SPLITTING SCHEMES AND SEGREGATION IN REACTION CROSS-DIFFUSION SYSTEMS[*]

J. A. CARRILLO[†], S. FAGIOLI[‡], F. SANTAMBROGIO[§], AND M. SCHMIDTCHEN[†]

Abstract. One of the most fascinating phenomenon observed in reaction diffusion systems is the emergence of segregated solutions, i.e., population densities with disjoint supports. We analyze such a reaction cross-diffusion system. In order to prove existence of weak solutions for a wide class of initial data without restriction of their supports or their positivity, we propose a variational splitting scheme combining ODEs with methods from optimal transport. In addition, this approach allows us to prove conservation of segregation for initially segregated data even in the presence of vacuum.

Key words. cross-diffusion, reaction diffusion, splitting schemes, variational schemes, segregation, pattern formation

AMS subject classifications. 35K57, 35A15, 47N60

DOI. 10.1137/17M1158379

1. Introduction. In this work we consider the following reaction diffusion system for the evolution on an interval Ω of two species with population densities $\rho, \eta \geq 0$:

$$
(1) \qquad \begin{cases} \partial_t \rho = \partial_x \left(\rho \partial_x \chi'(\rho + \eta) \right) + \rho F_1(\rho, \eta) + \eta G_1(\rho, \eta), \\ \partial_t \eta = \partial_x \left(\eta \partial_x \chi'(\rho + \eta) \right) + \eta F_2(\rho, \eta) + \rho G_2(\rho, \eta), \end{cases}
$$

where $\chi : [0, \infty) \to [0, \infty)$ is a C^1 superlinear function modeling nonlinear diffusion and F_i and G_i, $i = 1, 2$, model the reaction phenomena. Systems of this type appear naturally in mathematical biology. A fundamental biological phenomenon in interactions among different biological species is the *inhibition* or *activation* of growth whenever two populations occupy the same habitat. One species may promote or suppress the proliferation of the other species. In models involving cells or bacteria, the limited growth of different cell types can be attributed to volume or size constraints of the individual cells forming the different populations. The diffusive part in (1) was originally introduced in the seminal papers [30, 8, 7] and exhibits an intriguing phenomenon: segregated densities remain separated at all times.

In fact, nonlinear diffusions are natural ways to include volume filling effects in mathematical biology models; see [42, 20] for the case of the classical Keller–Segel system. Nonlinear diffusions help to avoid blow-up in these aggregation models in a biologically meaningful way and lead generically to asymptotic stabilization. In the

[†]Department of Mathematics, Imperial College London, SW7 2AZ London, UK (carrillo@imperial.ac.uk, m.schmidtchen15@imperial.ac.uk).

[‡]DISIM - Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila, 67100 L'Aquila (AQ), Italy (simone.fagioli@univaq.it).

[§]Laboratoire de Mathématiques d'Orsay, Université Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay Cedex, France (filippo.santambrogio@math.u-psud.fr).

absence of reactions, the system leads to the nonlinear diffusion equation

$$(2) \qquad \partial_t \sigma = \partial_x \left(\sigma \partial_x \chi'(\sigma) \right) = \partial_x^2 \beta(\sigma),$$

with $\sigma = \rho + \eta$, in which $\chi'(\sigma)$ models the resistance to compression of the whole group of individuals σ. The natural assumption in (2) in order for this to be a diffusion equation is that $\beta'(s) > 0$ for $s > 0$, possibly degenerating at $s = 0$, or equivalently, that $\chi''(s) > 0$ for $s > 0$. The particularly relevant case of $\chi(s) = s^2/2$ can be understood as the mean-field limit of interacting particles with very localized repulsion; see [40, 16, 12, 35].

   Reaction diffusion models similar to (1) appear in tissue growth models, where cell adhesion and volume effects are important factors determining cell sorting in heterogeneous cell populations (see [38, 23, 17, 22]) and zebrafish lateral line patterning [51]. They are also basic building blocks for a variety of cancer invasion models in the literature [26, 43, 47, 32, 6, 3, 4, 5, 29, 28, 45] in which the coupling with other biologically meaningful modeling factors, such as extracellular matrix, enzymatic activators, and other substances, is taken into account. These works usually involve drift terms due to long range attraction and/or repulsion between individuals, leading to related mathematical difficulties with respect to (1); see, for instance, [34].

   The nonlinear diffusion equation (2) is well studied (see [48]), and it can be understood as a gradient flow (see [41, 2, 49, 44]) of the energy functional

$$\mathcal{E}(\sigma) = \begin{cases} \displaystyle\int_\Omega \chi(\sigma)\,\mathrm{d}x, & \chi(\sigma) \in L^1(\Omega), \\ +\infty & \text{otherwise} \end{cases}$$

in the metric space of probability measures endowed with a suitable topology induced by the L^2 Monge–Kantorovich distance, denoted by d_2. The well-posedness of solutions to the nonlinear diffusion equation (2) was obtained in [41] by means of the so-called Jordan–Kinderlehrer–Otto (JKO) scheme (cf. [33]), which is a particular case of the minimizing movement scheme by De Giorgi; see [1] and the references therein. The idea of such a scheme is to recursively construct a sequence by solving a minimization problem in a certain metric space (X, d), corresponding to the space of probability measures endowed with d_2 in our case. Given some initial condition $\bar\sigma$ for (2) and a fixed time step $0 < \tau < 1$, we set $\sigma_\tau^0 = \bar\sigma$ and then recursively define

$$(3) \qquad \sigma_\tau^{n+1} \in \operatorname*{argmin}_{\sigma \in X} \left\{ \frac{1}{2\tau} d^2(\sigma, \sigma_\tau^n) + \mathcal{E}(\sigma) \right\}$$

for $n \in \mathbb{N}$. The seminal work of McCann [37] shows that $\mathcal{E}(\sigma)$ is displacement convex, or geodesically convex, on the metric space of probability measures on the line endowed with d_2 as soon as β is nondecreasing on $(0, +\infty)$, or equivalently χ convex, which is called the McCann's condition for short. We also refer the reader to [2, Chap. 9], [24, p. 26], and [50, Chap. 17] for this classical notion, and to [13] for related issues. Upon choosing a proper time interpolation σ_τ, it can be proven that the sequence $\{\sigma_\tau\}_\tau$ converges to a weak solution of (2); see [49, 2, 50, 44] and the references therein.

   In this work, we propose a variational splitting scheme in order to construct weak solutions to the reaction diffusion system (1). More precisely, we solve in an inner time step the diffusive part of the system by the JKO scheme related to the nonlinear diffusion equation (2), and then we transport both densities ρ, η through the flow generated by the equation for the total population σ. Note that in this step the total

and individual masses of the populations are unchanged in time. In the second inner time step of the splitting scheme, we solve the system of ODEs parameterized by the spatial variable $x \in \Omega$, leading to the final approximation of our new population densities after a time step. This variational splitting scheme will be written in detail in section 2. The splitting between reaction and diffusion steps follows naturally from the numerical analysis viewpoint, as it has already been used for variations of Keller–Segel models, where the diffusion step is solved by the JKO scheme [11, 25] in the case of a single population density coupled with a system of reaction diffusion equations.

Our main result shows the convergence of the splitting variational scheme towards weak solutions of the system (1). The main mathematical difficulty here arises from the cross-diffusion term allowing for segregation fronts to form in the solutions. This phenomenon was proven in [8] in the case of initial data with separated supports for the populations. More precisely, while [30] constructs a source solution of the system without reactions, similar to the well-known Barenblatt–Pattle profiles [48] for nonlinear diffusions, [8] constructs a solution to the system without reactions by formulating it as a free boundary problem for a single effective equation and by characterizing the segregation front through this free boundary. This approach can only work in the case when the supports of both populations are at a positive distance from each other initially. Later [6, 9] combined both the nonlinear diffusion and the reaction to obtain a system related to (1), showing a similar segregation phenomenon by regularization techniques. However, their approach heavily relies on the absence of vacuum, as they assume that σ_0 is bounded below by a positive constant.

These remarkable results have severe consequences—initially smooth solutions lose their regularity when both densities meet each other. In fact, they immediately become discontinuous at the contact interface. This phenomenon legitimizes our functional space choice of bounded functions of bounded variation; see the next section for a precise notion of weak solution and assumptions on the initial data.

In contrast to [8, 6, 9], we show the convergence of our variational splitting scheme for general initial data even in the presence of vacuum and for general nonlinearities. Moreover, we recover the results in [8, 6] about segregation fronts, even in the case of initial vacuum, by showing that initial data which are initially segregated remain segregated at all times. An important technical point in our proof relies on displacement convexity of an auxiliary functional that allows us to obtain further regularity on the approximate solutions in order to pass to the limit in the nonlinear diffusion terms. This auxiliary functional imposes a slightly more restricted set of nonlinear diffusions, satisfying some integrability condition at the origin; see the precise conditions in the next section.

The rest of the paper is organized as follows. In section 2 we introduce the variational splitting scheme, present the main result, and explain the strategy of the proof. Section 3 is dedicated to deriving all estimates necessary for proving not only the existence theorem but also the segregation theorem.

