Nonlocal cross-interaction systems on graphs: Energy landscape and dynamics

Georg Heinze*, †, Jan-Frederik Pietschmann†, Markus Schmidtchen‡

April 21, 2022

Abstract

We explore the dynamical behavior and energetic properties of a model of two species that interact nonlocally on finite graphs. The authors recently introduced the model in the context of nonquadratic Finslerian gradient flows on generalized graphs featuring nonlinear mobilities. In a continuous and local setting, this class of systems exhibits a wide variety of patterns, including mixing of the two species, partial engulfment, or phase separation. This work showcases how this rich behavior carries over to the graph structure. We present analytical and numerical evidence thereof.

1. Introduction

In recent years there has been a growing interest in dynamics on graphs in the applied mathematics community. A particular emphasis lies on systems exhibiting a discrete gradient-flow structure, cf. [29, 21, 28, 17]. In [26], the authors proposed a system of two nonlocally interacting species on a graph, extending the recent work [22], which can be cast into a Finslerian gradient-flow framework. The result pushes the frontiers of current developments in the direction of coupled multi-species systems with nonlinear mobilities. The dynamics are posed on (possibly infinite) graphs where either species may inhabit each node. Each species is represented by a probability measure on the set of vertices. The set of probability measures on the graph is then equipped with a, not necessarily quadratic, extended quasi-metric. Subsequently, the interaction system is shown to exhibit a gradient-flow structure. The results in [26] can be seen as extensions of previous results from several angles:

• The on-lattice dynamics of [22] are extended to multiple species, nonlinear mobilities and non-quadratic Finslerian structures.

• The Finslerian structure of the continuous \( p \)-Wasserstein spaces unveiled in [1] are carried over to a (discrete) graph setting.

• The research on distances induced by concave mobilities on the set of probability measures and the gradient flows in this space [27, 20, 11] is translated into a (discrete) graph setting for two species.

*corresponding author
†Fakultät für Mathematik, Technische Universität Chemnitz Reichenhainer Straße 41, Chemnitz, Germany. ([georg.heinze,jfpietschmann]@math.tu-chemnitz.de)
‡Institute of Scientific Computing, Technische Universität Dresden, Zellescher Weg 12-14, 01069 Dresden, Germany. (markus.schmidtchen@tu-dresden.de).
The driving motivation of the extension to several species is the well-known property of these interaction systems to exhibit a rich variety of patterns depending on the intraspecific and the interspecific interactions. In [25], a link is drawn between systems of interacting particles with two species of the form

$$\frac{dX_t^{(i)}}{dt} = -\frac{1}{N} \sum_{l=1}^{N} \nabla K^{(i)}(X_t^{(i)} - X_t^{(1)}) - \frac{1}{N} \sum_{l=1}^{N} \nabla K^{(2)}(X_t^{(i)} - X_t^{(2)}),$$

where \( l = 1, \ldots, N \) numbers consecutively the individuals belonging to species \( i = 1, 2 \) and the potentials \( K^{(ik)} \) encode the interactions between individuals of the same species (self-interactions) and members of opposing species (cross-interactions), respectively. Exploiting the 2-Wasserstein gradient flow structure solutions to the continuous counterpart, i.e.,

$$\partial_t \rho_t^{(i)} = \nabla \cdot (\rho_t^{(i)} \nabla (K^{(i)} * \rho_t^{(1)} + K^{(2)} * \rho_t^{(2)})),$$

are constructed for the nonlocal interaction energy

$$\mathcal{E}(\rho) = \frac{1}{2} \sum_{i,k=1}^{2} \int_{\mathbb{R}^d} K^{(ik)} * \rho_t^{(i)} d\rho_t^{(k)}, \quad \text{on} \ \mathbb{R}^d,$$

cf. [25]. Here, \( \rho_t^{(i)}, i = 1, 2 \), denote the probability densities of the two species. The formation of a rich variety of patterns in this class of systems is widely known, for instance in the context of predator-prey dynamics [19], pigment-cell interactions in zebrafish [34, 35], or cell-adhesion and tissue-growth [2, 14, 32, 31]. As pointed out earlier, the model proposed in [26] features nonlinear mobilities depending on the choice of potentials.

The rest of this paper is organized as follows. In Section 2 we introduce the rigorous setting and the notation used throughout the article. Section 3 is dedicated to a discussion of the dynamical properties of the system and stationary states, as well as their relation to minimizers of the interaction energy. Finally, Section 4 is devoted to illustrations of the dynamics using finite graphs and the influence of the graph parameters \( \mu, \eta \), as well as the exponent \( p \), and the scaling parameter \( \beta \). Moreover, we observe aggregation and phase separation on the graph depending on the choice of potentials.

2. Framework and notation

This section provides the necessary details on the setting and the model. For more details and the rigorous interpretation as a gradient flow, we refer the reader to [22, 26].
We start by introducing the graph setting. The vertices are given by a positive Radon measure \( \mu \in \mathcal{M}^+(\mathbb{R}^d) \), called the base measure. In this work, we will consider only finite graphs, i.e. base measures of the form

\[
\mu = \sum_{i=1}^{N} \mu_i \delta_{x_i},
\]

for some \( N \in \mathbb{N}, \mu_1, \ldots, \mu_N \in (0, \infty) \) and \( x_1, \ldots, x_N \in \mathbb{R}^d \). We only consider the case \( N \geq 2 \) as otherwise there are no dynamics.

The edges of the graph are defined as the set

\[
G := \{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d : x \neq y, \eta(x, y) > 0\},
\]

where \( \eta : \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty) \) is a nonnegative weight function. We consider undirected graphs, i.e. we assume \( \eta \) to be symmetric in the sense that \( \eta(x, y) = \eta(y, x) \) for \( x, y \in \mathbb{R}^d \). Further, we assume the restriction \( \eta|_G \) to be continuous.

For \( x_1, x_k \in \text{supp } \mu \), we shall use the notation \( \eta_{xk} = \eta(x_1, x_k) \) and similarly for other functions.

We denote by \( \mathcal{P}_\mu(\mathbb{R}^d) \) the space of probability measures \( \rho \) on \( \mathbb{R}^d \), which satisfy \( \rho \ll \mu \), and by \( \mathcal{M}_\mu(G) \) the space of Radon measures \( j \) on \( G \) which satisfy \( j \ll \mu \otimes \mu \).

Next we introduce nonlocal analogues of gradient and divergence. Given \( \varphi : \mathbb{R}^d \to \mathbb{R} \), its nonlocal gradient is \( \nabla \varphi(x, y) := \varphi(y) - \varphi(x) \) while for \( j \in \mathcal{M}(\mathbb{R}^d) \) we define the nonlocal divergence \( \nabla \cdot j \) by

\[
\int_{\mathbb{R}^d} \varphi(x) d\nabla \cdot j(x) := -\frac{1}{2} \int_{G} \nabla \varphi(x, y) \eta(x, y) d\eta(x, y), \quad \forall \varphi : \mathbb{R}^d \to \mathbb{R}.
\]  

(3)

We see that if \( j \) is symmetric, i.e. \( d\eta(x, y) = d\eta(y, x) \) for all \( x, y \in \mathbb{R}^d \), then \( \nabla \cdot j = 0 \). In the antisymmetric case \( d\eta(x, y) = -d\eta(y, x) \) for all \( x, y \in \mathbb{R}^d \), \( \nabla \cdot j \) simplifies to

\[
\int_{\mathbb{R}^d} \varphi(x) d\nabla \cdot j(x) = -\int_{\mathbb{R}^d} \varphi(x) d\eta(x, y) d\eta_j(x, y), \quad \forall \varphi : \mathbb{R}^d \to \mathbb{R}.
\]

Our dynamics is driven by the nonlocal cross-interaction energy given as

\[
E(\rho) = \frac{1}{2} \sum_{i,k=1}^{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} K^{(ik)}(x, y) \rho^{(i)}(x) \rho^{(k)}(y),
\]  

(4)

where \( K^{(ik)} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}, i, k = 1, 2 \) are continuous and satisfy

\[
\forall x, y \in \mathbb{R}^d \text{ it holds } K^{(ik)}(x, y) = K^{(ik)}(y, x) \quad \text{(K1)}
\]

To admit a gradient flow w.r.t. the nonlocal operators defined above, even on generalized graphs, the continuity assumption needs to be strengthened (see [26]) to

\[
\exists L_K \in (0, \infty) \text{ such that } \forall (x, y), (x', y') \in \mathbb{R}^d \times \mathbb{R}^d, \text{ and } i, k = 1, 2, \text{ we have}
\]

\[
\left| K^{(ik)}(x, y) - K^{(ik)}(x', y') \right| \leq L_K \left( |(x, y) - (x', y')| \lor |(x, y) - (x', y')|^p \right). \quad \text{(K2)}
\]

However, for the following considerations, this assumption is not required and will therefore be dropped. In our setting, (4) thus becomes

\[
E(\rho) = \frac{1}{2} \sum_{i,k=1}^{2} \sum_{i,k=1}^{N} K^{(ik)}_{i_k} \rho^{(i)}(x_i) \rho^{(k)}(x_k) \mu_i \mu_k, \quad \rho \in (\mathcal{P}_\mu(\mathbb{R}^d))^2,
\]  

(5)

where we used \( K^{(ik)}_{i_k} := K^{(ik)}(x_i, x_k) \) for \( i, k = 1, 2 \) and \( i, k = 1, \ldots, N \).
The dynamical system we are interested in (which can rigorously be understood as a gradient flow in the EDE sense [26]) driven by the nonlocal cross-interaction energy in the nonlocal framework is then given by

\[
\partial_t \rho_t^{(i)} = - (\nabla \cdot j_t^{(i)}),
\]

\[
 j_t^{(i)}(x,y) = m(\rho_t^{(i)}(x),\rho_t^{(i)}(y)) \left[ (v_t^{(i)})_+(x,y) \right]^{q-1}
 - m(\rho_t^{(i)}(x),\rho_t^{(i)}(y)) \left[ (v_t^{(i)})_-(x,y) \right]^{q-1},
\]

\[ (2 \text{ revisited}) \]

where \( i = 1, 2, \) \( 1 < q = p/(p-1) < \infty, \) \( m \) is a suitable mobility function, and where we abused notation by writing \( d j_t^{(i)} = f_t^{(i)} d\mu \otimes \mu. \)

Note the the equations contain a possibly non-linear mobility function which we define as follows.

**Definition 2.1.** Given two thresholds \( R, S \in (0, \infty), \) we call \( m \in C((0, R) \times (0, S)) \) a mobility function if it is concave and strictly positive in \((0, R) \times (0, S).\) For \((r, s) \in [0, R) \times [0, S)\) we denote \( m(r, s) = \lim_{\gamma \to S} m(r, s) \) and \( m(R, s) = \lim_{r \to R} m(r, s). \) We call such a mobility \( m \) upwind-admissible if for every \( s \geq 0 \) we have \( m(0, s) = 0. \)

In the sequel we require \( m \) to satisfy the technical assumption

\[
R < S < \infty \text{ or } \forall r, s \geq 0 : \lim_{A \to \infty} \frac{1}{A} m(\lambda r, \lambda s) = 0 \quad \text{or} \quad \forall s \geq 0 : m(r, s) = 0. \tag{A}
\]

Given such a mobility \( m \) and \( \rho^{(i)} \in \mathcal{P}_\mu(\mathbb{R}^d), i = 1, 2, \) we define

\[
m^{(i)} := m \left( \frac{d\rho^{(i)}}{d\mu}, \frac{d\rho^{(i)}}{d\mu} \right).
\]

As with \( \eta, \) we will often use the notation \( m_{t_k}^{(i)} := m^{(i)}(x_t, x_k). \)

**Remark 2.2.** The assumption \( m(0, s) = 0 \) is crucial as it ensures the non-negativity of \( \rho. \) However, note that this assumption excludes, among the choices, the choice \( m \equiv 1. \)

The assumption \((A)\) ensures that any family \((\rho_t)_{t \in [0,T]}\) solving \((2) \text{ revisited}\) and satisfying \( \text{supp} \rho_0 \subset \text{supp} \mu \) stays supported on \( \text{supp} \mu \) for all times \([26, \text{Proposition 2.35}].\) Since we only consider counting measures \( \mu, \) \( \text{supp} \rho_t \subset \text{supp} \mu \) is equivalent to \( \rho_t \ll \mu. \) Our main examples \( m(r, s) = r^{\theta_1} (1-s)^{\theta_2}; \) \( \theta_1, \theta_2 \in [0, 1] \) all satisfy \((A).\)

