ_1 and ρ_2, and assume that condition (15)(i) is satisfied. Then, the lack of cross diffusion (which only appears when both ρ_1 and ρ_2 are positive) pushes the two patterns towards 'spreading'. The only effect that could prevent the diffusive behaviour to dominate is the cross-interaction. Now, assume that (15)(ii) is satisfied, for instance in the form (18). Such a condition somehow ensures that the cross-interaction potential is not exerting any long-range confining effect on the particles to compensate the diffusive effect, because it only acts at small distances. When the two patterns touch each other, the cross-interaction potential will only interfere within the mixing area (still because of (15)(ii)), still not enough to produce a 'global confinement' and prevent the whole solution from decaying.

The above remark pinpoints an important fact about the minimization problem for $\mathcal{F}$: when the two species are separated and far from each other, the only coupling mechanism between them is the cross-interaction term ruled by the potential K. Hence, if we impose the conditions ensuring that the two species would feature a

nontrivial global minimum in absence of any coupling ($\epsilon < A$ in the above example), we immediately see that the functional can be diminished by decreasing the distance between the two species. Hence, if the two species are separated at the global minimum state, common sense would suggest that their supports are glued together. Proving that this is actually what happens is the main goal of this paper, and is tackled in the next sections from different viewpoints.

3.2. **Structure of Energy Minimizers.** The previous interpretation of the supports of ρ_1 and ρ_2 staying separated (possibly gluing together at one edge of their boundaries) is supported by the next result, which refers to a specific case, namely K being sufficiently smooth with unique maximum at zero, and $S_i = \sigma_i K$, $i = 1, 2$, with positive real numbers σ_i. Under these conditions, provided that $\sigma_1 + \sigma_2 > 2$, we can prove that the supports of the two components must necessarily have an empty intersection at local minimisers of $\mathcal{F}$, which corresponds to segregation. The next theorem is true also in multiple dimensions, but for simplicity we only state it in one space dimension.

Theorem 3.2. *Assume that $K \in C(\mathbb{R})$ with a unique maximum at zero and being strictly radially decreasing in a neighbourhood. Let $S_i = \sigma_i K$ for some $\sigma_1, \sigma_2 > 0$ with $\sigma_1 + \sigma_2 > 2$. Let $(\rho_1^\infty, \rho_2^\infty)$ be a local minimizer of $\mathcal{F}$ on $\mathcal{M}$, and*

$$\mathcal{S} := supp(\rho_1^\infty) \cap supp(\rho_2^\infty).$$

Then $\mathcal{S}$ has zero Lebesgue measure.

Proof. Assume that $\mathcal{S}$ has positive Lebesgue measure, then we can choose a subset $\mathcal{O}$ of positive measure and some constant $C > 0$ such that

$$\rho_i^\infty(x) \geq C, \qquad i = 1, 2, \text{ for almost every } x \in \mathcal{O}.$$

For a fixed $d > 0$ (e.g. half the diameter of $\mathcal{O}$) we can choose $z_1, z_2 \in \mathcal{O}$ with distance greater or equal than d and $0 < \gamma < d/4$ sufficiently small such that the balls $B_\gamma(z_i)$ of radius γ around z_1 and z_2 intersected with $\mathcal{O}$ have positive Lebesgue measure and are disjoint. Define

$$u_1(x) = \begin{cases} |B_\gamma(z_1) \cap \mathcal{O}|^{-1} & \text{if } x \in B_\gamma(z_1) \cap \mathcal{O} \\ -|B_\gamma(z_2) \cap \mathcal{O}|^{-1} & \text{if } x \in B_\gamma(z_2) \cap \mathcal{O} \\ 0 & \text{else} \end{cases}$$

and $u_2 = -u_1$. Since $u_1, u_2 \in L^2(\mathbb{R})$ have mean zero, for $\delta > 0$ sufficiently small, $\rho_i^\infty \pm \delta u_i$ is nonnegative and has the same mass as ρ_i^∞. Hence $(\rho_1^\infty \pm \delta u_1, \rho_2^\infty \pm \delta u_2)$ is admissible for the minimization of $\mathcal{F}$ with $\delta << 1$. It is hence straightforward to see that at a minimizer the first variation of $\mathcal{F}$ in direction (u_1, u_2) vanishes. With the homogeneity of $\mathcal{F}$, we find

$$\mathcal{F}[\rho_1^\infty + \delta u_1, \rho_2^\infty + \delta u_2] = \mathcal{F}[\rho_1^\infty, \rho_2^\infty] + \delta^2 \mathcal{F}[u_1, u_2].$$

Now, due to $u_1 + u_2 = 0$ we obtain

$$\begin{aligned} \mathcal{F}[u_1, u_2] &= -\int_{\mathbb{R}^d} \left(\frac{\sigma_1}{2} u_1 K * u_1 + \frac{\sigma_2}{2} u_2 K * u_2 + u_1 K * u_2 \right) \, dx \\ &= -\frac{1}{2}(\sigma_1 + \sigma_2 - 2) \int_{\mathbb{R}^d} u_1 K * u_1 \, dx. \end{aligned}$$

And we have

$$\begin{aligned}\int_{\mathbb{R}^d} u_1 K * u_1 \, dx =& |B_\gamma(z_1)\cap\mathcal{O}|^{-2}\int_{B_\gamma(z_1)\cap\mathcal{O}}\int_{B_\gamma(z_1)\cap\mathcal{O}} K(x-y)\,dx\,dy\\ &+|B_\gamma(z_2)\cap\mathcal{O}|^{-2}\int_{B_\gamma(z_2)\cap\mathcal{O}}\int_{B_\gamma(z_2)\cap\mathcal{O}} K(x-y)\,dx\,dy\\ &-\frac{2}{|B_\gamma(z_1)\cap\mathcal{O}|\,|B_\gamma(z_2)\cap\mathcal{O}|}\int_{B_\gamma(z_1)\cap\mathcal{O}}\int_{B_\gamma(z_2)\cap\mathcal{O}} K(x-y)\,dx\,dy\\ \geq& 2\left(\inf_{|z|<2\gamma} K(z) - \sup_{|z|>d-2\gamma} K(z)\right).\end{aligned}$$

For γ sufficiently small, the local strict radial decrease around zero implies that the infimum is larger than the supremum above (note that $2\gamma < d-2\gamma$) and thus, $\int_{\mathbb{R}^d} u_1 K * u_1 \, dx$ is positive. Hence, $\mathcal{F}[u_1,u_2] < 0$ and $(\rho_1^\infty,\rho_2^\infty)$ cannot be a minimizer of $\mathcal{F}$. □

In the case of one weaker interaction, e.g. $\sigma_2 < 1$ one can give an improved result on the support by following the proof of [31], where only the attractive part of the energy with an upper bound on the total density $\rho_1+\rho_2$ was considered. The idea is to write the energy as a functional of $\rho_1+\rho_2$ and ρ_2 and apply Riesz-Sobolev symmetrization to these functions in order to obtain states of lower energy. Since such a symmetrization leaves the L^2-norm of $\rho_1+\rho_2$ unchanged it can be applied directly to our case:

Theorem 3.3. *Assume that $K \in C(\mathbb{R})$ is strictly radially decreasing. Let $\sigma_1,\sigma_2 > 0$ with $\sigma_1+\sigma_2 > 2$ and $\sigma_2 < 1$. Let $(\rho_1^\infty,\rho_2^\infty)$ be a global minimiser with compact support of $\mathcal{F}$ on $\mathcal{M}$ with $m_1 = m_2$. Then $\rho_1^\infty+\rho_2^\infty$ is radially symmetric around its center of mass X_0 and there exist $b > a > 0$ such that*

$$supp(\rho_1^\infty) = [X_0-a, X_0+a], \quad supp(\rho_2^\infty) = [X_0-b, X_0-a]\cup[X_0+a, X_0+b]. \tag{19}$$

Of course the phase separation does not mean that the supports are at positive distance, due to the attractive cross interactions we expect that the supports are glued together. This is indeed true in general for stationary solutions. The proof is the same as the corresponding result for a single species in [22, Lemma 4.1], we hence only provide a sketch of the proof.

Proposition 3.2. *Let $\sigma_1,\sigma_2 > 0$, K be strictly decreasing with respect to the radial variable and $(\rho_1^\infty,\rho_2^\infty) \in \mathcal{M}\cap BV(\mathbb{R})^2$ be a weak solution of (13) according to Definition 3.1. Then $supp(\rho_1^\infty+\rho_2^\infty)$ is a connected interval.*

Sketch of the Proof. Let us just mention the main idea of the proof: Assume there exists an interval $[a,b]$ with $a < b$ not in the support of $\rho_1^\infty+\rho_2^\infty$, but such that the intersections of $(-\infty,a)$ and $(b,+\infty)$ with the support of $\rho_1^\infty+\rho_2^\infty$ are both nonempty. Then we can construct a smooth velocity V field equal to $+1$ on $(-\infty,a)$ and -1 on (b,∞). Consider then the infinitesimal change of the energy on the curve $(u_1(s,x),u_2(s,x))$ solving (locally) the Cauchy problem

$$\begin{cases}\partial_s u_i + \partial_x(u_i V) = 0\\ u_i(0,x) = \rho_i^\infty(x)\end{cases}$$

for $i = 1,2$. Since $\frac{d}{ds}\mathcal{F}[u_1(s),\rho_2^\infty]|_{s=0} = \frac{d}{ds}\mathcal{F}[\rho_1^\infty,u_2(s)]|_{s=0} = 0$ in view of the fact that $(\rho_1^\infty,\rho_2^\infty)$ is a stationary state according to Definition 3.1, due to the computation

$$0 = \frac{\epsilon}{2}\int_{-\infty}^{a}\partial_x(\rho_1^\infty+\rho_2^\infty)^2 dx - \frac{\epsilon}{2}\int_b^{+\infty}\partial_x(\rho_1^\infty+\rho_2^\infty)^2 dx,$$

with a few manipulations we obtain

$$\int_{-\infty}^{a} \rho_1^\infty S_1' * \rho_1^\infty dx + \int_{-\infty}^{a} \rho_1^\infty K' * \rho_2^\infty dx + \int_{-\infty}^{a} \rho_2^\infty S_2' * \rho_2^\infty dx$$
$$+ \int_{-\infty}^{a} \rho_2^\infty K' * \rho_1^\infty dx = \int_{b}^{+\infty} \rho_1^\infty S_1' * \rho_1^\infty dx$$
$$+ \int_{b}^{+\infty} \rho_1^\infty K' * \rho_2^\infty dx + \int_{b}^{+\infty} \rho_2^\infty S_2' * \rho_2^\infty dx + \int_{b}^{+\infty} \rho_2^\infty K' * \rho_1^\infty dx.$$

Now, since all the involved kernels S_1, S_2, K are even, with a few manipulations similar to those in [22, Lemma 4.1], we get that the support of $\rho_1^\infty + \rho_2^\infty$ must be empty in at least one between $(-\infty, a)$ and $(b, +\infty)$, which is a contradiction. This proves that $(\rho_1^\infty, \rho_2^\infty)$ cannot be a stationary state according to Definition 3.1. □

We can provide the same result for an arbitrary minimizer of the interaction energy with a similar proof, respectively the Lagrangian version:

Proposition 3.3. *Let $\sigma_1, \sigma_2 > 0$, K be strictly decreasing with respect to the radial variable and $(\rho_1^\infty, \rho_2^\infty) \in \mathcal{M}$ be a minimizer of $\mathcal{F}$. Then $supp(\rho_1^\infty + \rho_2^\infty)$ is a connected interval.*

Proof. Suppose that the difference between $supp(\rho_1^\infty + \rho_2^\infty)$ and its convex hull contains an interval $[a, b]$. Let $0 < d < \frac{b-a}{2}$ and define $(\tilde{\rho}_1, \tilde{\rho}_2) \in \mathcal{M}$ via

$$\tilde{\rho}_i(x) = \begin{cases} \rho_i^\infty(x - d) & \text{if } x < a + d \\ \rho_i^\infty(x + d) & \text{if } x > b - d \\ \rho_i^\infty(x) & \text{else .} \end{cases}$$

Then we have

$$\int_{\mathbb{R}} (\tilde{\rho}_1 + \tilde{\rho}_2)^2 \, dx = \int_{\mathbb{R}} (\rho_1^\infty + \rho_2^\infty)^2 \, dx,$$

while the attractive interaction terms in the energy are smaller for $(\tilde{\rho}_1, \tilde{\rho}_2)$. Hence, $\mathcal{F}[\tilde{\rho}_1, \tilde{\rho}_2] < \mathcal{F}[\rho_1^\infty, \rho_2^\infty]$, which is a contradiction. □

Despite the result of compact support for $\rho_1^\infty + \rho_2^\infty$, which holds for arbitrary positive interactions, we expect mixing in case of dominant cross-attraction, i.e. $\sigma_1 < 1$ and $\sigma_2 < 1$ similar to [31].

Note that Proposition 3.3 is not a special case of Proposition 3.2, since we cannot argue that $\rho_1^\infty + \rho_2^\infty$ are Lipschitz-continuous, hence we cannot apply the result Proposition 3.1. In order to do so we will need to verify the Lipschitz property, for which Proposition 3.3 appears to be crucial. In this direction we now provide a result closing the gap between energy minimizers and stationary solutions. At least for $\sigma_1 + \sigma_2 > 2$ and $\sigma_2 < 1$ the assumptions of the following theorem are satisfied:

Theorem 3.4. *Assume $K \in C^1(\mathbb{R})$ to be strictly radially decreasing. Let $\sigma_1, \sigma_2 > 0$ and let $(\rho_1^\infty, \rho_2^\infty)$ be a global minimiser with compact support of $\mathcal{F}$ on $\mathcal{M}$. Moreover, assume that for some $M > 1$ there exist $a_1 < a_2 < \ldots < a_M$ such that $supp(\rho_1^\infty + \rho_2^\infty) = \cup_{i=1}^{M-1}[a_i, a_{i+1}]$ and that for each $i = 1, \ldots, M-1$ we have*

$$\rho_k^\infty > 0, \rho_\ell^\infty = 0 \text{ on } (a_i, a_{i+1}), \tag{20}$$

with $\{k, \ell\} = \{1, 2\}$. Then $\rho_1^\infty + \rho_2^\infty$ is Lipschitz-continuous.

