Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 16–08

A FITNESS-DRIVEN CROSS-DIFFUSION SYSTEM FROM POPULATION DYNAMICS AS A GRADIENT FLOW

STANISLAV KONDRATYEV, LÉONARD MONSAINGEON AND DMITRY VOROTNIKOV

Abstract: We consider a fitness-driven model of dispersal of N interacting populations, which was previously studied merely in the case $N = 1$. Based on some optimal transport distance recently introduced, we identify the model as a gradient flow in the metric space of Radon measures. We prove existence of global non-negative weak solutions to the corresponding system of parabolic PDEs, which involves degenerate cross-diffusion. Under some additional hypotheses and using a new multicomponent Poincaré-Beckner functional inequality, we show that the solutions converge exponentially to an ideal free distribution in the long time regime.

Keywords: cross-diffusion, optimal transport, gradient flow, ideal free distribution, entropy-entropy production inequality.
Math. Subject Classification (2010): 35K51, 35K65, 35Q92, 49Q20, 58B20, 92D40.

1. Introduction

Living organisms tend to form distributional patterns but not to be arranged either uniformly or randomly. This spatial heterogeneity plays a crucial role in ecological theories and their practical applications. It should be taken into account when modeling epidemics, ecological catastrophes, competition, adaptation, maintenance of species diversity, parasitism, population growth and decline, social behaviour, and so on [21]. In order to understand the way the species distribute themselves it is important to pay attention to their dispersal strategies.

In this paper we study a system of PDEs for several interacting populations whose dispersal strategy is determined by a local intrinsic characteristic of organisms called *fitness* (cf. [10, 11]), essentially the signed difference between

Received March 21, 2016.
SK and DV were partially supported by CMUC – UID/MAT/00324/2013, funded by the Portuguese Government through FCT/MCTES and co-funded by the European Regional Development Fund through the Partnership Agreement PT2020. LM was supported by the Portuguese National Science Foundation through through FCT fellowship SFRH/BPD/ 88207/2012 and by the UT Austin/Portugal CoLab project *Phase Transitions and Free Boundary Problems.*

available resources and their consumption by the individuals. The fitness manifests itself as a growth rate, and simultaneously affects the dispersal as the species move along its gradient towards the most favorable environment. The equilibrium when the fitnesses of all populations vanish can be referred to as the *ideal free distribution* [14, 13], since no net movement of individuals occurs in this case. We are thus going to study the system

$$\partial_t u_i = -\operatorname{div}(u_i \nabla f_i) + u_i f_i, \qquad x \in \Omega, t > 0, \qquad i = 1, \ldots, N, \tag{1.1}$$

of N interacting species located in a bounded domain $\Omega \subset \mathbb{R}^d$. For prescribed resources $\mathbf{m} = (m_i(x))$ we assume a generic linear relation between the population densities $\mathbf{u} = (u_i(t, x))$ and their corresponding fitnesses $\mathbf{f} = (f_i(t, x))$:

$$\mathbf{f} = \mathbf{m} - A\mathbf{u}. \tag{1.2}$$

We assume that both the matrix A and the vector $\mathbf{m}$ depend on $x \in \Omega$, thus our model is spatially heterogeneous. Formula (1.2) expresses the idea that the fitness is determined by the difference between the available resources $\mathbf{m}$ and the animals' consumption $A\mathbf{u}$.

The mathematical difficulties which we will face when studying the parabolic system (1.1)-(1.2) come from the fact that it involves both cross-diffusion (for $N > 1$) and degenerate diffusion. In the case of merely one population ($N = 1$), the fitness-driven dispersal model (1.1), (1.2) was suggested in [27, 10] and studied in [12, 18] (see also [4]). Related fitness-driven two-species models were investigated in [5, 25] where one population uses the fitness-driven dispersal strategy and the other diffuses freely or does not move at all. In the case when A is a constant matrix, $\mathbf{m} \equiv 0$, and the second (reaction) term $u_i f_i$ in (1.1) is omitted, system (1.1), (1.2) is equivalent to the degenerate cross diffusion system which was recently analyzed in [1] with an application to seawater intrusion. Another population dynamics model which involves cross-diffusion is the Shigesada, Kawasaki and Teramoto model

$$\partial_t u_i = \Delta\left(u_i\left(d_i + \sum_{i=1}^{N} a_{ij} u_j\right)\right) + u_i\left(\left(c_i - \sum_{i=1}^{N} b_{ij} u_j\right)\right), \tag{1.3}$$

where the coefficients are non-negative constants. It has been extensively studied (mostly for $N = 2$) from the point of view of well-posedness and long-time behaviour (see, e.g., [9, 6, 17, 16, 26, 24] and the references therein). Note that the constants d_i in (1.3) are usually assumed to be strictly positive, hence this problem is not as degenerate as our system (1.1), (1.2).

On the other hand, being inspired by the ideas of the Monge-Kantorovich optimal transport theory [32, 33], we have recently constructed in [18] a new distance on the space of non-negative finite Radon measures. The same distance was almost simultaneously introduced in the independent works [8, 22] (see also subsequent [23, 7, 15]). This metric generates a formal (infinite dimensional) Riemannian structure on this space, and provides first- and second-order differential calculus in the spirit of Otto [29]. With this differential calculus at hand we were able to identify in [18] the scalar model as a metric gradient flow, which allowed us to prove long-time convergence to the ideal free distribution with explicit exponential rates. The goal of this paper is to extend our previous results to the multispecies case $N > 1$: we will observe that the problem (1.1), (1.2) can be interpreted as a formal gradient flow of some driving entropy functional on the Cartesian product of N spaces of non-negative Radon measures with respect to this geometric structure (provided that the matrix $A(x)$ is symmetric). Roughly speaking, the entropy $\mathcal{E}(\mathbf{u}) \approx \|\mathbf{u} - \mathbf{u}^\infty\|^2_{L^2(\Omega)} \approx \|\mathbf{f}\|^2_{L^2(\Omega)}$ will quantify the deviation from the ideal free distribution $\mathbf{u}^\infty$, characterized by $\mathbf{f} = 0$. In this framework and under some general structural assumptions on A and $\mathbf{m}$, we will prove existence of non-negative weak solutions to our problem (which to the best of our knowledge was known so far only in the scalar case [12]), and show that, at least for subcritical initial entropies, all the species persist and exponentially converge to the ideal free distribution. All our arguments will have a strong optimal transport flavor, but, due to the multicomponent nature of the problem preventing our entropy functional from being geodesically convex, the abstract results for metric gradient flows in [2] do not apply directly. As a consequence some technical work will be needed to justify the formal Riemannian computations and a priori estimates (in particular some chain rules in weighted spaces), and we will argue using several approximations and regularizations in a more standard PDE framework.

The paper is organized as follows: In Section 2 we impose precise structural assumptions and state our results. In Section 3 we describe the optimal transport distance on the product measure space $\mathcal{M}^+(\Omega)^N$, discuss the induced formal Riemannian structure and calculus, and highlight the gradient-flow structure of the system. We also present two formal but crucial computations, consisting of a priori entropy and gradient estimates to be derived more rigorously later on. Section 4 is devoted to the existence of weak solutions, whose proof will involve three successive regularizations. The last Section 5 contains the proof of the long-time convergence, based on a new vectorial Poincaré-Beckner type inequality which serves as an entropy-entropy production inequality for our gradient flow. The Appendix contains a technical functional analytic lemma.

2. Conventions and main results

Throughout the whole paper we assume that $\Omega \subset \mathbb{R}^d$ is an open, connected, bounded domain with sufficiently smooth boundary. We always denote vector functions with values in $\mathbb{R}^N$ by bold letters, e.g., $\mathbf{f} = (f_i(t,x))$. We assume that we are given a function $\mathbf{m} = (m_i(x)) : \overline{\Omega} \to \mathbb{R}^N$ and a symmetric positive-definite matrix function $A = (a_{ij}(x)) : \overline{\Omega} \to \mathbb{R}^{N\times N}$, and without further mention we shall always assume the uniform ellipticity condition

$$0 < \lambda_A \leq A(x) \leq \Lambda_A, \qquad x \in \overline{\Omega}$$

for some structural constants $\lambda_A \leq \Lambda_A$. We assume that A and $\mathbf{m}$ are sufficiently smooth. Note that we do not assume that all the components of A and $\mathbf{m}$ are non-negative.

All the integrals are always implicitly written with respect to the Lebesgue measures $\mathrm{d}x, \mathrm{d}t$, or $\mathrm{d}x\mathrm{d}t$. The symbol $\mathcal{M}^+(\Omega)$ denotes the space of (non-negative) Radon measures on Ω. Parentheses denote the scalar product in $L^2(\Omega)$, $L^2(\Omega)^N$ or $L^2(\Omega)^{N\times d}$. The symbol $\mathcal{C}_w(\mathcal{J}; X)$ stands for the space of weakly continuous (resp., narrowly continuous) curves with defined on $\mathcal{J} \subset \mathbb{R}$ and with values in $X = L^2(\Omega)$ (resp., $X = \mathcal{M}^+(\Omega)$).

We study the system

$$\begin{cases} \partial_t u_i = -\operatorname{div}(u_i \nabla f_i) + u_i f_i, & x \in \Omega,\ t > 0,\ i = 1, \dots, N, \\ u_i \frac{\partial f_i}{\partial \nu} = 0, & x \in \partial\Omega,\ t > 0, \\ u_i(0,x) = u_i^0(x), & x \in \Omega. \end{cases} \tag{2.1}$$

As already mentioned in the introduction and without further mention, we always denote the fitness by

$$\mathbf{f} = \mathbf{m} - A\mathbf{u},$$

and the ideal free distribution $\mathbf{u}^\infty(x)$ is obtained by solving $\mathbf{f} = 0$, i-e

$$\mathbf{u}^\infty := A^{-1}\mathbf{m} \qquad \Leftrightarrow \qquad \mathbf{f} = 0.$$

Note that at this stage $\mathbf{u}^\infty$ can have negative components and may therefore be biologically irrelevant (but it will be non-negative later on with extra structural conditions on $A, \mathbf{m}$), and that $\mathbf{u}^\infty$ is trivially a steady state of (2.1) with $f_i \equiv 0$.