Since \( \mu \) is a counting measure each curve \((\rho_t)_{t \in [0,T]} \subset \mathcal{P}_\mu(\mathbb{R}^d)\) can be expressed as

\[
\rho_t^{(i)} = \sum_{i=1}^N \rho_t^{(i)} \mu_i \delta_{x_i},
\]

for suitable functions \( \rho_t^{(i)} : [0, T] \to [0, (\mu_x)^{-1}] \) satisfying \( \sum_{i=1}^N \rho_t^{(i)} \mu_i = 1. \) By slight abuse of notation, we will sometimes write \( \rho^{(i)}(x) := \rho^{(i)}([x]). \) Using these definitions and choosing test functions \( \varphi \) such that \( \varphi(x_k) = \delta_{k=1}, \) in \((3),\) system \((2 \text{ revisited})\) becomes

\[
\frac{d}{dt} \rho_t^{(i)} = \sum_{k=1}^N \left( m_{x_k}^{(i)} \left[ (v_{x_k}^{(i)})_- \right]^{q-1} - m_{x_k}^{(i)} \left[ (v_{x_k}^{(i)})_+ \right]^{q-1} \right) \eta_{x_k} \mu_{x_k}, \tag{6}
\]

for \( i = 1, 2 \) and \( i = 1, \ldots, N, \) where

\[
v_{x_k}^{(i)} = - \beta^{(i)} \sum_{k=1}^N \nabla \left( K^{(ik)} \ast \rho^{(k)} \right)_{x_k} = \beta^{(i)} \sum_{k=1}^N \sum_{A=1}^N \left( K^{(ik)}_{x_A} - K^{(ik)}_{x_A} \right) \rho^{(k)}_{x_A} \mu_{x_A}. \tag{7}
\]
3. Dynamics, energy minimizers, and equilibria on finite graphs

In the first part of this section, we study properties of the interaction kernels yielding aggregation and segregation of energy minimizers, respectively. Subsequently, we link the equilibria of the dynamics to critical points of the energy functional, (4).

3.1. Energy minimizers

Our first result provides conditions on the interaction kernels that lead to aggregation or segregation, respectively, of minimizers of the energy functional (5). It is useful to introduce the notation

\[ D^{(ik)}(x, y) := K^{(ik)}(x, x) - K^{(ik)}(x, y). \]  

Note that when \( K^{(ik)} \) is an attractive kernel, then it satisfies \( D^{(ik)} < 0 \) on \( \mathbb{R}^d \times \mathbb{R}^d \), while repulsive kernels \( K^{(ik)} \) are such that \( D^{(ik)} > 0 \) on \( \mathbb{R}^d \times \mathbb{R}^d \).

**Theorem 3.1** (Conditions for energy optimality of aggregation). Let \( \mu = \sum_{i=1}^{N} \mu_i \delta_{x_i} \) for \( x_1, \ldots, x_N \in \mathbb{R}^d \) and let \( \rho = (\rho^{(1)}, \rho^{(2)}) \) with \( \rho^{(i)} = \sum_{i=1}^{N} \rho^{(i)}_i \mu_i \delta_{x_i} \). Then, the following holds:

a) If the kernels satisfy for every \( x, y \in \mathbb{R}^d \) with \( x \neq y \)

\[ D^{(11)}(x, x) < 0, \quad D^{(22)}(x, x) < 0, \]  

and

\[ D^{(11)}(x, y)D^{(22)}(x, y) > \left(D^{(12)}(x, y)\right)^2, \]  

then any energy-minimizing distribution is of the form \( \rho_\ast = (\delta_{x_{t_1}}, \delta_{x_{t_2}}) \) for two indices \( t_1, t_2 \in \{1, \ldots, N\} \).

b) For either \( i = 1 \) or \( i = 2 \) and every \( x, y \in \mathbb{R}^d \) with \( x \neq y \), there holds

\[ D^{(ii)}(x, y) < 0 \]  

and

\[ K^{(ii)}(x, x) = K^{(ii)}(y, y), \]  

then any energy minimizer \( \rho_\ast \) is such that species \( \rho_\ast^{(i)} \) is aggregated, i.e., \( \rho_\ast^{(i)}(x_i) = \delta_{x_i} \) for an index \( t_i \in \{1, \ldots, N\} \).

c) If for both \( i = 1, 2 \) we have (11) and (12), then again \( \rho_\ast = (\delta_{x_{t_1}}, \delta_{x_{t_2}}) \) for \( t_1, t_2 \in \{1, \ldots, N\} \). If in addition \( K^{(12)} \) satisfies for every \( x, y \in \mathbb{R}^d \) with \( x \neq y \)

\[ D^{(12)}(x, y) < 0, \]  

then \( t_1 = t_2 \). In this case, any choice \( t_1 = t_2 \in \{1, \ldots, N\} \) provides the minimal energy. Conversely, if for every \( x, y \in \mathbb{R}^d \) with \( x \neq y \) we have

\[ D^{(12)}(x, y) > 0, \]  

then \( t_1 \neq t_2 \). In this case, there exist at least two energy minimizing distributions, the second being obtained by interchanging \( t_1 \) and \( t_2 \).

**Remark 3.2.** Note that (12) holds in the important special case, where, for \( i = 1, 2 \), there exist suitable kernels \( \bar{K}^{(ii)} : \mathbb{R}^d \to \mathbb{R} \) such that

\[ K^{(ii)}(x, y) = \bar{K}^{(ii)}(y - x), \]  

for any \( x, y \in \mathbb{R}^d \).
Proof of Theorem 3.1. We prove each case separately.

Ad. 1. We argue by constructing an appropriate competitor. To this end, we fix two indices \( \ell_0, k_0 \in \{1, \ldots, N\} \) and define a shift \( s = (s^{(1)}, s^{(2)})^T \), with \( s^{(i)} \) satisfying \( -\rho^{(i)}_{k_0} \mu_{k_0} \leq s^{(i)} \leq \rho^{(i)}_{\ell_0} \mu_{\ell_0} \). Then, we obtain a modified probability measure, \( \tilde{\rho} \), by setting

\[
\tilde{\rho}^{(i)}(x_i) = \tilde{\rho}_i^{(i)} \mu_i = \begin{cases} 
\rho^{(i)}_{k_0} \mu_{k_0} + s^{(i)}, & i = k_0, \\
\rho^{(i)}_{\ell_0} \mu_{\ell_0} - s^{(i)}, & i = \ell_0, \\
\rho_i^{(i)} \mu_i, & i \notin \{\ell_0, k_0\}.
\end{cases}
\]

Then, using the notation \( D^{(ik)}_{\ell k} = D^{(ik)}(x_i, x_k) \) and defining for \( i, k = 1, 2 \)

\[
a_{ik} := \frac{1}{2} \left( D^{(ik)}_{\ell_0 k_0} + D^{(ik)}_{k_0 \ell_0} \right), \quad b_i := \frac{1}{2} \sum_{k=1}^N \sum_{i=1}^N \mu_i \rho_i^{(k)} \left( D^{(ik)}_{\ell_0} - D^{(ik)}_{k_0} \right),
\]

we see that the difference of the energies of \( \tilde{\rho} \) and \( \rho \) is given by

\[
\mathcal{E}(\tilde{\rho}) - \mathcal{E}(\rho) = \sum_{i=1}^2 \left( b_i s^{(i)} + \frac{1}{2} \sum_{k=1}^N a_{ik} s^{(i)} s^{(k)} \right).
\]

Thus, the optimal competitor is the solution to the constrained quadratic optimization problem

\[
\begin{aligned}
\text{minimize} & \quad f(s) = \frac{1}{2} s^T A s + b^T s \\
\text{s.t.} & \quad -\rho^{(1)}_{k_0} \mu_{k_0} \leq s^{(1)} \leq \rho^{(1)}_{\ell_0} \mu_{\ell_0}, \\
& \quad -\rho^{(2)}_{k_0} \mu_{k_0} \leq s^{(2)} \leq \rho^{(2)}_{\ell_0} \mu_{\ell_0},
\end{aligned}
\]

where \( A = (a_{ij})_{i,j=1,2} \) and \( b = (b_1, b_2)^T \) as in (14). Since \( s = (0,0)^T \) is admissible, any minimizer of (16) does not increase the energy.

Now, aggregation occurs whenever the minimizer lies on a corner point of the admissible rectangle. This happens when the problem is strictly concave, i.e., whenever \( A \) is negative definite. This is the case if

\[
a_{11} + a_{22} < -\sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2}.
\]

Hence, \( a_{ii} < 0 \) is required for both \( i = 1, 2 \) and additionally \( a_{11} a_{22} > a_{12}^2 \), which holds if and only if (9), (10) are satisfied.

Ad. 2. Requiring (11) in conjunction with (12) for either \( i = 1 \) or \( i = 2 \), we obtain \( \rho^{(i)}_{s^i} = \delta_{x_{si}} \). Indeed, due to the continuity of \( K^{(ik)} \), equations (12) and (13) are equivalent to

\[
K^{(ii)}(x, x) < K^{(ii)}(y, z), \quad \text{for all } x, y, z \in \mathbb{R}^d \text{ with } y \neq z.
\]

To see this, we fix \( \rho \) as before and, without loss of generality, we assume \( i = 1 \). Next, we choose

\[
\lambda_\ast \in \arg \min_{\kappa \in \{1, \ldots, N\}} \left( \sum_{i=1}^N K_{\kappa i}^{(12)} \rho^{(1)}_{\kappa} \mu_i \right),
\]

which exists, but need not be unique. We define a competitor \( \tilde{\rho} \) via \( \tilde{\rho}^{(1)} = \delta_{\lambda} \) and \( \tilde{\rho}^{(2)} = \rho^{(2)} \). Then, we have

\[
\mathcal{E}(\tilde{\rho}) = \frac{1}{2} \sum_{i,k=1}^N K^{(ik)}_{i k} \rho_i^{(1)} \rho_k^{(k)} \mu_i \mu_k
\]

\[
= \frac{1}{2} \sum_{i,k=1}^N \left( K^{(11)}_{i k} \rho_i^{(1)} \rho_k^{(1)} + K^{(12)}_{i k} \rho_i^{(2)} \rho_k^{(2)} + 2K^{(12)}_{i k} \rho_i^{(1)} \rho_k^{(2)} \right) \mu_i \mu_k
\]

\[
\geq \frac{1}{2} K^{(11)}_{i i} + \sum_{i=1}^N K^{(12)}_{i i} \mu_i + \frac{1}{2} \sum_{i,k=1}^N K^{(22)}_{i k} \rho_i^{(2)} \mu_i \mu_k = \mathcal{E}(\tilde{\rho}).
\]

6
We achieve equality if and only if \( \rho^{(i)} = \delta_i \) for some \( i \in \{1, \ldots, N\} \), which also satisfies (19). Hence, any energy minimizer is such that species 1 is aggregated. The same argument can be made for \( i = 2 \), if \( K^{(22)} \) satisfies (18).

Ad. 3. If (18) holds for both \( i = 1 \) and \( i = 2 \), then a calculation similar to (20) shows that \( p_\star = (\delta_{t_1}, \delta_{t_2}) \)

\[(t_1, t_2) \in \arg \min_{(i, \kappa) \in \{1, \ldots, N\} \times \{1, \ldots, N\}} K^{(12)}_{i \kappa},\]

defines an energy-minimizing distribution. Finally, to distinguish between the cases \( t_1 = t_2 \), and \( t_1 \neq t_2 \), we substitute \( \rho_\star = (\delta_{x_{t_1}}, \delta_{x_{t_2}}) \) into the definition of \( E(\rho) \) to find

\[E(\rho_\star) = \frac{1}{4} \left( K^{(11)}_{i t_1} + K^{(22)}_{t_2 t_2} + 2K^{(12)}_{i t_2} \right).\]

Due to (12), the first two terms are independent of the choice of \( t_1 \) and \( t_2 \). Thus, if \( D^{(12)} < 0 \), then (18) holds and the last term is optimal for \( t_1 = t_2 \). Conversely, when \( D^{(12)} < 0 \), then, necessarily it must hold \( t_1 \neq t_2 \) to achieve optimality in the last term.

\[\square\]

**Remark 3.3** (Energy optimality of segregation). A sufficient condition yielding energy optimality of a fully segregated distribution without assuming aggregation of the two species individually, is

\[
\min_{1 \leq i, k \leq N} \left[ K^{(12)}_{i t} - K^{(12)}_{i k} \right] > \frac{1}{2} \sum_{i=1}^{2} \left( \max_{1 \leq i, \kappa \leq N} K^{(i i)}_{i \kappa} - \min_{1 \leq i, \kappa \leq N} K^{(i i)}_{i \kappa} \right).
\]

However, for this to hold for arbitrary graphs requires, using the continuity of the kernels, that the self-interaction kernels \( K^{(ii)} \), \( i = 1, 2 \) have to be constants.