Proof. Let i be such that $k = 1$ and $\ell = 2$ (the opposite case is analogous). Moreover, let $\varphi \in C(\mathbb{R})$ be an arbitrary continuous function supported in a compact subset (a_i, a_{i+1}). Then by standard arguments on the first variation we see that $\mathcal{F}[\varphi, \rho_2^\infty] = 0$, which implies

$$\rho_1^\infty + \rho_2^\infty = \rho_1^\infty = \frac{1}{\epsilon} \int_R K(x - y)[\sigma_1 \rho_1^\infty(y) + \rho_2^\infty(y)] \, dy,$$

on (a_i, a_{i+1}). The continuous differentiability of K implies that $\rho_1^\infty + \rho_2^\infty$ is Lipschitz continuous (with modulus independent of i) on (a_i, a_{i+1}). Hence, it suffices to show that $\rho_1^\infty + \rho_2^\infty$ is continuous on the finite set of points $\{a_i\}$ in order to obtain Lipschitz continuity on $\mathbb{R}$. Due to the uniform Lipschitz continuity on the subintervals, the left- and right-sided limits exist for each subinterval.

First, let $\rho_k^\infty > 0$ in (a_1, a_2) and assume $\lim_{x\to a_1^+} \rho_k^\infty(x) > 0$. For $\delta_1, \delta_2 > 0$ being sufficiently small define

$$\tilde\rho_k(x) = \begin{cases} \rho_k^\infty - \delta_1 & \text{for } x \in (a_1, a_1 + \delta_2) \\ \rho_k^\infty + \delta_1 & \text{for } x \in (a_1, a_1 - \delta_2) \\ \rho_k^\infty(x) & \text{else} \end{cases} \qquad \text{and} \qquad \tilde\rho_\ell = \rho_\ell^\infty.$$

Then it is straight-forward to show that

$$\mathcal{F}[\tilde\rho_1, \tilde\rho_2] = \mathcal{F}[\rho_1^\infty, \rho_2^\infty] - \epsilon\delta_1\delta_2 \lim_{x\to a_1^+} \rho_k^\infty(x) + \mathcal{O}(\delta_1^2\delta_2 + \delta_1\delta_2^2),$$

which strictly smaller than $\mathcal{F}[\rho_1^\infty, \rho_2^\infty]$ for δ_1, δ_2 sufficiently small and hence a contradiction to $(\rho_1^\infty, \rho_2^\infty)$ being a minimiser. Thus,

$$\lim_{x\to a_1^+} (\rho_1^\infty(x) + \rho_2^\infty(x)) = \lim_{x\to a_1^+} \rho_k^\infty(x) = 0.$$

A completely analogous proof shows continuity at $x = a_M$. Finally consider the limit at $x = a_i$, $1 < i < M$, if $\lim_{x\to a_i^+} \rho_k^\infty(x) = \lim_{x\to a_i^+} \rho_\ell^\infty(x)$. Define a similar perturbation in a neighbourhood $(a_i - \delta_2, a_i + \delta_2)$ increasing the smaller and decreasing the larger density. With the same kind of expansion as at a_1 one can obtain a state with lower energy, which is a contradiction. □

4. Existence of segregated states via Krein-Rutman theorem

Hinted by the results in the previous section, we now provide reasonable sufficient conditions for the existence of segregated stationary states for system (13). We remark that the system (13) is translation invariant. Therefore, we shall fix the center of mass to zero for simplicity.

Definition 4.1. *We call a stationary solution (ρ_1, ρ_2) to (13) according to Definition 3.1 a* symmetric segregated steady state *if there exist $L_1, L_2 > 0$ such that*

$$supp(\rho_1) = [-L_1, L_1] =: I_1 \quad \text{and} \quad supp(\rho_2) = [-L_2, -L_1] \cup [L_1, L_2] =: I_2\,,$$

where ρ_1, ρ_2 are even functions, and C^1-regular inside their respective supports, such that $w = \rho_1 + \rho_2$ is monotone decreasing on $[0, L_2]$ with $w(L_2) = 0$.

Recall that Definition 3.1 prescribes $w(x) := \rho_1(x) + \rho_2(x)$ being Lipschitz continuous on $\mathbb{R}$.

Our strategy to prove existence of segregated steady states is a non-trivial extension of the strategy proposed in [22], and adapted in [23]. The main idea is to fix L_1 and L_2 and then look at the stationary equations for ρ_1 and ρ_2 as eigenvalue conditions for a suitable integral operator. The existence of the eigenvectors will be proven by using the strong version of the Krein-Rutman theorem.

Theorem 4.1 (Krein-Rutman)**.** *Let X be a Banach space, $K \subset X$ be a solid cone, such that $\lambda K \subset K$ for all $\lambda \geq 0$ and K has a nonempty interior K^o. Let T be a compact linear operator on X, which is strongly positive with respect to K, i.e. $T[u] \in K^o$ if $u \in K \setminus \{0\}$. Then,*

(i) *the spectral radius $r(T)$ is strictly positive and $r(T)$ is a simple eigenvalue with an eigenvector $v \in K^o$. There is no other eigenvalue with a corresponding eigenvector $v \in K$.*
(ii) *$|\lambda| < r(T)$ for all other eigenvalues $\lambda \neq r(T)$.*

We shall work with the following class of kernels:

$$S_1, S_2, K \text{ are symmetric and strictly decreasing on } [0, +\infty), \tag{21}$$

which, together with our nonnegativity assumption (A3) imply in particular that all the kernels are supported on $\mathbb{R}$.
Moreover, having fixed $L_1 < L_2$, we assume for all $x \in (0, L_1)$ that

$$S_1(x - L_1) - S_1(x + L_1) > K(x - L_1) - K(x + L_1), \tag{22}$$

$$\text{and } S_2(x - L_1) - S_2(x + L_1) < K(x - L_1) - K(x + L_1), \tag{23}$$

for all $x \in (L_1, L_2)$. Assumptions (22) and (23) are met for instance in the significant case $S_i = \sigma_i K$, $\sigma_1 > 1 > \sigma_2$. We assume further that

$$S_1'(L_1) < K'(L_1). \tag{24}$$

Let ρ_1 and ρ_2 be symmetric steady states, then we can rephrase (13) as

$$\begin{cases} \epsilon(\rho_1 + \rho_2) - S_1 * \rho_1 - K * \rho_2 = C_1 & \text{for } x \in I_1 \\ \epsilon(\rho_1 + \rho_2) - S_2 * \rho_2 - K * \rho_1 = C_2 & \text{for } x \in I_2, \end{cases} \tag{25}$$

for suitable constants $C_1, C_2 > 0$. Assuming segregation, we rewrite (25) as

$$\begin{cases} x \in I_1: \quad \epsilon\rho_1(x) = \displaystyle\int_{\mathbb{R}} S_1(x-y)\rho_1(y)dy + \int_{\mathbb{R}} K(x-y)\rho_2(y)dy + C_1 \\ \\ x \in I_2: \quad \epsilon\rho_2(x) = \displaystyle\int_{\mathbb{R}} S_2(x-y)\rho_2(y)dy + \int_{\mathbb{R}} K(x-y)\rho_1(y)dy + C_2 \end{cases}. \tag{26}$$

Let $\bar{w} = \rho_1(L_1) = \rho_2(L_1)$. Let $p(x) = -\rho_1'(x)$ be restricted to the interior of I_1 and $q(x) = -\rho_2'(x)$ to the interior of I_2. By differentiating (26) we obtain

$$\begin{cases} x \in I_1: \epsilon p(x) = \displaystyle\int_{I_1} S_1(x-y)p(y)dy + \int_{I_2} K(x-y)q(y)dy + \bar{w}A_1(x) \\ \\ x \in I_2: \epsilon q(x) = \displaystyle\int_{I_2} S_2(x-y)q(y)dy + \int_{I_1} K(x-y)p(y)dy + \bar{w}A_2(x) \end{cases}, \tag{27}$$

where

$$A_1(x) = S_1(x - L_1) - S_1(x + L_1) + K(x + L_1) - K(x - L_1), \tag{28}$$
$$A_2(x) = K(x - L_1) - K(x + L_1) + S_2(x + L_1) - S_2(x - L_1). \tag{29}$$

A symmetrization yields for $x > 0$, that

$$\begin{aligned} \epsilon p(x) &= \int_0^{L_1} \big(S_1(x-y) - S_1(x+y)\big)p(y)dy \\ &\quad + \int_{L_1}^{L_2} \big(K(x-y) - K(x+y)\big)q(y)dy + \bar{w}A_1(x), \\ \epsilon q(x) &= \int_{L_1}^{L_2} \big(S_2(x-y) - S_2(x+y)\big)q(y)dy \\ &\quad + \int_0^{L_1} \big(K(x-y) - K(x+y)\big)p(y)dy + \bar{w}A_2(x). \end{aligned}$$

To simplify notation, define the non-negative functions, for $x, y > 0$,

$$\bar{G}(x, y) = G(x - y) - G(x + y), \text{ for } G = S_1, S_2, K.$$

Then we can rewrite system (27) as

$$\epsilon p(x) = \int_0^{L_1} \bar{S}_1(x,y)p(y)dy + \int_{L_1}^{L_2} \bar{K}(x,y)q(y)dy + \bar{w}A_1(x),$$
$$\epsilon q(x) = \int_{L_1}^{L_2} \bar{S}_2(x,y)q(y)dy + \int_0^{L_1} \bar{K}(x,y)p(y)dy + \bar{w}A_2(x).$$

The result of this section is stated in the following Proposition.

Proposition 4.1. *Let assumptions (A1), (A3), (21), (22), (23), and (24) be satisfied. Let $0 < L_1 < L_2$ be fixed. Then, there exists a unique (up to mass normalisation) symmetric segregated steady state (ρ_1, ρ_2) to (13) with $\epsilon = \epsilon(L_1, L_2) > 0$.*

Proof. Consider the Banach space

$$X_{L_1,L_2} = \left\{ P = (p, q; w) \in C^1([0, L_1]) \times C^1([L_1, L_2]) \times \mathbb{R}, \mid p(0) = 0 \right\},$$

equipped with the $W^{1,\infty}$-norm for the first two components and with the standard one-dimensional Euclidean norm for the third component. We use the notation $P = (p, q; w)$ for all $P \in X_{L_1,L_2}$, with $p \in C^1([0, L_1])$, $q \in C^1([L_1, L_2])$, and $w \geq 0$. For a given $P \in X_{L_1,L_2}$ define $T_{L_1,L_2}[P] \in X_{L_1,L_2}$ as

$$T_{L_1,L_2}[P] = (f, g; w') \in C^1([0, L_1]) \times C^1([L_1, L_2]) \times \mathbb{R}, \text{ with}$$

$$f(x) = \int_0^{L_1} \bar{S}_1(x,y)p(y)dy + \int_{L_1}^{L_2} \bar{K}(x,y)q(y)dy + wA_1(x), \qquad x \in [0, L_1]$$
$$g(x) = \int_{L_1}^{L_2} \bar{S}_2(x,y)q(y)dy + \int_0^{L_1} \bar{K}(x,y)p(y)dy + wA_2(x), \qquad x \in [L_1, L_2]$$
$$w' = \int_{L_1}^{L_2} \left(\int_{L_1}^{L_2} \bar{S}_2(x,y)q(y)dy + \int_0^{L_1} \bar{K}(x,y)p(y)dy + wA_2(x) \right) dx.$$

The operator T_{L_1,L_2} is compact on the Banach space X_{L_1,L_2} as all the involved kernels are C^2 on compact intervals, and hence by Arzelá's theorem the image of the unit ball in X_{L_1,L_2} is pre-compact. It is easy to show that the set

$$K_{L_1,L_2} = \{ P = (p, q, w) \in X_{L_1,L_2} \mid p \geq 0, q \geq 0, w \geq 0 \},$$

is a solid cone in X_{L_1,L_2} and that

$$\begin{aligned} H_{L_1,L_2} \quad = \quad & \{ P = (p, q, w) \in K_{L_1,L_2} \mid p'(0) > 0,\ p(x) > 0\ \forall x \in]0, L_1], \\ & q(x) > 0\ \forall x \in [L_1, L_2],\ w > 0 \} \subset K^o_{L_1,L_2}, \end{aligned}$$

where $K^o_{L_1,L_2}$ denotes the interior of K_{L_1,L_2}. We show that T is strongly positive on the solid cone. First observe that, due to (22) and (23), all components of $T_{L_1,L_2}[(p, q; w)](x)$ are nonnegative if $(p, q; w) \in K_{L_1,L_2}$ with $(p, q; w) \neq 0$. Moreover, setting $f(x)$ as the first component of $T_{L_1,L_2}[(p, q; w)](x)$ as above, we get

$$\begin{aligned} \frac{d}{dx} f(x) \mid_{x=0} &= \int_0^{L_1} \bar{S}_{1,x}(0,y)p(y)dy + \int_{L_1}^{L_2} \bar{K}_x(0,y)q(y)dy + wA_1'(0) \\ &= \int_0^{L_1} \big(S_1'(-y) - S_1'(y)\big)p(y)dy + \int_{L_1}^{L_2} \big(K'(-y) - K'(y)\big)q(y)dy \\ &\quad + w\Big(S_1'(-L_1) - S_1'(L_1) + K'(L_1) - K'(-L_1)\Big) \\ &= -2\int_0^{L_1} S_1'(y)p(y)dy - 2\int_{L_1}^{L_2} K'(y)q(y)dy \ - 2wS_1'(L_1) + 2wK'(L_1) > 0, \end{aligned}$$

since $(p,q;w) \neq 0$. Hence, $T_{L_1,L_2}[(p,q,w)]$ belongs to H_{L_1,L_2} and therefore the operator is strongly positive on the cone K_{L_1,L_2}. The strong version of the Krein-Rutman theorem (Theorem 4.1) then gives the existence of an eigenvalue $\epsilon = \epsilon(L_1,L_2)$ such that

$$T_{L_1,L_2}[(p,q,\bar{w})] = \epsilon \cdot (p,q,w),$$

with eigenspace generated by a vector $(q,p,w) \in K^o_{L_1,L_2}$. Now, let $\rho_2(x) = \int_x^{L_2} q(x)dx$. The second and third equation in the eigenvalue condition imply $\bar{w} = \rho_2(L_1)$. Let $\rho_1(x) = \bar{w} + \int_x^{L_1} p(x)dx$. Clearly, $\rho_1(L_1) = \rho_2(L_1)$. Then, a computation similar to the one that led to the definition of the operator T_{L_1,L_2} after differentiating (26) implies that ρ_1 and ρ_2 are symmetric segregated steady states to (13). □