2. Preliminaries and main result. As already mentioned, our main aim is to study the existence of weak solutions for the following one-dimensional two species cross-diffusion and reaction system:

$$(4) \quad \begin{cases} \partial_t \rho = \partial_x \left(\rho \partial_x \chi'(\rho + \eta) \right) + \rho F_1(\rho, \eta) + \eta G_1(\rho, \eta) & \text{in } [0, T] \times \Omega, \\ \partial_t \eta = \partial_x \left(\eta \partial_x \chi'(\rho + \eta) \right) + \eta F_2(\rho, \eta) + \rho G_2(\rho, \eta) & \text{in } [0, T] \times \Omega, \\ \rho(t, x) \partial_x \chi'(\rho + \eta)(t, x) = 0 & \text{on } [0, T] \times \partial\Omega, \\ \eta(t, x) \partial_x \chi'(\rho + \eta)(t, x) = 0 & \text{on } [0, T] \times \partial\Omega, \\ \rho(\cdot, 0) = \rho_0, \quad \eta(\cdot, 0) = \eta_0 & \text{in } \Omega, \end{cases}$$

where $\Omega \subset \mathbb{R}$ is an open bounded interval, $T > 0$. Moreover, χ denotes an internal energy density, and F_i and G_i, $i = 1, 2$, model the reaction phenomena. Throughout, we use the notation $L_+^1(\Omega)$ to denote the set of nonnegative Lebesgue integrable functions. As mentioned before, the space of bounded functions with bounded variation is a natural functional setting.

DEFINITION 2.1 (space of bounded functions with bounded variation). *Let* $f : \bar{\Omega} \to \mathbb{R}$. *We define its variation with respect to a partition* $P := \{x_1 < x_2 < \cdots < x_{|P|}\} \subset \Omega$ *by*

$$V_P(f) := \sum_{i=1}^{|P|-1} |f(x_{i+1}) - f(x_i)|.$$

We call f *a function of bounded variation if its total variation* $\sup_P V_P(f) < \infty$ *is finite. Here the supremum is taken over all partitions of* Ω. *We denote by* $BV(\Omega)$ *the set of functions whose variation is bounded. Equipped with the norm*

$$\|f\|_{BV} := \sup_P V_P(f),$$

the set $BV(\Omega)$ *is a vector space.*

We shall see in the remainder of this section that the vector space of bounded function with bounded variation is a good choice to construct solutions. In our analysis we will exploit the following property.

LEMMA 2.2 ($BV(\Omega) \cap L^\infty(\Omega)$ is an $\mathbb{R}$-algebra). *The vector space* $BV(\Omega) \cap L^\infty(\Omega)$ *equipped with the pointwise multiplication*

$$\left(BV(\Omega) \cap L^\infty(\Omega)\right)^2 \ni (f, g) \mapsto fg \in BV(\Omega) \cap L^\infty(\Omega)$$

is a real algebra.

The proof of the previous result is standard. Notice that in one dimension, BV-regularity implies boundedness. We prefer to write $BV(\Omega) \cap L^\infty(\Omega)$ for the sake of clarity and possible future generalizations.

2.1. Metric structure. Consider an open, bounded interval $\Omega \subseteq \mathbb{R}$, and denote by $\mathcal{M}_+(\Omega)$ the set of positive and finite measures. Throughout this paper we make use of the notation

$$\mathcal{P}^m(\Omega) := \left\{\mu \in \mathcal{M}_+(\Omega) \,\middle|\, \mu(\Omega) = m\right\},$$

that is, the set of positive measures with mass $m > 0$. Consider a measure $\mu \in \mathcal{P}^m(\Omega)$ and a Borel map $\mathcal{T} : \mathbb{R} \to \mathbb{R}$. We denote by $\nu = \mathcal{T}_\#\mu \in \mathcal{P}^m(\Omega)$ the push-forward measure of μ through $\mathcal{T}$, defined by

$$\int_\Omega f(y) \, d\mathcal{T}_\#\mu(y) = \int_\Omega f(\mathcal{T}(x)) \, d\mu(x),$$

for all Borel functions f on Ω. We call $\mathcal{T}$ a *transport map* pushing μ to ν. We endow the space $\mathcal{P}^m(\Omega)$ with the p-Wasserstein distance, $p \geq 1$,

$$d_p^p(\mu_1, \mu_2) = \inf_{\gamma \in \Pi^m(\mu_1, \mu_2)} \left\{ \int_{\Omega \times \Omega} |x - y|^p \, d\gamma(x, y) \right\}.$$

Here, $\Pi^m(\mu_1, \mu_2)$ is the set of all transport plans γ between μ_1 and μ_2, that is, the set of positive measures of fixed mass $\gamma(\Omega \times \Omega) = m^2$, defined on the product space such that $\pi_\#^i \gamma = \mu_i$ for $i = 1, 2$, where π^i denotes the projection operator on the ith component of the product space. If μ_1 is absolutely continuous with respect to the Lebesgue measure, the optimal transport plan γ is unique and can be written as $\gamma = (\mathrm{id}, \mathcal{T})_\# \mu$. In addition there exists a *Kantorovich potential* φ that is linked to the transport map $\mathcal{T}$ in the following way:

$$(5) \qquad \mathcal{T}(x) = (\mathrm{id} - \partial_x \varphi)(x).$$

We refer the reader to [49, 2, 50, 44] and the references therein for a good account of the properties of transport distances and the state of the art in gradient flows/steepest descents of functionals in metric spaces of probability measures. While transport distances are an incredibly powerful tool for dealing with transport PDEs exhibiting a gradient flow structure, it is not applicable in the presence of source terms. This is because it is defined only for two measures of the same mass.

To resolve this shortcoming we will make use of the bounded-Lipschitz distance d_{BL}, classically used for the derivation of the Vlasov equation; see [39, 46, 19, 18, 31] and the references therein. The bounded-Lipschitz distance d_{BL}, also frequently called flat metric, is defined as follows:

$$(6)$$
$$d_{\mathrm{BL}}(\mu, \nu) := \sup \left\{ \int_\Omega f \, \mathrm{d}\,(\mu - \nu) \mid \|f\|_{L^\infty(\Omega)}, \|f'\|_{L^\infty(\Omega)} \le 1 \right\} = \|\mu - \nu\|_{(W^{1,\infty}(\Omega))^*}.$$

Note that the bounded-Lipschitz distance may be defined in any space dimension by the properties below. Thus, we state all required properties in our setting. Since our problem is posed on the product space, we extend in the canonical way the metric setting to the space $\mathcal{M}_+ \times \mathcal{M}_+$, which is the product space of nonnegative measures. For $d \in \{d_{\mathrm{BL}}, d_p\}$, $1 \le p < \infty$, we define the product metric (still denoted d) as

$$(7) \qquad d(U, \tilde{U}) := d(\rho, \tilde{\rho}) + d(\eta, \tilde{\eta}),$$

where $U = (\rho, \eta), \tilde{U} = (\tilde{\rho}, \tilde{\eta}) \in \mathcal{M}_+ \times \mathcal{M}_+$.

PROPOSITION 2.3 (properties of d_{BL}). *Let $\mu, \nu \in L_+^1(\Omega)$ be two densities. Then the following properties hold true:*
  (i) $d_{\mathrm{BL}}(\mu, \nu) \le \|\mu - \nu\|_{L^1(\Omega)}$,
  (ii) $d_{\mathrm{BL}}(\mu, \nu) \le d_1(\mu, \nu)$ *whenever* $\mu(\Omega) = \nu(\Omega)$.

Proof. Let $\mu, \nu \in L_+^1(\Omega)$ be arbitrary, and let $f \in W^{1,\infty}(\Omega)$ with $\|f\|_{W^{1,\infty}(\Omega)} \le 1$. Hence, we may write

$$\left| \int_\Omega f \, \mathrm{d}(\mu - \nu) \right| \le \int_\Omega |f| \, \mathrm{d}|\mu - \nu| = \|\mu - \nu\|_{L^1(\Omega)}.$$

Taking the supremum over all such functions f, we get the first statement by using definition (6). For the second statement, we additionally assume that $\mu(\Omega) = \nu(\Omega)$. We recall the dual definition of d_1,

$$d_1(\mu, \nu) = \sup \left\{ \int_\Omega f \, \mathrm{d}(\mu - \nu) \mid f \in \mathrm{Lip}(\Omega), \text{ s.t. } \|f\|_{\mathrm{Lip}(\Omega)} \le 1 \right\},$$

where $\mathrm{Lip}(\Omega)$ is the set of Lipschitz-continuous functions; see [49, 44]. Then property (ii) is a consequence of this formulation, and this concludes the proof. $\square$

COROLLARY 2.4. *Let $U_i = (\mu_i, \nu_i) \in \mathcal{M}_+(\Omega)^2$ for $i = 1, 2, 3$. Furthermore, assume $\mu_i, \nu_i \in L^1(\Omega)$ for $i = 1, 2$, and assume as well $\mu_2(\Omega) = \mu_3(\Omega)$ and $\nu_2(\Omega) = \nu_3(\Omega)$. Then,*

$$d_{\mathrm{BL}}(U_1, U_3) \leq \|U_1 - U_2\|_{L^1(\Omega)} + d_1(U_2, U_3).$$

Throughout, two quantities are crucial for our analysis: the sum of the two densities, σ, and the ratio, r, between one density, say ρ, and the sum

$$\sigma = \rho + \eta \quad \text{and} \quad r = \frac{\rho}{\sigma}.$$

Note that, at this stage, the quotient can only be written formally. A straightforward computation shows that these functions formally satisfy the system of PDEs