We will show in Section 3 that (2.1) is the gradient flow of the entropy

$$\begin{aligned}\mathcal{E}(\mathbf{u}) &= \frac{1}{2}\int_\Omega A(\mathbf{u}-\mathbf{u}^\infty)\cdot(\mathbf{u}-\mathbf{u}^\infty)\\ &= \frac{1}{2}\int_\Omega A^{-1}\mathbf{f}\cdot\mathbf{f} = \frac{1}{2}\int_\Omega(\mathbf{u}^\infty-\mathbf{u})\cdot\mathbf{f},\quad \mathbf{u}\in L^2(\Omega)^N \qquad (2.2)\end{aligned}$$

with respect to some optimal transport distance.

Definition 2.1. *Let* $\mathbf{u}^0\in L^2(\Omega)^N$, $u_i^0\geq 0$. *A vector function*

$$\mathbf{u}\in\mathcal{C}_w([0,\infty);L^2(\Omega)^N)\cap L^2_{loc}([0,+\infty);H^1(\Omega)^N),$$

$u_i\geq 0$, *is called a* non-negative weak solution *to problem* (2.1) *provided*

$$\frac{d}{dt}\int_\Omega \mathbf{u}\cdot\mathbf{w} = \sum_{i=1}^N\int_\Omega\left(u_i\nabla f_i\cdot\nabla w_i + u_i f_i w_i\right),\quad \forall\mathbf{w}\in(\mathcal{C}^1(\overline{\Omega}))^N,$$

in the sense of distributions $\mathcal{D}'(0,\infty)$, *and the initial condition* $\mathbf{u}(0)=\mathbf{u}^0$ *is satisfied weakly in the space* $L^2(\Omega)^N$.

Theorem 1 (existence of non-negative weak solutions)**.** *Let* $\mathbf{u}^0\in L^2(\Omega)^N$ *with* $u_i^0\geq 0$. *There exists a non-negative weak solution*

$$\mathbf{u}\in L^\infty(0,+\infty;L^2(\Omega)^N)\bigcap\mathcal{C}_w([0,+\infty);L^2(\Omega)^N)\bigcap L^2_{loc}([0,+\infty);H^1(\Omega)^N)$$

to problem (2.1), *satisfying the* Entropy-Dissipation-Inequality

$$\mathcal{E}(\mathbf{u}(t_1)) + \sum_i\int_{t_0}^{t_1}\int_\Omega u_i(|\nabla f_i|^2+|f_i|^2)\leq\mathcal{E}(\mathbf{u}(t_0)) \qquad (2.3)$$

for a.e. $0\leq t_0\leq t_1$.

Remark 2.1. *We were not able to prove any uniqueness result due to the lack of geodesic convexity, which usually gives contractivity in the metric sense and thus uniqueness; we therefore believe that any hypothetical proof of uniqueness cannot come from standard mass transport arguments and should rely on some PDE approach.*

To fix the ideas and improve readability, we assume that A does not depend on x. The generalization to the x-dependent case would require technical work and employs the fact that the last sum in the expansion $\nabla f_i =$

$\nabla m_i - \sum_j a_{ij}\nabla u_j - \sum_j u_j \nabla a_{ij}$ is of lower order with respect to the penultimate one, but all the arguments below would carry through with minor modifications.

With an additional assumption, we obtain long-time convergence $\mathbf{u}(t) \to \mathbf{u}^\infty$ to the ideal free distribution with survival of all the species:

Theorem 2 (long-time behavior)**.** *Let* $\mathbf{u}^0$ *be as in Theorem 1, and assume in addition that the structural hypothesis* (5.1) *holds. Then* $\mathbf{u}^\infty(x) > 0$ *(componentwise), and there exists* $E^* \equiv E^*(A, \mathbf{m}) > 0$ *such that, for any subcritical initial datum* $\mathcal{E}(\mathbf{u}^0) < E^*$*, our solution* $\mathbf{u}$ *converges exponentially to* $\mathbf{u}^\infty$ *as*

$$\mathcal{E}(\mathbf{u}(t)) \leq e^{-\gamma t}\mathcal{E}(\mathbf{u}^0) \tag{2.4}$$

for all $t \geq 0$ *and some* $\gamma \equiv \gamma(\mathbf{u}^0, \mathbf{m}, A, \Omega) > 0$*.*

Note that our coercivity assumption $A \geq \lambda_A$ controls $\mathcal{E}(\mathbf{u}) \geq \frac{\lambda_A}{2}\|\mathbf{u} - \mathbf{u}^\infty\|^2_{L^2(\Omega)}$, thus the entropy decay (2.4) immediately implies L^2 convergence. The consequences and interpretation of the additional assumption (5.1) will be discussed later on in Section 5. In Theorem 2 we had to restrict to subcritical entropies $\mathcal{E}(\mathbf{u}^0) < E^*$ for technical reasons, but we conjecture that (2.4) holds for any $\mathbf{u}^0 \geq 0$ (unless some component of $\mathbf{u}^0$ is identically zero). Indeed our proof of the long-time convergence works provided that no extinction occurs, say $\|u_i(t)\|_{L^p(\Omega)} \geq c_i > 0$ for all $i \in \{1, \dots, N\}$, $t \geq 0$, and some $p \geq 1$. For both the ODE dynamics (i-e when all densities and resources are constant in space) and for the one-animal PDE dynamics [12, 18] this is true, but due to the cross-diffusion we were not able to prove the non-extinction in the general case. In other words, our subcriticality assumption in Theorem 2 is a technical workaround guaranteeing that our solution stays away from a finite number of certain partial extinction regimes. These regimes correspond to the situations when some (or all) populations go extinct, and the survivors compose a (lower-dimensional) ideal free distribution. This allows us to provide an explicit value for E^* depending only on the structure of the problem, see Section 5 for the details.

3. The gradient-flow structure and a priori estimates

The celebrated Benamou-Brenier formula was originally established in [3] to characterize the quadratic Monge-Kantorovich-Wasserstein distance as a dynamical evolution problem, and is restricted to *conservative* optimal transport of measures with fixed mass (typically probability measures). In [18] we constructed an optimal transport distance $\mathtt{d}$ on the space of *arbitrary* non-negative Radon measures $\mathcal{M}^+(\Omega)$, based on a modified dynamical Benamou-Brenier formula allowing for mass variations. More precisely, for any $u_0, u_1 \in \mathcal{M}^+(\Omega)$ the distance reads

$$\mathtt{d}^2(u_0, u_1) = \min_{(u,g)} \int_0^1 \int_\Omega (|\nabla g_t(x)|^2 + |g_t(x)|^2) \mathrm{d}u_t(x) \mathrm{d}t,$$

where the infimum is taken among narrowly continuous curves

$$t \mapsto u_t \in \mathcal{C}_w([0,1]; \mathcal{M}^+(\Omega))$$

with endpoints u_0, u_1 such that the non conservative continuity equation

$$\partial_t u_t + \operatorname{div}(u_t \nabla g_t) = u_t g_t$$

holds in the sense of distributions $\mathcal{D}'((0,1) \times \Omega)$. Biologically, one can think of g as fitness: in the continuity equation above the individuals u reproduce or die with rate g equal to the local fitness, and move along the velocity field ∇g towards the most favorable environment. Our construction was originally derived in the whole space $\Omega = \mathbb{R}^d$, but immediately extends to general domains imposing natural zero-flux boundary conditions on the velocity fields ∇g_t on $\partial\Omega$. In addition to nice geometrical and topological properties (completeness, existence of geodesics, lower semi-continuity with respect to weak-$*$ convergence, characterization of Lipschitz curves...) the metric $\mathtt{d}$ gives a formal Riemannian structure *à la Otto* [29] on the space $\mathcal{M}^+(\Omega)$, endowing the tangent plane

$$T_u\mathcal{M}^+ = \{\partial_t u = \zeta : \qquad \zeta = -\operatorname{div}(u\nabla g) + ug, \quad g \in H^1(\mathrm{d}u)\}$$

with the norm

$$\|\zeta\|^2_{T_u\mathcal{M}^+} = \|g\|^2_{H^1(\mathrm{d}u)} = \int_\Omega (|\nabla g|^2 + |g|^2) \mathrm{d}u \tag{3.1}$$

and scalar product

$$\langle \zeta_1, \zeta_2 \rangle_{T_u\mathcal{M}^+} := (g_1, g_2)_{H^1(\mathrm{d}u)} = \int_\Omega (\nabla g_1 \cdot \nabla g_2 + g_1 g_2) \mathrm{d}u.$$

Here tangent vectors $\partial_t u = \zeta \in T_u\mathcal{M}^+$ are always identified with scalar potentials $g \in H^1(\mathrm{d}u)$ through the elliptic equation $-\operatorname{div}(u\nabla g) + ug = \zeta$ (supplemented with zero-flux boundary conditions on $\partial\Omega$ if needed). In particular this allows to compute metric gradients for functionals $\mathcal{F}(u) = \int_\Omega F(x,u)$ on $\mathcal{M}^+$ as

$$\operatorname{grad}_{\mathtt{d}} \mathcal{F}(u) = -\operatorname{div}\left(u\nabla \frac{\delta F}{\delta u}\right) + u\frac{\delta F}{\delta u}, \tag{3.2}$$

where $\frac{\delta F}{\delta u} = \partial_u F(x,u)$ stands for the first variation with respect to u and $\nabla = \nabla_x$ is the usual gradient in space. We refer to [18] for further details and explanations.