Remark 4.1. The above result shows that a segregated steady state (ρ_1,ρ_2) exists for arbitrary positive constants L_1 and L_2 such that $\mathrm{supp}(\rho_1) = [-L_1,L_1]$ and $\mathrm{supp}(\rho_2) = [-L_2,-L_1]\cup[L_1,L_2]$ for some diffusion constant $\epsilon > 0$ which depends on L_1 and L_2. Clearly, we obtain a one-parameter family of solutions upon multiplication of (ρ_1,ρ_2) by an arbitrary positive constant. Such a constant is uniquely determined by the total mass $m = m_1 + m_2$ with $m_1 = \int_{\mathbb{R}} \rho_1(x)dx$ and $m_2 = \int_{\mathbb{R}} \rho_2(x)dx$. The two values m_1 and m_2 are not determined explicitly here. We expect the value of L_1 to be determined uniquely once L_2 and the two masses m_1 and m_2 are fixed. Most importantly, this approach does not provide an explicit information on the value of the diffusion constant ϵ. A similar situation occurs in the problem studied in [23] for the one species equation with general power-law diffusion. Reasoning as in [22], we expect that the diffusion constant ϵ obeys a monotone increasing law of the form $\epsilon = \bar{\epsilon}(L_2)$. If such a conjecture is true, then for any given ϵ in the range of the map $\bar{\epsilon}$ and for any given $m_1, m_2 > 0$ there exist unique $L_1, L_2 > 0$ and a unique (up to x-translations) segregated state (ρ_1,ρ_2) with $m_1 = \int_{\mathbb{R}} \rho_1(x)dx$, $m_2 = \int_{\mathbb{R}} \rho_2(x)dx$, $\mathrm{supp}(\rho_1) = [-L_1,L_1]$, and $\mathrm{supp}(\rho_2) = [-L_2,-L_1]\cup[L_1,L_2]$. Such a conjecture will be addressed in full generality in a future study. In the next section we are able to prove that *such a statement is true for a small enough diffusion coefficient.*

5. Existence and uniqueness of steady states for small diffusion

Here we prove existence and uniqueness of a symmetric segregated steady state for fixed masses, and a fixed *small* diffusion coefficient. We refer to Definition 4.1 to denote segregated states. Similar to [21], we formulate the problem in the pseudo-inverse formalism and then use an implicit function theorem argument. Throughout this section we shall require

$$\min\{S_1''(0),\, S_2''(0),\, K''(0)\} > 0. \tag{30}$$

The main result of this section reads as follow.

Theorem 5.1 (Existence of segregated steady states for small diffusion)**.** *Assume that the interaction kernels fulfil the additional assumption* (30)*. Then, there exists a constant ϵ_0 such that for all $\epsilon \in (0,\epsilon_0)$ the stationary equation* (13) *admits a unique solution in the sense of Definition 4.1 with fixed masses $m_1 = \int \rho_1 dx$ and $m_2 = \int \rho_2 dx$.*

5.1. **Formulation in the pseudo-inverse variable.** Assume that (ρ_1,ρ_2) is a segregated state with the structure as in Definition 4.1, and

$$\int \rho_1(x)dx = z_1\,, \qquad \int \rho_2(x)dx = 1 - z_1.$$

Hence, $w = \rho_1 + \rho_2$ has unit mass and is supported on $[-L_2, L_2]$. Let

$$F(x) = \int_{-\infty}^{x} w(y)dy.$$

Let $u : [0,1) \to \mathbb{R}$ be the pseudo-inverse of F

$$u(z) = \inf\{x \in \mathbb{R} : \ F(x) \geq z\}.$$

Set $u_i(z) = u(z)\mathbf{1}_{J_i}(z)$, $i = 1, 2$, with

$$\mathrm{supp}(u_1) = \left[\frac{1}{2} - \frac{z_1}{2}, \frac{1}{2} + \frac{z_1}{2}\right] =: J_1,$$

$$\mathrm{supp}(u_2) = \left[0, \frac{1}{2} - \frac{z_1}{2}\right] \cup \left[\frac{1}{2} + \frac{z_1}{2}, 1\right] =: J_2.$$

Then (13) can be rewritten as

$$\begin{cases} \epsilon \partial_x w(u_1(z)) = \int_{J_1} S_1'\big(u_1(z) - u_1(\zeta)\big)d\zeta + \int_{J_2} K'\big(u_1(z) - u_2(\zeta)\big)d\zeta\,, & z \in J_1\,, \\ \epsilon \partial_x w(u_2(z)) = \int_{J_2} S_2'\big(u_2(z) - u_2(\zeta)\big)d\zeta + \int_{J_1} K'\big(u_2(z) - u_1(\zeta)\big)d\zeta\,, & z \in J_2\,. \end{cases}$$

Recall that ρ_1 and ρ_2 are symmetric, which implies that the pseudo-inverse u satisfies $u(1-z) = -u(z)$. Moreover, since w is strictly positive on $(-L_2, L_2)$, u is strictly increasing on $(0,1)$ and Lipschitz continuous on the compact subintervals of $(0,1)$. Due to w being zero at $z = 0$ and $z = 1$, u is expected to have an infinite slope at the boundary.

Since $w = \rho_1 + \rho_2$ and since ρ_1, ρ_2 have disjoint supports, we have

$$\begin{cases} \dfrac{\epsilon}{2}\partial_z(\partial_z u_1(z))^{-2} = \displaystyle\int_{J_1} S_1'\big(u_1(z) - u_1(\zeta)\big)d\zeta \\ \qquad\qquad\qquad + \displaystyle\int_{J_2} K'\big(u_1(z) - u_2(\zeta)\big)d\zeta, \quad z \in J_1, \\ \dfrac{\epsilon}{2}\partial_z(\partial_z u_2(z))^{-2} = \displaystyle\int_{J_2} S_2'\big(u_2(z) - u_2(\zeta)\big)d\zeta \\ \qquad\qquad\qquad + \displaystyle\int_{J_1} K'\big(u_2(z) - u_1(\zeta)\big)d\zeta, \quad z \in J_2. \end{cases} \tag{31}$$

The idea is to solve (31) for small ϵ, hinted by the fact that the case $\epsilon = 0$ has the unique solution $u_1 \equiv u_2 \equiv 0$, which corresponds to ρ_1 and ρ_2 being two Dirac's deltas with masses z_1 and $1 - z_1$ respectively. Similar to [20], we expect that the support of w is small for small ϵ. This suggests the linearisation $u_i = \delta v_i$, $i = 1, 2$, with v_1, v_2 being odd functions defined on J_1, J_2 respectively. A simple scaling argument suggests that $\delta = \epsilon^{\frac{1}{3}}$, and therefore

$$\begin{cases} z \in J_1 & \frac{\delta}{2}\partial_z(\partial_z v_1)^{-2} = \int_{J_1} S_1'\big(\delta(v_1 - v_1(\zeta))\big) + \int_{J_2} K'\big(\delta(v_1 - v_2(\zeta))\big), \\ z \in J_2 & \frac{\delta}{2}\partial_z(\partial_z v_2)^{-2} = \int_{J_2} S_2'\big(\delta(v_2 - v_2(\zeta))\big) + \int_{J_1} K'\big(\delta(v_2 - v_1(\zeta))\big). \end{cases}$$

Multiplying the first equation by $\delta\partial_z v_1$ and the second one by $\delta\partial_z v_2$, we get

$$\delta^2\partial_z(\partial_z v_1)^{-1} = \partial_z \int_{J_1} S_1\big(\delta(v_1(z) - v_1(\zeta))\big) + \partial_z \int_{J_2} K\big(\delta(v_1(z) - v_2(\zeta))\big),$$

$$\delta^2\partial_z(\partial_z v_2)^{-1} = \partial_z \int_{J_2} S_2\big(\delta(v_2(z) - v_2(\zeta))\big) + \partial_z \int_{J_1} K\big(\delta(v_2(z) - v_1(\zeta))\big).$$

Taking the primitives w.r.t. z, we obtain for $z \in J_1$, respectively $z \in J_2$ that

$$\frac{\delta^2}{\partial_z v_1} = \int_{J_1} S_1\big(\delta(v_1(z) - v_1(\zeta))\big)d\zeta + \int_{J_2} K\big(\delta(v_1(z) - v_2(\zeta))\big)d\zeta + \alpha_1, \tag{32}$$

$$\frac{\delta^2}{\partial_z v_2} = \int_{J_2} S_2\big(\delta(v_2(z) - v_2(\zeta))\big)d\zeta + \int_{J_1} K\big(\delta(v_2(z) - v_1(\zeta))\big)d\zeta + \alpha_2, \tag{33}$$

with integration constants α_1, α_2, which are obtained by substituting $z = 1$ into (33)

$$\alpha_2 = -\int_{J_2} S_2\big(\delta(v_2(1) - v_2(\zeta))\big)d\zeta - \int_{J_1} K\big(\delta(v_2(1) - v_1(\zeta))\big)d\zeta, \tag{34}$$

and imposing the continuity condition for $(\partial_z u)^{-1}$ in $\tilde{z} = (1 + z_1)/2$

$$\begin{aligned}
\alpha_1 \;=\; & \int_{J_2} S_2\big(\delta\big(v_2((1+z_1)/2) - v_2(\zeta)\big)\big) - S_2\big(\delta(v_2(1) - v_2(\zeta))\big)\, d\zeta \\
& + \int_{J_1} K\big(\delta\big(v_2((1+z_1)/2) - v_1(\zeta)\big)\big) - K\big(\delta(v_2(1) - v_1(\zeta))\big)\, d\zeta \\
& - \int_{J_1} S_1\big(\delta\big(v_1((1+z_1)/2) - v_1(\zeta)\big)\big)d\zeta \\
& - \int_{J_2} K\big(\delta\big(v_1((1+z_1)/2) - v_2(\zeta)\big)\big)d\zeta.
\end{aligned}$$

The symmetry requirements imply

$$\int_{J_1} v_1(\zeta)d\zeta = \int_{J_2} v_2(\zeta)d\zeta = 0.$$

Now we multiply (32) by $\delta\partial_z v_1$, (33) by $\delta\partial_z v_2$, integrate w.r.t. z and introduce G_i, H such that $G_i' = S_i$, $i = 1, 2$, and $H' = K$. Here G_1, G_2, H can be chosen odd with $G_1(0) = G_2(0) = H(0)$. The integration constants can be recovered by prescribing

$$v_1(1/2) = 0\,,\; v_2(1) = \lambda\,,\; v_1(\tilde{z}) = v_2(\tilde{z}) = \mu\,.$$

After some manipulations we obtain

$$\begin{aligned}
z - \frac{1}{2} \;=\; & \delta^{-3}\Big[\int_{J_1} G_1\big(\delta(v_1(z) - v_1(\zeta))\big) - \delta v_1(z) S_1\big(\delta(\mu - v_1(\zeta))\big)\, d\zeta \\
& \qquad + \int_{J_2} H\big(\delta(v_1(z) - v_2(\zeta))\big) - \delta v_1(z) K\big(\delta(\mu - v_2(\zeta))\big)\, d\zeta\Big] \\
& + \delta^{-2} v_1(z)\Big[\int_{J_2} S_2\big(\delta(\mu - v_2(\zeta))\big) - S_2\big(\delta(\lambda - v_2(\zeta))\big)\, d\zeta \\
& \qquad + \int_{J_1} K\big(\delta(\mu - v_1(\zeta))\big) - K\big(\delta(\lambda - v_1(\zeta))\big)\, d\zeta\Big]\,,\; z \in [1/2, \tilde{z}]\,,
\end{aligned} \tag{35}$$

as well as

$$\begin{aligned}
z - \tilde{z} \;=\; & \delta^{-3}\Big[\int_{J_2} G_2\big(\delta(v_2(z) - v_2(\zeta))\big) - \delta v_2(z) S_2\big(\delta(\lambda - v_2(\zeta))\big)\, d\zeta \\
& \qquad + \int_{J_1} H\big(\delta(v_2(z) - v_1(\zeta))\big) - \delta v_2(z) K\big(\delta(\lambda - v_1(\zeta))\big)\, d\zeta\Big] \\
& - \delta^{-3}\Big[\int_{J_2} G_2\big(\delta(\mu - v_2(\zeta))\big) - \delta\mu S_2\big(\delta(\lambda - v_2(\zeta))\big)\, d\zeta \\
& \qquad + \int_{J_1} H\big(\delta(\mu - v_1(\zeta))\big) - \delta\mu K\big(\delta(\lambda - v_1(\zeta))\big)\, d\zeta\Big]\,, \quad z \in [\tilde{z}, 1]\,.
\end{aligned} \tag{36}$$

A solution (v_1, v_2) to (35), (36) should be extended to $[0, 1/2]$ to obtain an odd profile on the whole interval $[0, 1]$.

Remark 5.1. The functional system (35)-(36) has been obtained manipulating the original system (32)-(33) through integrations w.r.t. the independent variable z and dividing by δ^2. Therefore, seeking for a solution to (35)-(36) for $\delta = 0$ somehow involves the computation of the *first order term* in the expansion of the r.h.s. of (32)-(33) with respect to δ^2. This computation will be performed in the

next subsection. In some sense, in terms of the linearisation $u_i = \delta v_i$, we are computing the first order term w.r.t. δ of u_i near $\delta = 0$.