### 3.2. Dynamics and stationary states

Let us proceed to studying properties of solutions to the system ((2) revisited). We are particularly interested in the link between energy minimizers and stationary states of the gradient system. To this end, recall that the velocity \( v = (v^{(1)}, v^{(2)}) \) is defined as

\[v^{(i)}_{i \kappa} = -\beta^{(i)} \sum_{k=1}^{N} \nabla \left( K^{(i i)}_{i k} \ast \rho^{(k)} \right) = \beta^{(i)} \sum_{k=1}^{N} \sum_{j=1}^{N} \left( K^{(i k)}_{i j} - K^{(j k)}_{i j} \right) P_{i j}^{(k)} \mu_j.\]

Therefore, by the symmetry of \( \eta \), for arbitrary \( f : \text{supp} \mu \to \mathbb{R}^n \), \( n \in \mathbb{N} \), and for \( i = 1, 2 \) we obtain

\[
\sum_{i=1}^{N} f_i \mu_i \frac{d}{dt} \rho^{(i)}_t = \sum_{i, \kappa=1}^{N} f_i \left( m^{(i)}_{i \kappa} \left[ (v^{(i)})_+ \right]^{q-1} - m^{(i)}_{i \kappa} \left[ (v^{(i)})_+ \right]^{q-1} \right) \eta_{i \kappa} \mu_i \mu_t.
\]

Choosing \( f \equiv 1 \), we have that \( f_\kappa - f_i = 0 \) for all \( \kappa, i \in \{1, \ldots, N\} \) and thus conservation of mass is guaranteed.

The second question is, whether the center of mass is preserved by the dynamics, which corresponds to the choice \( f \equiv x \). The following result provides a counterexample.
We expect the (nonpositive) dissipation to vanish at solutions to the stationary equation, which, indeed, is equivalent to

\[ \frac{d}{dt} x_c = 2 \left( \frac{2}{3} (K(2, -1) - K(-1, -1)) + \frac{1}{3} (K(2, 2) - K(-2, 2)) \right)_- \]

which vanishes if and only if

\[ \frac{2}{3} (K(2, -1) - K(-1, -1)) + \frac{1}{3} (K(2, 2) - K(-2, 2)) = 0. \]

Since we can choose these terms independently without violating (K1) or even (K2), we find that, in general, \( x_c \) is not conserved in time. Indeed, this statement is true even for a single species. The reason the preservation of the center of mass breaks down is the lack of symmetry of the upwind scheme in conjunction with the choice of the underlying graph as well as the structure of the kernel \( K \).

However, supposing \( K(x, x) = K(y, y) \), for all \( x, y \in \mathbb{R}^d \), and upon replacing the upwind relation \( j(x, y) = \rho(x)v_+(x, y) - \rho(y)v_-(x, y) \) by a symmetric relation of the form \( j(x, y) = \theta(\rho(x), \rho(y)v(x, y)) \), with \( \theta(x, y) = \theta(y, x) \), the center of mass is conserved in time. The same is true for volume schemes without upwind [13, Lemma 4.3].

On the other hand, center of mass conservation results are also valid in the particular case of finite volume upwind schemes, i.e. when symmetry is provided by the graph (see, e.g., [18, Lemma 3.2]).

Another important choice is \( f_i = \sum_{k=1}^2 \sum_{l,k=1}^N K_{i,k}^{(ik)} \rho_{A}^{(k)} \mu_A \), which, if we sum over \( i = 1, 2 \), yields the dissipation of the energy \( E(p) \). We have

\[
\frac{d}{dt} E(p) = \sum_{i,l,k=1}^N \sum_{l,k=1}^N \left( \frac{1}{\beta^{(i)}} m_{i,k}^{(i)} (v^{(i)}_{i,k})_+^q - m_{i,k}^{(i)} (v^{(i)}_{i,k})_-^q \right) \eta_{i,k} \mu_k \mu_i \leq 0.
\]

We expect the (nonpositive) dissipation to vanish at solutions to the stationary equation

\[ 0 = \sum_{k=1}^N \left( m_{i,k}^{(i)} (v^{(i)}_{i,k})_+^q - m_{i,k}^{(i)} (v^{(i)}_{i,k})_-^q \right) \eta_{i,k} \mu_k, \quad i = 1, \ldots, N, \]

which, indeed, is equivalent to

\[ \sum_{k=1}^N m_{i,k}^{(i)} (v^{(i)}_{i,k})_+^q - \sum_{k=1}^N m_{i,k}^{(i)} (v^{(i)}_{i,k})_-^q = \sum_{k=1}^N m_{i,k}^{(i)} (v^{(i)}_{i,k})_+^q - \sum_{k=1}^N m_{i,k}^{(i)} (v^{(i)}_{i,k})_-^q, \quad i = 1, \ldots, N. \]

From its definition (7), we see that the velocity \( v^{(i)}(x, y) \) is antisymmetric with respect to its two arguments \( x, y \in \mathbb{R}^d \) and also additive, i.e., for \( x, y, z \in \mathbb{R}^d \) we have

\[ v^{(i)}(x, z) = v^{(i)}(x, y) + v^{(i)}(y, z). \]

This observation allows for the following characterization:
4.1. Dynamics on a two-point space

To gain a better understanding of the nature of the gradient flow, we analyze the dynamics on a two-point space, supp μ = {x₁, x₂}, which shall act as a starting point for simulations, below. In addition, we assume that for \( i, k = 1, 2 \) and all \( x, y \in \mathbb{R}^d \) we have
\[
K^{(ik)}(x, x) = K^{(ik)}(y, y),
\]
and that for \( i, k = 1, 2 \) the mobility satisfies
\[
m^{(i)}_{ik} = 0 \iff \rho^{(i)}_i = 0.
\]

The starting point of our considerations is an explicit formulation of all possible stationary states assuming \( \eta_{ik} > 0 \) (as otherwise, there is no interaction between the vertices). We recall the definition (8) of the differences \( D^{(ik)}(x, y) \), which reads on the graph as
\[
D^{(ik)}_{tx} := K^{(ik)}_t - K^{(ik)}_{tx}.
\]
By (27) and the symmetry \((K1)\), we have
\[
D_{12}^{(ik)} = K_{11}^{(ik)} - K_{12}^{(ik)} = K_{22}^{(ik)} - K_{21}^{(ik)} = D_{21}^{(ik)}.
\]
This motivates the notation \(D^{(ik)} := D_{12}^{(ik)} = D_{21}^{(ik)}\) that is used for the remainder of this subsection and in Appendix A.

**Decoupled case.** Let us first consider the case \(D^{(12)} = 0\), in which case the equations for the two species are decoupled. Then, it suffices to consider only one species \(i \in \{1, 2\}\). By Lemma 3.5 and (27) there are two possible ways of obtaining a stationary state:

a) \(v_{i\kappa}^{(i)} = 0\),

b) \(v_{i\kappa}^{(i)} > 0, \rho_{i}^{(i)} = 0\),

where \(i, \kappa \in \{1, 2\}\) and \(i \neq \kappa\). The corresponding stationary states are characterized explicitly in Proposition 4.1.

**Proposition 4.1 (Two-point stationary states - decoupled case).** Let \(N = 2\), let (26) hold, and let \(D^{(12)} = 0\). Then, for \(i, \kappa \in \{1, 2\}\) with \(i \neq \kappa\), we have

a) If \(D^{(ii)} = 0\), then \(v_{i\kappa}^{(i)} = 0\) holds for every \(\rho^{(i)}(x_1) = 1 - \rho^{(i)}(x_2) \in [0, 1]\).
   Conversely, if \(D^{(ii)} \neq 0\), then \(v_{i\kappa}^{(i)} = 0\) is achieved only at \(\rho^{(i)}(x_1) = \rho^{(i)}(x_2) = \frac{1}{2}\).

b) \(D^{(ii)} < 0\) is equivalent to being able to satisfy \(v_{i\kappa}^{(i)} > 0\) and \(\rho_{i}^{(i)}(x_i) = 0\) simultaneously for either \(i = 1\) or \(i = 2\).

The proof of this proposition and the remaining results in this subsection is straightforward, yet technical, and can be found in the appendix A.

Having derived stationary states and conditions for their existence, next we describe their stability.

**Proposition 4.2 (Dynamics - Decoupled case).** Let \(N = 2\), let (26) and (27) hold, and let \(D^{(12)} = 0\). Let \(\rho_0^{(i)}(x_i) = 1 - \rho_0^{(i)}(x_\kappa) = \frac{1}{2} + \varepsilon\) with \(\varepsilon \in [0, 1/2]\) and \(i, \kappa \in \{1, 2\}, i \neq \kappa\). Then, it holds

1. If \(D^{(ii)} > 0\), then \(\rho_{i}^{(i)}(x_i) \to 1/2\) as \(t \to \infty\) for every \(\varepsilon \in [0, 1/2]\).

2. If \(D^{(ii)} < 0\), then
   - If \(\varepsilon = 0\), then \(\rho_{i}^{(i)}(x_i) = \rho_{i}^{(i)}(x_\kappa) = 1/2\), for all \(t \geq 0\).
   - If \(\varepsilon \in (0, 1/2)\), then \(\rho_{i}^{(i)}(x_i) \to 1\) and \(\rho_{i}^{(i)}(x_\kappa) \to 0\), as \(t \to \infty\).

**Remark 4.3.** Proposition 4.1 shows that, in the decoupled case, species \(i\) aggregates on a single vertex if and only if \(D^{(ii)} \leq 0\). By Proposition 4.2 we see that these aggregated states are asymptotically stable, whenever \(D^{(ii)} > 0\). In particular, this holds for attractive potentials \(K^{(ii)}\). Conversely, states in which the mass of species \(i\) is distributed equally on the two vertices are always a stationary. However, this type of stationary state is only stable, when \(D^{(ii)} < 0\), which is in particular true for repulsive potentials \(K^{(ii)}\).

**Coupled case.** Next, let us assume \(D^{(12)} \neq 0\), i.e., there is a coupling. Then, by Lemma 3.5, there are four different conditions leading to stationary states:

a) \(v_{i\kappa}^{(1)} = v_{i\kappa}^{(2)} = 0\),

b) \(v_{i\kappa}^{(i)} > 0\) and \(\rho_{i}^{(i)} = 0\) as well as \(v_{i\kappa}^{(k)} = 0\),

c) \(v_{i\kappa}^{(1)} > 0\) and \(v_{i\kappa}^{(2)} > 0\) as well as \(\rho_{i}^{(1)} = \rho_{i}^{(2)} = 0\),

d) \(v_{i\kappa}^{(1)} > 0\) and \(v_{i\kappa}^{(2)} < 0\) as well as \(\rho_{i}^{(1)} = 0\) and \(\rho_{i}^{(2)} = 1\),

for \(i, k, \kappa \in \{1, 2\}\) with \(i \neq k\) and \(i \neq \kappa\). In the simplified setting, these stationary states can also be characterized explicitly as shown in the following proposition.
**Proposition 4.4** (Stationary states - Coupled case). Let $N = 2$, let (26) and (27) hold, and let $D^{(12)} \neq 0$. Then, for $i, \kappa \in \{1, 2\}$ with $i \neq \kappa$, we have

a) A probability measure satisfying a) exists.
If $D^{(11)}D^{(22)} \neq (D^{(12)})^2$, then it is unique and given by
$$
\rho_a^{(1)}(x_1) = \rho_a^{(1)}(x_2) = \frac{1}{2}, \quad \text{and} \quad \rho_a^{(2)}(x_1) = \rho_a^{(2)}(x_2) = \frac{1}{2}.
$$