Since we want to deal here with multicomponent variables $\mathbf{u} = (u_1, \dots, u_N)$ we endow $(\mathcal{M}^+(\Omega))^N$ with the natural product distance

$$\mathtt{D}^2(\mathbf{u},\mathbf{v}) = \sum_{i=1}^N \mathtt{d}^2(u_i, v_i),$$

giving the natural Riemannian metrics

$$\langle \boldsymbol{\zeta}^1, \boldsymbol{\zeta}^2 \rangle_{T_{\mathbf{u}}(\mathcal{M}^+)^N} = \left(\mathbf{g}^1, \mathbf{g}^2\right)_{H^1(\mathrm{d}\mathbf{u})} = \sum_{i=1}^N \int_\Omega (\nabla g_i^1 \cdot \nabla g_i^2 + g_i^1 g_i^2)\mathrm{d}u_i$$

in the tangent space $T_{\mathbf{u}}(\mathcal{M}^+)^N = \bigoplus_{i=1}^N T_{u_i}\mathcal{M}^+$. Here we identify again the tangent vectors $\boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_N)$ with potentials $\boldsymbol{g} = (g_1, \dots g_N)$ via the elliptic equations $-\operatorname{div}(u_i \nabla g_i) + u_i g_i = \zeta_i$ (with homogeneous Neumann boundary conditions). The metric derivatives with respect to $\mathtt{D}$ can be simply computed applying (3.2) component by component, i-e gradients of functionals $\mathcal{F}(\mathbf{u}) = \int_\Omega F(x, u_1, \dots, u_N)$ read

$$\operatorname{grad}_{\mathtt{D}} \mathcal{F}(\mathbf{u}) = \left(-\operatorname{div}\left(u_i \nabla \frac{\delta F}{\delta u_i}\right) + u_i \frac{\delta F}{\delta u_i}\right)_{i=1\dots N}.$$

For the particular case $\mathcal{E}(\mathbf{u}) = \frac{1}{2}\int_\Omega A(\mathbf{u}-\mathbf{u}^\infty)\cdot(\mathbf{u}-\mathbf{u}^\infty)$ and with the previous notation $\mathbf{f} = \mathbf{m} - A\mathbf{u}$, we have $\frac{\delta F}{\delta u_i} = (A(\mathbf{u}-\mathbf{u}^\infty))_i = (A\mathbf{u}-\mathbf{m})_i = -f_i$, thus our system of PDEs can indeed be written as the gradient flow

$$\frac{d\mathbf{u}}{dt} = -\operatorname{grad}_{\mathtt{D}} \mathcal{E}(\mathbf{u}) \quad \Leftrightarrow \quad \partial_t u_i = -\operatorname{div}(u_i \nabla f_i) + u_i f_i. \tag{3.3}$$

At this stage let us derive two formal estimates, which will be crucial for the subsequent analysis. Here we ignore all the regularity issues, and we shall make these estimates rigorous throughout the several regularized problems involved in the proof of existence. The first estimate is the Entropy-Dissipation-Inequality (2.3), and is inherent to the gradient-flow structure. Indeed from (3.3) we should have along reasonably smooth solutions that

$$\frac{d\mathcal{E}}{dt} = \left\langle \operatorname{grad}_{\mathtt{D}} \mathcal{E}(\mathbf{u}), \frac{d\mathbf{u}}{dt} \right\rangle_{T_{\mathbf{u}}(\mathcal{M}^+)^N} = -\| \operatorname{grad}_{\mathtt{D}} \mathcal{E}(\mathbf{u})\|^2_{T_{\mathbf{u}}(\mathcal{M}^+)^N}.$$

Given the above definition of the tangent norms and the explicit computation of the metric gradient in terms of the fitness, this reads in our setting

$$\mathcal{E}(\mathbf{u}(t_1)) + \sum_i \int_{t_0}^{t_1} \int_\Omega u_i(|\nabla f_i|^2 + |f_i|^2) = \mathcal{E}(\mathbf{u}(t_0)) \quad \text{for all } 0 \le t_0 \le t_1.$$

This is often referred to as the Entropy-Dissipation-*Equality* for the obvious reasons, and implies of course (2.3). However the latter *inequality* is well known to still completely characterize metric gradient flows [2], and will turn out to be more flexible and easier to obtain rigorously along the various approximations.

The second fundamental estimate, which will serve as a technical tool, is the a priori gradient estimate

$$\|\nabla \mathbf{u}\|^2_{L^2(0,T;L^2(\Omega)^{N\times d})} \le C(1+T)$$

and can be viewed as a "flow interchange" estimate as introduced in [28]. Indeed the estimate formally follows from computing the dissipation of the Boltzmann entropy $\mathcal{H}(u_i) = \int_\Omega \{u_i \log u_i - u_i + 1\}$ along solutions of our PDE, which is the gradient flow of the driving functional $\mathcal{E}$. More precisely, testing $\log u_i$ in (2.1) (recall that our weak solutions will be non-negative) we compute

$$\begin{aligned} \frac{d}{dt}\mathcal{H}(u_i) = \int_\Omega \log u_i \, \partial_t u_i = \int_\Omega \log u_i \big\{ -\operatorname{div}(u_i \nabla f_i + u_i f_i) \big\} \\ = \int \nabla u_i \cdot \nabla f_i + \int_\Omega u_i f_i \log u_i. \end{aligned}$$

Writing $f_i = m_i - (A\mathbf{u})_i$ in the last gradient term, summing over i's, and integrating in time, this can be rearranged as

$$\sum_i \left\{ \mathcal{H}(u_i(T)) + \int_0^T \int_\Omega \nabla u_i \cdot \nabla (A\mathbf{u})_i \right\} \leq \sum_i \left\{ \mathcal{H}(u_i(0)) \right. \\ \left. + \int_0^T \int_\Omega \nabla u_i \cdot \nabla m_i + \frac{1}{2}\int_0^T \int_\Omega u_i |f_i|^2 + \frac{1}{2}\int_0^T \int_\Omega u_i |\log u_i|^2 \right\}.$$

Observe from $\mathcal{E}(\mathbf{u}) \geq \frac{\lambda_A}{2}\|\mathbf{u}-\mathbf{m}\|_{L^2(\Omega)^N}$ and the previous EDI (2.3) that $\mathbf{u}(t)$ is bounded in $L^2(\Omega)^N$, thus the subquadratic terms $\mathcal{H}(u_i(0))$ and $\int_\Omega u_i|\log u_i|^2$ are controlled uniformly in time in the right-hand side. Exploiting next $\mathcal{H}(u_i(T)) \geq 0$ and the coercivity $A \geq \lambda_A$, a suitable use of Young's inequality finally gives

$$\frac{\lambda_A}{2}\|\nabla \mathbf{u}\|^2_{L^2(0,T;L^2(\Omega)^N)} \\ \leq C(1+T) + \frac{2}{\lambda_A}\|\nabla \mathbf{m}\|^2_{L^2(\Omega)^N} T + \sum_i \frac{1}{2}\int_0^T \int_\Omega u_i |f_i|^2 \\ \leq C(1+T).$$

Here we used (2.3) to bound the dissipation term $\int_0^\infty \int_\Omega u_i |f_i|^2 \leq \mathcal{E}(\mathbf{u}^0) \leq C$. We also implicitly assumed that A is a constant matrix, otherwise some extra lower terms appear but the gradient estimate is still true.

4. Existence of weak solutions

Our construction of weak solutions will involve three levels of approximation, indexed by the regularization parameters $\varepsilon, \delta \to 0$ and $M \to \infty$. More precisely, we shall consider the regularized problems

$$\begin{cases} \partial_t u_i + \varepsilon(\mathcal{A}\mathbf{u})_i = -\operatorname{div}(\tilde{u}_i \nabla f_i) + \tilde{u}_i f_i + \delta \Delta u_i, \\ u_i(0,x) = u_i^0(x) \end{cases} \quad \text{with } \mathbf{f} = \mathbf{m} - A\mathbf{u}. \tag{4.1}$$

Here

$$\tilde{u}_i = \max(0, \min(M, u_i))$$

is the truncation between 0 and $M \gg 1$. The operator $\mathcal{A}$ is a suitable elliptic operator of higher order to be precised shortly together with its associated boundary conditions, and will essentially allow to consider the second-order cross diffusion as a compact perturbation of $\varepsilon\mathcal{A}$ for fixed $\varepsilon > 0$. Note that in the original PDEs the terms $u_i \nabla f_i$ only belong to $L^1(\Omega)$ if $\mathbf{u} \in H^1(\Omega)^N$, while the truncations $\tilde{u}_i \nabla u_i \in L^2(\Omega)$ behave much better for fixed $M < \infty$. Lastly, the $\delta\Delta$ term will help to gain coercivity and control the degenerate cross-diffusion.

We shall first take $\varepsilon \to 0$, then $M \to \infty$, and finally $\delta \to 0$ in this order. The most delicate limit will be $M \to \infty$, when we loose the regularity $\tilde{u}_i \nabla f_i \in L^2(\Omega)$ to the more delicate $u_i \nabla f_i \in L^1(\Omega)$. This step will also require the rigorous justification of the formal computations and chain rules from Section 3, and will be the most involved. It is worth pointing out that solutions will become non-negative $u_i(t, x) \geq 0$ only after taking $\varepsilon \to 0$, which will then carry through the next limits $M \to \infty, \delta \to 0$. For convenience we will work in finite time intervals $[0, T]$. All our estimates will give local-in time control, and we will retrieve in the end a global solution by standard diagonal extraction. In order to keep the notation light we omit the ε, M, δ indexes as often as possible, and throughout the manuscript we will precise the dependence of the solution on the regularizing parameters when needed.

Our first step is to prove existence of solutions to (4.1) for small $\varepsilon, \delta > 0$ and large $M < \infty$ in any fixed time interval $[0, T]$. In order to give a precise meaning to this problem, consider the Hilbert triple (see the Appendix for the abstract definition)

$$(H^r(\Omega))^N, L^2(\Omega)^N, (H^r)^*(\Omega)^N) \qquad \text{for fixed } r > 1 + \frac{d}{2}.$$

Denote by $\mathcal{A}$ the Riesz isometry between the spaces $(H^r)^N$ and $((H^r)^*)^N$. We recall the Sobolev embedding $H^r(\Omega) \subset \mathcal{C}^1(\overline{\Omega})$, which is compact.