5.2. **A functional equation.** System (35), (36) can be viewed as a functional equation in the following form. Introduce the Banach space

$$\Omega = \Big\{ (v_1, \mu, v_2, \lambda) \in L^\infty([1/2, \tilde z]) \times \mathbb{R} \times L^\infty([\tilde z, 1]) \times \mathbb{R} \,:\, v_1 \text{ is right continuous at } 1/2 \text{ and left continuous at } \tilde z\,, \; v_2 \text{ is right continuous at } \tilde z \text{ and left continuous at } 1, \; v_1(1/2) = 0\,, \quad v_1(\tilde z) = v_2(\tilde z) = \mu\,, \quad v_2(1) = \lambda \Big\}.$$

And consider the standard norm on Ω

$$|||(v_1, \mu, v_2, \lambda)||| = \|v_1\|_{L^\infty} + \|v_2\|_{L^\infty} + |\mu| + |\lambda|.$$

For $\alpha > 0$, consider the norm

$$|||(v_1, \mu, v_2, \lambda)|||_\alpha := |||(v_1, \mu, v_2, \lambda)\,||| + \sup_{z \in [\tilde z, 1]} \frac{|\lambda - v_2(z)|}{(1-z)^\alpha},$$

and set $\Omega_\alpha := \{(v_1, \mu, v_2, \lambda) \in \Omega \,:\, |||(v_1, \mu, v_2, \lambda)|||_\alpha < +\infty\}$.
We now define the convex subset $U_\alpha \subset \Omega$

$$U_\alpha := \{(v_1, \mu, v_2, \lambda) \in \Omega_\alpha \,:\, v_1 \text{ and } v_2 \text{ are increasing}\}\,.$$

In order to formulate (35), (36) only on $[1/2, 1]$, we again use the *symmetrised* function

$$\bar G(x; y) := G(x - y) + G(x + y)\,, \quad x, y > 0\,,$$

for given $G : \mathbb{R} \to \mathbb{R}$. Moreover, we use the notation $\tilde J_i = J_i \cap [1/2, 1]\,, \quad i = 1, 2.$
Let $(v_1, \mu, v_2, \lambda) \in \Omega$, and $\delta > 0$. Define

$$\begin{aligned} F_1[(v_1, \mu, v_2, \lambda); \delta](z) &:= \frac{1}{2} - z \\ &+\delta^{-3} \left[\int_{\tilde J_1} \bar G_1(\delta v_1(z); \delta v_1(\zeta)) - \delta v_1(z) \bar S_1(\delta\mu; \delta v_1(\zeta))\, d\zeta \right. \\ &\qquad \left. + \int_{\tilde J_2} \bar H(\delta v_1(z); \delta v_2(\zeta)) - \delta v_1(z) \bar K(\delta\mu; \delta v_2(\zeta))\, d\zeta \right] \\ &+\delta^{-2} v_1(z) \left[\int_{\tilde J_2} \bar S_2(\delta\mu; \delta v_2(\zeta)) - \bar S_2(\delta\lambda; \delta v_2(\zeta))\, d\zeta \right. \\ &\qquad \left. + \int_{\tilde J_1} \bar K(\delta\mu; \delta v_1(\zeta)) - \bar K(\delta\lambda; \delta v_1(\zeta))\, d\zeta \right], \quad z \in \tilde J_1\,, \end{aligned} \tag{37}$$

$$\begin{aligned} F_2[(v_1, \mu, v_2, \lambda); \delta](z) &:= \tilde z - z \\ &+\delta^{-3} \left[\int_{\tilde J_2} \bar G_2(\delta v_2(z); \delta v_2(\zeta)) - \delta v_2(z) \bar S_2(\delta\lambda; \delta v_2(\zeta))\, d\zeta \right. \\ &\qquad \left. + \int_{\tilde J_1} \bar H(\delta v_2(z); \delta v_1(\zeta)) - \delta v_2(z) \bar K(\delta\lambda; \delta v_1(\zeta))\, d\zeta \right] \\ &-\delta^{-3} \left[\int_{\tilde J_2} \bar G_2(\delta\mu; \delta v_2(\zeta)) - \delta\mu \bar S_2(\delta\lambda; \delta v_2(\zeta))\, d\zeta \right. \\ &\qquad \left. + \int_{\tilde J_1} \bar H(\delta\mu; \delta v_1(\zeta)) - \delta\mu \bar K(\delta\lambda; \delta v_1(\zeta))\, d\zeta \right], \quad z \in \tilde J_2\,. \end{aligned} \tag{38}$$

Here F_1 and F_2 also depend on the z-variable. For simplicity, we drop this dependence throughout the rest of the section. Substituting $z = \tilde{z}$ in (37) and $z = 1$ in (38), we define

$$m[(v_1, \mu, v_2, \lambda); \delta] := F_1[(v_1, \mu, v_2, \lambda); \delta](\tilde{z}),$$
$$\ell[(v_1, \mu, v_2, \lambda); \delta] := F_2[(v_1, \mu, v_2, \lambda); \delta](1),$$

by substituting $v_1(\tilde{z}) = \mu$ and $v_2(1) = \lambda$ in both expressions.

Define the extension of F_1 and F_2 to $\delta = 0$ by using Taylor expansion together with the symmetry properties of the involved kernels. Let

$$C_1 = -S_1''(0)|J_1| - K''(0)|J_2|\,, \quad C_2 = -S_2''(0)|J_2| - K''(0)|J_1|\,.$$

Due to the assumptions on S_1, S_2, and K, the constants $S_1''(0)$, $S_2''(0)$ and $K''(0)$ are non-positive. After some tedious calculations we obtain the natural definition

$$F_1[(v_1, \mu, v_2, \lambda); 0](z) = \frac{1}{2} - z - \frac{C_1}{6} v_1(z)^3 + \frac{1}{2}\Big[C_1\mu^2 + C_2(\lambda^2 - \mu^2)\Big] v_1(z)\,, \quad (39)$$

$$m[(v_1, \mu, v_2, \lambda); 0] = \frac{1}{2} - \tilde{z} + \left(\frac{C_1}{3} - \frac{C_2}{2}\right)\mu^3 + \frac{C_2}{2}\mu\lambda^2\,,$$

$$F_2[(v_1, \mu, v_2, \lambda); 0](z) = \tilde{z} - z - \frac{C_2}{6} v_2(z)^3 + \frac{C_2}{2}\lambda^2 v_2(z) + \frac{C_2}{6}\mu^3 - \frac{C_2}{2}\lambda^2\mu\,, \quad (40)$$

$$\ell[(v_1, \mu, v_2, \lambda); 0] = \tilde{z} - 1 + \frac{C_2}{6}\mu^3 + \frac{C_2}{3}\lambda^3 - \frac{C_2}{2}\mu\lambda^2\,.$$

Our goal is to solve

$$F_1[(v_1, \mu, v_2, \lambda); \delta] = m[(v_1, \mu, v_2\lambda); \delta] = F_2[(v_1, \mu, v_2, \lambda); \delta] = \ell[(v_1, \mu, v_2, \lambda); \delta] = 0\,,$$

for $\delta > 0$ small enough. First we will solve the case $\delta = 0$ and then prove that a solution still exists when δ is close to zero. The solution for $\delta = 0$ is already partially explicit from formulas (39), (40). We only need to determine μ and λ. To do so, we set the m and ℓ components above equal to zero and obtain

$$\tilde{z} - \frac{1}{2} = \left(\frac{C_1}{3} - \frac{C_2}{2}\right)\mu^3 + \frac{C_2}{2}\mu\lambda^2\,,\ 1 - \tilde{z} = \frac{C_2}{6}\mu^3 + \frac{C_2}{3}\lambda^3 - \frac{C_2}{2}\mu\lambda^2\,, \quad (41)$$

which should be solved under the constraint $\mu < \lambda$. The second equation in (41) can be rewritten as

$$(\mu - \lambda)^2(2\lambda + \mu) = 6\,(1 - \tilde{z})/C_2,$$

which describes a cubic hyperbola, asymptotic to the straight lines $\mu - \lambda = 0$ and $2\lambda + \mu = 0$ in the (μ, λ) plane. Consider the branch of such a curve in the region $0 < \mu < \lambda$, intersecting the $\mu = 0$ axis at $(\mu, \lambda) = (0, \lambda_0)$ with

$$\lambda_0 = \left[3\,(1 - \tilde{z})\,/C_2\right]^{1/3}.$$

Such a branch describes a monotone increasing function $\lambda = \bar{\lambda}(\mu)$ on $\mu > 0$, asymptotic to $\lambda = \mu$ as $\mu \to +\infty$. Now, summing up the two equations in (41) we get the following additional condition on λ and μ:

$$\lambda = \tilde{\lambda}(\mu) := \left(\frac{3}{2C_2} + \frac{C_2 - C_1}{C_2}\mu^3\right)^{1/3}.$$

The function $\lambda = \tilde{\lambda}(\mu)$ is also monotone, and attains the value

$$\tilde{\lambda}(0) = \lambda_1 := \left(\frac{3}{2C_2}\right)^{1/3} > \lambda_0\,.$$

On the other hand, $\tilde{\lambda}(\mu)$ is asymptotic to the straight line $\lambda = ((C_2 - C_1)/C_2)^{1/3}\mu$ as $\mu \to +\infty$. Since $((C_2 - C_1)/C_2)^{1/3} < 1$, the two curves $\lambda = \bar{\lambda}(\mu)$ and $\lambda = \tilde{\lambda}(\mu)$

intersect at exactly one point in the region $0 < \mu < \lambda$. Hence, μ and λ are uniquely determined.

At this stage, the solution (v_1, v_2) can be easily recovered as the pseudo-inverse variables associated to the densities $\tilde\rho_1$ and $\tilde\rho_2$

$$\tilde\rho_1(x) = \left[\frac{1}{2}\left(C_1\mu^2 + C_2(\lambda^2-\mu^2)\right) - \frac{C_1}{2}x^2\right]_+ \mathbf{1}_{[-\mu,\mu]}\,,$$
$$\tilde\rho_2(x) = \frac{C_2}{2}(\lambda^2 - x^2)_+ \mathbf{1}_{[-\lambda,-\mu]\cup[\mu,\lambda]}\,,$$

which corresponds to two Barenblatt profiles centered at $x = 0$ (with possibly different masses and supports) such that the resulting profile $w = \tilde\rho_1 + \tilde\rho_2$ is continuous at $x = \pm\mu$. This proves that

$$F_1[(v_1,\mu,v_2,\lambda);0] = m[(v_1,\mu,v_2,\lambda);0] = F_2[(v_1,\mu,v_2,\lambda);0] = \ell[(v_1,\mu,v_2,\lambda);0] = 0$$

has a unique solution $(v_1, \mu, v_2, \lambda) \in \Omega$. From now on we call such solution $(v_{0,1}, \mu_0, v_{0,2}, \lambda_0)$.

Remark 5.2 (Inner and outer species)**.** In the above computation we arbitrarily assigned ρ_1 to the role of the 'inner' species in the segregated state. All of the above procedure also works with ρ_2 as inner species, and therefore we should speak of two possible segregated states rather than just one. The only parameters characterizing each species with respect to this pattern are $S_1''(0)$ and $S_2''(0)$, which model self-attraction forces. These constants affect the constants C_1 and C_2 above, which are non-negative. When $C_2 < C_1$, the self-attraction force of the first species is stronger than that of the second one, which suggests that ρ_1 will concentrate faster than ρ_2, thus producing a pattern in which the first species occupies the inner region whereas the second species occupies the outer region. On the other hand, a steady state with the reversed order can still be constructed under the condition $C_2 < C_1$, but we conjecture it to be unstable. This is supported by numerical simulations shown in Section 6, in which we see that the inner species is the one featuring the 'more concentrated' Barenblatt-type profile. We observe that the slope of the function $\lambda = \tilde\lambda(\mu)$ above is negative if $C_2 < C_1$ and positive if $C_1 < C_2$. This implies that the values of μ and λ solving (41) are smaller when $C_2 < C_1$ compared to the case $C_2 > C_1$. Hence, ρ_1 and ρ_2 have smaller support when $C_2 < C_1$. We deduce that the 'correct' steady state is 'more concentrated' than the unstable one.

Remark 5.3 (Non-symmetric segregation)**.** As shown by the numerical simulations in Section 6, segregation may emerge via a non-symmetric structure, in which $\rho_1 + \rho_2$ is supported on a connected interval and ρ_1 and ρ_2 feature exactly one jump discontinuity, see figure 2 below. This is typically the case for instance when the two species are initially separated. The mathematical proof of the existence of non-symmetric segregated steady states can be carried out similarly to the symmetric case. We omit the details and restrict to the symmetric case for simplicity.

5.3. **Solution via implicit function theorem for $\delta > 0$.** In this section we adapt the strategy of [21, Section 4] to our two-species problem. In order to solve problem (35), (36) for $\delta > 0$ small enough, we analyse the operator

$$U_{1/2} \ni (v_1,\mu,v_2,\lambda) \mapsto \mathcal{F}[(v_1,\mu,v_2,\lambda);\delta] := (F_1, m, F_2, \ell)[(v_1,\mu,v_2,\lambda)] \in \Omega_1.$$

Our first goal is to prove that $\mathcal{F}$ is Frechét continuously differentiable in a neighborhood of $(v_{1,0}, \mu_0, v_{2,0}, \lambda_0)$, the solution of $\mathcal{F}[(v_1,\mu,v_2,\lambda);0] = 0$. By Taylor expansion in (35), (36) w.r.t. δ around $\delta = 0$ one sees that $\mathcal{F}$ has a continuous first derivative with respect to δ. We omit the details.