If $D^{(11)}D^{(22)} = (D^{(12)})^2$, then the probability measures
$$
\rho_a^{(1)}(x_1) = \frac{1}{2}(1 + r), \quad \rho_a^{(1)}(x_2) = \frac{1}{2}(1 - r),
$$
$$
\rho_a^{(2)}(x_1) = \frac{1}{2}\left(1 - \frac{D^{(12)}}{D^{(22)}r}\right), \quad \rho_a^{(2)}(x_2) = \frac{1}{2}\left(1 + \frac{D^{(12)}}{D^{(22)}r}\right),
$$
where $r \in \left[-1 \lor \frac{D^{(22)}}{D^{(12)}}, 1 \land \frac{D^{(22)}}{D^{(12)}}\right]$ all satisfy a).

b) A probability measure satisfying b) exists if and only if for one $i, k \in \{1, 2\}, k \neq i$ we have
$$
|D^{(12)}| \leq |D^{(kk)}|.
$$
and
$$
D^{(ii)} < \frac{(D^{(12)})^2}{D^{(kk)}}.
$$

Then, it is given by
$$
\rho_{b_i}^{(i)}(x_i) = 1, \quad \rho_{b_i}^{(i)}(x_k) = 0,
$$
$$
\rho_{b_k}^{(k)}(x_i) = \frac{1}{2}\left(1 - \frac{D^{(12)}}{D^{(kk)}}\right), \quad \rho_{b_k}^{(k)}(x_k) = \frac{1}{2}\left(1 + \frac{D^{(12)}}{D^{(kk)}}\right).
$$
In this case, there exists a second stationary state of the same form since $i$ and $k$ are interchangeable.

c) A probability measure satisfying c) exists if and only if
$$
D^{(11)} < -D^{(12)}, \quad D^{(22)} < -D^{(12)}.
$$
Then, it is given by
$$
\rho_c^{(1)}(x_i) = 1, \rho_c^{(1)}(x_k) = 0 \text{ and } \rho_c^{(2)}(x_i) = 1, \rho_c^{(2)}(x_k) = 0.
$$
In this case, since $i$ and $k$ can be swapped, a second stationary state of the same type exists.

d) A probability measure satisfying d) exists if and only if
$$
D^{(11)} < D^{(12)}, \quad D^{(22)} < D^{(12)}.
$$
Then, it is given by
$$
\rho_d^{(1)}(x_i) = 1, \rho_d^{(1)}(x_k) = 0 \text{ and } \rho_d^{(2)}(x_i) = 0, \rho_d^{(2)}(x_k) = 1.
$$
Again, there is a second stationary state of the same form as $i$ and $k$ can be swapped.

Having established these existence conditions, we make the following observations:

**Corollary 4.5.** Let $N = 2$, let (26) and (27) hold, and let $D^{(12)} \neq 0$. Then, $\rho_a$ is the only stationary state if and only if
$$
D^{(11)} > 0, D^{(22)} > 0, \quad D^{(11)}D^{(22)} > (D^{(12)})^2.
$$
All existing stationary states are of the form $\rho_a$ as in (28) if and only if
$$
D^{(11)} > 0, D^{(22)} > 0, \quad D^{(11)}D^{(22)} \geq (D^{(12)})^2.
$$
Conversely, it remains to consider the case of stability of the respective stationary states. In this case, it is important to keep in mind that since the distributions are given by 

\[ \rho(b) = \frac{1}{2}, \quad \rho(a) = \frac{1}{2} \]

Remark 4.6. Note that (33) turns the corresponding minimizing problem (16) into a strictly convex problem with unique minimizer \(\rho_a\), while (34) turns (16) into a convex problem with minimizers \(\rho_{ab}\).

Corollary 4.7. Let \(N = 2\), let (26) and (27) hold, and let \(D(12) \neq 0\). If \(\hat{\rho}_c = \hat{\rho}_d = 0\), then we also have \(\hat{\rho}_{b1} = \hat{\rho}_{b2} = 0\).

Conversely, \(\hat{\rho}_{b1} = \hat{\rho}_{b2} = 0\) implies that either \(\hat{\rho}_c = \hat{\rho}_d = 0\) holds or that we have \(\hat{\rho}_a = 0\) and \(D(iii) = -|D(12)|\) for one \(i \in \{1, 2\}\) and one \(\alpha \in \{a, b\}\).

Having derived stationary states and conditions for their existence, we now want to learn more about their stability. Let us study the stability of the different stationary states in the case \(D(12) \neq 0\) and \(D(11)D(22) \neq (D(12))^2\). In this case, up to symmetry there are only four distinct stationary states. As seen in Proposition 4.4, the stationary distributions are given by

\[
\begin{align*}
\rho_a^{(1)}(x_i) &= \frac{1}{2}, \quad \rho_a^{(2)}(x_i) = \frac{1}{2}, \\
\rho_{b1}^{(i)}(x_i) &= 1, \quad \rho_{b1}^{(k)}(x_i) = \frac{1}{2} \left(1 - \frac{D(12)}{D(22)}\right), \\
\rho_c^{(1)}(x_i) &= 1, \quad \rho_c^{(2)}(x_i) = 1, \\
\rho_{d1}^{(1)}(x_i) &= 1, \quad \rho_{d1}^{(2)}(x_i) = 0.
\end{align*}
\]

Here it is important to keep in mind that since \(\epsilon = 1\) and \(\epsilon = 2\) are both possible, \(\rho_{b1}, \rho_{b2}, \rho_c\) and \(\rho_d\) each characterize a pair of two distinct stationary states.

Proposition 4.8 (Stability I). Let \(N = 2\), let (26) and (27) hold, and assume \(D(12) \neq 0\) as well as \(D(11)D(22) \neq (D(12))^2\). Then, it holds

a) \(\rho_a\) is asymptotically stable if and only if it is the only stationary state.

b) If \(\rho_{b1}\) with \(i \in \{1, 2\}\) are stationary, then they are both asymptotically stable if and only if \(\rho_{b2}\) with \(k = 3 - i\) is not stationary.

c) If \(\rho_c\) are stationary, then they are asymptotically stable.

d) If \(\rho_d\) are stationary, then they are asymptotically stable.

Proof. This is a consequence of the Remarks A.3, A.4 and A.5, which give a more nuanced insight into the stability of the respective stationary states.

It remains to consider the case \(D(11)D(22) = (D(12))^2\) where the following holds true.

Proposition 4.9 (Stability II). Let \(N = 2\), let (26) and (27) hold, and assume \(D(12) \neq 0\) as well as \(D(11)D(22) = (D(12))^2\). Then, the stationary states \(\rho_a\) are stable, but not asymptotically stable.

1. If \(D(iii) > 0\) for \(i = 1, 2\), then no other stationary state exists.

2. If \(D(iii) < 0\) for \(i = 1, 2\) and \(D(12) < 0\), then the stationary states \(\rho_c\) are asymptotically stable.

3. If \(D(iii) < 0\) for \(i = 1, 2\) and \(D(12) > 0\), then the stationary states \(\rho_d\) are asymptotically stable.

Proof. This is a consequence of the Remarks A.3, A.4 and A.5, which give a more nuanced insight into the stability of the respective stationary states.

Remark 4.10. Summarizing the results for the coupled case, we see the following:
• If both self-interactions are attractive (which implies $D^{(i)} < 0$ for $i = 1, 2$), then Proposition 4.4 shows that the stationary states with the lowest energy are such that the species are both aggregated. If $D^{(12)} < 0$, the lowest energy is achieved for aggregation of both species at the same vertex. If, however, $D^{(12)} > 0$, the lowest energy is achieved for aggregation at opposite vertices. This is true both in the regime of dominant self-interactions, i.e., when $D^{(11)}D^{(22)} \geq (D^{(12)})^2$, and in the regime of dominant cross-interactions, i.e. when $D^{(11)}D^{(22)} < (D^{(12)})^2$.

• In the regime of dominant self-interactions $D^{(11)}D^{(22)} \geq (D^{(12)})^2$ Proposition 4.4 implies that the sign of the self-interactions, which is the same for the two species, dictates whether the stationary states with the lowest energy aggregate on a vertex ($D^{(i)} < 0$ for $i = 1, 2$) or are distributed uniformly, ($D^{(i)} > 0$ for $i = 1, 2$).

• In the regime of dominant cross-interaction ($D^{(11)}D^{(22)} < (D^{(12)})^2$) Proposition 4.4 yields that the cross-interactions dictate that the optimal stationary states are such that at least one species is aggregated – even for two repulsive self-interactions.

• Regarding the stability, Proposition 4.8 shows that when $D^{(11)}D^{(22)} \neq (D^{(12)})^2$, then optimal stationary states are asymptotically stable. Further, it shows that when aggregation of both species at the same vertex as well as aggregation of the species at opposite vertices are stationary, then all these states are asymptotically stable. Finally, we see that when $D^{(11)}D^{(22)} \neq (D^{(12)})^2$, then stationary states that are not energetically optimal or such that both species are aggregated, are unstable.

• Proposition 4.4 shows that when the self- and cross-interactions are balanced ($D^{(11)}D^{(22)} \neq (D^{(12)})^2$), then there exists a continuum of stationary states. By Proposition and 4.9, these are stable but not asymptotically stable when the self-interactions are repulsive. Conversely, when the self-interactions are attractive, then they are unstable, and there exist two asymptotically stable stationary states, where both species are aggregated either at the same vertex ($D^{(12)} < 0$) or opposite vertices ($D^{(12)} > 0$).

**Example 4.11 (Volume filling mobility).** The volume filling mobility $m^{(i)}_{ik} = \rho^{(i)}(1 - \rho^{(i)})$, is covered by the considerations of this section, if $\mu_1, \mu_2 \leq 1$. On the other hand, if $\mu_1, \mu_2 > 2$, then the set of admissible distributions for the dynamics is reduced. For example, in the case $\mu_1 = \mu_2 = 2$, the only admissible distribution is $\rho^{(1)}(x_1) = \rho^{(1)}(x_2) = \rho^{(2)}(x_1) = \rho^{(2)}(x_2) = 1/2$. If $1 < \mu_1, \mu_2 < 2$, then there are more admissible distributions, but the species still cannot aggregate in a single vertex, even if the interaction kernels are attractive, i.e., if $D^{(ik)} < 0$ for $i, k = 1, 2$. If $(\mu_1)^{-1} + (\mu_2)^{-1} < 1$, then no admissible distribution exists.

If $\mu_1, \mu_2 = 1$, then
\[
m^{(i)}_{ik} = \begin{cases} \rho^{(i)}(1 - \rho^{(i)}), & \text{if } \rho^{(i)} \in [0, 1], \\ -\infty, & \text{if } \rho^{(i)} \notin [0, 1], \end{cases}
\]
is also an admissible, though rather unnatural mobility map. With this mobility $\rho_c$ as well as $\rho_A$ are always stationary states, since the mobility vanishes at these points, even if the self interaction kernels are repulsive.

**Remark 4.12 (Visualization).** In Figures 1 - 5, we illustrate the effects of varying different parameters of the two-point model on the energy landscapes (heat maps), stationary states (light blue circles) and corresponding energy level sets (light blue lines), as well as dynamics (quivers) of the two-point space. Here, the mass $\rho^{(1)}(x_1) = 1 - \rho^{(1)}(x_\kappa)$ is displayed on the x-axis, while the mass $\rho^{(2)}(x_1) = 1 - \rho^{(2)}(x_\kappa)$ is displayed on the y-axis. We focus only on the case of linear mobility $m^{(i)}_{ik} = \rho^{(i)}$ for $i = 1, 2$.

• In Figure 1 we vary the cross-interactions while staying in the regime of strong attractive self-interactions ($D^{(11)}D^{(22)} < 0, D^{(11)}D^{(22)} > (D^{(12)})^2$). We observe that it the cross-interaction is attractive (A), the stationary states $\rho_c$ are optimal. If the cross-interaction is zero (B), both $\rho_c$ and $\rho_A$ are optimal. If the cross-interaction is repulsive (C), $\rho_A$ is optimal.
Figure 1.: Fixing $\beta^{(1)} = \beta^{(2)}$, $p = 2$ and $m_{ik}^{(i)} = \rho_{i}^{(i)}$, we illustrate the energy landscapes (heat map), stationary states (light blue circles), their corresponding energy level sets (light blue lines), and dynamics (quivers) for different cross-interaction, while staying in the regime $D^{(11)}/D^{(22)} < 0$, $D^{(11)} D^{(22)} > (D^{(12)})^2$.