(8)
$$\begin{cases} \partial_t \sigma = \partial_x \left(\sigma \partial_x \chi'(\sigma)\right) + \sigma \left(r(\tilde{F}_1 + \tilde{G}_2) + (1 - r)\left(\tilde{G}_1 + \tilde{F}_2\right)\right), \\ \partial_t r = \partial_x r \partial_x \chi'(\sigma) + r(1 - r)\left(\tilde{F}_1 - \tilde{F}_2\right) + (1 - r)^2 \tilde{G}_1 - r^2 \tilde{G}_2, \end{cases}$$

where we used $\tilde{F}_i(\sigma, r) = F_i(r\sigma, (1 - r)\sigma)$ and $\tilde{G}_i(\sigma, r) = G_i(r\sigma, (1 - r)\sigma)$, for $i = 1, 2$, to denote the reaction terms in the transformed variables. In order to simplify the analysis in section 3, let us introduce the more concise notation,

$$A_1(r, \sigma) := \tilde{F}_1 + \tilde{G}_2, \quad A_2(r, \sigma) := \tilde{G}_1 + \tilde{F}_2, \quad \text{and} \quad A_3(r, \sigma) := \tilde{F}_1 - \tilde{F}_2.$$

Note that these functions are Lipschitz and bounded, as they are linear combinations of $BV \cap L^\infty$ functions. Thus, the transformed system (8) can be rewritten in the more compact form

(9)
$$\begin{cases} \partial_t \sigma = \partial_x \left(\sigma \partial_x \chi'(\sigma)\right) + \sigma \left(rA_1 + (1 - r)A_2\right), \\ \partial_t r = \partial_x r \partial_x \chi'(\sigma) + r(1 - r)A_3 + (1 - r)^2 \tilde{G}_1 - r^2 \tilde{G}_2. \end{cases}$$

Let us note here that taking $F_i = G_i = 0$ in (1) yields the nonlinear cross-diffusion system

(10)
$$\begin{cases} \partial_t \rho = \partial_x \left(\rho \partial_x \chi'(\sigma)\right), \\ \partial_t \eta = \partial_x \left(\eta \partial_x \chi'(\sigma)\right), \end{cases}$$

where the sum σ satisfies the nonlinear diffusion equation (2).

For our purposes, we need some properties on the internal energy density. The role of these properties will become clearer after the statement of our main result.

DEFINITION 2.5 (internal energy density). *A function $\chi : [0, \infty] \to \mathbb{R}$ is an internal energy density if*
   (NL-i) $\chi \in C^0([0, \infty], \mathbb{R}) \cap C^2((0, \infty), \mathbb{R})$ *with* $\chi'' > 0$;
   (NL-ii) $\lim_{h \downarrow 0} \chi'(h) = 0$; *and*
   (NL-iii) *the integrals* $\kappa(x) := \int_1^x \frac{\chi''(s)}{s} \, ds$ *and* $K(\sigma) := \int_0^\sigma \kappa(x) \, dx$ *exist.*

As mentioned in the introduction, the space of probability measures endowed with d_2 has proven to be an exceptional choice of metric space. In this case the minimizer in (3) satisfies the optimality condition

(11)
$$\frac{\varphi}{\tau} + \chi'(\sigma^{n+1}) = const.;$$

cf., for instance, [44, Proposition 7.20]. Here φ denotes the associated Kantorovich potential transporting σ^{n+1} to σ^n in $\mathcal{P}^{m_1+m_2}$, with m_1 and m_2 the masses of ρ and η [44, Theorem 1.17]. Notice that this optimality is only obtained in those references for nondegenerate problems, where $\beta'(0^+) > 0$. However, similar approximation techniques, such as those developed in [41, p. 156] and [24, p. 27], allow one to overcome the difficulties associated with degenerate diffusions, where $\beta'(0^+) = 0$. In the rest of this paper, we will proceed as if we were dealing with nonlinearities leading to nondegenerate diffusions, since by this standard approximation procedure, the same result can be obtained for the degenerate ones. We are now ready to introduce our notion of weak solutions.

DEFINITION 2.6 (notion of weak solutions).   *A couple* $\rho, \eta \in C\big(0, T; BV(\Omega) \cap L^\infty(\Omega)\big)$ *is a weak solution to system* (4) *if* $\sigma = \rho + \eta \in L^2(0, T; H^1(\Omega))$ *and there holds*

$$\int_\Omega (\rho(t) - \rho(s)) \, \zeta \, \mathrm{d}x = \int_s^t \int_\Omega -\rho \partial_x \chi'(\rho + \eta) \partial_x \zeta + (\rho F_1(\rho, \eta) + \eta G_1(\rho, \eta)) \, \zeta \, \mathrm{d}x \, \mathrm{d}\bar\tau,$$

$$\int_\Omega (\eta(t) - \eta(s)) \, \xi \, \mathrm{d}x = \int_s^t \int_\Omega -\eta \partial_x \chi'(\rho + \eta) \partial_x \xi + (\eta F_2(\rho, \eta) + \rho G_2(\rho, \eta)) \, \xi \, \mathrm{d}x \, \mathrm{d}\bar\tau$$

for $0 \le s < t \le T$ *and any two test functions* $\zeta, \xi \in C_c^\infty(\Omega)$.

Remark 2.7.

1. Notice that the functional

$$\mathcal{K}(\sigma) := \int_\Omega K(\sigma) \, \mathrm{d}x,$$

associated to χ, as in Definition 2.5, satisfies the McCann condition since $K''(s) = \frac{\chi''(s)}{s} > 0$ and therefore is displacement convex. These facts will be crucial in section 3.3 in order to prove Lemma 3.10. We state here that it is necessary to obtain additional regularity from the dissipation of this functional on the nonlinear diffusion term, thus allowing us to pass to the limit in the approximating sequence.

2. Let us note that we can also allow for nonlocal reaction terms, i.e.,

$$F_i = W_{1,i} \star \rho + W_{2,i} \star \eta,$$

and similarly for G_i, $i = 1, 2$. The only assumption we need to impose on the kernels is that they are smooth and integrable. These models are found in modeling pattern formation, for instance, the kernel-based Turing pattern system [36] or the nonlinear aggregation-diffusion system [51].

3. Similarly, we can—at least formally—interpret the system (10) as a gradient flow of the functional

$$\mathcal{E}(\rho, \eta) = \int_\Omega \chi(\rho + \eta) \, \mathrm{d}x$$

in the product Wasserstein space.

2.2. Splitting scheme. We are now ready to introduce our splitting scheme for (4). Let some initial data $\rho_0, \eta_0 \in BV(\Omega) \cap L^\infty(\Omega)$ be given such that there exists a function $r_0 \in BV(\Omega)$ such that

$$\sigma_0 := \rho_0 + \eta_0 \quad \text{and} \quad \frac{\rho_0}{\sigma_0} = r_0\big|_{\{\sigma_0 > 0\}},$$

and $0 \leq r_0 \leq 1$. Furthermore we assume F_i and G_i, $i = 1, 2$, are bounded and Lipschitz with respect to ρ and η, and we impose $G_1(0, \cdot) \geq 0$ and $G_2(\cdot, 0) \geq 0$ to ensure positivity of solutions.

We fix $0 < \tau < 1$ and $n \in \{0, \dots, N\}$, with $N \in \mathbb{N}$ such that $N\tau = T$. We then recursively construct the piecewise constant approximation to the system as follows. We impose

$$(\rho^0, \eta^0) = (\rho_0, \eta_0),$$

and then construct $U^{n+1} = (\rho^{n+1}, \eta^{n+1})$ by the following scheme. We split the equation into a reaction step and a diffusion step on the time interval $[t^n, t^{n+1})$, with $t^n = n\tau$, for all $0 \leq n \leq N$.

2.2.1. Reaction step. The reaction phase consists of solving the system of ODEs

$$\begin{cases} \partial_t \sigma = \Sigma(\sigma, r) := \sigma\big(rA_1 + (1 - r)A_2\big), \\ \partial_t r = R(\sigma, r) := r(1 - r)A_3 + (1 - r)^2 \tilde{G}_1 - r^2 \tilde{G}_2, \\ \sigma(t^n) = \sigma^n \quad \text{and} \quad r(t^n) = r^n \end{cases}$$

in the time interval $[t^n, t^{n+1})$. We then set

(12a) $\qquad \rho^{n+1/2} := r\sigma\big|_{t=(n+1)\tau} \quad \text{and} \quad \eta^{n+1/2} := (1 - r)\sigma\big|_{t=(n+1)\tau}.$

A straightforward computation reveals

(12b) $$\begin{aligned} \partial_t(r\sigma) &= r\sigma \tilde{F}_1 + (1 - r)\sigma \tilde{G}_1, \\ \partial_t((1 - r)\sigma) &= (1 - r)\sigma \tilde{F}_2 + r\sigma \tilde{G}_2, \end{aligned}$$

formally verifying that $r\sigma$ and $(1 - r)\sigma$ solve the reaction part of (4). Note that we solve the system rigorously in the subsequent section.