For fixed $\delta > 0$, the Jacobian of $\mathcal{F}[\cdot;\delta]$ w.r.t. the first four variables has the block structure

$$D\mathcal{F} = \begin{pmatrix} \frac{\partial F_1}{\partial v_1} & \frac{\partial F_1}{\partial \mu} & \frac{\partial F_1}{\partial v_2} & \frac{\partial F_1}{\partial \lambda} \\ \frac{\partial m}{\partial v_1} & \frac{\partial m}{\partial \mu} & \frac{\partial m}{\partial v_2} & \frac{\partial m}{\partial \lambda} \\ \frac{\partial F_2}{\partial v_1} & \frac{\partial F_2}{\partial \mu} & \frac{\partial F_2}{\partial v_2} & \frac{\partial F_2}{\partial \lambda} \\ \frac{\partial \ell}{\partial v_1} & \frac{\partial \ell}{\partial \mu} & \frac{\partial \ell}{\partial v_2} & \frac{\partial \ell}{\partial \lambda} \end{pmatrix},$$

where partial derivatives with respect to v_1 and v_2 are meant to be Frechét derivatives. We now compute such terms. Consider perturbations $(w_1, a, w_2, b) \in \Omega_{1/2}$ such that

$$(v_{0,1}, \mu_0, v_{0,2}, \lambda_0) + (w_1, a, w_2, b) \in U_{1/2}\,. \tag{42}$$

We notice that (42) is satisfied if $|||(w_1, a, w_2, b)|||_{1/2}$ is small enough, since $v_{0,1}$ and $v_{0,2}$ have their gradient bounded from below by a positive constant. Upon extending w_1 and w_2 odd to $[1-\tilde{z}, \tilde{z}]$ and $[0, 1-\tilde{z}] \cup [\tilde{z}, 1]$ respectively, one has

$$\int_{J_1} w_1(\zeta)d\zeta = \int_{J_2} w_2(\zeta)d\zeta = 0.$$

Now we compute the partial derivatives of $\mathcal{F}$. For simplicity, we drop the $_0$ indices to denote the $\delta = 0$ state and avoid indicating the respective interval for the variable z.

Let $(w_1, a, w_2, b) \in \Omega_{1/2}$. By extending all involved functions to $[0, 1/2]$ in the usual symmetric form, we get for $\delta > 0$

$$\begin{aligned}
&\frac{\partial F_1}{\partial v_1}[(v_1, \mu, v_2, \lambda); \delta](w_1) \\
&\quad = \delta^{-2}\left[\int_{J_1} S_1\big(\delta(v_1(z) - v_1(\zeta))\big)(w_1(z) - w_1(\zeta))\right. \\
&\qquad\qquad -w_1(z)S_1\big(\delta(\mu - v_1(\zeta))\big) + \delta v_1(z)S_1'\big(\delta(\mu - v_1(\zeta))\big)w_1(\zeta)\,d\zeta \\
&\qquad\left. + \int_{J_2} K\big(\delta(v_1(z) - v_2(\zeta))\big)w_1(z) - w_1(z)K\big(\delta(\mu - v_2(\zeta))\big)d\zeta\right] \\
&\qquad + \delta^{-2}w_1(z)\left[\int_{J_2} S_2\big(\delta(\mu - v_2(\zeta))\big) - S_2\big(\delta(\lambda - v_2(\zeta))\big)d\zeta\right. \\
&\qquad\qquad\left. + \int_{J_1} K\big(\delta(\mu - v_1(\zeta))\big) - K\big(\delta(\lambda - v_1(\zeta))\big)d\zeta\right],
\end{aligned}$$

and in the limit $\delta \searrow 0$ we obtain

$$\frac{\partial F_1}{\partial v_1}[(v_1, \mu, v_2, \lambda); 0](w_1) = -\frac{C_1}{2}w_1(z)\big(v_1(z)^2 - \mu^2\big) - \frac{C_2}{2}w_1(z)\big(\mu^2 - \lambda^2\big). \tag{43}$$

This limit is so far just formal. The same holds for the $\delta \searrow 0$ limits computed below. However, we shall prove later that these are actually rigorous limits. Similarly,

$$\begin{aligned}
&\frac{\partial F_1}{\partial v_2}[(v_1, \mu, v_2, \lambda); \delta](w_2) \\
&\quad = \delta^{-2}\int_{J_2} -K\big(\delta(v_1(z) - v_2(\zeta))\big)w_2(\zeta) + v_1(z)K'\big(\delta(\mu - v_2(\zeta))\big)w_2(\zeta)d\zeta \\
&\qquad + \delta^{-2}v_1(z)\int_{J_2} -S_2'\big(\delta(\mu - v_2(\zeta))\big) + S_2'\big(\delta(\lambda - v_2(\zeta))\big)w_2(\zeta)d\zeta.
\end{aligned}$$

One can easily see that, in the $\delta \searrow 0$ limit one gets

$$\frac{\partial F_1}{\partial v_2}[(v_1, \mu, v_2, \lambda); 0](w_2) = 0\,.$$

We now compute for $\delta > 0$,

$$\begin{aligned}&\frac{\partial F_1}{\partial \mu}[(v_1,\mu,v_2,\lambda);\delta](a)\\&\quad= \delta^{-1}v_1(z)\left[\int_{J_1} S_1'\big(\delta(\mu - v_1(\zeta))\big)d\zeta + \int_{J_2} S_2'\big(\delta(\mu - v_2(\zeta))\big)d\zeta\right]a,\end{aligned}$$

and a Taylor expansion arguments shows

$$\frac{\partial F_1}{\partial \mu}[(v_1,\mu,v_2,\lambda);0](a) = (C_1 - C_2)v_1(z)\mu\, a\,.$$

Similarly,

$$\begin{aligned}&\frac{\partial F_1}{\partial \lambda}[(v_1,\mu,v_2,\lambda);\delta](b)\\&\quad= \delta^{-2}v_1(z)\left[-\delta\int_{J_2} S_2'\big(\delta(\lambda - v_2(\zeta))\big)d\zeta - \delta\int_{J_1} K'\big(\delta(\lambda - v_1(\zeta))\big)d\zeta\right]b\,,\end{aligned}$$

with the $\delta \searrow 0$ limit

$$\frac{\partial F_1}{\partial \lambda}[(v_1,\mu,v_2,\lambda);0](b) = C_2 v_1(z)\lambda b.$$

Turning to the second row of $D\mathcal{F}$ we get

$$\begin{aligned}&\frac{\partial m}{\partial v_1}[(v_1,\mu,v_2,\lambda);\delta](w_1)\\&\quad= \delta^{-3}\int_{J_1}\Big[-\delta S_1\big(\delta(\mu - v_1(\zeta))\big) + \delta^2\mu S_1'\big(\delta(\mu - v_1(\zeta))\big)\Big]w_1(\zeta)d\zeta\\&\qquad+ \delta^{-2}\mu\int_{J_1}\Big[-\delta K'\big(\delta(\mu - v_1(\zeta))\big) + \delta K'\big(\delta(\lambda - v_1(\zeta))\big)\Big]w_1(\zeta)d\zeta,\end{aligned}$$

and a simple symmetry argument shows that

$$\frac{\partial m}{\partial v_1}[(v_1,\mu,v_2,\lambda);0](w_1) = 0.$$

Next we compute

$$\begin{aligned}&\frac{\partial m}{\partial \mu}[(v_1,\mu,v_2,\lambda);\delta](a)\\&\quad= -\delta^{-1}\mu a\int_{J_1} S_1'\big(\delta(\mu - v_2(\zeta))\big)d\zeta - \delta^{-1}\mu a\int_{J_2} K'\big(\delta(\mu - v_2(\zeta))\big)d\zeta\\&\qquad+ \delta^{-2}a\int_{J_2} S_2\big(\delta(\mu - v_2(\zeta))\big) - S_2\big(\delta(\lambda - v_2(\zeta))\big)d\zeta\\&\qquad+ \delta^{-2}a\int_{J_1} K\big(\delta(\mu - v_1(\zeta))\big) - K\big(\delta(\lambda - v_1(\zeta))\big)d\zeta\\&\qquad+ \delta^{-1}\mu a\int_{J_2} S_2'\big(\delta(\mu - v_2(\zeta))\big)d\zeta + \delta^{-1}\mu a\int_{J_1} K'\big(\delta(\mu - v_1(\zeta))\big)d\zeta,\end{aligned}$$

and a simple computation shows in the $\delta \searrow 0$ limit,

$$\frac{\partial m}{\partial \mu}[(v_1,\mu,v_2,\lambda);0](a) = \left(C_1\mu^2 - \frac{3}{2}C_2\mu^2 + \frac{C_2}{2}\lambda^2\right)a\,.$$

Similar to the above, we have

$$\begin{aligned}
&\frac{\partial m}{\partial v_2}[(v_1,\mu,v_2,\lambda);\delta](w_2)\\
&\quad=\delta^{-3}\int_{J_2}\Big[-\delta K\big(\delta(\mu-v_2(\zeta))\big)+\delta\mu K'\big(\delta(\mu-v_2(\zeta))\big)\Big]w_2(\zeta)d\zeta\\
&\qquad+\delta^{-2}\mu\int_{J_2}\Big[-\delta S_2'\big(\delta(\mu-v_2(\zeta))\big)+\delta S_2\big(\delta(\lambda-v_2(\zeta))\big)\Big]w_2(\zeta)d\zeta\,,\\
&\text{with }\frac{\partial m}{\partial v_2}[(v_1,\mu,v_2,\lambda);0](w_2)=0.
\end{aligned}$$

We conclude the second row of the Jacobian by computing

$$\begin{aligned}
&\frac{\partial m}{\partial \lambda}[(v_1,\mu,v_2,\lambda);\delta](b)\\
&\quad=-\delta^{-1}\mu b\int_{J_2}S_2'\big(\delta(\lambda-v_2(\zeta))\big)d\zeta-\delta^{-1}\mu b\int_{J_1}K'\big(\delta(\lambda-v_2(\zeta))\big)d\zeta,
\end{aligned}$$

with the $\delta\searrow 0$ limit being

$$\frac{\partial m}{\partial \lambda}[(v_1,\mu,v_2,\lambda);0](b)=C_2\lambda\mu.$$

Now looking at the third row of $D\mathcal{F}$, we compute

$$\begin{aligned}
&\frac{\partial F_2}{\partial v_1}[(v_1,\mu,v_2,\lambda);\delta](w_1)\\
&\quad=\delta^{-3}\int_{J_1}-\delta K\big(\delta(v_2(z)-v_1(\zeta))\big)w_1(\zeta)+\delta^2 v_2(z)K'\big(\delta(\lambda-v_1(\zeta))\big)w_1(\zeta)d\zeta\\
&\qquad-\delta^{-3}\int_{J_1}-\delta w_1(\zeta)K\big(\delta(\mu-v_1(\zeta))\big)+\delta^2\mu w_1(\zeta)K'\big(\delta(\lambda-v_1(\zeta))\big)d\zeta,\qquad(44)
\end{aligned}$$

and for $\delta\searrow 0$ this shows

$$\frac{\partial F_2}{\partial v_1}[(v_1,\mu,v_2,\lambda);0](w_1)=0.$$

Then, we have

$$\begin{aligned}
&\frac{\partial F_2}{\partial \mu}[(v_1,\mu,v_2,\lambda);\delta](a)\\
&\quad=-\delta^{-3}\int_{J_2}\delta S_2\big(\delta(\mu-v_2(\zeta))\big)a-\delta S_2\big(\delta(\lambda-v_2(\zeta))\big)a\,d\zeta\\
&\qquad-\delta^{-3}\int_{J_1}\delta K\big(\delta(\mu-v_1(\zeta))\big)a-\delta K\big(\delta(\lambda-v_1(\zeta))\big)a\,d\zeta,
\end{aligned}$$

with the $\delta\searrow 0$ limit

$$\frac{\partial F_2}{\partial \mu}[(v_1,\mu,v_2,\lambda);0](a)=\frac{C_2}{2}(\mu^2-\lambda^2)a\,.$$

We continue computing the third row of $D\mathcal{F}$ with

$$\begin{aligned}
&\frac{\partial F_2}{\partial v_2}[(v_1,\mu,v_2,\lambda);\delta](w_2)\\
&\quad=\delta^{-3}\int_{J_2}\delta S_2\big(\delta(v_2(z)-v_2(\zeta))\big)(w_2(z)-w_2(\zeta))-\delta w_2(z)S_2\big(\delta(\lambda-v_2(\zeta))\big)\\
&\qquad\qquad+\delta^2 v_2(z)w_2(\zeta)S_2'\big(\delta(\lambda-v_2(\zeta))\big)\,d\zeta\\
&\qquad+\delta^{-3}\int_{J_1}\delta w_2(z)K\big(\delta(v_2(z)-v_1(\zeta))\big)-\delta w_2(z)K\big(\delta(\lambda-v_1(\zeta))\big)\,d\zeta\\
&\qquad-\delta^{-3}\int_{J_2}-\delta w_2(\zeta)S_2\big(\delta(\mu-v_2(\zeta))\big)+\delta^2\mu w_2(\zeta)S_2'\big(\delta(\lambda-v_2(\zeta))\big)\,d\zeta,
\end{aligned}$$

and the $\delta\searrow 0$ limit

$$\frac{\partial F_2}{\partial v_2}[(v_1,\mu,v_2,\lambda);0](w_2)=\frac{C_2}{2}(\lambda^2-v_2^2(z))w_2(z)\,.$$

To conclude the third row of the Jacobian, we have

$$\begin{aligned}
&\frac{\partial F_2}{\partial\lambda}[(v_1,\mu,v_2,\lambda);\delta](b)\\
&\quad=-\delta^{-3}\int_{J_2}\delta^2(v_2(z)-\mu)bS_2'\big(\delta(\lambda-v_2(\zeta))\big)d\zeta\\
&\qquad-\delta^{-3}\int_{J_1}\delta^{-2}(v_2(z)-\mu)bK'\big(\delta(\lambda-v_1(\zeta))\big)d\zeta\,,
\end{aligned}$$

and the $\delta\searrow 0$ limit is

$$\frac{\partial F_2}{\partial\lambda}[(v_1,\mu,v_2,\lambda);0](b)=C_2(v_2(z)-\mu)\lambda b\,.$$

Finally, we analyse the last row of the Jacobian of $\mathcal{F}$.