Figure 2.: Fixing $\beta^{(1)} = \beta^{(2)}$, $p = 2$ and $m_{ik}^{(i)} = \rho_{i}^{(i)}$, we display the energy landscapes (heat map), stationary states (light blue circles), their corresponding energy level sets (light blue lines), and dynamics (quivers) in the different regimes $D^{(11)}/D^{(22)} \lessgtr (D^{(12)})^2$ for repulsive (top row) and attractive (bottom row) self-interactions.

- In Figure 2 we compare the different regimes $D^{(11)}/D^{(22)} \lessgtr (D^{(12)})^2$ for repulsive (top row) and attractive (bottom row) self-interactions. In (A) and (D) the cross-interaction is stronger than the self-interactions, which leads to $\rho_{b}$ being optimal, both for repulsive and attractive self-interactions. In (B) the cross- and self-repulsions balance each other out, leading to a family of stationary states of the form $\rho_{a}$ on the line $\rho^{(1)}(x) = 1 - \rho^{(2)}(x)$. Conversely, in (E) the cross-repulsions and self-attractions balance each other out on the line $\rho^{(1)}(x) = \rho^{(2)}(x)$, thus also leading to a family of stationary states of the form $\rho_{a}$. Outside this line, the self-attractions and cross-repulsions draw the dynamics towards $\rho_{a}$. In (C) the self-repulsions dominate the energy landscape, thus pushing the dynamics towards the only stationary state, $\rho_{a}$. Conversely, in (F) the self-attractions dominate, leading to $\rho_{c}$ and $\rho_{d}$ being asymptotically stable, while $\rho_{a}, \rho_{b_{1}}$, and $\rho_{b_{2}}$ are unstable.

- In Figure 3 we vary the strength of the self-interactions, which have opposite signs, while fixing repulsive cross-interactions. In (A) the self-interaction of species 1 is repulsive, leading to $\rho_{b_{1}}$ being asymptotically
Figure 3.: Fixing $\beta^{(1)} = \beta^{(2)}$, $p = 2$ and $m^{(i)}_{ik} = \rho^{(i)}$, we illustrate the energy landscapes (heat map), stationary states (light blue circles), their corresponding energy level sets (light blue lines), and dynamics (quivers), for self-interactions of opposite signs and varying strength, while the cross-interactions are fixed and repulsive.

Figure 4.: Fixing $p = 2$ and $m^{(i)}_{ik} = \rho^{(i)}$, $D^{(11)} = D^{(22)} = 1$ and $D^{(12)} = 0$, we display the energy landscapes (heat map), stationary states (light blue circles), and dynamics (quivers) for varying $\beta^{(1)}$ and $\beta^{(2)}$.

Figure 5.: Fixing $m^{(i)}_{ik} = \rho^{(i)}$, $D^{(11)} = D^{(22)} = 1$, $D^{(12)} = 0$ and $\beta^{(1)} = \beta^{(2)}$, we display the energy landscapes (heat map), stationary states (light blue circles), and dynamics (quivers) for varying $p$.

- In Figure 4 we display the effects of varying $\beta^{(1)}$ and $\beta^{(2)}$. We observe that the energy landscape and stationary states remain unchanged, while the ratio between $\partial_t \rho^{(1)}$ and $\partial_t \rho^{(2)}$ changes with the ratio $\beta^{(1)}$ and $\beta^{(2)}$, achieving symmetry between the species, when $\beta^{(1)} = \beta^{(2)}$. 

stable. Conversely, in (C) the self-interaction of species 2 is repulsive, leading to $\rho_{b2}$ being asymptotically stable. In (B) both self-interactions vanish, leading to $\rho_d$ being asymptotically stable, similar to Figure 2(A,D). 

• In Figure 4 we display the effects of varying $\beta^{(1)}$ and $\beta^{(2)}$. We observe that the energy landscape and stationary states remain unchanged, while the ratio between $\partial_t \rho^{(1)}$ and $\partial_t \rho^{(2)}$ changes with the ratio $\beta^{(1)}$ and $\beta^{(2)}$, achieving symmetry between the species, when $\beta^{(1)} = \beta^{(2)}$. 

15
In Figure 5 we illustrate the effects of varying $p$. Similar to Figure 4, we observe that the energy landscape and stationary states remain unchanged, while $\partial_i \rho^{(1)}$ and $\partial_i \rho^{(2)}$ change with $p$.

4.2. The effects of the cut-off function $\eta$ on a 3-point space

Consider three points on a straight line, i.e., $\mu = \delta_{x_1} + \delta_{x_2} + \delta_{x_3}$, where $x_1 = -r_1$, $x_2 = 0$, $x_3 = r_2$ for some $r_1, r_2 > 0$. Let $K^{(i)}(x, y) = |x - y|$ for $i = 1, 2$ and $K^{(12)}(x, y) = -\alpha |x - y|$, for some $\alpha \geq 0$. Besides, assume $m^{(i)}_{ik} = \rho^{(i)}_k$ for $i = 1, 2$ and $i, k = 1, 2, 3$. It is easy to check that the energy minimum is achieved, when one species is aggregated at $x_1$, while the other is aggregated at $x_3$.

However, when we impose the initial distribution $\rho^{(1)}_1 = \rho^{(1)}_3 = 0$, $\rho^{(1)}_2 = 1$ and $\rho^{(1)}_1 = 1 - \rho^{(1)}_3 = \frac{1}{2}(1 + \delta)$, $\rho^{(2)}_2 = 0$, for some $\delta \in (0, 1)$, then the velocities given by

$$v^{(i)}_{ik} = \sum_{k=1}^{2} \sum_{\ell=1}^{3} (K^{(i)}_{k\ell} - K^{(i)}_{k,\ell}) \rho^{(k)}_{\ell},$$

can be calculated to

$$v^{(1)}_{12} = r_1(\alpha + 1), \quad v^{(1)}_{13} = r_1 - r_2 + \alpha \delta (r_1 + r_2), \quad v^{(1)}_{23} = r_2(\alpha - 1),$$

$$v^{(2)}_{12} = r_2(-\alpha - \delta), \quad v^{(2)}_{13} = \alpha r_2 - \alpha r_1 - \delta (r_1 + r_2), \quad v^{(2)}_{23} = r_2(\alpha - \delta).$$

Based on this, the time derivatives become

$$\partial_i \rho^{(i)}_1 = \frac{1}{\beta^{(i)}} \sum_{k=1}^{3} \eta_{ik} \left[(v^{(i)}_{ik})_+ - (v^{(i)}_{ik})_- \rho^{(i)}_1 \right],$$

whence

$$\partial_1 \rho^{(1)}_1 = [r_1(1 + \alpha \delta)]_+ \eta_{12},$$

$$\partial_2 \rho^{(1)}_1 = - [r_1(1 + \alpha \delta)]_- \eta_{12} - [r_2(1 - \alpha \delta)]_- \eta_{23},$$

$$\partial_3 \rho^{(1)}_1 = [r_2(1 - \alpha \delta)]_+ \eta_{23},$$

$$\partial_2 \rho^{(2)}_1 = \frac{1}{2\beta} \left[(r_1 - r_2) + \alpha \delta (r_1 + r_2)\right]_- (1 - \delta) \eta_{13}$$

$$- \frac{1}{2\beta} \left[(r_1 - r_2) + \alpha \delta (r_1 + r_2)\right]_+ (1 + \delta) \eta_{13} - \frac{1}{2\beta} [r_1(\alpha + \delta)]_- (1 + \delta) \eta_{12},$$

$$\partial_3 \rho^{(2)}_1 = \frac{1}{2\beta} [r_1(\alpha + \delta)]_- (1 + \delta) \eta_{12} + \frac{1}{2\beta} [r_2(\alpha - \delta)]_- (1 - \delta) \eta_{23},$$

$$\partial_3 \rho^{(2)}_1 = \frac{1}{2\beta} \left[(r_1 - r_2) + \alpha \delta (r_1 + r_2)\right]_+ (1 + \delta) \eta_{13}$$

$$- \frac{1}{2\beta} \left[(r_1 - r_2) + \alpha \delta (r_1 + r_2)\right]_- (1 - \delta) \eta_{13} - \frac{1}{2\beta} [r_2(\alpha - \delta)]_- (1 - \delta) \eta_{23}.$$ (35)

Let us consider some interesting particular cases of this situation and look at the effect of different supports of $\eta$.

Cutting off the interaction between $x_1$ and $x_3$. First, consider the situation, where $\eta(x, y) = \mathbb{1}_{\{|x-y|<r\}}$ for some $r < r_1 + r_2$. By (35) we see that then $\partial_1 \rho^{(1)}_1 = \partial_3 \rho^{(2)}_1 = 0$ if $|\delta| \leq \alpha \wedge \alpha^{-1}$, i.e., we have a family of energetically non-optimal stationary states.

On the other hand, when $\eta_{13} > 0$ we require $\delta = \alpha \frac{r_1 + r_2}{r_2}$, together with $|\delta| \leq \alpha \wedge \alpha^{-1}$ to obtain a stationary state, two conditions which are mutually exclusive for some choices of $r_1, r_2$ and $\alpha$.

In particular, we see that even for very weak cross-repulsion, i.e., for $\alpha \ll 1$, we may not obtain aggregation of both species in $x_2$ as one would observe it for a particle system with strong self-attraction and weak cross-repulsion.

Cutting off $x_3$. Assume $r_2 > r_1$ and let $\eta(x, y) = \mathbb{1}_{\{|x-y|<r\}}$ for some $r_1 < r < r_2$. Then, the interaction of the node $x_1$ with $x_1$ and $x_2$ is completely cut off. Thus, the only remaining condition for our distribution to be stationary is $\delta \geq -\alpha \wedge -\alpha^{-1}$, leading to an even larger family of stationary states.
4.3. Geometric effects on the dynamic in a 4-Point space

In order to elucidate the geometric effects let us consider the following, rather minimal, example of a four point space. It is given by the base measure \( \mu = \sum_{i=1}^{4} \delta_{x_i} \), which is embedded in \( \mathbb{R}^2 \). Here the nodes are given by

\[
x_1 = (0, 0), \quad x_2 = (1, 0), \quad x_3 = (1, 1), \quad \text{and} \quad x_4 = (\varepsilon, 1),
\]

with \( \varepsilon \in [0, 1] \). For the sake of the argument let us assume the two species are concentrated on the two neighboring nodes such that, initially,

\[
\rho_0^{(1)} = \delta_{x_1}, \quad \text{and} \quad \rho_0^{(2)} = \delta_{x_1},
\]

respectively. For simplicity let \( m_{ik}^{(i)} = \rho_{ik}^{(i)} \) for \( i = 1, 2, \, \ell, \kappa = 1, 2, 3, \) and let \( p = 2 \). In addition, we assume attractive intra-specific interactions of the form

\[
K^{(ii)}(x, y) = |x - y|_1, \quad i = 1, 2,
\]

and mutually repulsive inter-specific interactions of the form

\[
K^{(12)}(x, y) = -\alpha |x - y|_1,
\]

for some \( \alpha \geq 0 \). Here, \( |x|_1 = |x^1| + |x^2| \). We also assume that the cross-interaction is symmetric, i.e., \( \beta = (1, 1) \). Intuitively, for this type of cross-interactions, we expect that the second species pushes the first one onto the opposing node, \( x_3 \). However, we now show that this is not the case. On the contrary, let us compute

\[
\partial_t \rho_3^{(1)} = \sum_{i=1}^{4} (v_{3,i}^{(1)})_+ \rho_3^{(1)} - (v_{3,i}^{(1)})_+ \rho_3^{(1)}. \tag{36}
\]

Recalling that, initially,

\[
v_{1k}^{(1)} = \sum_{i=1}^{4} \left[ K_{i,k}^{(11)} - K_{i,k}^{(11)} \right] \rho_{i,k}^{(1)} + \sum_{i=1}^{4} \left[ K_{i,k}^{(12)} - K_{i,k}^{(12)} \right] \rho_{i,k}^{(2)}
\]

\[
= \left[ K_{i,k}^{(11)} - K_{i,k}^{(11)} \right] \rho_{i,k}^{(1)} + \left[ K_{i,k}^{(12)} - K_{i,k}^{(12)} \right] \rho_{i,k}^{(2)}
\]

\[
= \left[ K_{i,k}^{(11)} - K_{i,k}^{(11)} \right] + \left[ K_{i,k}^{(12)} - K_{i,k}^{(12)} \right],
\]

as all mass is concentrated on the two nodes, \( x_1, x_4 \). With the above choice of interaction potentials, we obtain

\[
v_{1k}^{(1)} = |x_i - x_4| - |x_k - x_4| - \alpha |x_i - x_1| + \alpha |x_k - x_1|. \tag{37}
\]

Substituting (37) into (36) this yields

\[
\partial_t \rho_3^{(1)} = \sum_{i=1}^{4} (v_{3,i}^{(1)})_+ \rho_3^{(1)} - (v_{3,i}^{(1)})_+ \rho_3^{(1)} = (v_{3,4}^{(1)})_-
\]

\[
= [1 - \varepsilon - 2\alpha + \alpha (1 + \varepsilon)]_-. \]

In particular, for \( 0 \leq \alpha < 1 \) no mass flows, even if the node \( \rho^{(1)} \) is concentrated on, \( x_4 \), is moved closer to \( x_3 \). In fact, only if the cross repulsion is strong enough, i.e., \( \alpha > 1 \), mass flows onto \( x_3 \). What is more, this remains true even if \( \varepsilon = 0 \), a scenario in which the dynamics of the associated particle system would lead to two particles, one for each species, that move into opposing directions.