2.2.2. Diffusion step. After the reaction phase, we solve

(12c) $\qquad U^{n+1} \in \displaystyle\operatorname*{argmin}_{U \in \mathcal{P}^{m_1}(\Omega) \times \mathcal{P}^{m_2}(\Omega)} \left\{ \frac{1}{2\tau} d_2^2 \left(U, U^{n+1/2} \right) + \mathcal{E}(U) \right\},$

where $m_1 = \rho^{n+1/2}(\Omega)$ and $m_2 = \eta^{n+1/2}(\Omega)$. Let $\mathcal{T}$ denote the associated optimal transport map, i.e., $\mathcal{T}_\# U^{n+1} = U^{n+1/2}$. We define

(12d) $\qquad \sigma^{n+1} := \rho^{n+1} + \eta^{n+1} \quad \text{and} \quad r^{n+1} := r^{n+1/2} \circ \mathcal{T}.$

While the definition of σ^{n+1} is somewhat natural, the definition of r^{n+1} seems a bit surprising. Let us note here that, indeed, $r^{n+1} = \rho^{n+1}/\sigma^{n+1}$ where $\sigma^{n+1} > 0$. For the precise argument we refer the reader to (15) in section 3.2.

2.2.3. Combination of both steps—construction of a solution. Throughout this paper we refer to (12a) as the *reaction step* and to (12c) as the *diffusion step*, respectively.

DEFINITION 2.8 (piecewise constant interpolation). *Let $(r^n, \sigma^n)_{n \in \mathbb{N}}$ be the sequence obtained from the splitting scheme. Then we define the piecewise constant interpolations by*

$$r_\tau(t, x) = r^n(x) \quad and \quad \sigma_\tau(t, x) = \sigma^n(x),$$

as well as by

$$\rho_\tau(t, x) = r^n(x)\sigma^n(x) \quad and \quad \eta_\tau(t, x) = (1 - r^n(x))\sigma^n(x),$$

for all $(x, t) \in \Omega \times [t^n, t^{n+1})$. Furthermore, we write $U_\tau := (\rho_\tau, \eta_\tau)$.

We will say that two densities $\rho, \eta \in BV(\Omega) \cap L^\infty(\Omega)$ are segregated if the intersection of the interior of their supports is empty. We are now ready to state our main result.

THEOREM 2.9 (convergence to weak solutions). *Let $\rho_0, \eta_0 \in BV(\Omega) \cap L^\infty(\Omega)$, and assume there exists a function $r_0 \in BV(\Omega)$ such that $r_0 = \rho_0/(\rho_0 + \eta_0)$ on $\{\rho_0 + \eta_0 > 0\}$ and $0 \le r_0 \le 1$. Furthermore, we assume that the nonlinearity χ satisfies the assumptions in Definition 2.5, and that the reaction terms F_i and G_i, $i = 1, 2$, are bounded and Lipschitz with respect to ρ and η with $G_1(0, \cdot) \ge 0$ and $G_2(\cdot, 0) \ge 0$. Then, upon the extraction of a subsequence, the piecewise constant interpolations $(\rho_\tau)_{\tau>0}$ and $(\eta_\tau)_{\tau>0}$ converge to a weak solution of system (4) in the sense of Definition 2.6. Moreover, if initially the two densities ρ_0, η_0 are segregated, then, in the absence of cross-reactions, the limit densities $\rho(t, \cdot), \eta(t, \cdot)$ remain segregated at all times.*

3. Proof of the main result. This section is dedicated to proving the main result of the paper: the convergence of the approximation obtained by the splitting scheme to a solution of the system. It is organized as follows. In section 3.1 and 3.2 we establish the crucial BV-estimates and L^∞-bounds. In section 3.3 we combine the estimates from the previous sections in order to get uniform estimates for a whole iteration. Finally, in section 3.4 we show how to extract a convergent subsequence and identify its limit as a weak solution to system (4).

3.1. Estimates for the reaction step. Since the right-hand sides $\Sigma(\sigma, r), R(\sigma, r)$ are Lipschitz continuous in both components, we note that the solution of the reaction system is unique.

PROPOSITION 3.1 (L^∞-estimates of the reaction step). *Let (r^n, σ^n) be given by our splitting scheme. Then there holds*

$$0 \le \sigma^{n+1/2} \le \|\sigma^n\|_{L^\infty} \exp(c\tau) \quad and \quad 0 \le r^{n+1/2} \le 1$$

for some constant $c > 0$ independent of n and any $x \in \Omega$.

Proof. First, we show that $r^{n+1/2} \in [0, 1]$ holds. Assume the contrary; i.e., there exists an $x \in \Omega$ such that $r^{n+1/2}(x) < 0$ or $r^{n+1/2}(x) > 1$. If $r^{n+1/2}(x) < 0$, then, by continuity, there exists a time $t^\star(x) \in (t^n, t^{n+1})$ such that $r(t^\star, x) = 0$ and $\partial_t r(t^\star, x) < 0$. However, this is absurd, as

$$0 > \partial_t r(t^\star, x) = R\big(\sigma(t^\star, x), r(t^\star, x)\big) = \tilde{G}_1\big(\sigma(t^\star, x), 0\big) \ge 0.$$

Analogously, it can be shown that $r^{n+1}(x) \le 1$. Finally, for the positivity of σ we can use a similar argument. Let us assume $\sigma^n(x) \ge 0$ and $\sigma^{n+1/2}(x) < 0$ for some $x \in \Omega$. Then, there exists another $t^\star$ such that

$$0 > \partial_t \sigma(t^\star, x) = \Sigma\big(\sigma(t^\star, x), r(t^\star, x)\big) = 0,$$

which is clearly a contradiction. For the L^∞-bound we simply apply Gronwall's lemma and the fact that $r \in [0, 1]$. $\square$

Since we control the L^∞-norm of both r and σ, the local solution provided by the Cauchy–Lipschitz theorem is indeed global. Next we address the BV-estimates during the reaction step.

PROPOSITION 3.2 (bounded variation of $r^{n+1/2}$ and $\sigma^{n+1/2}$). *Let us consider* (r^n, σ^n) *as initial data for our splitting scheme. Then, the reaction step is BV-stable in the following sense:*

$$\|r^{n+1/2}\|_{BV} + \|\sigma^{n+1/2}\|_{BV} \leq (\|r^n\|_{BV} + \|\sigma^n\|_{BV}) \exp(c\tau)$$

for some positive c, depending only on the Lipschitz constants of F_i, G_i and the L^∞-bounds on F_i, G_i, $i = 1, 2$.

Proof. Using the transformed system (9), r and σ satisfy the following equations in the reaction step:

$$\partial_t \sigma = \Sigma(\sigma, r) \quad \text{and} \quad \partial_t r = R(\sigma, r).$$

Upon integrating in time we get

$$\sigma(t) = \sigma(s) + \int_s^t \Sigma\big(\sigma(\bar\tau), r(\bar\tau)\big) \, \mathrm{d}\bar\tau \quad \text{and} \quad r(t) = r(s) + \int_s^t R\big(\sigma(\bar\tau), r(\bar\tau)\big) \, \mathrm{d}\bar\tau.$$

Now, let $P \subset \Omega$ be an arbitrary partition. We compute the variation of σ and r with respect to P and obtain

$$Q(t) := V_P(\sigma(t)) + V_P(r(t)) \leq V_P(\sigma(s)) + V_P(r(s)) + \int_s^t V_P(\Sigma(\bar\tau)) + V_P(R(\bar\tau)) \, \mathrm{d}\bar\tau,$$

whence

$$Q(t) \leq Q(s) + c \int_s^t Q(\bar\tau) \, \mathrm{d}\bar\tau,$$

where c depends only on the L^∞-bounds and the Lipschitz-continuity of A_i for $\in \{1, 2, 3\}$ and the L^∞-bounds of σ and r. Applying Gronwall's lemma, we finally obtain

$$Q(t) \leq Q(s) \exp\big(c(t - s)\big).$$

Passing to the supremum first on the right-hand side and then on the left-hand side yields the result. $\qquad\square$

3.2. Estimates for the diffusion step. This section is devoted to establishing BV-estimates and L^∞-bounds in the diffusion step. To this end, we will make use of the following lemma.

LEMMA 3.3 (same optimality conditions). *Let (ρ^{n+1}, η^{n+1}) be given by the JKO step for $\mathcal{E}$; cf. (12c). Moreover, let*

$$\sigma^* \in \operatorname*{argmin}_{\sigma \in \mathcal{P}^{m_1 + m_2}(\Omega)} \left\{ \frac{1}{2\tau} d_2^2(\sigma, \sigma^{n+1/2}) + \mathcal{E}(\sigma) \right\},$$

and let $\mathcal{T}$ be the associated optimal transport map, $\mathcal{T}_\# \sigma^* = \sigma^{n+1/2}$.