The weak form of (4.1) (with a certain implicit higher-order Neumann boundary condition which is of no importance to us) is the following Cauchy problem

$$\mathbf{u}' + \varepsilon \mathcal{A}\mathbf{u} = Q(\mathbf{u}), \qquad \mathbf{u}|_{t=0} = \mathbf{u}^0, \tag{4.2}$$

where the first equality is understood as an ODE in the space $((H^r)^*)^N$ and the initial datum should be taken in the sense of $\mathcal{C}([0, T]; (L^2)^N)$. The operator $Q : (H^1)^N \to ((H^1)^*)^N$ is determined by duality as

$$\langle Q(\mathbf{u}), \mathbf{w} \rangle = \sum_{i=1}^{N} \int_\Omega \left(\tilde{u}_i \nabla f_i \cdot \nabla w_i + \tilde{u}_i f_i w_i - \delta \nabla u_i \cdot \nabla w_i \right),$$

$\mathbf{w} \in (H^1)^N$. Note that, for fixed $M > 0$, the operator Q has sublinear growth and is continuous from $\mathcal{C}^1(\overline{\Omega})^N$ to $((H^r)^*)^N$. By Lemma 5.1 with $X = H^r(\Omega)^N$,

$V = \mathcal{C}^1(\overline{\Omega})^N$, $Y = L^2(\Omega)^N$, there exists a solution $\mathbf{u}$ to (4.2) in the class

$$L^2(0,T;(H^r)^N) \cap H^1(0,T;((H^r)^*)^N) \cap \mathcal{C}([0,T];(L^2)^N) \tag{4.3}$$

for any fixed $T > 0$.

4.1. The limit $\varepsilon \to 0$. In this section $M < \infty$ and $\delta > 0$ are fixed and only ε varies. In order to send $\varepsilon \to 0$ we first rigorously derive an entropy estimate. To this end, we test (4.2) with $-\mathbf{f}$ (this test function is legitimate since $\mathbf{f}(t,x) = \mathbf{m}(x) - A\mathbf{u}(t,x)$ belongs to the intersection (4.3)), and arrive at

$$\frac{d}{dt}\mathcal{E}(\mathbf{u}) + \varepsilon(\mathbf{u}, A\mathbf{u} - \mathbf{m})_{(H^r)^N} + \delta(\nabla\mathbf{u}, \nabla A\mathbf{u} - \nabla\mathbf{m})_{(L^2)^{N\times d}} \\ + \sum_{i=1}^N \int_\Omega \tilde{u}_i(|f_i|^2 + |\nabla f_i|^2) = 0. \tag{4.4}$$

Integrating in time from 0 to any $\tau \le T$, exploiting the coercivity $A \ge \lambda_A \,\mathrm{Id}$, and applying Cauchy's inequality we find that

$$\mathcal{E}(\mathbf{u}(\tau)) + \frac{1}{2}\varepsilon\lambda_A\|\mathbf{u}\|^2_{L^2(0,\tau;(H^r)^N)} + \frac{1}{2}\delta\lambda_A\|\nabla\mathbf{u}\|^2_{L^2(0,\tau;(L^2)^{N\times d})} \\ + \sum_i \int_0^\tau \int_\Omega \tilde{u}_i(|\nabla f_i|^2 + |f_i|^2) \\ \le \mathcal{E}(\mathbf{u}^0) + T\frac{\varepsilon}{2\lambda_A}\|\mathbf{m}\|^2_{(H^r)^N} + T\frac{\delta}{2\lambda_A}\|\nabla\mathbf{m}\|^2_{(L^2)^{N\times d}}. \tag{4.5}$$

Recalling that $\mathcal{E}(\mathbf{u})$ controls the L^2 norm, this implies the a priori estimates

$$\|\mathbf{u}\|^2_{L^\infty(0,T;(L^2)^N)} \le C(1+T), \tag{4.6}$$

$$\|\mathbf{u}\|^2_{L^2(0,T;(H^1)^N)} \le C\delta^{-1}(1+T), \tag{4.7}$$

$$\|\mathbf{u}\|^2_{L^2(0,T;(H^r)^N)} \le C\varepsilon^{-1}(1+T) \tag{4.8}$$

for small ε, δ, where the various constants C are independent of $M, \varepsilon, \delta, T$. Testing (4.2) by arbitrary $\mathbf{w} \in (H^r)^N$ and employing (4.7) and (4.8), we deduce

$$\|\mathbf{u}'\|_{L^2(0,T;((H^r)^*)^N)} \le C_1,$$

where $C_1 \equiv C_1(\delta, M, T)$ is independent of ε. By the Banach-Alaoglu theorem and Aubin-Lions-Simon lemma, we can find a sequence $\mathbf{u}^{\varepsilon_k}$ of solutions to (4.2) with $\varepsilon = \varepsilon_k \to 0$ (for fixed M, δ) such that

$$\mathbf{u}^{\varepsilon_k} \rightharpoonup \mathbf{u} \text{ weakly in } L^2(0,T;(H^1)^N) \text{ and weakly-* in } L^\infty(0,T;(L^2)^N),$$

$$\mathbf{u}^{\varepsilon_k} \to \mathbf{u} \text{ strongly in } L^2(0,T;(L^2)^N) \text{ and in } \mathcal{C}([0,T];((H^1)^*)^N),$$

$$(\mathbf{u}^{\varepsilon_k})' \rightharpoonup \mathbf{u}' \text{ weakly in } L^2(0,T;((H^r)^*)^N).$$

By the classical continuity property [20] of Nemytskii truncations,

$$\tilde{\mathbf{u}}^{\varepsilon_k} \to \tilde{\mathbf{u}} \text{ strongly in } L^2(0,T;(L^2)^N),$$

and because $\mathbf{f}^{\varepsilon_k} = \mathbf{m} - A\mathbf{u}^{\varepsilon_k} \to \mathbf{m} - A\mathbf{u} = \mathbf{f}$ strongly in $L^2(0,T;(L^2)^N)$ and weakly in $L^2(0,T;H^1)$ the products

$$\tilde{u}_i^{\varepsilon_k} \nabla f_i^{\varepsilon_k} \rightharpoonup \tilde{u}_i \nabla f_i \quad \text{and} \quad \tilde{u}_i^{\varepsilon_k} f_i^{\varepsilon_k} \rightharpoonup \tilde{u}_i f_i \qquad \text{weakly in } L^1(0,T;L^1)$$

converge as strong-weak limits. Passing to the limit $\varepsilon_k \to 0$ in (4.2), we see that $\mathbf{u}$ solves the problem

$$\mathbf{u}' = Q(\mathbf{u}), \quad \mathbf{u}|_{t=0} = \mathbf{u}^0, \tag{4.9}$$

which is nothing but the weak form of the problem

$$\begin{cases} \partial_t u_i = -\operatorname{div}(\tilde{u}_i \nabla f_i) + \tilde{u}_i f_i + \delta \Delta u_i, & x \in \Omega, \\ \tilde{u}_i \frac{\partial f_i}{\partial \nu} - \delta \frac{\partial u_i}{\partial \nu} = 0, & x \in \partial\Omega, \\ u_i(0,x) = u_{0i}(x). \end{cases} \tag{4.10}$$

By density it is easy to check that $\mathbf{u}' \in L^2(0,T;((H^1)^*)^N)$, and the first equality in (4.9) holds in the space $(H^1)^*)^N$ for a.e. t. By standard Lions-Magenes interpolation results [35, Lemma 2.2.7] we have moreover $\mathbf{u} \in \mathcal{C}([0,T];(L^2)^N)$.

We now show non-negativity of our weak solution $\mathbf{u} = \lim \mathbf{u}^{\varepsilon_k}$ to (4.10) and derive a priori $L^2(0,T;(H^1)^N)$ estimates uniformly in M, δ, which will allow to take the limit $M \to \infty, \delta \to 0$ in the next sections. After the previous limit $\varepsilon \to 0$ the solution $\mathbf{u}$ belongs at this stage to $\mathcal{C}([0,T];(L^2)^N) \cap L^2(0,T;(H^1)^N)$ for fixed $M, \delta > 0$. Therefore we can take again $-\mathbf{f}$ as a test function and repeat the previous computations (4.4)(4.5) with now $\varepsilon = 0$, and we get similarly

$$\mathcal{E}(\mathbf{u}(t_1)) + \sum_i \int_{t_0}^{t_1} \int_\Omega \tilde{u}_i (|\nabla f_i|^2 + |f_i|^2) \le \mathcal{E}(\mathbf{u}(t_0)) + T \frac{\delta}{2\lambda_A} \|\nabla \mathbf{m}\|^2_{(L^2)^{N\times d}} \tag{4.11}$$

for all $0 \leq t_0 \leq t_1 \leq T$. This is of course an approximation of the Entropy Dissipation Inequality (2.3), which will pass to the successive limits $M \to \infty$ and $\delta \to 0$ later on.

In order show that $u_i \geq 0$ we take the admissible test function $v_i := \min\{u_i, 0\} \in L^2(0,T;H^1)$ in (4.10). By the classical Serrin's chain rule $\nabla v_i = \chi_{[u_i<0]}\nabla u_i$ we get for all components $i = 1 \dots N$

$$\frac{d}{dt}\left(\frac{1}{2}\int_\Omega |v_i|^2\right) = \int_\Omega \tilde{u}_i \nabla f_i \cdot \nabla v_i + \int_\Omega \tilde{u}_i f_i v_i - \delta \int_\Omega |\nabla v_i|^2 \leq 0,$$

where we used that by definition the truncation

$$\tilde{u}_i = \max\{0, \min\{u_i, M\}\} = 0$$

wherever $v_i = \min\{u_i, 0\} \neq 0$ and $\nabla v_i = \nabla u_i \chi_{[u_i<0]} \neq 0$ so that the first two integrands in the middle term are identically zero. Since we consider non-negative initial data $u_i^0 \geq 0$ we have $v_i(0,.) = 0$, thus $v_i(t,.) = 0$ for all later times and

$$v_i = \min\{0, u_i\} \equiv 0 \qquad \Rightarrow \; u_i(t,x) \geq 0 \text{ a.e. in } (0,T)\times\Omega.$$

From now on we slightly abuse the notation and still write $\tilde{u}_i = \min\{u_i, M\} \geq 0$ for the upper truncation only, which is justified since we just proved that $u_i \geq 0$.