A similar computation reveals that, initially,

\[
\partial_t \rho_1^{(2)} = -(-1 + \alpha - 2\alpha \varepsilon)_+ - (-2 - \alpha \varepsilon + \alpha - \varepsilon)_+ - (-\varepsilon - \alpha \varepsilon)_+.
\]

This shows that the second species remains concentrated on the node \( x_1 \) whenever \( 0 \leq \alpha \leq 1 \) and \( \varepsilon \geq 0 \). Moreover, for cross-interaction strengths that are stronger than the self-interactions, \( \alpha > 1 \), it is still possible for
Figure 6.: We display the dynamics of the non-symmetric ($\varepsilon \in (0, 1)$) four-point space with the initial mass of $\rho^{(2)}$ concentrated on $x_4$ (green, bottom row) and the initial mass of $\rho^{(1)}$ concentrated on $x_1$ (red, top row). Only if the cross-repulsion strength is strong enough, $\alpha > 1$, can the second species drive the first one out of the node it is inhabiting. Against any intuition based on the associated particle system, this is not the case if the repelling cross-interaction is weak.

The second species to remain stationary; however, now, the size of $\varepsilon$ begins to play a role made precise by the above equation. For instance, if $\alpha = 3/2$ and $\varepsilon = 1/4$, the second species remains located on the first node while it pushes the first species from the fourth node to the third node, which also follows from the form of $\partial_t \rho^{(1)}_3$, initially. The behavior is displayed in Figure 6. If $\varepsilon$ is sufficiently small or zero but $\alpha > 1$, the above equation yields a change in density of the second species on $x_1$. Similarly, the first species will escape to node $x_3$, which can be seen in Figure 7. This relatively minimal example goes to show that the dynamics on the graph are very different from the particle system associated to the energy functional (1).
\[ t = 0 \quad t = 6 \quad t = 11 \quad t = 17 \quad t = 20 \]

Figure 7.: In the symmetric setting, two species that are initially concentrated on neighbouring vertices of the
four-point space will drive out one another in the case of sufficiently strong cross-repulsion, \( \alpha > 1 \). Only if the cross-repulsion strength is strong enough, \( \alpha > 1 \), can the second species drive the first one out of the node it is inhabiting and vice versa. This phenomenon is quite counter-intuitive as both species are self-attractive. Nonetheless, the mass of both species is evenly split. Against any intuition based on the associated particle system, this is not the case if the repelling cross-interaction is weak.

4.4. Pattern formation on large graphs

In the preceding section we have studied two species whose intra-specific interaction is of attractive nature while
the cross-interaction is purely repulsive. With the intuition gained in the first example, let us consider a larger
graph of 10-by-10 nodes arranged in a two-dimensional array with \( m(r, s) = r \) and \( p = 2 \). Focusing only on
kernels of the form \( K^{(ik)}(x, y) = K^{(ik)}(x - y) \), we assume the cross-interactions to be repulsive and to be given
by the potentials

\[ K^{(12)}(x) = (10 - 20|x|)_+ \]

Potentials with a compact support like the cross-interaction potential are typically chosen whenever interactions
are nonlocal, but the sensing radius is imposed. As for the above choice, this sensing radius is \( r = 1/2 \), a distance
beyond which no repulsive interactions are present.

Concerning the self-interactions, we choose the potentials proposed by Kolokolnikov and Evers, cf. [23], which
are of the form

\[ K^{(ii)}(x) = -a \log(|x|) + b \frac{|x|^2}{2}, \]

i.e., consisting of a locally repulsive part and a globally attractive part, respectively. This potential has the
remarkable property that it leads to compactly supported steady states of constant density. In fact, for a single
species, there holds

\[ \rho^{(i)} \nabla \tilde{K}^{(ii)} \ast \rho^{(i)} = 0, \]

or, on the support of \( \rho^{(i)} \), we have

\[ \nabla \tilde{K}^{(ii)} \ast \rho^{(i)} = 0. \]

Taking the divergence yields

\[ 0 = \Delta \tilde{K}^{(ii)} \ast \rho^{(i)} = -\frac{a}{2\pi} \delta \ast \rho^{(i)} + b = b - \frac{a}{2\pi} \rho^{(i)}, \]
In the situation with symmetric nodes and non-symmetric cross-repulsion (i.e., $\beta \neq 1$), one of the two species which are initially concentrated on neighboring edges of the four-point space will drive out the other in the case of sufficiently strong cross-repulsion, $\alpha > 1$. Again, against particle system intuition, this is not the case if the repelling cross-interaction is weak.

\[ \rho^{(i)} = 2\pi b/a, \]

where $\rho^{(i)} > 0$. For our purpose we choose

\[ \tilde{K}^{(ii)}(x) = -\log(|x|) - \frac{|x|^2}{400}. \] (38)

In Figure 9, we observe that starting from randomly distributed mass on the nodes ($t = 0$), the mass immediately begins to aggregate, ($t = 1$ and $t = 10$). As time passes the profile approaches the stationary profile ($t = 5000$).

Based on this numerical experiment we alter the self-interaction potential slightly in that we introduce a sensing radius, too. Instead of (38), we truncate the self-interaction potentials and consider

\[ \tilde{K}^{(ii)}(x) = \left( -a \log(|x|) + b|x|^2 \right) \mathbb{1}_{\{|x| < s\}} + \left( -a \log(s) + bs^2 \right) \mathbb{1}_{\{|x| \geq s\}} \]
where $a = 1$, $b = 1/400$, and the sensing radius is chosen as $0 < s = 0.2$. Outside of the interaction region, $|x| > s$, the potential is extended by a constant. Concerning the cross-interaction terms, we use compactly supported, repulsive potentials of the form

$$\tilde{K}^{(12)}(x) = (4 - 20|x|)_+,$$

i.e., the cross-interaction radius is $s = 0.5$. In Figure 10 we give a representative example of the dynamics these interaction rules give rise to. As in the previous example, we initialize both densities randomly. As time evolves, we observe a self-sorting phenomenon as regions of co-habitation disappear, and islands of individual occupancy emerge, cf. Figure 10, $t = 200$.

![Figure 10](image1)

Figure 10: Starting from random initial data (left panel), the two species (red and green, respectively) have a sensing radius with respect to both, members of their own species and members of the opposing species. We observe emergence behavior in the form of self-sorting, i.e., the final configuration consists of two species with disjoint habitats.

### 4.5. Nonlinear mobilities and nonquadratic exponents

In this section, we want explore the effects of changing the mobility $m$ as well as the exponent $p = q/(q-1)$ in (6). To this end, we study the case of 10-by-10 nodes arranged in a two-dimensional lattice, which is fully connected and where the base measure is given by $\mu_i = 1/20$ on all vertices. To observe the effects of different mobilities most clearly, we choose attractive interactions given by $K^{(i)}(x) = 20(1 - \exp(|x|))$ for $i, k = 1, 2$. We consider the linear mobility $m(r, s) = r$ with $p \in \{1.65, 2, 5\}$ as well as the volume filling mobility $m(r, s) = r(1 - s)$ for $p = 2$. For the comparison of the different models, we choose the same random initial distribution for all cases.

In Figure 11, for all $p$ we observe aggregation of both species to a single vertex as time progresses. Comparing the different $p$, we find that an increase of $p$ slows down the convergence speed towards the stationary state. This is due of the power $q = 1/(p-1)$ in (6). If $q < 1$ is small, velocities near $0$ increase, speeding up the convergence near the stationary state. Conversely, if $q > 1$ is large, velocities near $0$ decrease further, slowing down the convergence near the stationary state.

In contrast to the aggregation in Figure 11, in Figure 12 we see that the volume filling mobility truncates the set of admissible states, eliminating configurations, which assign to much mass of any species $i$ to a single vertex. Because $\mu_i = 1/20$ for all $i \in \{1, \ldots, 100\}$, the maximum threshold is $\rho_i^{(i)} = 1/20$ for all $i \in \{1, \ldots, 100\}$ and $i = 1, 2$. Due to the geometry of the graph and the kernel structure, the total mass of the stationary state in the volume filling case is spread to 21 vertices.
Figure 11.: We display the evolution of both species (red and green, respectively) for attractive self- and cross-interactions, the mobility $m(r,s) = r$ and $p \in \{1.65, 2, 5\}$, in all cases starting from the same random initial data (left panel). For all $p$ we observe aggregation of both species to a single vertex as time progresses. Comparing the three values of $p$, we see that an increase of $p$ slows down aggregation speed.
time.

\[ t = 0 \]
\[ t = 10 \]
\[ t = 300 \]
\[ t = 1000 \]
\[ t = 2000 \]

Figure 12.: Starting from random initial data (left panel), we display the evolution of both species (red and green, respectively) for attractive self- and cross-interactions, the mobility \( \mu_t = 1/20 \) for all \( t \in \{1, \ldots, 100\} \) and \( p = 2 \). We observe that the aggregation is confined by the upper bound \( \rho_t^{(i)} \leq \mu_i \) for \( i = 1, 2 \), which results in the formation of a patch of populated vertices.

Acknowledgements

GH acknowledges support by the Studienstiftung des deutschen Volkes in form of a PhD stipend.

A. Appendix

We now provide proofs of the claims made in Subsection 4.1. Most arguments rely on a study of the sign of the velocity field, which, in its essence, informs the direction in which mass is being transported.

Decoupled case.

The following simplification proves useful for the subsequent analysis. Using \( \rho_t^{(i)}(x_i) = 1 - \rho_t^{(i)}(x_\kappa) \), we have

\[
v^{(i)}_{i\kappa} = \sum_{k=1}^{2} \sum_{A=1}^{2} \left( K^{(ik)}_{\kappa A} - K^{(ik)}_{k A} \right) \rho^{(k)} A \mu_A = \sum_{A=1}^{2} \left( K^{(ii)}_{\kappa A} - K^{(ii)}_{k A} \right) \rho^{(i)} A \mu_A
\]

\[
= -D^{(ii)} \nabla \rho^{(i)}_{i\kappa} = D^{(ii)} \left( 2\rho^{(i)}(x_i) - 1 \right) = D^{(ii)} s^{(i)},
\]

where we set

\[
s^{(i)} := 2\rho^{(i)}(x_i) - 1 \in [-1, 1].
\]

Proof of Proposition 4.1. We have seen that for \( D^{(12)} = 0 \) there are two possible ways of obtaining a stationary state:

a) \( v^{(i)}_{i\kappa} = 0 \),

b) \( v^{(i)}_{i\kappa} > 0, \rho^{(i)}(x_i) = 0 \),

where \( i, \kappa \in \{1, 2\} \) with \( i \neq \kappa \).

We see that a) always holds if \( D^{(ii)} = 0 \), since by (39) we have \( v^{(i)}_{i\kappa} = D^{(ii)} s^{(i)} \). Conversely, if \( D^{(ii)} \neq 0 \), then

\[
0 = D^{(ii)} \left( 2\rho^{(i)}(x_i) - 1 \right) \text{ immediately implies } \rho^{(i)}(x_i) = \frac{1}{2}.
\]
As for statement b), let us assume \( D^{(ii)} < 0 \).

Inserting \( \rho^{(i)}(x_i) = 1 \) in to (39), we obtain \( v_{ik}^{(i)} = D^{(ii)} \). Hence, b) holds if only if \( D^{(ii)} < 0 \). Here, it is important to keep in mind that \( i = 1 \) and \( i = 2 \) are both possible, leading to aggregation at either \( x_1 \) or \( x_2 \). 