Then there holds $\sigma^* = \rho^{n+1} + \eta^{n+1}$ and $\mathcal{T}_\rho = \mathcal{T}_\eta = \mathcal{T}_\sigma$ on $\mathrm{supp}(\rho^{n+1}) \cup \mathrm{supp}(\eta^{n+1})$ for the associated optimal transport maps. Moreover, the optimality conditions read

$$\frac{\delta\mathcal{E}}{\delta\rho} + \frac{\varphi_\rho}{\tau} = c_1, \quad \frac{\delta\mathcal{E}}{\delta\eta} + \frac{\varphi_\eta}{\tau} = c_2, \quad with \ \frac{\delta\mathcal{E}}{\delta\rho} = \frac{\delta\mathcal{E}}{\delta\eta} = \chi'(\sigma^{n+1}),$$

on $\mathrm{supp}(\rho^{n+1}) \cup \mathrm{supp}(\eta^{n+1})$ as well as $\varphi_\rho = \varphi_\eta = \varphi$ up to an additive constant.

Proof. Let

$$(\rho^{n+1}, \eta^{n+1}) \in \mathrm{argmin}\left\{\frac{1}{2\tau}d_2^2(U, U^{n+1/2}) + \mathcal{E}(U)\right\},$$

and let $\mathcal{T}_\rho, \mathcal{T}_\eta$ be the associated transport maps, i.e., $\mathcal{T}_{\rho\#}\rho^{n+1} = \rho^{n+1/2}$ and $\mathcal{T}_{\eta\#}\eta^{n+1} = \eta^{n+1/2}$. Note that the optimality conditions take the form

$$\chi'(\rho^{n+1} + \eta^{n+1}) + \frac{\varphi_\rho}{\tau} = c_\rho \quad \text{on } \mathrm{supp}(\rho^{n+1}),$$
$$\geq c_\rho \quad \text{elsewhere}$$

for ρ, and

$$\chi'(\rho^{n+1} + \eta^{n+1}) + \frac{\varphi_\eta}{\tau} = c_\eta \quad \text{on } \mathrm{supp}(\eta^{n+1}),$$
$$\geq c_\eta \quad \text{elsewhere}$$

for η, respectively. Thus, the optimal maps are given by

$$\mathcal{T}_\rho = \mathrm{id} + \tau\partial_x\chi'(\rho^{n+1} + \eta^{n+1}) \quad \text{on } \mathrm{supp}(\rho^{n+1}),$$
$$\mathcal{T}_\eta = \mathrm{id} + \tau\partial_x\chi'(\rho^{n+1} + \eta^{n+1}) \quad \text{on } \mathrm{supp}(\eta^{n+1}).$$

In particular, note that $\mathcal{T}_\rho \equiv \mathcal{T}_\eta =: \mathcal{T}$ on $\mathrm{supp}(\rho^{n+1}) \cap \mathrm{supp}(\eta^{n+1})$. We claim that $\sigma^{n+1} := \rho^{n+1} + \eta^{n+1}$ is a minimizer of the problem

$$(13) \qquad\qquad \min\left\{\frac{1}{2\tau}d_2^2(\sigma, \sigma^{n+1/2}) + \mathcal{E}(\sigma)\right\}.$$

To this end, suppose there exists $\sigma^* \in \mathcal{P}^{m_1+m_2}(\Omega)$ such that

$$\frac{1}{2\tau}d_2^2(\sigma^*, \sigma^{n+1/2}) + \mathcal{E}(\sigma^*) < \frac{1}{2\tau}d_2^2(\sigma^{n+1}, \sigma^{n+1/2}) + \mathcal{E}(\sigma^{n+1}).$$

Then there holds

$$
\frac{1}{2\tau}d_2^2(U^{n+1}, U^{n+1/2}) + \mathcal{E}(U^{n+1})
$$

$$
= \frac{1}{2\tau}\left(\int_\Omega |x - \mathcal{T}_\rho(x)|^2 \rho^{n+1}(x)\,\mathrm{d}x + \int_\Omega |x - \mathcal{T}_\eta(x)|^2 \eta^{n+1}(x)\,\mathrm{d}x\right) + \mathcal{E}(U^{n+1})
$$

$$
= \frac{1}{2\tau}\left(\int_{\mathrm{supp}(\rho^{n+1})} |x - \mathcal{T}_\rho(x)|^2 \rho^{n+1}(x)\,\mathrm{d}x\right.
$$

$$
\left. + \int_{\mathrm{supp}(\eta^{n+1})} |x - \mathcal{T}_\eta(x)|^2 \eta^{n+1}(x)\,\mathrm{d}x\right) + \mathcal{E}(U^{n+1})
$$

$$
= \frac{1}{2\tau}\left(\int_{\mathrm{supp}(\rho^{n+1})\backslash(\mathrm{supp}(\rho^{n+1})\cap\mathrm{supp}(\eta^{n+1}))} |x - \mathcal{T}_\rho(x)|^2 \rho^{n+1}(x)\,\mathrm{d}x\right.
$$

$$
+ \int_{\mathrm{supp}(\eta^{n+1})\backslash(\mathrm{supp}(\rho^{n+1})\cap\mathrm{supp}(\eta^{n+1}))} |x - \mathcal{T}_\eta(x)|^2 \eta^{n+1}(x)\,\mathrm{d}x
$$

$$
\left. + \int_{(\mathrm{supp}(\rho^{n+1})\cap\mathrm{supp}(\eta^{n+1}))} |x - \mathcal{T}(x)|^2 \sigma^{n+1}(x)\right) + \mathcal{E}(U^{n+1})
$$

$$
= \frac{1}{2\tau}\left(\int_{\mathrm{supp}(\rho^{n+1})\backslash(\mathrm{supp}(\rho^{n+1})\cap\mathrm{supp}(\eta^{n+1}))} |x - \mathcal{T}_\rho(x)|^2 \sigma^{n+1}(x)\,\mathrm{d}x\right.
$$

$$
+ \int_{\mathrm{supp}(\eta^{n+1})\backslash(\mathrm{supp}(\rho^{n+1})\cap\mathrm{supp}(\eta^{n+1}))} |x - \mathcal{T}_\eta(x)|^2 \sigma^{n+1}(x)\,\mathrm{d}x
$$

$$
\left. + \int_{(\mathrm{supp}(\rho^{n+1})\cap\mathrm{supp}(\eta^{n+1}))} |x - \mathcal{T}(x)|^2 \sigma^{n+1}(x)\right) + \mathcal{E}(U^{n+1})
$$

$$
\geq \frac{1}{2\tau}d_2^2(\sigma^{n+1}, \sigma^{n+1/2}) + \mathcal{E}(\sigma^{n+1})
$$

$$
> \frac{1}{2\tau}d_2^2(\sigma^*, \sigma^{n+1/2}) + \mathcal{E}(\sigma^*)
$$

$$
= \frac{1}{2\tau}\int_\Omega |x - \mathcal{S}(x)|^2 \sigma^*(x)\,\mathrm{d}x + \mathcal{E}(\sigma^*),
$$

where $\mathcal{S}$ is the optimal transport map such that $\mathcal{S}_\#\sigma^* = \sigma^{n+1/2}$. We may write $\sigma^* = \rho^* + \eta^*$ such that $\mathcal{S}_\#\rho^* = \rho^{n+1/2}$ and $\mathcal{S}_\#\eta^* = \eta^{n+1/2}$. Then we may write

$$
\frac{1}{2\tau}d_2^2(U^{n+1}, U^{n+1/2}) + \mathcal{E}(U^{n+1})
$$

$$
> \frac{1}{2\tau}\int_\Omega |x - \mathcal{S}(x)|^2 \rho^*(x)\,\mathrm{d}x + \frac{1}{2\tau}\int_\Omega |x - \mathcal{S}(x)|^2 \eta^*(x)\,\mathrm{d}x + \mathcal{E}(\rho^* + \eta^*)
$$

$$
\geq \frac{1}{2\tau}d_2^2(U^*, U^{n+1/2}) + \mathcal{E}(U^*),
$$

which is absurd, for it implies that U^{n+1} is not a minimizer. Hence, we conclude that σ^{n+1} is a minimizer of problem (13) and, as a by-product, $\mathcal{T}_\rho \equiv \mathcal{T}_\eta$ on $\mathrm{supp}(\rho^{n+1}) \cup \mathrm{supp}(\eta^{n+1}) = \mathrm{supp}(\sigma^{n+1})$.

Finally, let us address the construction of the partition $\sigma^* = \rho^* + \eta^*$. To this end, we set $A := \mathrm{supp}(\sigma^*) \subset \Omega$ and recall that this set is compact. Therefore, its complement in Ω is open, and we may write it as

$$
A^c = \bigcup_{i\in\mathbb{N}} I_i,
$$

i.e., as the countable union of open intervals. Taking the closure of these intervals, we see that

$$B := \Omega \setminus \bigcup_{i \in \mathbb{N}} \mathrm{cl}(I_i)$$

satisfies $A = B$ up to a set of measure zero. However, since $\sigma^{n+1/2}$ is absolutely continuous, the transport map cannot map intervals with positive mass onto a point. Thus, restricting $\mathcal{S}$ to B even makes the map injective, and we can readily define its inverse. Setting $\rho^* = \mathcal{S}|_{B}^{-1}{}_{\#} \rho^{n+1/2}$ and $\eta^* = \mathcal{S}|_{B}^{-1}{}_{\#} \eta^{n+1/2}$ is a suitable decomposition of σ^*, which concludes the proof. $\square$

PROPOSITION 3.4 (L^∞ stability of the diffusive step). *Let $(r^{n+1/2}, \sigma^{n+1/2})$ be given by the splitting scheme* (12a). *Then these quantities satisfy*

$$0 \le \sigma^{n+1} \le \|\sigma^{n+1/2}\|_{L^\infty} \quad and \quad 0 \le r^{n+1} \le 1,$$

after the diffusion step for any $0 \le n \le N$.