$$\begin{aligned}
&\frac{\partial\ell}{\partial v_1}[(v_1,\mu,v_2,\lambda);\delta](w_1)\\
&\quad=\delta^{-3}\int_{J_1}\Big[-\delta K\big(\delta(\lambda-v_1(\zeta))\big)+\delta K\big(\delta(\mu-v_1(\zeta))\big)\Big]w_1(\zeta)d\zeta\\
&\qquad-\delta^{-3}\int_{J_1}\Big[-\delta^2\lambda K'\big(\delta(\lambda-v_1(\zeta))\big)+\delta^2\mu K'\big(\delta(\lambda-v_(\zeta))\big)\Big]w_1(\zeta)d\zeta,
\end{aligned}$$

and the $\delta\searrow 0$ limit can be easily computed to be

$$\frac{\partial\ell}{\partial v_1}[(v_1,\mu,v_2,\lambda);0](w_1)=0.$$

Then, we have

$$\begin{aligned}
&\frac{\partial\ell}{\partial\mu}[(v_1,\mu,v_2,\lambda);\delta](a)\\
&\quad=\delta^{-3}a\int_{J_2}-\delta S_2\big(\delta(\mu-v_2(\zeta))\big)+\delta S_2\big(\delta(\lambda-v_2(\zeta))\big)\,d\zeta\\
&\qquad+\delta^{-3}a\int_{J_1}-\delta K\big(\delta(\mu-v_1(\zeta))\big)+\delta K\big(\delta(\lambda-v_1(\zeta))\big)\,d\zeta,
\end{aligned}$$

with the $\delta\searrow 0$ limit being

$$\frac{\partial\ell}{\partial\mu}[(v_1,\mu,v_2,\lambda);\delta](a)=\frac{C_2}{2}(\mu^2-\lambda^2)a.$$

We then continue with

$$\begin{aligned}&\frac{\partial \ell}{\partial v_2}[(v_1,\mu,v_2,\lambda);\delta](w_2)\\&\quad=\delta^{-3}\int_{J_2}-\delta w_2(\zeta)S_2\big(\delta(\lambda-v_2(\zeta))\big)+\delta w_2(\zeta)S_2\big(\delta(\mu-v_2(\zeta))\big)\,d\zeta\\&\qquad+\delta^{-3}\int_{J_2}\delta^2\lambda w_2(\zeta)S_2'\big(\delta(\lambda-v_2(\zeta))\big)-\delta^2\mu w_2(\zeta)S_2'\big(\delta(\lambda-v_2(\zeta))\big)\,d\zeta,\end{aligned}$$

and

$$\frac{\partial \ell}{\partial v_2}[(v_1,\mu,v_2,\lambda);0](w_2)=0.$$

The last term is

$$\begin{aligned}&\frac{\partial \ell}{\partial \lambda}[(v_1,\mu,v_2,\lambda);\delta](b)\\&\quad=\delta^{-3}b\int_{J_2}-\delta^2\lambda S_2'\big(\delta(\lambda-v_2(\zeta))\big)+\delta^2\mu S_2'\big(\delta(\lambda-v_2(\zeta))\big)\,d\zeta\\&\qquad+\delta^{-3}b\int_{J_1}-\delta^2\lambda K'\big(\delta(\lambda-v_1(\zeta))\big)+\delta^2\mu K'\big(\delta(\lambda-v_1(\zeta))\big)\,d\zeta,\end{aligned}$$

and the $\delta\searrow 0$ limit is

$$\frac{\partial \ell}{\partial \lambda}[(v_1,\mu,v_2,\lambda);0](b)=C_2(\lambda^2-\mu\lambda)b\,.$$

The above computations show for δ small enough, that $D\mathcal{F}[(v_{0,1},\mu_0,v_{0,2},\lambda_0);\delta]$ is a bounded linear operator from Ω into itself, and that $D\mathcal{F}$ is continuous at $\delta=0$ in the operator norm. This is easily seen via Taylor expansion, using bounds on the L^∞ norms, and symmetry properties.

Lemma 5.1. *$D\mathcal{F}[(v_{0,1},\mu_0,v_{0,2},\lambda_0);\delta]$ is a bounded linear operator from $\Omega_{1/2}$ to Ω_1 for $\delta>0$ small enough.*

Proof. Due to the structure of the spaces Ω_α, we only need to check the following. Define for $z\in J_2$ and $\delta\geq 0$,

$$\begin{aligned}g_2(\delta;z)\quad=\quad&\frac{\partial F_2}{\partial v_1}[(v_1,\mu,v_2,\lambda);\delta](w_1)+\frac{\partial F_2}{\partial \mu}[(v_1,\mu,v_2,\lambda);\delta](a)\\&+\frac{\partial F_2}{\partial v_2}[(v_1,\mu,v_2,\lambda);\delta](w_2)+\frac{\partial F_2}{\partial \lambda}[(v_1,\mu,v_2,\lambda);\delta](b)\\&-\frac{\partial \ell}{\partial v_1}[(v_1,\mu,v_2,\lambda);\delta](w_1)-\frac{\partial \ell}{\partial \mu}[(v_1,\mu,v_2,\lambda);\delta](a)\\&-\frac{\partial \ell}{\partial v_2}[(v_1,\mu,v_2,\lambda);\delta](w_2)-\frac{\partial \ell}{\partial \lambda}[(v_1,\mu,v_2,\lambda);\delta](b)\,.\end{aligned}$$

We need to prove that

$$\sup_{|||(w_1,a,w_2,b)|||_{1/2}\leq 1}\frac{1}{1-z}\big|g_2(\delta;z)-g_2(0;z)\big|\searrow 0,\qquad\text{as }\delta\searrow 0.$$

Consider all the above terms separately, and for notational reasons omit the dependence on (v_1,μ,v_2,λ). By Taylor expansion, and using simple symmetry properties, we get

$$\begin{aligned}&\frac{1}{1-z}\left[\left(\frac{\partial F_2}{\partial v_1}(\delta)-\frac{\partial F_2}{\partial v_1}(0)\right)(w_1)-\left(\frac{\partial \ell}{\partial v_1}(\delta)-\frac{\partial \ell}{\partial v_1}(0)\right)(w_1)\right]\\&\quad=-\delta\int_{J_1}w_1(\zeta)\left[\frac{K''(\bar{x}(\zeta))}{2}\frac{(\lambda-v_1(z))^2}{1-z}+\frac{K'''(\tilde{x}(\zeta))}{6}\frac{(v_2(z)-\lambda)^3}{1-z}\right]d\zeta,\end{aligned}\tag{45}$$

for some $\bar{x}(\zeta)$ and $\tilde{x}(\zeta)$ in the interval $[0, \lambda\delta]$. Now, it can be proved (cf. a similar argument in [21, Lemma 4.1]) that $(\lambda - v_2(z))/(1-z)^{1/2}$ is uniformly bounded in $[\tilde{z}, 1]$. This provides the desired estimate for the related term in (45). One can check that

$$\frac{\partial F_2}{\partial \mu}(\delta) - \frac{\partial \ell}{\partial \mu}(\delta) - \frac{\partial F_2}{\partial \mu}(0) + \frac{\partial \ell}{\partial \mu}(0) \equiv 0.$$

We now estimate

$$\begin{aligned}
&\frac{1}{1-z}\left[\left(\frac{\partial F_2}{\partial v_2}(\delta) - \frac{\partial F_2}{\partial v_2}(0)\right)(w_2) + \left(\frac{\partial F_2}{\partial \lambda}(\delta) - \frac{\partial F_2}{\partial \lambda}(0)\right)(b)\right.\\
&\qquad\left. - \left(\frac{\partial \ell}{\partial v_2}(\delta) - \frac{\partial \ell}{\partial v_2}(0)\right)(w_2) - \left(\frac{\partial \ell}{\partial \lambda}(\delta) - \frac{\partial \ell}{\partial \lambda}(0)\right)(b)\right]\\
&= \frac{1}{1-z}\left\{\int_{J_2} w_2(\zeta)\left[\frac{1}{2}(v_2(z)-\lambda)^2 S_2''\big(\delta(\lambda - v_2(\zeta))\big) + \frac{\delta}{6}(\lambda - v_2(z))^3 S_2'''(\bar{x})\right] d\zeta\right.\\
&\qquad + w_2(z)\int_{J_2} \delta^{-2} S_2\big(\delta(v_2(z) - v_2(\zeta))\big) - S_2\big(\delta(\lambda - v_2(\zeta))\big)\, d\zeta\\
&\qquad + w_2(z)\int_{J_1} \delta^{-2} K\big(\delta(v_2(z) - v_1(\zeta))\big) - K\big(\delta(\lambda - v_1(\zeta))\big)\, d\zeta\\
&\qquad + b\delta^{-1}(\lambda - v_2(z))\left[\int_{J_2} S_2'\big(\delta(\lambda - v_2(\zeta))\big) d\zeta + \int_{J_1} K'\big(\delta(\lambda - v_1(\eta))\big) d\zeta\right]\\
&\qquad \left. - C_2\lambda b(v_2(z) - \lambda) + \frac{C_2}{2} w_2(z)(v_2(z)^2 - \lambda^2)\right\},
\end{aligned}$$

and some tedious Taylor expansions imply

$$\begin{aligned}
&\frac{1}{1-z}\left[\left(\frac{\partial F_2}{\partial v_2}(\delta) - \frac{\partial F_2}{\partial v_2}(0)\right)(w_2) + \left(\frac{\partial F_2}{\partial \lambda}(\delta) - \frac{\partial F_2}{\partial \lambda}(0)\right)(b)\right.\\
&\qquad\left. - \left(\frac{\partial \ell}{\partial v_2}(\delta) - \frac{\partial \ell}{\partial v_2}(0)\right)(w_2) - \left(\frac{\partial \ell}{\partial \lambda}(\delta) - \frac{\partial \ell}{\partial \lambda}(0)\right)(b)\right]\\
&\quad = O(\delta^2)(1 + w_2(z))\left(\frac{(\lambda - v_2(z))^2}{1-z} + \frac{(\lambda - v_2(z))^4}{1-z}\right)\\
&\qquad + \frac{(w_2(z) - b)(v_2(z) - \lambda)}{1-z} O(\delta).
\end{aligned}$$

This proves the assertion. □

Lemma 5.2. *$D\mathcal{F}[(v_{0,1}, \mu_0, v_{0,2}, \lambda_0); 0]$ is a linear isomorphism between $\Omega_{1/2}$ and Ω_1 for $\delta > 0$ small enough.*

Proof. We observe that the Jacobian of $\mathcal{F}$ at $\delta = 0$ has the structure

$$D\mathcal{F}[(v_1, \mu, v_2, \lambda); 0] = \begin{pmatrix} \frac{\partial F_1}{\partial v_1} & (C_1 - C_2)v_1\mu & 0 & C_2 v_1 \mu \\ 0 & C_1\mu^2 - \frac{3}{2}C_2\mu^2 + \frac{C_2}{2}\lambda^2 & 0 & C_2\lambda\mu \\ 0 & \frac{C_2}{2}(\mu^2 - \lambda^2) & \frac{\partial F_2}{\partial v_2} & C_2(v_2 - \mu)\lambda \\ 0 & \frac{C_2}{2}(\mu^2 - \lambda^2) & 0 & C_2(\lambda^2 - \mu\lambda) \end{pmatrix}.$$

Given $(h_1, \alpha, h_2, \beta) \in \Omega_1$, we have to prove that

$$(w_1, a, w_2, b)^T = D\mathcal{F}[(v_1, \mu, v_2, \lambda); 0](h_1, \alpha, h_2, \beta)^T, \tag{46}$$

admits a unique solution $(w_1, a, w_2, b) \in \Omega_{1/2}$ with

$$|||(w_1, a, w_2, b)|||_{1/2} \le C\, |||\, (h_1, \alpha, h_2, \beta)|||_1,$$

for some $C > 0$ independent of $(h_1, \alpha, h_2, \beta)$.

As a first step, we claim that $\frac{\partial F_1}{\partial v_1}$ is invertible as a map from L^∞ to L^∞ at $\delta = 0$. To see this, we use (43). For $h_1(z) = \frac{\partial F_1}{\partial v_1}(w_1)$ we get

$$\|w_1\|_{L^\infty(J_1)} = 2 \left\| \Big(C_1(v_1(z)^2 - \mu^2) + C_2(\mu^2 - \lambda^2) \Big)^{-1} \right\|_{L^\infty(J_1)} \|h_1\|_{L^\infty(J_1)},$$

and the assertion follows. Therefore, the proof will be completed if we can show that the sub-matrix

$$\begin{pmatrix} C_1\mu^2 - \frac{3}{2}C_2\mu^2 + \frac{C_2}{2}\lambda^2 & 0 & C_2\lambda\mu \\ \frac{C_2}{2}(\mu^2 - \lambda^2) & \frac{\partial F_2}{\partial v_2} & C_2(v_2 - \mu)\lambda \\ \frac{C_2}{2}(\mu^2 - \lambda^2) & 0 & C_2(\lambda^2 - \mu\lambda) \end{pmatrix},$$

is an invertible operator in the components (a, w_2, b). First, we prove that

$$C_2\lambda(\lambda - \mu)\left(C_1\mu^2 - C_2\mu^2 + \frac{C_2}{2}\lambda^2 + \frac{C_2}{2}\mu\lambda \right) \neq 0, \tag{47}$$

which is equivalent to

$$C_2\lambda(\lambda - \mu)\left(C_1\mu^2 + \frac{C_2}{2}(\lambda - \mu)(\lambda + 2\mu) \right) \neq 0.$$

This is always satisfied since $\lambda > \mu$. Condition (47) implies that the linear system

$$\begin{aligned} \left[C_1\mu^2 - \frac{3}{2}C_2\mu^2 + \frac{C_2}{2}\lambda^2 \right] a + C_2\lambda\mu b &= \alpha, \\ \frac{C_2}{2}(\mu^2 - \lambda^2)a + C_2(\lambda^2 - \mu\lambda)b &= \beta, \end{aligned}$$

has a unique solution (a, b). Now we only need to determine w_2. By subtracting the last two rows of the linear system (46), and by some simple manipulation, we obtain

$$w_2(z) - b = \frac{2}{\lambda + v_2(z)} \left[\frac{b}{2}(\lambda - v_2(z)) + \frac{2(h_2(z) - \beta)}{C_2(\lambda - v_2(z))} \right].$$

Since $(\lambda - v_2(z))/(1 - z)$ is uniformly positive on $[\tilde{z}, 1]$ (cf. a similar proof in [21, Lemma 1.4]), we obtain the desired assertion, by dividing the above identity by $\sqrt{1 - z}$. □

We are now ready to prove the main theorem of this section, Theorem 5.1, as well as one of the most important results in this paper.