Proof of Proposition 4.2. We have seen that for \( D^{(ii)} = 0 \) every state \( \rho^{(i)} \) is stationary, and that for \( D^{(ii)} > 0 \) the only stationary state is \( \rho^{(i)}_u = \frac{1}{2} \), which is therefore attractive, proving the first statement.

Concerning the second statement, let us consider \( D^{(ii)} < 0 \). First, let \( \epsilon = 0 \). Since \( \rho^{(i)}_u = \frac{1}{2} \) is stationary, \( \rho^{(i)}_t(x_i) = \frac{1}{2} \), for \( t = 1, 2 \) and all \( t \geq 0 \). Next, consider \( \epsilon > 0 \). Since \( v_{ik}^{(i)} = D^{(ii)} s^{(i)} \) and \( D^{(ii)} < 0 \), we immediately obtain \( v_{ik}^{(i)} s^{(i)} = D^{(ii)} (s^{(i)})^2 < 0 \) for all \( s^{(i)} \neq 0 \), i.e. by (39) \( v_{ik}^{(i)} \) and \( s^{(i)} \) must have opposite signs. Thus, \( v_{ik}^{(i)} < 0 \), which means mass flows from \( x_k \) to \( x_i \), leading to an aggregation of mass on \( x_i \). Therefore, \( \rho^{(i)}_t = \frac{1}{2} \) is unstable and the stationary states \( \rho^{(i)}_s \in \{0, 1\} \) are asymptotically stable for any \( \epsilon < 1/2 \).

Proof of Proposition 4.4. We recall that by Lemma 3.5, there are four different conditions leading to stationary states:

a) \( v_{ik}^{(1)} = v_{ik}^{(2)} = 0 \),

b) \( v_{ik}^{(i)} > 0 \) and \( \rho^{(i)}(x_i) = 0 \) as well as \( v_{ik}^{(k)} = 0 \),

c) \( v_{ik}^{(1)} > 0 \) and \( v_{ik}^{(2)} > 0 \) as well as \( \rho^{(1)}(x_i) = \rho^{(2)}(x_i) = 0 \),

d) \( v_{ik}^{(1)} > 0 \) and \( v_{ik}^{(2)} < 0 \) as well as \( \rho^{(1)}(x_i) = 0 \) and \( \rho^{(2)}(x_i) = 1 \),

for \( i, k, \iota, \kappa \in \{1, 2\} \) with \( i \neq k \) and \( \iota \neq \kappa \), which we assume for the remainder of this section. Similar to (39), we can rewrite \( v^{(i)} \) as

\[
v_{ik}^{(i)} = D^{(ii)}(2\rho^{(i)}(x_i) - 1) + D^{(12)}(2\rho^{(k)}(x_i) - 1) = D^{(ii)} s^{(i)} + D^{(12)} s^{(k)},
\]

where \( s^{(i)} \) and \( s^{(k)} \) are given by (40).

a) Using \( D^{(12)} \neq 0 \), we obtain

\[
s^{(1)} = -\frac{D^{(22)}}{D^{(12)}} s^{(2)}, \quad s^{(2)} = -\frac{D^{(11)}}{D^{(12)}} s^{(1)}.
\]

Hence, if \( D^{(11)} = 0 \) or \( D^{(22)} = 0 \), then we immediately find \( s^{(1)} = s^{(2)} = 0 \), which corresponds to \( \rho^{(1)}_t \mu_t = \rho^{(2)}_t \mu_t = \frac{1}{2} \). If both \( D^{(11)} \neq 0 \) and \( D^{(22)} \neq 0 \), we additionally have

\[
s^{(1)} = -\frac{D^{(12)}}{D^{(11)}} s^{(2)}, \quad s^{(2)} = -\frac{D^{(12)}}{D^{(22)}} s^{(1)},
\]

which we can combine with (42) to obtain

\[
D^{(11)} D^{(22)} s^{(i)} = \left(D^{(12)} \right)^2 s^{(i)},
\]

for \( i = 1, 2 \). Therefore, if \( D^{(11)} D^{(22)} \neq (D^{(12)})^2 \) we also find \( s^{(1)} = s^{(2)} = 0 \). Otherwise all equations in (42) and (43) are equivalent and define a family of stationary states as claimed.
Here, if $D^{(kk)} = 0$, then, by (41) $v^{(k)} = 0 \iff D^{(12)} = 0$. Hence, we require $D^{(kk)} \neq 0$. Then, $v^{(k)} = 0$ gives

$$s^{(k)} = \frac{D^{(12)}}{D^{(kk)}}. \quad (44)$$

The fact that $s^{(k)} \in [-1, 1]$ implies

$$|D^{(12)}| \leq |D^{(kk)}|. \quad (45)$$

This is a necessary condition for the stationary state to exist. The inequality $v^{(i)} < 0$ gives us the additional condition

$$D^{(ii)} < -D^{(12)}s^{(k)} = \left(\frac{D^{(12)}}{D^{(kk)}}\right)^2. \quad (46)$$

It is important to keep in mind that we can exchange $i$ and $k$, leading to an exchange of $i$ and $k$ in (44), (45) and (46). Also, we can exchange $i$ and $\kappa$, which corresponds to exchanging $s^{(i)}$ for $-s^{(i)}$, $s^{(k)}$ for $-s^{(k)}$, $v^{(i)}$ for $-v^{(i)}$ and $v^{(k)}$ for $-v^{(k)}$, but does not affect any of the conditions (44), (45) and (46). Hence, for any given combination of interaction kernels, there exist 0, 2 or 4 stationary states of the form b).

c) and d) Combining c), (40) and (41) immediately gives us the existence conditions $D^{(ii)} < -D^{(12)}, i = 1, 2$. Similarly, combining d), (40) and (41) yields the existence conditions $D^{(ii)} < D^{(12)}, i = 1, 2$.

**Proof of Corollary 4.7.** Each of the conditions (33) and (34) immediately implies that (30) is violated no matter the choice of $i$ and $k$. Additionally, it implies $D^{(ii)} \geq |D^{(12)}| \geq \pm D^{(12)}$ for at least one $i \in \{1, 2\}$ and thus also the violation of both (31) and (32). Conversely, if both (31) and (32) are violated, then $D^{(kk)} \geq |D^{(12)}| > 0$ must hold for at least one $k \in \{1, 2\}$, which implies (29). Therefore, (30) must not be satisfied, i.e., we must have $D^{(11)}D^{(22)} \geq (D^{(12)})^2$ and consequently $D^{(ii)} > 0$ for $i \neq k$.

**Proof of Corollary 4.7.** For the first claim observe that (31) together with (32) imply

$$D^{(11)} < -|D^{(12)}|, \quad D^{(22)} < -|D^{(12)}|.$$  

This implies (29) as well as

$$D^{(11)}D^{(22)} > (D^{(12)})^2, \quad D^{(11)} < 0, \quad D^{(22)} < 0,$$

which yields (30) for both $k = 1$ and $k = 2$.

For the converse case we have

$$|D^{(11)}| \geq |D^{(12)}|, \quad |D^{(22)}| \geq |D^{(12)}|, \quad D^{(11)} < \left(\frac{D^{(12)}}{D^{(22)}}\right)^2, \quad D^{(22)} < \left(\frac{D^{(12)}}{D^{(11)}}\right)^2. \quad (47)$$

From this we immediately see that $\sgn (D^{(11)}) = \sgn (D^{(22)})$ as otherwise either the third or the fourth inequality in (47) is violated. Additionally taking into account the first and second inequality in (47), we find that $\sgn (D^{(11)}) = \sgn (D^{(22)}) = -1$ as otherwise we would have a contradiction. This gives us

$$D^{(11)} \leq -|D^{(12)}|, \quad D^{(22)} \leq -|D^{(12)}|,$$

which, when combined with (47), shows the claim.

**Lemma A.1.** Let $N = 2$, let (26) and (27) hold, and assume $D^{(12)}_{12} \neq 0$. Then, for the stationary states given in Proposition 4.4, we have

1. If $D^{(11)}D^{(22)} = (D^{(12)})^2$, then $E(\rho_b) = E(\rho_a)$ for $r \in \left[-1 \vee -\frac{D^{(22)}}{D^{(12)}} - \frac{r}{1 - r} \frac{D^{(22)}}{D^{(12)}}\right]$.
2. If for $i \in \{1, 2\}$ $\dot{\rho}_b = 0$ for $i \in \{1, 2\}$, then $E(\rho_b) < E(\rho_a)$.
3. If $\rho_c$ is a stationary state, then $E(\rho_c) < E(\rho_a)$.
4. If $\rho_d$ is a stationary state, then $E(\rho_d) < E(\rho_a)$.
5. If $\rho_{b_i} = \rho_{a_i} = 0$ for $\alpha \in \{c, d\}$ and $i \in \{1, 2\}$, then $E(\rho_{a_i}) < E(\rho_{b_i})$.
6. If $\rho_c$ and $\rho_d$ are stationary states, then we have $\text{sgn}(E(\rho_d) - E(\rho_c)) = -\text{sgn}(D^{(12)})$.

**Proof.** In Theorem 3.1 we have identified conditions implying that aggregation of one or both species is energetically optimal. Now we want to specify these findings in the 2-point setting. To this end we calculate the signed differences between the energies of the explicit stationary states derived in Proposition 4.4. Setting $d_{\alpha_1 \alpha_2} := E(\rho_{\alpha_2}) - E(\rho_{\alpha_1})$, for different combinations of $\alpha_1, \alpha_2 \in \{a, a_r, b_1, b_2, c, d\}$. A short computation yields

$$d_{aa} = \frac{1}{4} \left( D^{(11)} - \frac{(D^{(12)})^2}{D^{(22)}} \right), \quad d_{ab} = \frac{1}{4} \left( D^{(ii)} - \frac{(D^{(ij)})^2}{D^{(kk)}} \right),$$

$$d_{ac} = \frac{1}{4} \left( D^{(11)} + 2D^{(12)} + D^{(22)} \right), \quad d_{ad} = \frac{1}{4} \left( D^{(11)} - 2D^{(12)} + D^{(22)} \right),$$

$$d_{bc} = \frac{1}{4} \left( D^{(kk)} + D^{(12)} \right), \quad d_{bd} = \frac{1}{4} \left( D^{(kk)} - D^{(12)} \right), \quad d_{cd} = -D^{(12)}.$$

With this, $D^{(11)}D^{(22)} = (D^{(12)})^2$ immediately yields $d_{aa} = 0$.

Applying the conditions which imply that $\rho_{b_i}, \rho_c$ and $\rho_d$ define stationary states to (48), we make the following observations:

- We have seen that $\rho_{b_i}$ is a stationary state if both (29) and (30) hold. However, (30) immediately yields $d_{ab} < 0$. Hence, if there is a stationary state of the form $\rho_{b_i}$, then it admits a lower energy than $\rho_a$.

- We have found that $\rho_c$ defines two stationary states if (31) holds, which immediately implies that $d_{ac} < 0$. Similarly, we find by (32) that $d_{ad} < 0$ if $\rho_d$ defines a stationary state. Thus, $\rho_c$ and $\rho_d$ respectively admit lower energies than $\rho_a$, if they are stationary states.

- If (29) and (30) as well as (31) hold, then we immediately obtain that $D^{(kk)} < 0$. Analogously, (29) and (30) together with (32) also imply $D^{(kk)} < 0$. Thus, if $\rho_{b_i}$ as well as either $\rho_c$ or $\rho_d$ are stationary states, then aggregation of both species admits a lower energy than aggregation of only one species.

- Finally, if (31) and (32) are simultaneously satisfied, $\rho_c$ admits a lower energy than $\rho_d$ if we have cross-attraction, while $\rho_d$ admits a lower energy than $\rho_c$ if we have cross-repulsion. Further, (31) and (32) in conjunction imply (29) and (30) for both $i = 1, k = 2$ and $i = 2, k = 1$. In this situation all candidates are stationary states.

\[ \Box \]

**Remark A.2** (Stability of the energy minimizer). If $D^{(11)}D^{(22)} \neq (D^{(12)})^2$ and $\rho_a$ is not the unique stationary state, then there exist exactly two $(D^{(12)} \neq 0)$ or exactly four $(D^{(12)} = 0)$ points minimizing the energy cf. Corollary 4.7 and Lemma A.1. Since these are distinct from each other and since the energy cannot increase along trajectories of the gradient flow by (22), the two are asymptotically stable.