Proof. Choose $x_0 \in \mathrm{argmax}(\sigma^{n+1})$. Then, by the optimality condition (11), we have

$$\chi'(\sigma^{n+1}) + \frac{\varphi}{\tau} = c,$$

and we get $x_0 \in \mathrm{argmin}(\varphi)$, where we used the fact that $\chi'' \ge 0$; cf. Definition 2.5(NL-iii). Hence, $\varphi''(x_0) \ge 0$ and, consequently, by passing to the derivative in (5) we get

(14) $$\mathcal{T}'(x_0) = 1 - \partial_{xx}\varphi(x_0) \le 1,$$

where $\mathcal{T}$ is the transport map from σ^{n+1} to $\sigma^{n+1/2}$. After a change of variables we get

$$\sigma^{n+1}(x) \le \|\sigma^{n+1}\|_{L^\infty} = \sigma^{n+1}(x_0) = \mathcal{T}'(x_0)\sigma^{n+1/2}(\mathcal{T}(x_0)) \overset{(14)}{\le} \sigma^{n+1/2}(\mathcal{T}(x_0))$$
$$\le \|\sigma^{n+1/2}\|_{L^\infty}$$

for any $x \in \Omega$. For the nonnegativity we observe that $\mathcal{T}' \ge 0$ by Brenier's theorem [14, 15, 49, 50]. Note that here we exploited the fact that the problem is posed in one spatial dimension. Thus,

$$\sigma^{n+1}(x) = \sigma^{n+1/2} \circ \mathcal{T}(x)\mathcal{T}'(x) \ge 0.$$

Finally, the bounds for r^{n+1} follow from its definition (cf. (12d)), as the composition with a monotone function does not change the infimum and the supremum of a function. This concludes the proof. $\square$

PROPOSITION 3.5 (bounded variation of r^{n+1} and σ^{n+1}). *Let $(r^{n+1/2}, \sigma^{n+1/2})$ be given. After the diffusion step, they satisfy the estimates*

$$\|\sigma^{n+1}\|_{BV} \le \|\sigma^{n+1/2}\|_{BV} \quad and \quad \|r^{n+1}\|_{BV} \le \|r^{n+1/2}\|_{BV}.$$

Proof. The result for the BV-norm of the minimizer, σ^{n+1}, is shown analogously to the proof of Theorem 1.1; cf. [27]. Now we need to show that the BV-norm of the ratio r does not increase. Recall the definition of r^{n+1} (cf. (12d)) as

$$r^{n+1} := r^{n+1/2} \circ \mathcal{T},$$

where $\mathcal{T}$ is the transport map such that $\rho^{n+1/2} = \mathcal{T}_{\#}\rho^{n+1}$ and $\sigma^{n+1/2} = \mathcal{T}_{\#}\sigma^{n+1}$. Note that it is indeed the same function and there holds $\mathcal{T}' \geq 0$, by Brenier's theorem [14, 15, 49, 50].

Now, let $P \subset \Omega$ be any partition of Ω. There holds

$$V_P(r^{n+1}) = V_P(r^{n+1/2} \circ \mathcal{T}) = V_{P'}(r^{n+1/2}) \leq \|r^{n+1/2}\|_{BV},$$

where P' is another partition induced by the monotone map $\mathcal{T}$. Taking the supremum over all partitions P, we get

$$\|r^{n+1}\|_{BV} \leq \|r^{n+1/2}\|_{BV}.$$

Finally, note that, indeed,

$$r^{n+1}(x) = \frac{\rho^{n+1}(x)}{\sigma^{n+1}(x)}$$

on $\mathrm{supp}(\sigma^{n+1})$, as we shall see now. According to Lemma 3.3, the same transport map $\mathcal{T}$ pushes ρ^{n+1} onto $\rho^{n+1/2}$ and σ^{n+1} onto $\sigma^{n+1/2}$. As a consequence, the densities satisfy

$$\rho^{n+1}(x) = \rho^{n+1/2}\big(\mathcal{T}(x)\big)\,\mathcal{T}'(x) \quad \text{and} \quad \sigma^{n+1}(x) = \sigma^{n+1/2}\big(\mathcal{T}(x)\big)\,\mathcal{T}'(x),$$

whence

$$(15) \qquad r^{n+1}(x) := r^{n+1/2} \circ \mathcal{T}(x) = \frac{\rho^{n+1/2}}{\sigma^{n+1/2}} \circ \mathcal{T}(x) = \frac{\rho^{n+1}}{\sigma^{n+1}}(x). \qquad \square$$

3.3. Combined estimates for an entire splitting step. We have now garnered all information necessary to pass to the limit. Let us combine the estimates from the previous section in the following lemma.

LEMMA 3.6 (*BV-estimates and L^{∞}-estimates*). *The sequence $(r_\tau, \sigma_\tau)_{\tau > 0}$ obtained by the splitting scheme is uniformly bounded in $L^{\infty}(0, T; L^{\infty} \cap BV)$. More precisely, there holds*

$$\sup_{t \in [0,T]} \|r_\tau\|_{L^{\infty}} \leq C \quad and \quad \sup_{t \in [0,T]} \|\sigma_\tau\|_{L^{\infty}} \leq C$$

and

$$\sup_{t \in [0,T]} \|r_\tau\|_{BV} \leq C \quad and \quad \sup_{t \in [0,T]} \|\sigma_\tau\|_{BV} \leq C$$

for some positive constant $C < \infty$ depending only on T and the initial data and not on τ.

Proof. The uniform L^{∞}-bounds are a consequence of combining Propositions 3.1 and 3.4. We use these uniform L^{∞}-bounds in the estimates for the BV-norm; cf. Propositions 3.2 and 3.5. Combining both for the reaction and diffusion step, we also obtain the uniform BV-bounds. $\square$

As a result of Lemmas 3.6 and 2.2, we obtain the following corollary.

COROLLARY 3.7 (BV-estimates and L^∞-estimates for ρ, η). *The sequences of approximated densities $(\rho_\tau)_{\tau>0}$, $(\eta_\tau)_{\tau>0}$ are uniformly bounded, i.e.,*

$$\|\rho_\tau\|_{L^\infty(0,T;L^\infty \cap BV)} \leq C \quad and \quad \|\eta_\tau\|_{L^\infty(0,T;L^\infty \cap BV)} \leq C,$$

where the constant $C > 0$ is independent of the parameter τ.

Finally, we need to prove an estimate on the cross-diffusion term to be able to pass to the limit later. This estimate is achieved in Lemma 3.10, which is preceded by two technical lemmas. We exploit the existence of an auxiliary functional guaranteed by Definition 2.5. Note that in the absence of the reaction part, this would indeed be an entropy in the classical sense, i.e., it would be decayed along solutions. Since we are interested in a uniform estimate, we begin by proving a control of this functional during each reaction phase.

LEMMA 3.8 (control of the auxiliary functional in the reaction step). *$\mathcal{K}$ increases at most at a constant rate independent of n. More precisely, there holds*

$$\mathcal{K}(\sigma^{n+1/2}) \leq \mathcal{K}(\sigma^n) + c\tau,$$

for any $n \in \mathbb{N}$.

Proof. A straightforward computation yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega K(\sigma) \, \mathrm{d}x = \int_\Omega \kappa(\sigma)\partial_t\sigma \, \mathrm{d}x = \int_\Omega \kappa(\sigma)\sigma(rA_1 + (1-r)A_2) \, \mathrm{d}x \leq c,$$

using the uniform L^∞-bound on σ and the fact that $|\Omega| < \infty$. Hence,

$$\mathcal{K}(\sigma^{n+1/2}) \leq \mathcal{K}(\sigma^n) + c\tau,$$

where c is independent of n. $\qquad\square$

Next, we address the diffusion step. As mentioned earlier, the auxiliary functional, $\mathcal{K}$, is an entropy for the diffusion part, and from its dissipation we obtain the necessary regularity, as asserted in the following lemma.

LEMMA 3.9 (H^1-bound for σ^{n+1}). *The minimizer of the JKO step satisfies the estimate*

$$\tau\|\partial_x\sigma^{n+1}\|_{L^2(\Omega)}^2 \leq \left(\mathcal{K}(\sigma^{n+1/2}) - \mathcal{K}(\sigma^{n+1})\right)$$

for each $0 \leq n \leq N$.