Proof of Theorem 5.1. The results in this section, in particular Lemma 5.1 and Lemma 5.2, together with the implicit function theorem on Banach spaces (see e. g. [32, Theorem 15.1]), imply that the functional equation $\mathcal{F}[(v_1, \mu, v_2, \lambda); \delta] = 0$ has a solution for δ small enough. Here $\mathcal{F}$ is defined in Subsection 5.2 (see in particular (37)) and at the beginning of Subsection 5.3. Hence, we obtain a solution (v_1, v_2) to (35)-(36). The computations in Subsection 5.1 imply that (u_1, u_2) with $u_i = \delta v_i$, $i = 1, 2$, is a solution to (31) once we achieve enough regularity for the v_i. This follows easily from (35)-(36). Indeed, since δ is very small and since S_1, S_2, and K_0 are bounded away from zero, we can easily recover the $\partial_z v_i$ after differentiating (35)-(36) with respect to z. In particular, we find that the $\partial_z v_i$ are bounded away from zero. Hence, we can divide the resulting equations by $\partial_z v_1$ and $\partial_z v_2$ respectively. Similar to the above, we can once again differentiate (using the C^1 regularity of the v_i) and obtain (31) for (u_1, u_2). The usual change of variable transforming pseudo inverse variables to densities shows that (ρ_1, ρ_2) solves (13) where $\rho_i = \partial_x F_i$, $i = 1, 2$, and F_i is the pseudo-inverse of u_i for $i = 1, 2$. □

6. Numerical Simulations

Here we present some examples of sorting phenomena and mixing by solving (6) numerically in one space dimension. We use a particle method introduced for equations in gradient flow form in [29, 26], which is equivalent to a finite difference scheme for the pseudo-inverse equation, see also [44] and [37] for scalar conservation laws.

In this section we denote by ρ and η the two densities and by u and v the corresponding pseudo-inverse functions, that are solutions of the following system

$$\begin{cases} \frac{du}{dt} = -\epsilon\frac{1}{2}\left((u_z)^{-2}\right)_z - \epsilon\eta_x(u) + \int_0^1 \left(S_1'(u(z)-u(\zeta)) + K'(u(z)-v(\zeta))\right)d\zeta, \\ \frac{dv}{dt} = -\epsilon\frac{1}{2}\left((v_z)^{-2}\right)_z - \epsilon\rho_x(v) + \int_0^1 \left(S_2'(v(z)-v(\zeta)) + K'(v(z)-u(\zeta))\right)d\zeta, \end{cases}$$

for $z \in [0,1]$. This is in order to avoid confusion w.r.t. the indices for discretization. Clearly, the masses of ρ and η are normalized to one. The main issue for the equations above is how to treat the cross-diffusion part numerically. We intentionally left the cross-diffusion terms above in the form of 'external potentials'. Given $N \in \mathbb{N}$ we consider a partition of the interval $[0,1]$, $\{z_i\}_{i=1}^N$ and call $u(z_i) = u_i$, $v(z_i) = v_i$, and $m = \frac{1}{N}$. The discretization in space the reads

$$d_t u_i = \frac{m}{2}\frac{(u_{i+1}-u_i)^2 - (u_i-u_{i-1})^2}{(u_{i+1}-u_i)^2} - \eta_x(u_i) + m\sum_{j=1}^N \left(S_1'(u_i-u_j) + K'(u_i-v_j)\right),$$

$$d_t v_i = \frac{m}{2}\frac{(v_{i+1}-v_i)^2 - (v_i-v_{i-1})^2}{(v_{i+1}-v_i)^2} - \rho_x(v_i) + m\sum_{j=1}^N \left(S_2'(v_i-v_j) + K'(v_i-u_j)\right).$$

for $i = 1, ..., N$. Integrating the above ODE system we get two families of particles $\{u_i(t)\}_{i=1}^N$, $\{v_i(t)\}_{i=1}^N$, for $t \in [0,T]$ and reconstruct the discrete densities as follows

$$\rho^N(t,x) = \sum_{i=1}^N \frac{2m}{u_{i+1}(t) - u_{i-1}(t)} \mathbf{1}_{\left\{\left(u_{i-\frac{1}{2}}(t), u_{i+\frac{1}{2}}(t)\right)\right\}},$$

$$\eta^N(t,x) = \sum_{i=1}^N \frac{2m}{v_{i+1}(t) - v_{i-1}(t)} \mathbf{1}_{\left\{\left(v_{i-\frac{1}{2}}(t), v_{i+\frac{1}{2}}(t)\right)\right\}}.$$

Given the two initial conditions ρ^0 and η^0 the initial positions for the ODE system are determined via the atomization

$$u_1^0 = \sup_{x\in\mathbb{R}}\left\{\int_{-\infty}^x \rho^0(y)dy < \frac{1}{N}\right\} \quad , \quad u_i^0 = \sup_{x\in\mathbb{R}}\left\{\int_{u_{i-1}^0}^x \rho^0(y)dy < \frac{1}{N}\right\},$$

$$v_1^0 = \sup_{x\in\mathbb{R}}\left\{\int_{-\infty}^x \eta^0(y)dy < \frac{1}{N}\right\} \quad , \quad v_i^0 = \sup_{x\in\mathbb{R}}\left\{\int_{v_{i-1}^0}^x \eta^0(y)dy < \frac{1}{N}\right\},$$

for $i = 2, ..., N$. The cross-diffusion part is reconstructed at each time iteration using the discretized density. For the nonlocal part we choose

$$S_i = \sigma_i K \ , \ \sigma_i > 0 \ , \ i = 1,2 \ , \text{ and } \sigma_1 + \sigma_2 > 2 \ ,$$

where K is a normalized Gaussian potential. For a diffusion coefficient $\epsilon = 1$, we show segregation phenomena in Figures 1, 2 for two different choices of initial conditions. Two different types of segregation are possible. In Figure 1 the initial data for the two species are perfectly matching, this produces symmetric segregated states in the large time limit. In Figure 2 the two initial data are shifted, this produces a non symmetric segregation in which the two species form two adjacent patterns with connected support. Mixing is shown in Figures 3, 4 for the diffusion dominated regime, namely $\sigma_1 + \sigma_2 < 2$ and diffusion coefficient $\epsilon > 1$. Again, different

situations may arise depending on the initial data. In the former case (perfectly overlapping initial conditions) the two species are almost entirely overlapping for large times, whereas in the latter case they overlap in a proper subset of the support of $\rho + \eta$. In all simulations we have $N = 50$ and final time $T = 2$.

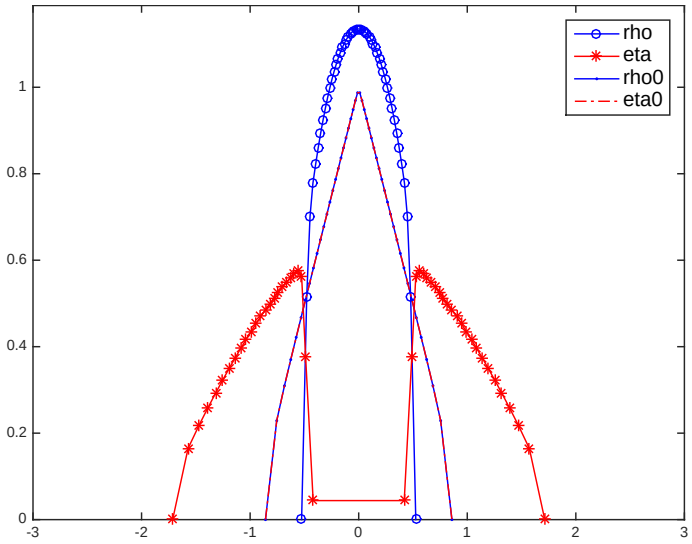


FIGURE 1. Symmetric segregation for $\epsilon = 1$, $\sigma_1 = 10$, $\sigma_2 = 1.5$, $\rho_0(x) = \eta_0(x) = (1 - |x|)_+$

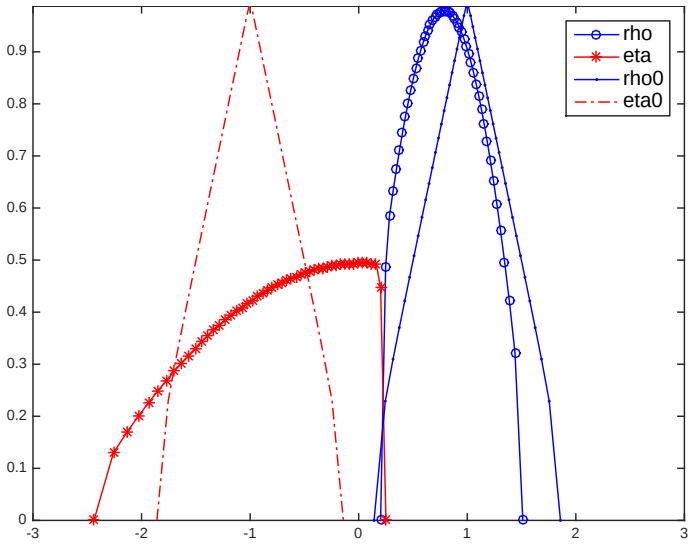


FIGURE 2. Segregation for two densities that are initially disjointed, with $\epsilon = 1$, $\sigma_1 = 10$, $\sigma_2 = 1.5$, $\rho_0(x) = (1 - |x - 1|)_+$, $\eta_0(x) = (1 - |x + 1|)_+$

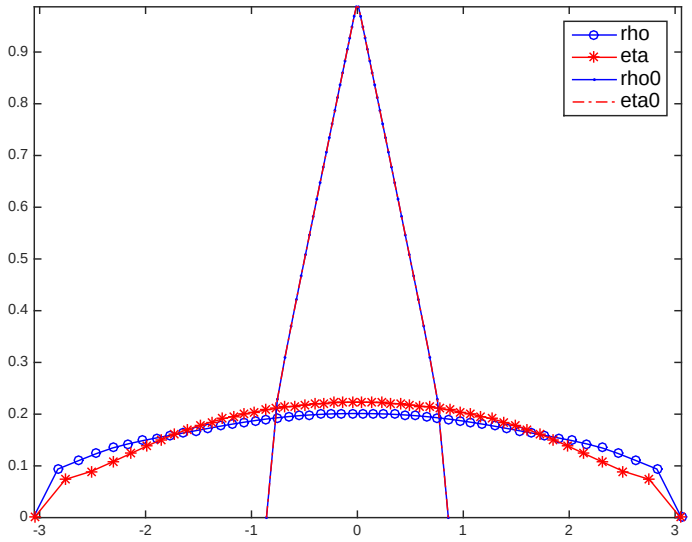


FIGURE 3. Diffusion dominated regime for $\epsilon = 3$, $\sigma_1 = 0.1$, $\sigma_2 = 0.8$, $\rho_0(x) = \eta_0 = (x) = (1 - |x|)_+$

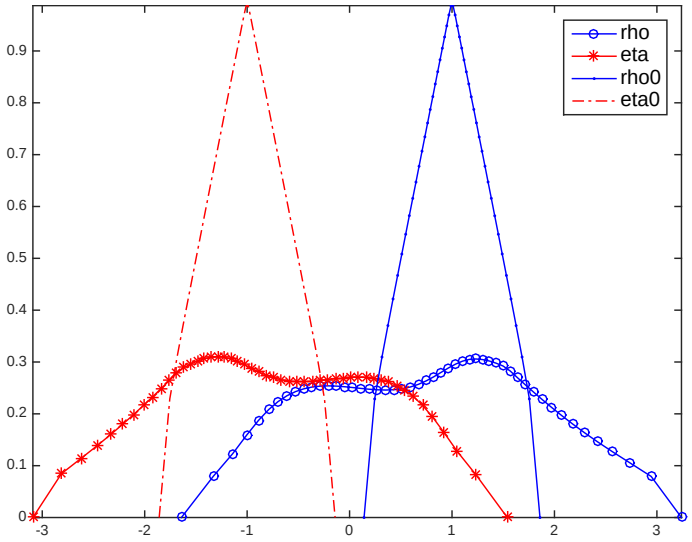


FIGURE 4. Mixing phenomena in the diffusion dominated regime. Here $\epsilon = 3$, $\sigma_1 = 0.1$, $\sigma_2 = 0.8$, $\rho_0(x) = (1 - |x - 1|)_+$, $\eta_0(x) = (1 - |x + 1|)_+$

ACKNOWLEDGMENTS

MDF and SF were supported by the EU Erasmus Mundus MathMods programme, www.mathmods.eu, by the GNAMPA (Italian group of Analysis, Probability, and Applications) projects *Geometric and qualitative properties of solutions to elliptic and parabolic equations*, respectively *Analisi e stabilità per modelli di equazioni alle derivate parziali nella matematica applicata*. MDF and SF acknowledge hospitality by the Institut für Angewandte Mathematik: Analysis und Numerik, Westfälische Wilhelms-Universität Münster, which hosted them in 2012

when the first ideas of this work came out. This work was done while MB and AS were PIs of the Excellence Cluster *Cells in Motion* (CiM), funded by the DFG.

References

[1] G. Aletti, G. Naldi, and G. Toscani. First-order continuous models of opinion formation. *SIAM J. Appl. Math.*, 67(3):837–853 (electronic), 2007.

[2] W. Alt. Degenerate diffusion equation with drift functionals modelling aggregation. *Nonlin. Anal.*, 9:811–836, 1985.

[3] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[4] C. Appert-Rolland, P. Degond, and S. Motsch. Two-way multi-lane traffic model for pedestrians in corridors. *Netw. Heterog. Media*, 6(3):351–381, 2011.

[5] J. Bedrossian. Global minimizers for free energies of subcritical aggregation equations with degenerate diffusion. *Applied Mathematics Letters*, 24:1927–1932, 2011.

[6] J. Bedrossian. Large mass global solutions for a class of l1-critical nonlocal aggregation equations and parabolic-elliptic Patlak-Keller-Segel models. *Comm. Partial Differential Equations*, 40:1119–1136, 2015.

[7] D. Benedetto, E. Caglioti, and M. Pulvirenti. A kinetic equation for granular media. *RAIRO Modél. Math. Anal. Numér.*, 31(5):615–641, 1997.