**Remark A.3** (Stability of $\rho_a$). Employing the assumption that all kernels are constant on the diagonal, we can rewrite the energy as

$$E(\rho) = D^{(11)} \left( \rho^{(1)}(x_1) - \frac{1}{2} \right)^2 + D^{(22)} \left( \rho^{(2)}(x_1) - \frac{1}{2} \right)^2$$

$$+ 2D^{(12)} \left( \rho^{(1)}(x_1) - \frac{1}{2} \right) \left( \rho^{(2)}(x_1) - \frac{1}{2} \right) + \sum_{i,k=1}^{2} \frac{1}{4} \left( k^{(i)k} + k^{(k)i} \right), \quad (49)$$
which shows that the level sets of the energy are quadrics centered around \((1/2, 1/2)\) and lets us identify \(\rho_a = (1/2, 1/2)\) as a critical point. Moreover, it allows for a quick way to verify that the hessian of \(E\) reads

\[
H_E = 2 \begin{pmatrix} D^{(11)} & D^{(12)} \\ D^{(12)} & D^{(22)} \end{pmatrix},
\]

which is positive definite if and only if (33) holds. Hence, \(\rho_a\) is a minimizer of the energy if and only if we have (33). By Corollary 4.5 this is equivalent to \(\rho_a\) being the only stationary state of the dynamics on the two-point space.

This also means that \(\rho_a\) is unstable, whenever \(D^{(11)}D^{(22)} \neq (D^{(12)})^2\) and a second stationary state other than \(\rho_a\) exists. What is more, the proof of Corollary 4.7 in conjunction with (49) show that \(\rho_a\) maximizes the energy if we have \(\hat{\rho}_c = \hat{\rho}_d = 0\).

**Remark A.4 (Stability of \(\rho_c\) and \(\rho_d\)).** If \(\hat{\rho}_a \neq \hat{\rho}_c = 0\), we have seen that \(\rho_c\) minimizes the energy and is therefore asymptotically stable by the Remark A.2. Analogously, \(\rho_d\) is asymptotically stable whenever \(\hat{\rho}_c \neq \hat{\rho}_d = 0\).

Let us now consider the case \(\hat{\rho}_c = \hat{\rho}_d = 0\). From Corollary 4.7 we know that in this case we also have \(\hat{\rho}_{b_1} = \hat{\rho}_{b_2} = 0\). By the proof of Corollary 4.7 we have \(D^{(11)} < 0\), \(D^{(22)} < 0\) and \(\left|D^{(12)}\right| < \sqrt{D^{(11)}D^{(22)}}\). Therefore, equation (49) tells us that the quadrics describing the level sets of \(E\) are ellipses and that the energy is maximized at \((1/2, 1/2)\). Since every stationary state other than \(\rho_a\) lies on the border of the feasible set, the energy level sets of these stationary states touch the border of the feasible set. In particular there is at least one stationary state of the form \(\rho_{b_i}\), \(i \in \{1, 2\}\) on each edge of the feasible set. This means that the set \(\{(\rho^{(1)}(x_1), \rho^{(2)}(x_1)) \in [0,1] \times [0,1] : E(\rho) < E(\rho_{b_1}) \land E(\rho_{b_2})\}\) consists of four different connected components, each located in one of the corners of the feasible set (see for example Figure 1). Since by (22) the energy along the trajectories of the gradient flow cannot increase and the trajectories are continuous, a trajectory starting in one of these components can never leave it and must therefore converge to the stationary state located in the corner of the feasible set, which is contained in the component. Hence, whenever \(\rho_c\) or \(\rho_d\) are stationary states, they are asymptotically stable.

**Remark A.5 (Stability of \(\rho_{b_i}\)).** From Remark A.2 and Lemma A.1 we know that for \(i \in \{1, 2\}\) the stationary states of the form \(\rho_{b_i}\) are asymptotically stable, if the only other existing stationary state is \(\rho_a\).

From Corollary 4.7 we know that \(\hat{\rho}_c = 0\) or \(\hat{\rho}_d = 0\), if \(\hat{\rho}_{b_1} = \hat{\rho}_{b_2} = 0\). Thus, we now consider the case, where \(\hat{\rho}_{b_i} = 0\) for \(i \in \{1, 2\}\) and either \(\hat{\rho}_c = 0\) or \(\hat{\rho}_d = 0\). For \(k \in \{1, 2\}, k \neq i\) as before, (29) in conjunction with (31) or (32) implies that we have \(D^{(kk)} < \pm D^{(12)} \leq \left|D^{(12)}\right| \leq \left|D^{(kk)}\right|\), i.e., \(D^{(kk)} < 0\). We introduce a small perturbation \(\rho^{(i)} = \rho^{(i)}(\varepsilon) = \rho^{(i)}_b(\varepsilon) + \varepsilon, \rho^{(k)}(x_1) = \rho^{(k)}_b(x_1) - \varepsilon, \) for \(\varepsilon\) sufficiently small. Upon comparing its energy with that of \(\rho_{b_i}\), we observe

\[
E(\rho) - E(\rho_{b_i}) = D^{(kk)} \varepsilon^2 < 0.
\]

By (22) the energy cannot increase along the gradient flow and the perturbed distribution converges to a different stationary state implying the instability of \(\rho_{b_i}\).

**References**

[1] Martial Agueh. Finsler structure in the p-wasserstein space and gradient flows. *Comptes Rendus Mathematique*, 350(1):35–40, 2012.

[2] Nicola J Armstrong, Kevin J Painter, and Jonathan A Sherratt. A continuum approach to modelling cell–cell adhesion. *Journal of theoretical biology*, 243(1):98–113, 2006.

[3] Rafael Bailo, José A Carrillo, and Jingwei Hu. Bound-preserving finite-volume schemes for systems of continuity equations with saturation. *arXiv preprint arXiv:2110.08186*, 2021.
[4] Julien Barré, Pierre Degond, Diane Peurichard, and Ewelina Zatorska. Modelling pattern formation through differential repulsion. *Networks & Heterogeneous Media*, 15(3):307–352, 2020.

[5] Judith Berendsen, Martin Burger, and Jan-Frederik Pietschmann. On a cross-diffusion model for multiple species with nonlocal interaction and size exclusion. *Nonlinear Analysis*, 159:10–39, 2017.

[6] Maria Bruna, Martin Burger, Helene Ranetbauer, and Marie-Therese Wolfram. Cross-diffusion systems with excluded-volume effects and asymptotic gradient flow structures. *Journal of Nonlinear Science*, 27(2):687–719, 2017.

[7] Martin Burger, José A Carrillo, Jan-Frederik Pietschmann, and Markus Schmidtchen. Segregation effects and gap formation in cross-diffusion models. *Interfaces and Free Boundaries*, 22(2):175–203, 2020.

[8] Martin Burger, Marco Di Francesco, J.-F. Pietschmann, and Bärbel Schlake. Nonlinear cross-diffusion with size exclusion. *SIAM J. Math. Anal.*, 42(6):2842–2871, 2010.

[9] Martin Burger, Marco Di Francesco, Simone Fagioli, and Angela Stevens. Sorting phenomena in a mathematical model for two mutually attracting/repelling species. *SIAM Journal on Mathematical Analysis*, 50(3):3210–3250, 2018.

[10] Martin Burger and Jun-Frederik Pietschmann. Flow characteristics in a crowded transport model. *Nonlinearity*, 29(11):3528, 2016.

[11] J.A. Carrillo, S. Lisini, G. Savaré, and D. Slépčev. Nonlinear mobility continuity equations and generalized displacement convexity. *Journal of Functional Analysis*, 258(4):1273–1309, 2010.

[12] José A Carrillo, Yanghong Huang, and Markus Schmidtchen. Zoology of a nonlocal cross-diffusion model for two species. *SIAM Journal on Applied Mathematics*, 78(2):1078–1104, 2018.

[13] José A Carrillo, Francois James, Frédéric Lagoutière, and Nicolas Vauchelet. The Filippov characteristic flow for the aggregation equation with mildly singular potentials. *Journal of Differential Equations*, 260(1):304–338, 2016.

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

[15] José Antonio Carrillo, Marco Di Francesco, Antonio Esposito, Simone Fagioli, and Markus Schmidtchen. Measure solutions to a system of continuity equations driven by newtonian nonlocal interactions. *Discrete & Continuous Dynamical Systems*, 40(2):1191, 2020.

[16] José Antonio Carrillo, Annachiara Colombi, and Marco Scianna. Adhesion and volume constraints via nonlocal interactions determine cell organisation and migration profiles. *Journal of Theoretical Biology*, 445:75–91, 2018.

[17] Shui-Nee Chow, Wen Huang, Yao Li, and Haomin Zhou. Fokker–Planck equations for a free energy functional or Markov process on a graph. *Archive for Rational Mechanics and Analysis*, 203(3):969–1008, 2012.

[18] François Delarue, Frédéric Lagoutière, and Nicolas Vauchelet. Convergence analysis of upwind type schemes for the aggregation equation with pointy potential. *Annales Henri Lebesgue*, 1(1):000–000, 2019.

[19] Marco Di Francesco and Simone Fagioli. A nonlocal swarm model for predators–prey interactions. *Mathematical Models and Methods in Applied Sciences*, 26(02):319–355, 2016.

[20] Jean Dolbeault, Bruno Nazaret, and Giuseppe Savaré. A new class of transport distances between measures. *Calculus of Variations and Partial Differential Equations*, 34(2):193–231, Jun 2008.

[21] Matthias Erbar and Jan Maas. Gradient flow structures for discrete porous medium equations. *Discrete Contin. Dyn. Syst.*, 34(4):1355–1374, 2014.
[22] Antonio Esposito, Francesco S Patacchini, André Schlichting, and Dejan Slepčev. Nonlocal-interaction equation on graphs: gradient flow structure and continuum limit. *Archive for Rational Mechanics and Analysis*, pages 1–62, 2021.

[23] Joep HM Evers, Razvan C Fetecau, and Theodore Kolokolnikov. Equilibria for an aggregation model with two species. *SIAM Journal on Applied Dynamical Systems*, 16(4):2287–2338, 2017.

[24] Simone Fagioli and Yahya Jaafra. Multiple patterns formation for an aggregation/diffusion predator-prey system. *Networks & Heterogeneous Media*, 16(3):377, 2021.

[25] Marco Di Francesco and Simone Fagioli. Measure solutions for non-local interaction PDEs with two species. *Nonlinearity*, 26(10):2777–2808, sep 2013.

[26] Georg Heinze, Jan-Frederik Pietschmann, and Markus Schmidtchen. Nonlocal cross-interaction systems on graphs: Nonquadratic finslerian structure and nonlinear mobilities, 2021.

[27] Stefano Lisini and Antonio Marigonda. On a class of modified wasserstein distances induced by concave mobility functions defined on bounded intervals. *manuscripta mathematica*, 133(1-2):197–224, June 2010.

[28] Jan Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261(8):2250–2292, 2011.

[29] Jan Maas. Entropic Ricci curvature for discrete spaces. In *Modern approaches to discrete curvature*, volume 2184 of *Lecture Notes in Math.*, pages 159–174. Springer, Cham, 2017.

[30] James Mason, Robert L Jack, and Maria Bruna. Macroscopic behaviour in a two-species exclusion process via the method of matched asymptotics. *arXiv preprint arXiv:2203.01038*, 2022.

[31] Hideki Murakawa and Hideru Togashi. Continuous models for cell–cell adhesion. *Journal of theoretical biology*, 374:1–12, 2015.

[32] Kevin J Painter, JM Bloomfield, JA Sherratt, and A Gerisch. A nonlocal model for contact attraction and repulsion in heterogeneous cell populations. *Bulletin of mathematical biology*, 77(6):1132–1165, 2015.

[33] Matthew J Simpson and Ruth E Baker. Corrected mean-field models for spatially dependent advection-diffusion-reaction phenomena. *Physical Review E*, 83(5):051922, 2011.

[34] Alexandria Volkening, Madeline R Abbott, N Chandra, B Dubois, F Lim, D Sexton, and Bjorn Sandstede. Modeling stripe formation on growing zebrafish tailfins. *Bulletin of mathematical biology*, 82(5):1–33, 2020.

[35] Alexandria Volkening and Björn Sandstede. Iridophores as a source of robustness in zebrafish stripes and variability in danio patterns. *Nature communications*, 9(1):1–14, 2018.