Proof. Let

$$(\rho^{n+1}, \eta^{n+1}) \in \mathrm{argmin}\left\{\frac{1}{2\tau}d_2^2(\cdot, U^{n+1/2}) + \mathcal{E}(U)\right\}.$$

Let $\sigma_s = (\mathcal{T}_s)_{\#}\sigma^{n+1}$ be the geodesic interpolation between $\sigma_s|_{s=0} = \sigma^{n+1}$ and $\sigma_s|_{s=1} = \sigma^{n+1/2}$, given by

$$\mathcal{T}_s = (1-s)\mathrm{id} + s\mathcal{T},$$

and

$$\mathcal{T} = \mathrm{id} - \partial_x\varphi,$$

for the associated Kantorovich potential, φ; cf. (5). As a consequence, the velocity field is given by

$$v_s = (\mathcal{T} - \mathrm{id}) \circ \mathcal{T}_s^{-1},$$

satisfying the continuity equation

$$\partial_s \sigma_s = \partial_x(\sigma_s v_s).$$

We differentiate the entropy along the geodesic and obtain

$$\frac{\mathrm{d}}{\mathrm{d}s} \int_\Omega K(\sigma_s)\,\mathrm{d}x = \int_\Omega \kappa(\sigma_s)\partial_s\sigma_s\,\mathrm{d}x = -\int_\Omega \kappa'(\sigma_s)\partial_x\sigma_s\sigma_s v_s\,\mathrm{d}x$$

$$= -\int_\Omega \chi''(\sigma_s)\partial_x\sigma_s v_s\,\mathrm{d}x = -\int_\Omega \partial_x\chi'(\sigma_s)v_s\,\mathrm{d}x.$$

Thus, at $s = 0$ the evolution of the entropy (9) becomes

$$\frac{\mathrm{d}}{\mathrm{d}s} \int_\Omega K(\sigma_s)\,\mathrm{d}x\bigg|_{s=0} = \int_\Omega \partial_x\chi'(\sigma^{n+1})\partial_x\varphi\,\mathrm{d}x.$$

Using the optimality condition (11), we obtain

$$\tau \int_\Omega \left|\partial_x\chi'(\sigma^{n+1})\right|^2 \mathrm{d}x = -\int_\Omega \partial_x\chi'(\sigma^{n+1})\partial_x\varphi\,\mathrm{d}x$$

$$= \frac{\mathrm{d}}{\mathrm{d}s} \int_\Omega K(\sigma_s)\,\mathrm{d}x\bigg|_{s=0}$$

$$\leq \mathcal{K}(\sigma^{n+1/2}) - \mathcal{K}(\sigma^{n+1}),$$

where the last inequality is a consequence of the geodesic convexity of the entropy; cf. Remark 2.7. This concludes the proof. $\qquad\square$

We combine the previous lemmas to obtain the desired estimate for a full iteration and finally for the piecewise constant interpolation, σ_τ.

LEMMA 3.10 (uniform $L^2((0,T) \times \Omega)$-bound for $\partial_x\sigma_\tau$). *There holds*

$$\|\partial_x\chi'(\sigma_\tau)\|_{L^2((0,T)\times\Omega)} \leq C$$

for some positive constant depending only on T.

Proof. This statement is a consequence of combining Lemmas 3.8 and 3.9 to get

$$\tau\|\partial_x\chi'(\sigma^{n+1})\|_{L^2(\Omega)}^2 \leq \mathcal{K}(\sigma^{n+1/2}) - \mathcal{K}(\sigma^{n+1})$$

$$\leq c\tau + \mathcal{K}(\sigma^n) - \mathcal{K}(\sigma^{n+1}).$$

Summing over $n = 0 \ldots N - 1$ gives

$$\|\partial_x\chi'(\sigma_\tau)\|_{L^2((0,T)\times\Omega)}^2 \leq cT + \mathcal{K}(\sigma^0) - \inf_\sigma \mathcal{K}(\sigma) \leq C,$$

which yields the statement of the lemma. $\qquad\square$

LEMMA 3.11 (total-square 2-Wasserstein distance estimates). *For every $n \in \{0, \ldots, N\}$ consider two consecutive steps for (12c), $U^{n+1/2} = (\rho^{n+1/2}, \eta^{n+1/2})$ and $U^{n+1} = (\rho^{n+1}, \eta^{n+1})$. Then there exists a constant C such that*

$$\frac{1}{2\tau} \sum_{n=0}^N d_2^2(U^{n+1/2}, U^{n+1}) \leq C.$$

Proof. Using the minimizing property of U^{n+1}, we have

$$\frac{1}{2\tau}d_2^2(U^{n+1/2}, U^{n+1}) \leq \mathcal{E}(U^{n+1/2}) - \mathcal{E}(U^{n+1}).$$

Adding and subtracting $\mathcal{E}(U^n)$ on the right-hand side, we obtain

$$\begin{aligned}
\mathcal{E}(U^{n+1/2}) &- \mathcal{E}(U^n) \\
&= \int_\Omega \chi(\sigma^{n+1/2})\,\mathrm{d}x - \int_\Omega \chi(\sigma^n)\,\mathrm{d}x \\
&= \int_\Omega \chi\left(\sigma^n + \int_{t^n}^{t^{n+1}} \sigma(rA_1 + (1-r)A_2)\,\mathrm{d}s\right)\mathrm{d}x - \int_\Omega \chi(\sigma^n)\,\mathrm{d}x \\
&\leq \int_\Omega \chi'(\xi(x)) \int_{t^n}^{t^{n+1}} \sigma(rA_1 + (1-r)A_2)\,\mathrm{d}s\,\mathrm{d}x
\end{aligned}$$

due to a Taylor expansion of χ around σ^n, where $\xi \in [\sigma^n, \sigma^{n+1/2}]$. Owing to the uniform L^∞-bounds on the σ and $\sigma^{n+1/2}$, we may further simplify

$$(16) \qquad \mathcal{E}(U^{n+1/2}) - \mathcal{E}(U^n) \leq \|\chi'\|_{L^\infty} \int_\Omega \int_{t^n}^{t^{n+1}} |\sigma(rA_1 + (1-r)A_2)|\,\mathrm{d}s\,\mathrm{d}x \leq C\tau,$$

where, in the last inequality, we used the L^∞-bounds on σ, r as well as on F_i, G_i. Hence, we may summarize

$$\frac{1}{2\tau}d_2^2(U^{n+1/2}, U^{n+1}) \leq \mathcal{E}(U^n) - \mathcal{E}(U^{n+1}) + C\tau$$

for a full time step. Finally, summing over n gives the following result:

$$\begin{aligned}
\frac{1}{2\tau}\sum_{n=1}^N d_2^2(U^{n+1/2}, U^{n+1}) &\leq \sum_{n=0}^N \left(\mathcal{E}(U^n) - \mathcal{E}(U^{n+1}) + C\tau\right) \\
&= \mathcal{E}(U^0) - \mathcal{E}(U_\tau^N) + N\,(C\tau) \\
&\leq \mathcal{E}(U^0) - \inf_{U \in \mathcal{M}_+ \times \mathcal{M}_+} \mathcal{E}(U) + CN\tau \\
&\leq \bar{C}. \qquad\qquad\qquad\qquad \square
\end{aligned}$$

3.4. Convergence. We now prove the convergence result.

PROPOSITION 3.12. *The piecewise constant interpolations defined in Definition 2.8 admit subsequences converging uniformly to absolutely continuous curves, $\bar\rho, \bar\eta$, with respect to d_{BL} with values in $\mathcal{M}_+(\Omega)$. Moreover, $\bar\rho$ and $\bar\eta$ are d_{BL}-continuous functions on $[0,T]$.*

Proof. The proof is based on the application of a generalized version of the Ascoli–Arzelà theorem; cf. [2, section 3].

We begin by establishing "almost continuity" of the approximation. To this end, let $0 \leq s < t \leq T$ be two time instances. Then there exist two uniquely determined integers, m, k, $m \leq k-1$, satisfying

$$s \in ((m-1)\tau, m\tau] \quad \text{and} \quad t \in ((k-1)\tau, k\tau],$$

such that

$$(17) \quad d_{\mathrm{BL}}(U_\tau(s), U_\tau(t)) \leq \sum_{n=m}^{k-1} d_{\mathrm{BL}}(U^n, U^{n+1}) \leq \left(\sum_{n=m}^{k-1} d_{\mathrm{BL}}^2(U^n, U^{n+1}) \right)^{\frac{1}{2}} |k-m|^{\frac{1}{2}},$$

where $U_\tau(t) = (\rho_\tau(t), \eta_\tau(t))$ as defined in Definition 2.8. It becomes apparent that we need to address the bounded-Lipschitz term next. To this end, we use the triangulation established in Corollary 2.4 to estimate it by the L^1-distance in the reaction step and the W_1-distance in the diffusion step. Hence,

$$\sum_{n=m}^{k-1} d_{\mathrm{BL}}^2(U^n, U^{n+1}) \leq 2 \sum_{n=0}^{N} \|U^n - U^{n+1/2}\|_{L^1(\Omega)}^2 + 2 \sum_{n=0}^{N} d_1^2(U^{n+1/2}, U^{n+1}).$$