In order to mimic the formal gradient estimate from Section 3, we would like to test $\log u_i$ in (4.10). However this is not rigorous because $\log u_i$ may not be an admissible test function. We use instead the truncated logarithm and Boltzmann entropy, defined as

$$\log_\zeta^M(z) := \begin{cases} \log\zeta & \text{if } 0 < z \leq \zeta \\ \log z & \text{if } \zeta \leq z < M \\ \log M & \text{else} \end{cases} \quad \text{and } H_\zeta^M(z) := \int_1^z \log_\zeta^M(s)\mathrm{d}s$$

for small $\zeta > 0$ ($M > 0$ is the same truncation level as before). Note that

$$0 \leq H_\zeta^M(z) \leq H(z) := z\log z - z + 1$$

with monotone pointwise convergence $H_\zeta^M(.) \nearrow H(.)$ as $\zeta \searrow 0$ and $M \nearrow \infty$ (the convergence is actually locally uniform). We stress at this point that all the next estimates will be uniform in δ, M, and all the constants C_T below will depend on the data and $T > 0$ only (if $\delta, \zeta > 0$ are small and $M > 0$ is large).

For fixed $0 < \zeta < M < \infty$ the functions $\log_\zeta^M, H_\zeta^M$ are globally Lipschitz, thus $\log_\zeta^M(u_i) \in L^2(0,T;H^1)$ is an admissible test function in (4.10) and, using the chain rule $\nabla \log_\zeta^M(u_i) = \frac{\nabla u_i}{u_i}\chi_{[\zeta<u_i<M]}$,

$$\frac{d}{dt}\left(\int_\Omega H_\zeta^M(u_i)\right) = \left\langle \partial_t u_i, \log_\zeta^M(u_i)\right\rangle_{(H^1)^*,H^1}$$
$$= \int\limits_{[\zeta<u_i<M]} \nabla f_i \cdot \nabla u_i + \int_\Omega \tilde{u}_i f_i \log_\zeta^M(u_i) - \delta \int\limits_{[\zeta<u_i<M]} \frac{|\nabla u_i|^2}{|u_i|}.$$

Integrating from $t=0$ to $t=T$, exploiting the monotonicity $0 \leq H_\zeta^M \leq H$, applying Young's inequality to $(f_i, \log_\zeta^M u_i)_{L^2(\mathrm{d}\tilde{u}_i)}$, and discarding the last non-positive term, we get

$$0 \leq \int_\Omega H_\zeta^M(u_i(T))$$
$$\leq \int_\Omega H_\zeta^M(u_i(0)) + \iint\limits_{Q_T\cap[\zeta<u_i<M]} \nabla f_i \cdot \nabla u_i + \iint\limits_{Q_T} \tilde{u}_i f_i \log_\zeta^M(u_i)$$

$$\leq \int_\Omega H(u_i(0)) + \iint\limits_{Q_T\cap[\zeta<u_i<M]} \nabla f_i \cdot \nabla u_i +$$
$$\frac{1}{2}\left(\iint\limits_{Q_T} \tilde{u}_i |f_i|^2 + \iint\limits_{Q_T} \tilde{u}_i \left|\log_\zeta^M(u_i)\right|^2\right)$$
$$\leq \mathcal{H}(u_i(0)) + \iint\limits_{Q_T\cap[\zeta<u_i<M]} \nabla f_i \cdot \nabla u_i + \frac{1}{2}\iint\limits_{Q_T} \tilde{u}_i|f_i|^2 + \frac{1}{2}\iint\limits_{Q_T} u_i \left|\log(u_i)\right|^2.$$

In the last line we also used $|\log_\zeta^M(z)| \leq |\log z|$ for small ζ and large M. Controlling the subquadratic terms $\mathcal{H}(u_i(0)), \iint u_i|\log u_i|^2$ by $u_i(0) \in L^2(\Omega)$ and $\|u_i\|_{L^\infty(0,T;L^2)} \leq C_T$, and exploiting the dissipation estimate (4.11) we get

$$0 \leq C_T + \iint\limits_{Q_T\cap[\zeta<u_i<M]} \nabla f_i \cdot \nabla u_i$$

for all $i = 1 \dots N$. This immediately passes to the limit $\zeta \searrow 0$, and recalling that by definition $f_i = m_i - (A\mathbf{u})_i$ we rewrite this limit as

$$\iint\limits_{Q_T} \nabla(A\mathbf{u})_i \cdot \nabla u_i \leq C_T + \iint\limits_{Q_T} \nabla m_i \cdot \nabla u_i - \iint\limits_{[u_i \geq M]} \nabla f_i \cdot \nabla u_i.$$

Summing over i's, taking advantage of the coercivity $A \geq \lambda_A$, and suitably applying Young's inequality to the last two terms, this easily gives

$$\iint\limits_{Q_T} |\nabla \mathbf{u}|^2 \leq C_T + C_A \sum_i \iint\limits_{[u_i \geq M]} |\nabla f_i|^2$$

and it is enough to show that the last term can be bounded uniformly in M, δ. To this end, observe by definition of $\tilde{u}_i = \min\{u_i, M\}$ that the dissipation estimate (4.11) immediately yields

$$\iint\limits_{[u_i \geq M]} |\nabla f_i|^2 = \iint\limits_{[u_i \geq M]} \frac{\tilde{u}_i}{M} |\nabla f_i|^2 \leq \frac{1}{M} \iint\limits_{Q_T} \tilde{u}_i |\nabla f_i|^2 \leq M^{-1} C_T \leq C_T$$

if $M \geq 1$. This finally gives the desired gradient estimate

$$\iint_{Q_T} |\nabla \mathbf{u}|^2 \leq C_T$$

uniformly in $M, \delta > 0$.

In order to estimate the time derivative, let us recall that the weak formulation of (4.10) holds in duality with all $\mathcal{C}^1(\overline{\Omega})$ (actually, even H^1) test functions. Since $\|\mathbf{u}\|_{L^\infty(0,T;(L^2)^N)} \leq C_T$ and we just proved that $\|\nabla \mathbf{u}\|_{L^2(0,T;(L^2)^{N \times d})} \leq C_T$ we see that the products $\tilde{u}_i \nabla f_i$ and $\tilde{u}_i f_i$ are bounded respectively in $L^2(0,T;L^1)$ and $L^\infty(0,T;L^1)$ uniformly in δ, M. We conclude by duality in (4.10) that $\partial_t u_i = -\operatorname{div}(\tilde{u}_i \nabla f_i) + \tilde{u}_i f_i + \delta \Delta u_i$ is bounded in $L^2(0,T;(\mathcal{C}^1(\overline{\Omega})^*)^N)$. To summarize, the solution $\mathbf{u} = \lim \mathbf{u}^{\varepsilon_k}$ of (4.10) has the regularity

$$\|\mathbf{u}'\|_{L^2(0,T;(\mathcal{C}^1(\overline{\Omega})^*)^N)} + \|\mathbf{u}\|_{L^\infty(0,T;(L^2)^N)} + \|\mathbf{u}\|_{L^2(0,T;(H^1)^N)} \leq C_T \tag{4.12}$$

uniformly in M, δ.

4.2. The limit $M \to \infty$. Here we want to pass to the limit $M \to +\infty$ in (4.9) and (4.10) for fixed $\delta > 0$.

Define the limit operator $Q_\infty : (H^1)^N \to ((\mathcal{C}^1(\overline{\Omega}))^*)^N$ by

$$\langle Q_\infty(\mathbf{u}), \mathbf{w}\rangle = \sum_{i=1}^N \int_\Omega (u_i \nabla f_i \cdot \nabla w_i + u_i f_i w_i - \delta \nabla u_i \cdot \nabla w_i),$$

$\mathbf{w} \in \mathcal{C}^1(\overline{\Omega})^N$. By (4.12), the Banach-Alaoglu theorem and Aubin-Lions-Simon lemma, there exists a sequence $M_k \to +\infty$ ($\delta > 0$ is fixed) such that, for the corresponding solutions $\mathbf{u}^{M_k}$, we have

$$\mathbf{u}^{M_k} \rightharpoonup \mathbf{u} \text{ weakly in } L^2(0,T;(H^1)^N) \text{ and weakly-* in } L^\infty(0,T;L^2),$$
$$\mathbf{u}^{M_k} \to \mathbf{u} \text{ strongly in } L^2(0,T;(L^2)^N) \text{ and in } \mathcal{C}([0,T];((H^1)^*)^N),$$
$$(\mathbf{u}^{M_k})' \rightharpoonup \mathbf{u}' \text{ weakly-* in } L^2(0,T;(\mathcal{C}^1(\overline{\Omega})^*)^N)$$

with

$$\|\mathbf{u}'\|_{L^2(0,T;(\mathcal{C}^1(\overline{\Omega})^*)^N)} + \|\mathbf{u}\|_{L^\infty(0,T;(L^2)^N)} + \|\mathbf{u}\|_{L^2(0,T;(H^1)^N)} \le C_T \tag{4.13}$$

(uniformly in δ). Because the truncation operator $z \mapsto \tilde{z} = \min\{z, M\}$ is 1-Lipschitz uniformly in M and $u_i^{M_k} \to u_i$ strongly in L^2, one readily checks that

$$\tilde{u}_i^{M_k} = \min\left\{u_i^{M_k}, M_k\right\} \to u_i \qquad \text{in } L^2(0,T;L^2).$$

Therefore the products pass to the strong-weak limit as before, $\tilde{u}_i^{M_k} \nabla f_i^{M_k} \rightharpoonup u_i \nabla f_i$ and $\tilde{u}_i^{M_k} f_i^{M_k} \rightharpoonup u_i f_i$ in $L^1(0,T;L^1)$, and $\mathbf{u}$ solves the weak formulation

$$\mathbf{u}' = Q_\infty(\mathbf{u}), \quad \mathbf{u}|_{t=0} = \mathbf{u}_0$$

of

$$\begin{cases} \partial_t u_i = -\operatorname{div}(u_i \nabla f_i) + u_i f_i + \delta \Delta u_i, & x \in \Omega, \\ u_i \frac{\partial f_i}{\partial \nu} - \delta \frac{\partial u_i}{\partial \nu} = 0, & x \in \partial\Omega, \\ u_i(0,x) = u_{0i}(x). \end{cases} \tag{4.14}$$

In other words, $\mathbf{u}$ solves

$$\frac{d}{dt}\int_\Omega \mathbf{u} \cdot \mathbf{w} = \sum_{i=1}^N \int_\Omega (u_i \nabla f_i \cdot \nabla w_i + u_i f_i w_i - \delta \nabla u_i \cdot \nabla w_i),$$

$\mathbf{w} \in (\mathcal{C}^1(\overline{\Omega}))^N$, in the sense of scalar distributions. Moreover, by the Lions-Magenes lemma [35, Lemma 2.2.6] we see that

$$\mathbf{u} \in \mathcal{C}_w([0,T];(L^2)^N),$$

hence the initial condition is taken in $\mathcal{C}_w([0,T];(L^2)^N)$.