[8] A. L. Bertozzi, T. Laurent, and J. Rosado. L^p theory for the multidimensional aggregation equation. *Comm. Pure Appl. Math.*, 64(1):45–83, 2011.

[9] M. Bertsch, R. Dal Passo, and M. Mimura. A free boundary problem arising in a simplified tumour growth model of contact inhibition. *Interfaces and Free Boundaries*, 12:235–250, 2010.

[10] M. Bertsch, M. E. Gurtin, and D. Hilhorst. On interacting populations that disperse to avoid crowding: the case of equal dispersal velocities. *Nonlinear Analysis, Theory, Methods and Applications*, 11:493–499, 1987.

[11] M. Bertsch, M. E. Gurtin, D. Hilhorst, and L. A. Peletier. On interacting populations that disperse to avoid crowding: preservation of segregation. *J. Math. Biology*, 23:1–13, 1985.

[12] M. Bertsch, D. Hilhorst, H. Izuhara, and M. Mimura. A nonlinear parabolic-hyperbolic system for contact inhibition of cell-growth. *Differential Equations and Applications*, 4:137–157, 2012.

[13] P. Biler. Growth and accretion of mass in an astrophysical model. *Applicationes Mathematicae*, 23:173–189, 1995.

[14] A. Blanchet, E. Carlen, and J. A. Carrillo. Functional inequalities, thick tails and asymptotics for the critical mass Patlak-Keller-Segel model. *Journal of Functional Analysis*, 261:2142–2230, 2012.

[15] A. Blanchet, J. A. Carrillo, and P. Laurencot. Critical mass for a Patlak-Keller-Segel model with degenerate diffusion in higher dimensions. *Calculus of Variations and Partial Differential Equations*, 35:133–168, 2009.

[16] A. Blanchet, J.A. Carrillo, D. Kinderlehrer, M. Kowalczyk, P. Laurencot, and S. Lisini. A hybrid variational principle for the Keller-Segel system in $\mathbb{R}^2$. *ESAIM: Mathematical Modelling and Numerical Analysis*, 49(6):1553–1576, 2015.

[17] A. Blanchet, J. Dolbeault, and B. Perthame. Two-dimensional Keller-Segel model: Optimal critical mass and qualitative properties of the solutions. *Electron. J. Differ. Equations, (2006)*, (44):1–32 (electronic), 2006.

[18] S. Boi, V. Capasso, and D. Morale. Modeling the aggregative behavior of ants of the species *polyergus rufescens*. *Nonlinear Anal. Real World Appl.*, 1(1):163–176, 2000. Spatial heterogeneity in ecological models (Alcalá de Henares, 1998).

[19] N. V. Brilliantov and T. Pöschel. *Kinetic theory of granular gases*. Oxford Graduate Texts. Oxford University Press, Oxford, 2004.

[20] M. Burger and M. Di Francesco. Large time behavior of nonlocal aggregation models with nonlinear diffusion. *Netw. Heterog. Media*, 3(4):749–785, 2008.

[21] M. Burger and M. Di Francesco. Large time behavior of nonlocal aggregation models with nonlinear diffusion. *Netw. Heterog. Media*, 3(4):749–785, 2008.

[22] M. Burger, M. Di Francesco, and M. Franek. Stationary states of quadratic diffusion equations with long-range attraction. *Commun. Math. Sci.*, 11:709–738, 2013.

[23] M. Burger, R. C. Fetecau, and Y. Huang. Stationary states and asymptotic behaviour of aggregation models with nonlinear local repulsion. *SIAM J. Appl. Dyn. Syst.*, 13:397–424, 2014.

[24] J. A. Carrillo, Y.-P. Choi, and M. Hauray. The derivation of swarming models: mean-field limit and Wasserstein distances. In: Collective dynamics from bacteria to crowds, 1-46 CISM Courses and Lect. 553, 2014.
[25] J. A. Carrillo, S. Fagioli, F. Santambrogio, and M. Schmidtchen. Splitting schemes & segregation in reaction-(cross-)diffusion systems. *arXiv: 1711.05434*, 2017.
[26] J. A. Carrillo, Y. Huang, F. S. Patacchini, and G. Wolansky. Numerical study of a particle method for gradient flows. *Kinetic and Related Models (KRM)*, 10:613–641, 2017.
[27] J. A. Carrillo, Y. Huang, and M. Schmidtchen. Zoology of a non-local cross-diffusion model for two species. *arXiv: 1705.03320*, 2017. to appear in SIAM J. Appl. Math.
[28] J. A. Carrillo, R. J. McCann, and C. Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.*, 179(2):217–263, 2006.
[29] J. A. Carrillo, F. S. Patacchini, P. Sternberg, and G. Wolansky. Convergence of a particle method for diffusive gradient flows in one dimension. *SIAM J. Math. Anal.*, 48:3708–3741, 2016.
[30] Y. Chen and T. Kolokolnikov. A minimal model of predator-swarm interaction. *J. R. Soc. Interface*, 11, 2014.
[31] M. Cicalese, L. De Luca, M. Novaga, and M. Ponsiglione. Ground states of a two phase model with cross and self attractive interactions. *SIAM J. Math. Anal.*, 48(5):3412–3443, 2016.
[32] K. Deimling. *Nonlinear Functional Analysis*. Springer, Berlin, 1985.
[33] L. Derbel and P.E. Jabin. The set of concentration for some hyperbolic models of chemotaxis. *J. of Hyperbolic Differ. Equ.*, 4:331–349, 2007.
[34] M. Di Francesco, A. Esposito, and S. Fagioli. Nonlinear degenerate cross-diffusion systems with nonlocal interaction. *Nonlinear Analysis*, 169:94–117, 2018.
[35] M. Di Francesco and S. Fagioli. Measure solutions for non-local interaction PDEs with two species. *Nonlinearity*, 26:2777–2808, 2013.
[36] M. Di Francesco and S. Fagioli. A nonlocal swarm model for predators–prey interactions. *Mathematical Models and Methods in Applied Sciences*, 26:319–355, 2016.
[37] M. Di Francesco, S. Fagioli, M. D. Rosini, and G. Russo. Follow-the-leader approximations of macroscopic models for vehicular and pedestrian flows. *arXiv:1610.06743*, 2016.
[38] M. Di Francesco and Y. Jaafra. Multiple large-time behavior of nonlocal interaction equations with quadratic diffusion. *arXiv: 1710.0821*, 2017.
[39] B. Düring, P. Markowich, J.-F. Pietschmann, and M.-T. Wolfram. Boltzmann and Fokker-Planck equations modelling opinion formation in the presence of strong leaders. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 465(2112):3687–3708, 2009.
[40] E. E. Espejo, A. Stevens, and J. J. L. Velázquez. Simultaneous finite time blow-up in a two-species model for chemotaxis. *Analysis (Munich)*, 29(3):317–338, 2009.
[41] H. Gajewski and Gröger K. On the basic equations for carrier transport in semiconductors. *J. Math. Anal. Appl.*, 113(1):12–35, 1986.
[42] H. Gajewski and K. Zacharias. Global behavior of a reaction-diffusion system modelling chemotaxis. *Math. Nachr.*, 195:77–114, 1998.
[43] E. Geigant and M. Stoll. Stability of peak solutions of a non-linear transport equation on the circle. *Electron. J. Differential Equations*, 157:41, 2012.
[44] L. Gosse and G. Toscani. Lagrangian numerical approximations to one-dimensional convolution-diffusion equations. *SIAM J. Sci. Comput.*, 28:1203–1227, 2006.
[45] M.A. Herrero and J.J.L. Velázquez. Chemotactic collapse for the Keller-Segel model. *J. Math. Biol.*, 35:177–194, 1996.
[46] D. Horstmann. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. I. *Jahresber. Deutsch. Math.-Verein.*, 105:103–165, 2003.
[47] D. Horstmann. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. II. *Jahresber. Deutsch. Math.-Verein.*, 106:51 – 69, 2004.
[48] D. Horstmann and M. Lucia. Nonlocal elliptic boundary value problems related to chemotactic movement of mobile species. In *Mathematical analysis on the self-organization and self-similarity*, RIMS Kôkyûroku Bessatsu, B15, pages 39–72. Res. Inst. Math. Sci. (RIMS), Kyoto, 2009.
[49] W. Jäger and S. Luckhaus. On explosions of solutions to a system of partial differential equations modelling chemotaxis. *Trans. Amer. Math. Soc.*, 329(2):819–824, 1992.
[50] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
[51] A. Jüngel. The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity*, 28(6):1963–2001, 2015.
[52] G. Kaib. *Ph.D-thesis, University of Münster*, 2016.
[53] G. Kaib. Stationary states of an aggregation equation with degenerate diffusion and bounded attractive potential. *SIAM J. Math. Anal.*, 49:272–296, 2017.

[54] K. Kang, I. Primi, and J.J.L. Velázquez. A 2d-model of cell sorting induced by propagation of chemical signals along spiral waves. *Comm. Partial Differential Equations*, 38:1069–1122, 2013.

[55] E.F Keller and L. Segel. Initiation of slime mold aggregation viewed as an instability. *J. theor. Biol.*, 26:399–415, 1970.

[56] I Kim and A. R. Mészáros. On nonlinear cross-diffusion systems: an optimal transport approach. *arXiv: 1705.02457*, 2017.

[57] P. Laurencot and B.-V. Matioc. A gradient flow approach to a thin film approximation of the Muskat problem. *Calculus of Variations and Partial Differential Equations*, 47:319–341, 2013.

[58] G. Lemon and J. King. A functional differential equation model for biological cell sorting due to differential adhesion. *Math. Models Methods Appl. Sci.*, 23:93–126, 2013.

[59] M. Liero and A. Mielke. Gradient structures and geodesic convexity for reaction-diffusion systems. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 371(2005):20120346, 28, 2013.

[60] A. Magni, C. Mantegazza, and M. Novaga. Motion by curvature on planar networks II. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 15(5):117–144, 2016.

[61] S. Matsukuma and A. Durston. Chemotactic cell sorting in dictyostelium discoideum. *J. Embryol. Exp. Morphol.*, 50:243–251, 1979.

[62] M. Mimura and M. Yamaguti. Pattern formation in interacting and diffusing systems in population biology. *Adv. Biophys.*, 15:19–65, 1982.

[63] D. Morale, V. Capasso, and K. Oelschläger. An interacting particle system modelling aggregation behavior:from individuals to populations. *Journal of Math. Biology*, 50:49–66, 2005.

[64] J. Moser. A sharp form of an inequality by n. trudinger. *Indiana Univ. Math. Journal*, 20(11):1077–1092, 1971.

[65] J. Moser. On a nonlinear problem in differential geometry. *Dynamical systems. Proc. Sympos., Univ. Bahia, Salvador, 1971. Academic Press, New York*, pages 273–280, 1973.

[66] T. Nagai. Blow-up of radially symmetric solutions to a chemotaxis system. *Adv. Math. Sci. Appl.*, 5:581–601, 1995.

[67] K. Oelschläger. Large systems of interacting particles and porous medium equation. *J. Diff. Eq.*, 88:294–346, 1990.

[68] K. J. Painter. Continuous models for cell migration in tissues and applications to cell sorting via differential chemotaxis. *Bulletin of Mathematical Biology*, 71:1117–1147, 2009.

[69] B. Perthame. *Parabolic equations in biology. Growth, reaction, movement and diffusion.* Springer, 2015.

[70] K. Post. A system of non-linear partial differential equations modeling chemotaxis with sensitivity functions. *Ph.D.-thesis, Humboldt University of Berlin*, 1999.

[71] I. Primi, A. Stevens, and Velázquez J.J.L. Mass-selection in alignment models with non-deterministic effects. *Comm. Partial Differential Equations*, 34(4-6):419–456, 2009.

[72] R. Schaaf. Stationary solutions of chemotaxis systems. *Trans. Amer. Math. Soc.*, 292(2):531–556, 1985.

[73] A. Stevens. The derivation of chemotaxis equations as limit dynamics of moderately interacting stochastic many-particle systems. *SIAM J. Appl. Math*, 61(1):183–212, 2000.

[74] Y. Sugiyama. Global existence in sub-critical cases and finite time blow-up in super critical cases to degenerate Keller-Segel systems. *Differential Integral Equations*, 19(8):841–976, 2006.

[75] Y. 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.

[76] K. Sznajd-Weron and J. Sznajd. Opinion evolution in closed community. *Int. J. Mod. Phys. C*, 11:1157–1166, 2000.

[77] C. M. Topaz, A. L. Bertozzi, and M.A. Lewis. A nonlocal continuum model for biological aggregation. *Bulletin of Mathematical Biology*, 68(7):1601–1623, 2006.

[78] B. Vasiev and C.J. Weijer. Modeling chemotactic cell sorting during Dictyostelium discoideum mound formation. *Biophys. J.*, 76:595–605, 1999.

[79] J. L. Vázquez. *The porous medium equation.* Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, Oxford, 2007. Mathematical Theory.

[80] A. Voss-Böhme and A. Deutsch. The cellular basis of cell sorting kinetics. *J. Theoret. Biol.*, 263:419–436, 2010.

[81] G. Wolansky. On steady distributions of self-attracting clusters under friction and fluctuations. *Arch. Rational Mech. Anal.*, 119(4):355–391, 1992.

[82] G. Wolansky. Multi-component chemotactic system in the absence of conflict. *European J. of Applied Mathematics*, 13(6):641–661, 2002.

[83] G. Wolansky. Chemotactic systems in the presence of conflicts: a new functional inequality. *J. Differential Equations*, 261(9):5119–5143, 2016.

[84] J. Zinsl and Matthes D. Transport distances and geodesic convexity for systems of degenerate diffusion equations. *Calc. Var. Partial Differential Equations*, 54(4):3397–3438, 2015.

Martin Burger, Angela Stevens - Angewandte Mathematik, Westfälische Wilhelms-Universität Münster, Einsteinstr. 62, 48149 Münster, Germany.

Marco Di Francesco, Simone Fagioli - Dipartimento di Ingegneria e Scienze dell'Informazione e Matematica, Università degli Studi dell'Aquila, Via Vetoio 1, 67100 Coppito, L'Aquila, Italy.