Before moving to the next limit $\delta \to 0$, we need to show that the dissipation estimate (4.11) also passes to the limit $M_k \to \infty$. This is not straightforward because of the cubic products $\tilde{u}_i(|\nabla f_i|^2+|f_i|^2)$, as $\tilde{u}_i^{M_k}$ does not a priori converge uniformly and $f_i^{M_k}$ should not converge strongly in $L^2(0,T;H^1)$. In order to circumvent this technical difficulty we use a variant of the Banach-Alaoglu theorem in varying $L^2(\mathrm{d}\mu^k)$ spaces:

Lemma 4.1 (compactness of vector-fields)**.** *Let $\mathcal{O} \subset \mathbb{R}^p$ be an open set, $\{\mu^k\}_{k\geq 0}$ a sequence of finite non-negative Radon measures narrowly converging to μ, and $\mathbf{v}^k$ a sequence of vector fields on $\mathcal{O}$. If*

$$\|\mathbf{v}^k\|_{L^2(\mathcal{O},\mathrm{d}\mu^k)} \leq C$$

then there exists $\mathbf{v} \in L^2(\mathcal{O},\mathrm{d}\mu)$ such that, up to extraction of some subsequence,

$$\forall\, \boldsymbol{\zeta} \in \mathcal{C}_c^\infty(\mathcal{O}): \qquad \lim_{k\to\infty}\int_{\mathcal{O}} \mathbf{v}^k\cdot\boldsymbol{\zeta}\,\mathrm{d}\mu^k = \int_{\mathcal{O}} \mathbf{v}\cdot\boldsymbol{\zeta}\,\mathrm{d}\mu$$

and

$$\|\mathbf{v}\|_{L^2(\mathcal{O},\mathrm{d}\mu)} \leq \liminf_{k\to\infty}\|\mathbf{v}^k\|_{L^2(\mathcal{O},\mathrm{d}\mu^k)}.$$

The proof of this fact by optimal transport techniques can be found in [2]; this lemma also follows from a variant of the Banach-Alaoglu theorem [18, Proposition 5.3]. We will apply this lemma component by component with $\mathcal{O} = (t_0,t_1)\times\Omega \subset \mathbb{R}^{d+1}$ for fixed $0 \leq t_0 \leq t_1 \leq T$ and the sequence of measures $\mathrm{d}\mu^k(t,x) := \tilde{u}_i^{M_k}(t,x)\mathrm{d}x\mathrm{d}t$, which converges narrowly to $\mathrm{d}\mu(t,x) = u_i(t,x)\mathrm{d}x\mathrm{d}t$ due to the strong $L^2(Q_T)$ convergence $\tilde{u}_i^{M_k} \to u_i$. The vector-fields of interest are of course the $\mathbb{R}^d$-valued $\nabla f_i^{M_k}$ and the scalar $f_i^{M_k}$. Indeed, for fixed $0 \leq t_0 \leq t_1 \leq T$ we have by (4.11) that

$$\|\nabla f_i^{M_k}\|^2_{L^2(\mathcal{O},\mathrm{d}\mu^k)} = \int_{t_0}^{t_1}\int_\Omega |\nabla f_i^{M_k}|^2\tilde{u}_i^{M_k} \leq \mathcal{E}(\mathbf{u}(0)) + C\delta T \leq C_T.$$

Extracting a subsequence if needed, we see that there is a vector-field $\mathbf{v}_i \in L^2(\mathcal{O},\mathrm{d}\mu)$ such that

$$\|\mathbf{v}_i\|^2_{L^2(\mathcal{O},\mathrm{d}\mu)} \leq \liminf_{k\to\infty}\|\nabla f_i^{M_k}\|^2_{L^2(\mathcal{O},\mathrm{d}\mu^k)},$$

and we claim that $\mathbf{v}_i = \nabla f_i$ in $L^2(\mathcal{O}, \mathrm{d}\mu)$. To see this, observe that from the weak convergence in Lemma 4.1 there holds

$$\lim_{k\to\infty} \int_{t_0}^{t_1} \int_\Omega \nabla f_i^{M_k} \cdot \zeta \tilde{u}_i^{M_k} = \int_{t_0}^{t_1} \int_\Omega \mathbf{v} \cdot \zeta u_i$$

for all $\zeta \in \mathcal{C}_c^\infty((t_0, t_1) \times \Omega; \mathbb{R}^d)$. On the other hand we already proved that $\tilde{u}_i^{M_k} \to u_i$ strongly in $L^2(0, T; L^2)$ and $\nabla f_i^{M_k} \rightharpoonup \nabla f_i$ weakly in $L^2(0, T; L^2)$, thus $\tilde{u}_i^{M_k} \nabla f_i^{M_k} \rightharpoonup u_i \nabla f_i$ weakly in $L^1(0, T; L^1)$. Therefore

$$\int_{t_0}^{t_1} \int_\Omega \nabla f_i \cdot \boldsymbol{\zeta} u_i = \int_{t_0}^{t_1} \int_\Omega \mathbf{v}_i \cdot \boldsymbol{\zeta} u_i$$

for all $\boldsymbol{\zeta} \in \mathcal{C}_c^\infty((t_0, t_1) \times \Omega; \mathbb{R}^d)$. By density of $\mathcal{C}_c^\infty$ we conclude that $\mathbf{v}_i = \nabla f_i$ in $L^2(\mathcal{O}, \mathrm{d}\mu)$, which shows in particular that the limit is independent of t_0, t_1 and the subsequence. Whence

$$\forall\, 0 \le t_0 \le t_1 : \qquad \int_{t_0}^{t_1} \int_\Omega u_i |\nabla f_i|^2 \le \liminf_{k\to\infty} \int_{t_0}^{t_1} \int_\Omega \tilde{u}_i^{M_k} |\nabla f_i^{M_k}|^2.$$

The argument is identical for the terms $\tilde{u}_i^{M_k} |f_i^{M_k}|^2$. In order to finally retrieve the Entropy-Dissipation-Inequality, observe that $\mathbf{u}^{M_k} \to \mathbf{u}$ in $L^2(0, T; (L^2)^N)$ implies that $\mathcal{E}(\mathbf{u}^{M_k}(t)) \to \mathcal{E}(\mathbf{u}(t))$ for almost every $t \in (0, T)$. Taking the $\liminf\limits_{M_k\to\infty}$ in (4.11) we obtain

$$\mathcal{E}(\mathbf{u}(t_1)) + \sum_i \int_{t_0}^{t_1} \int_\Omega u_i (|\nabla f_i|^2 + |f_i|^2) \le \mathcal{E}(\mathbf{u}(t_0)) + T \frac{\delta}{2\lambda_A} \|\nabla \mathbf{m}\|_{L^2}^2 \tag{4.15}$$

for almost every $0 \le t_0 \le t_1 \le T$.

4.3. The limit $\delta \to 0$. We are now ready to pass to the last limit $\delta \to 0$. By (4.13) we can find a sequence $\delta_k \to 0$ such that, for the corresponding solutions $\mathbf{u}^{\delta_k}$ to (4.14), we have

$$\mathbf{u}^{\delta_k} \rightharpoonup \mathbf{u} \text{ weakly in } L^2(0, T; (H^1)^N) \text{ and weakly-* in } L^\infty(0, T; (L^2)^N),$$
$$\mathbf{u}^{\delta_k} \to \mathbf{u} \text{ strongly in } L^2(0, T; (L^2)^N) \text{ and in } \mathcal{C}([0, T]; ((H^1)^*)^N),$$
$$(\mathbf{u}^{\delta_k})' \rightharpoonup \mathbf{u}' \text{ weakly-* in } L^2(0, T; (\mathcal{C}^1(\overline{\Omega})^*)^N)$$

for every fixed $T > 0$. Arguing as before for the products $u_i^{\delta_k} \nabla f_i^{\delta_k}$ and $u_i^{\delta_k} f_i^{\delta_k}$ we can take the strong-weak limits, and the $\delta \Delta u_i$ term in (4.14) goes weakly to zero due to the $L^2(0, T; H^1)$ bound. Similarly to the previous step, $\mathbf{u} \in$

$\mathcal{C}_w([0,T];(L^2)^N)$, and we can pass to the limit in the initial condition. Thus the limit $\mathbf{u}$ is a weak solution to the original problem (2.1). By standard diagonal extraction arguments it is easy to see that $\mathbf{u} = \lim \mathbf{u}^{\delta_k}$ can be chosen independent of the fixed time $T > 0$. Thus the above convergence holds locally in time in $[0,\infty)$, the weak solution is global, and has the desired regularity in any finite time interval.

As for the Entropy-Dissipation-Inequality, we can repeat the exact same argument as in the previous section and pass to the $\liminf\limits_{\delta_k \to 0}$ in (4.15) (with the last term $T\frac{\delta_k}{2\lambda_A}\|\nabla \mathbf{m}\|^2_{(L^2)^{N\times d}}$ vanishing in any finite time interval) to obtain (2.3).

5. Long-time convergence

This section is devoted to the proof of Theorem 2, and without further mention we assume that for any $I = \{i_1, \dots, i_r\} \subset \{1, \dots, N\}$, $i_1 < \dots < i_r$, and $j \notin I$ there holds

$$\begin{vmatrix} a_{i_1 i_1} & \cdots & a_{i_1 i_r} & m_{i_1} \\ \vdots & \ddots & \vdots & \vdots \\ a_{i_r i_1} & \cdots & a_{i_r i_r} & m_{i_r} \\ a_{j i_1} & \cdots & a_{j i_r} & m_j \end{vmatrix} (x) \geq \kappa \tag{5.1}$$

for some constant $\kappa \equiv \kappa(A, \mathbf{m}) > 0$.

Remark 5.1. *Letting $I = \emptyset$ in (5.1), we see that necessarily $m_j(x) \geq \kappa > 0$ for any j and $x \in \Omega$, which means that there are only positive resources.*

With this assumption, some elementary algebra shows in particular that the ideal free distribution

$$\mathbf{u}^\infty(x) = A^{-1}\mathbf{m}(x) \geq c_\kappa > 0$$

becomes now a biologically relevant (non-negative) coexistence steady state of (2.1) with $\mathbf{f} \equiv 0$. This particular distribution is clearly mathematically significant given the definition of the entropy (2.2), and we shall prove below that it also attracts the long-time dynamics as in Theorem 2. However, (5.1) also implies the existence of a finite number of non-negative *partial extinction* steady states constructed as follows: given any $I \subset \{1, \dots, N\}$ and recalling that by definition $\mathbf{f} = \mathbf{m} - A\mathbf{u}$, it is easy to see from (5.1) that the linear system

$$\begin{cases} u_i = 0 & (i \in I) \\ f_j = 0 & (j \notin I), \end{cases}$$

has a unique solution $\mathbf{u}^I(x)$ satisfying $u_j(x) \geq c_\kappa > 0$ for all $j \notin I$ (thus $\mathbf{u}^I(x) > 0$ componentwise). Note that those $\mathbf{u}^I$ are trivially steady states of (2.1) with $(u_i)_{i\in I} \equiv 0$ and $(f_j)_{j\notin I} \equiv 0$, and that the ideal free distribution $\mathbf{u}^\infty = \mathbf{u}^\emptyset$ is the unique coexistence state obtained by taking $I = \emptyset \Leftrightarrow \mathbf{f} = 0$. In fact our condition (5.1) is equivalent to the hypothesis that all the components $(u^I_j)_{j\notin J}$ of those steady states are positive and uniformly bounded away from zero, cf. [19] (by definition the other components $(u^I_i)_{i\in I}$ vanish identically).

The partial extinction set is then the collection of all such stationary solutions $\mathbf{u}^I(x)$ for all possible choices of $I \subset \{1, \dots, N\}$ with $I \neq \emptyset$, of which there is a finite combinatorial number $p_N = 2^N - 1$:

$$\begin{aligned}\mathbf{U}^{ext} &= \{\mathbf{u}^{ext,1}(x), \dots, \mathbf{u}^{ext,p_N}(x)\}\\ &= \{\mathbf{u}^I(x) : \emptyset \neq I \subset \{1, \dots, N\}\} \subset H^1(\Omega)^N.\end{aligned}$$

The *critical entropy* $E^* > 0$ appearing in Theorem 2 is then defined as the minimal entropy among all the partial extinction states,

$$E^* = \min\left\{\mathcal{E}(\mathbf{u}^{ext}) : \quad \mathbf{u}^{ext} \in \mathbf{U}^{ext}\right\}, \tag{5.2}$$

and depends only on $A, \mathbf{m}$ (and Ω).

Though biologically admissible, the partial extinction states are actually degenerate points in our analysis: whenever $u_i(x) \equiv 0$, the formal Riemannian structure from Section 3 degenerates since the i-th tangent plane $T_{u_i}\mathcal{M}^+ = \{0\}$ becomes trivial, see in particular the definition of tangent norms (3.1) in terms of $\|.\|_{H^1(\mathrm{d}u)}$ norms. As a consequence we will need to stay away from those points. This is particularly clear in the following functional inequality, which will be the key to proving the long-time convergence below and follows from a more general Poincaré-Beckner inequality established by us in [19]:

Theorem 3. *Let $\mathbf{U} \subset H^1(\Omega, \mathbb{R}^N)$ be a set of functions such that*

(i) $\mathbf{u} \geq 0$ for any $\mathbf{u} \in \mathbf{U}$;

(ii) no sequence $\{\mathbf{u}^k\}_{k\geq 0} \subset \mathbf{U}$ converges strongly in $L^q(\Omega)^N$ to any of the partial extinction states $\mathbf{u}^{ext} \in \mathbf{U}^{ext}$ for some $q \in [1, 2)$.

Then there exists a constant $C_\mathbf{U} > 0$ such that

$$\forall\, \mathbf{u} \in \mathbf{U} : \qquad \int_\Omega \sum_{i=1}^N |f_i|^2 \,\mathrm{d}x \leq C_\mathbf{U} \int_\Omega \sum_{i=1}^N u_i(|f_i|^2 + |\nabla f_i|^2) \,\mathrm{d}x. \tag{5.3}$$

Remark 5.2. *Apart from $\mathbf{U}$, the constant $C_\mathbf{U}$ also depends on the upper bounds for $|\mathbf{m}|$ and $|A|$, on λ_A, and on κ in (5.1).*

Condition (ii) means that the $\mathbf{U}$ must be separated from the finite set of partial extinction points $\mathbf{U}^{ext}$. Moreover, this Poincaré-type inequality can be reinterpreted as an *entropy-entropy production* inequality, as is common in the framework of gradient flows in Wasserstein spaces. Indeed from Section 3 and the formal Riemannian structure the right-hand side is nothing but the dissipation $\mathcal{D}(\mathbf{u}) = \|\operatorname{grad}_{\mathtt{D}} \mathcal{E}(\mathbf{u})\|^2_{T_\mathbf{u}\mathcal{M}^+}$, and recalling $\mathcal{E}(\mathbf{u}) \lesssim \|\mathbf{f}\|^2_{L^2(\Omega)}$ from the coercivity, the left hand side controls $\mathcal{E}(\mathbf{u}) = \mathcal{E}(\mathbf{u}) - \mathcal{E}(\mathbf{u}^\infty)$. Thus (5.3) gives the entropy-entropy dissipation control $\mathcal{D}(\mathbf{u}) \geq C(\mathcal{E}(\mathbf{u}) - \mathcal{E}(\mathbf{u}^\infty))$, which classically implies convergence in the entropy sense.

We are now in position of proving the long-time convergence:

Proof of Theorem 2: Recalling the definition of the critical entropy (5.2) and given a subcritical initial data

$$\mathcal{E}(\mathbf{u}^0) < E^*,$$

we introduce the set

$$\mathbf{U} := \left\{\mathbf{u} = (u_1, \dots, u_N) \colon u_i \geq 0, \quad u_i \in H^1(\Omega), \quad \mathcal{E}(\mathbf{u}) \leq \mathcal{E}(\mathbf{u}^0)\right\}$$

(depending only on $\mathbf{u}^0$, A, $\mathbf{m}$, and Ω). From the EDI (2.3) we see that we have invariance

$$\mathbf{u}(0) = \mathbf{u}^0 \in \mathbf{U} \quad \Rightarrow \quad \mathbf{u}(t) \in \mathbf{U} \text{ for a.a. } t \geq 0$$

along the time-evolution, and we claim that $\mathbf{U}$ meets the assumptions of Theorem 3. To see this, assume by contradiction that there is a sequence $\mathbf{u}^k \in \mathbf{U}$ such that $\mathbf{u}^k \to \mathbf{u}^{ext}$ strongly in $L^q(\Omega)$ for some partial extinction state $\mathbf{u}^{ext} \in \mathbf{U}^{ext}$ and some $q \in [1, 2)$. Since $\frac{\lambda_A}{2}\|\mathbf{u}^k - \mathbf{u}^\infty\|^2_{L^2(\Omega)} \leq \mathcal{E}(\mathbf{u}^k) \leq \mathcal{E}(\mathbf{u}^0)$ we see that $\{\mathbf{u}^k\}$ is bounded in $L^2(\Omega)$, and up to extraction of a subsequence we can therefore assume that $\mathbf{u}^k \rightharpoonup \mathbf{u}$ weakly in L^2 for some limit $\mathbf{u}$. By uniqueness of the limit we see that $\mathbf{u} = \mathbf{u}^{ext}$, and by lower semi-continuity

$$\mathcal{E}(\mathbf{u}^{ext}) \leq \liminf_{k\to\infty} \mathcal{E}(\mathbf{u}^k) \leq \mathcal{E}(\mathbf{u}^0) < E^*.$$

This is impossible by the definition (5.2) of E^*, which entails the claim.

From (2.2) and the coercivity $A \geq \lambda_A$ we recall that

$$\mathcal{E}(\mathbf{u}) \leq C_A \int_\Omega |\mathbf{f}|^2.$$

We can therefore apply Theorem 3 in the EDI (2.3) and conclude that there is $\gamma = \gamma_{\mathbf{U}} > 0$, depending only on $\mathcal{E}(\mathbf{u}^0)$ and the data, such that for a.e. $t \geq t_0 \geq 0$ there holds

$$\mathcal{E}(\mathbf{u}(t)) + \gamma \int_{t_0}^{t} \mathcal{E}(\mathbf{u}(s)) \mathrm{d}s$$

$$\leq \mathcal{E}(\mathbf{u}(t)) + \sum_i \int_{t_0}^{t} \int_\Omega (|\nabla f_i|^2 + |f_i|^2) u_i \leq \mathcal{E}(\mathbf{u}(t_0)).$$

Hence $t \mapsto \mathcal{E}(\mathbf{u}(t)) + \gamma \int_0^t \mathcal{E}(\mathbf{u}(s)) \mathrm{d}s$ is monotone non-increasing, and therefore

$$\frac{d\mathcal{E}}{dt} + \gamma \mathcal{E} \leq 0$$

in the sense of scalar distributions $\mathcal{D}'(0,\infty)$. This immediately implies the exponential decay (2.4) for a.e t by a standard Grönwall argument. Finally, since $\mathbf{u} \in \mathcal{C}_w([0,\infty); L^2(\Omega)^N)$, the function $t \mapsto \mathcal{E}(\mathbf{u}(t))$ is lower semicontinuous, and (2.4) extends to all $t \geq 0$. ■

Remark 5.3. *From the biological perspective, the distributions $\mathbf{u}^I$ ($I \neq \emptyset$) describe scenarios when some of the species $(u_i)_{i\in I}$ have died out, and the survivors $(u_j)_{j\notin J}$ compose a (lower-dimensional) ideal free distribution. It is important to point out that these partial ideal free distributions (which we need to avoid to secure the entropy-entropy production inequality) are unstable and repulsive: if a small L^∞ density of any of the extinct populations $(u_i)_{i\in I}$ is reintroduced, its fitness f_i will be positive and bounded away from zero (see* [19]*), therefore the environment is favorable to that species and it will unlikely go extinct again. That is why we conjecture that* (2.4) *holds for any $\mathbf{u}^0 \geq 0$ (unless some component of $\mathbf{u}^0$ is identically zero).*

Appendix

Assume that $X \subset Y$ are two Hilbert spaces, with continuous embedding operator $i : X \to Y$, and that $i(X)$ is dense in Y. The adjoint operator $i^* : Y^* \to X^*$ is continuous and, since $i(X)$ is dense in Y, one-to-one. Since i is one-to-one, $i^*(Y^*)$ is dense in X^*, and one may identify Y^* with a dense subspace of X^*. Due to the Riesz representation theorem, one may also identify Y with Y^*. We arrive at the chain of inclusions:

$$X \subset Y \equiv Y^* \subset X^* \tag{5.4}$$

with dense and continuous embeddings. Observe that in this situation, for $f \in Y, u \in X$, their scalar product in Y coincides with the $< X^*, X >$ duality

$$(f, u)_Y = \langle f, u \rangle_{X^*, X}. \tag{5.5}$$

Such triples (X, Y, X^*) are called *Hilbert triples* (sometimes also referred to as Gelfand or Lions triples), see, e.g., [31, 35] for more details.

Lemma 5.1. *Let*

$$X \subset Y \subset X^*$$

be a Hilbert triple. Let $\mathcal{A} : X \to X^*$ *be a linear continuous operator such that*

$$\langle \mathcal{A}u, u \rangle \geq \alpha \|u\|_X^2$$

for all $u \in X$ *and some common* $\alpha > 0$. *Let* V *be a Banach space such that*

$$X \subset V \subset Y$$

where the first embedding is compact and the second is continuous. Assume that both X *and* V *are separable. Let*

$$\mathcal{Q} : V \to X^*$$

be a continuous operator. Assume that

$$\|\mathcal{Q}(u)\|_{X^*} \leq C(1 + \|u\|_V) \tag{5.6}$$

for all $u \in X$. *Then the Cauchy problem*

$$u'(t) + \mathcal{A}u(t) = \mathcal{Q}(u(t)), \quad u|_{t=0} = u_0, \tag{5.7}$$

has a solution in the class

$$L^2(0, T; X) \cap H^1(0, T; X^*) \cap C([0, T]; Y) \tag{5.8}$$

for every $u_0 \in Y$.

We omit the proof since a more general statement is proven in [30], cf. also [35, Section 6.3], [34, Section 4].

References

[1] J. Alkhayal, S. Issa, M. Jazar, and R. Monneau. Existence result for degenerate cross-diffusion system with application to seawater intrusion. *ArXiv e-prints*, August 2014.

[2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient Flows: in Metric Spaces and in the Space of Probability Measures*. Basel: Birkhäuser Basel, 2008.

[3] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the monge-kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[4] Robert Stephen Cantrell, Chris Cosner, and Yuan Lou. Approximating the ideal free distribution via reaction-diffusion-advection equations. *J. Differential Equations*, 245(12):3687–3703, 2008.

[5] Robert Stephen Cantrell, Chris Cosner, Yuan Lou, and Chao Xie. Random dispersal versus fitness-dependent dispersal. *J. Differential Equations*, 254(7):2905–2941, 2013.

[6] Li Chen and Ansgar Jüngel. Analysis of a parabolic cross-diffusion population model without self-diffusion. *J. Differential Equations*, 224(1):39–59, 2006.

[7] Lenaic Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. Unbalanced optimal transport: geometry and Kantorovich formulation. *arXiv preprint arXiv:1508.05216*, 2015.

[8] Lenaic Chizat, Bernhard Schmitzer, Gabriel Peyré, and François-Xavier Vialard. An interpolating distance between optimal transport and Fischer-Rao. *arXiv preprint arXiv:1506.06430*, 2015.

[9] Y. S. Choi, Roger Lui, and Yoshio Yamada. Existence of global solutions for the Shigesada-Kawasaki-Teramoto model with strongly coupled cross-diffusion. *Discrete Contin. Dyn. Syst.*, 10(3):719–730, 2004.

[10] Chris Cosner. A dynamic model for the ideal-free distribution as a partial differential equation. *Theoretical Population Biology*, 67(2):101–108, 2005.

[11] Chris Cosner. Beyond diffusion: conditional dispersal in ecological models. In J. Mallet-Paret et al., editor, *Infinite Dimensional Dynamical Systems*, pages 305–317. Springer, 2013.

[12] Chris Cosner and Michael Winkler. Well-posedness and qualitative properties of a dynamical model for the ideal free distribution. *Journal of mathematical biology*, 69(6-7):1343–1382, 2014.

[13] Stephen D Fretwell. *Populations in a seasonal environment.* Princeton University Press, 1972.

[14] Stephen Dewitt Fretwell and Henry L Lucas. On territorial behavior and other factors influencing habitat distribution in birds I. Theoretical development. *Acta Biotheoretica*, 19(1):16–36, 1969.

[15] T. Galloüet and L. Monsaingeon. A JKO splitting scheme for Kantorovich-Fisher-Rao gradient flows. *ArXiv e-prints*, February 2016.

[16] A. Jüngel and N. Zamponi. Qualitative behavior of solutions to cross-diffusion systems from population dynamics. *ArXiv e-prints*, December 2015.

[17] Ansgar Jüngel. The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity*, 28(6):1963–2001, 2015.

[18] S. Kondratyev, L. Monsaingeon, and D. Vorotnikov. A new optimal transport distance on the space of finite Radon measures. *Adv. Differential Equations, to appear*, 2016.

[19] Stansislav Kondratyev, Léonard Monsaingeon, and Dmitry Vorotnikov. A new multicomponent Poincaré-Beckner inequality. *ArXiv e-prints, March 2016.*

[20] M. A. Krasnoselskii. *Topological methods in the theory of nonlinear integral equations.* Translated by A. H. Armstrong; translation edited by J. Burlak. A Pergamon Press Book. The Macmillan Co., New York, NY, USA, 1964.

[21] Pierre Legendre and Marie Josée Fortin. Spatial pattern and ecological analysis. *Vegetatio*, 80(2):107–138, 1989.

[22] Matthias Liero, Alexander Mielke, and Giuseppe Savaré. Optimal entropy-transport problems and a new Hellinger-Kantorovich distance between positive measures. *arXiv preprint arXiv:1508.07941*, 2015.

[23] Matthias Liero, Alexander Mielke, and Giuseppe Savaré. Optimal transport in competition with reaction: the Hellinger-Kantorovich distance and geodesic curves. *arXiv preprint arXiv:1509.00068*, 2015.

[24] Yuan Lou and Wei-Ming Ni. Diffusion, self-diffusion and cross-diffusion. *J. Differential Equations*, 131(1):79–131, 1996.

[25] Yuan Lou, Youshan Tao, and Michael Winkler. Approaching the ideal free distribution in two-species competition models with fitness-dependent dispersal. *SIAM J. Math. Anal.*, 46(2):1228–1262, 2014.

[26] Yuan Lou and Michael Winkler. Global existence and uniform boundedness of smooth solutions to a cross-diffusion system with equal diffusion rates. *Comm. Partial Differential Equations*, 40(10):1905–1941, 2015.

[27] Alec D MacCall. *Dynamic geography of marine fish populations*. Washington Sea Grant Program Seattle, 1990.

[28] Daniel Matthes, Robert J McCann, and Giuseppe Savaré. A family of nonlinear fourth order equations of gradient flow type. *Communications in Partial Differential Equations*, 34(11):1352–1397, 2009.

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

[30] V.B. Surya Prasath and Dmitry Vorotnikov. On time adaptive critical variable exponent vectorial diffusion flows and their applications in image processing. *ArXiv e-prints, March 2016.*

[31] R. Temam. *Navier-Stokes equations*, volume 2 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, Revised edition, 1979. Theory and numerical analysis, With an appendix by F. Thomasset.

[32] Cédric Villani. *Topics in optimal transportation*. American Mathematical Soc., 2003.

[33] Cédric Villani. *Optimal transport: old and new*. Springer Science & Business Media, 2008.

[34] Dmitry A. Vorotnikov. Weak solvability for equations of viscoelastic diffusion in polymers with variable coefficients. *Journal of Differential Equations*, 246(3):1038–1056, 2009.

[35] V. G. Zvyagin and D. A. Vorotnikov. *Topological approximation methods for evolutionary problems of nonlinear hydrodynamics*, volume 12 of *de Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter & Co., Berlin, Germany, 2008.

Stanislav Kondratyev
CMUC, Department of Mathematics, University of Coimbra, 3001-501 Coimbra, Portugal
E-mail address: kondratyev@mat.uc.pt

Léonard Monsaingeon
CAMGSD, Instituto Superior Técnico, Avenida Rovisco Pais, 1049-001 Lisbon, Portugal
E-mail address: leonard.monsaingeon@gmail.com

Dmitry Vorotnikov
CMUC, Department of Mathematics, University of Coimbra, 3001-501 Coimbra, Portugal
E-mail address: mitvorot@mat.uc.pt