For the reaction step, an argument similar to (16) yields

$$\|U^n - U^{n+1/2}\|_{L^1(\Omega)}^2 = \left| \int_\Omega |\rho^{n+1/2} - \rho^n| + |\eta^{n+1/2} - \eta^n| \right|^2$$

$$= \left| \int_\Omega \left| \int_{t^n}^{(n+1)\tau} r\sigma \tilde{F}_1 + (1-r)\sigma \tilde{G}_1 \, ds \right| \right.$$

$$\left. + \left| \int_{t^n}^{(n+1)\tau} (1-r)\sigma \tilde{F}_2 + r\sigma \tilde{G}_2 \, ds \right| \, dx \right|^2$$

$$\leq C\tau^2$$

by (12b), where C depends on the L^∞-bounds of r, σ as well as of F_i, G_i, $i = 1, 2$. Using the fact that the p-Wasserstein distances are ordered, we even have

$$\sum_{n=m}^{k-1} d_{\mathrm{BL}}^2(U^n, U^{n+1}) \leq 2 \sum_{n=0}^{N} \|U^n - U^{n+1/2}\|_{L^1(\Omega)}^2 + 2 \sum_{n=0}^{N} d_2^2(U^{n+1/2}, U^{n+1})$$

$$\leq C\tau,$$

where we used the total-square estimate, Lemma 3.11. Therefore, (17) becomes

$$d_{BL}(U_\tau(s), U_\tau(t)) \leq \left(\sum_{n=m}^{k-1} d_{\mathrm{BL}}^2(U^n, U^{n+1}) \right)^{\frac{1}{2}} |k-m|^{\frac{1}{2}}$$

$$\leq C\sqrt{\tau} \left(\frac{|t-s|}{\tau} + 1 \right)^{\frac{1}{2}}$$

$$\leq C(\sqrt{|t-s|} + \sqrt{\tau}).$$

Thus, we get the "almost $\frac{1}{2}$-Hölder continuity" for the curve $U_\tau(t)$, and we obtain the uniform narrow compactness on compact time intervals by using the refined version of the Ascoli–Arzelà theorem; cf. [2, section 3]. $\qquad \square$

COROLLARY 3.13 (strong convergence in $L^p(0, T; L^q(\Omega))$). *Let $1 \leq p, q < \infty$, and let $(\rho_\tau)_{\tau>0}$ and $(\eta_\tau)_{\tau>0}$ be the sequences of the piecewise constant interpolations as in Definition 2.8. Then there exist two functions $\rho, \eta \in L^p(0, T; L^q(\Omega))$ and subsequences, again denoted by $(\rho_\tau)_{\tau>0}$ and $(\eta_\tau)_{\tau>0}$, such that*

$$\rho_\tau \to \rho \quad and \quad \eta_\tau \to \eta$$

strongly in $L^p(0, T; L^q(\Omega))$ as $\tau \to 0$. Moreover, the convergence holds pointwise in time; i.e., for all $t \in [0, T]$,

$$\rho_\tau(t) \to \rho(t) \quad and \quad \eta_\tau(t) \to \eta(t)$$

hold strongly in $L^q(\Omega)$.

Proof. Note that it suffices to show the result for $(\rho_\tau)_{\tau>0}$, as the same argument applies for $(\eta_\tau)_{\tau>0}$. By Proposition 3.12 we can extract a subsequence, still denoted the same, such that

$$\rho_\tau(t) \rightharpoonup \rho,$$

in $\mathcal{M}_+(\Omega)$ for all $t \in [0, T]$. Furthermore, from the uniform $BV \cap L^\infty$-bounds, for any $t \in [0, T]$ the sequence converges strongly in $L^1(\Omega)$. Using the uniform L^∞-bounds and the dominated convergence theorem, we obtain the pointwise-in-time convergence in any $L^q(\Omega)$.

Now let us apply the same argument on the whole domain, $[0, T] \times \Omega$. The pointwise convergence and the uniform $L^\infty(0, T; L^\infty(\Omega))$-bound imply

$$\rho_\tau \to \rho,$$

strongly in $L^p(0, T; L^q(\Omega))$ by the dominated convergence theorem. This concludes the proof. □

LEMMA 3.14 (identification of the limit). *The sequence constructed in (12a), (12c) converges to a weak solution of (1).*

Proof. The proof consists of two parts: the diffusion part and the reaction part. We write them in their respective weak formulation and combine them to obtain the complete approximation of the weak formulation. Here we only show the argument for ρ, as the corresponding result for η is shown analogously. We begin with the diffusion part.

Diffusion part. We consider the two steps before and after the application of the JKO scheme (12c),

$$U^{n+1/2} = (\rho^{n+1/2}, \eta^{n+1/2}) \quad \text{and} \quad U^{n+1} = (\rho^{n+1}, \eta^{n+1}).$$

For a given test function $\zeta \in C_c^\infty(\Omega)$, let $\mathcal{T}$ be the optimal transport map from ρ^{n+1} to $\rho^{n+1/2}$. Upon integration over Ω, we get

$$\frac{1}{\tau} \int_\Omega \left(\rho^{n+1}(x) - \rho^{n+1/2}(x) \right) \zeta(x)\,\mathrm{d}x = \frac{1}{\tau} \int_\Omega \rho^{n+1}(x) \left(\zeta(x) - \zeta(\mathcal{T}(x)) \right)\,\mathrm{d}x.$$

Taylor expanding $\zeta(\mathcal{T}(x))$ around x yields

$$\begin{aligned}\zeta(\mathcal{T}(x)) &= \zeta(x + (\mathcal{T}(x) - x)) \\ &= \zeta(x) + \partial_x \zeta(x)(\mathcal{T}(x) - x) + \mathcal{O}\left(|\mathcal{T}(x) - x|^2 \right).\end{aligned}$$

Moreover, using the fact that $\mathcal{T}(x) = x - \partial_x \varphi(x)$, where φ is the Kantorovich potential associated to the optimal map $\mathcal{T}$, the integral above can be rewritten as

$$\begin{aligned}\frac{1}{\tau} &\int_\Omega \left(\rho^{n+1}(x) - \rho^{n+1/2}(x) \right) \zeta(x)\,\mathrm{d}x \\ &= -\frac{1}{\tau} \int_\Omega \rho^{n+1}(x) \left(\mathcal{T}(x) - x \right) \partial_x \zeta(x)\,\mathrm{d}x + O\left(d_2^2(\rho^{n+1}, \rho^{n+1/2}) \right) \\ &= \frac{1}{\tau} \int_\Omega \rho^{n+1}(x) \partial_x \varphi(x) \partial_x \zeta(x)\,\mathrm{d}x + O\left(d_2^2(\rho^{n+1}, \rho^{n+1/2}) \right).\end{aligned}$$

Thanks to the optimality condition, Lemma 3.3, we get

(18)
$$\frac{1}{\tau}\int_{\Omega}\left(\rho^{n+1} - \rho^{n+1/2}\right)\zeta\,\mathrm{d}x$$
$$= -\int_{\Omega}\rho^{n+1}\partial_x\chi'(\sigma^{n+1})\,\partial_x\zeta\,\mathrm{d}x + \mathcal{O}\left(d_2^2(\rho^{n+1}, \rho^{n+1/2})\right).$$

Reaction part. Note that

(19)
$$\int_{\Omega}\frac{\rho^{n+1/2} - \rho^n}{\tau}\zeta\,\mathrm{d}x = \int_{\Omega}\frac{r^{n+1/2}\sigma^{n+1/2} - r^n\sigma^n}{\tau}\zeta\,\mathrm{d}x$$
$$= \int_{\Omega}\int_{t^n}^{t^{n+1}}\frac{1}{\tau}\partial_t\big(r(\bar\tau)\sigma(\bar\tau)\big)\zeta\,\mathrm{d}\bar\tau\,\mathrm{d}x$$
$$= \int_{\Omega}\left(r^{n+1/2}\sigma^{n+1/2}\tilde{F}_1 + (1 - r^{n+1/2})\sigma^{n+1/2}\tilde{G}_1\right)\zeta\,\mathrm{d}x + \mathcal{O}(\tau)$$
$$= \int_{\Omega}\left(\rho^{n+1/2}F_1^{n+1/2} + \eta^{n+1/2}G_1^{n+1/2}\right)\zeta\,\mathrm{d}x + \mathcal{O}(\tau),$$

with the shortcuts $F_1^{n+1/2} = F_1(\rho^{n+1/2}, \eta^{n+1/2})$ and $G_1^{n+1/2} = G_1(\rho^{n+1/2}, \eta^{n+1/2})$, having used an approximation of (12b).

Combination of both steps. Let us combine the reaction step and the diffusion step of the splitting scheme. Upon summing up (18), (19), we obtain

(20)
$$0 = \frac{1}{\tau}\int_{\Omega}\zeta(x)(\rho^{n+1}(x) - \rho^n(x))\,\mathrm{d}x + \mathcal{O}(d_2^2(\rho^{n+1}, \rho^{n+1/2}))$$
$$+ \int_{\Omega}\rho^{n+1}\partial_x\chi'(\rho^{n+1}(x) + \eta^{n+1}(x))\partial_x\zeta(x)\,\mathrm{d}x$$
$$- \int_{\Omega}\left(\rho^{n+1/2}F_1(\rho^{n+1/2}, \eta^{n+1/2}) + \eta^{n+1/2}G_1(\rho^{n+1/2}, \eta^{n+1/2})\right)\zeta\,\mathrm{d}x$$
$$+ \mathcal{O}(\tau).$$

We rewrite this equation in terms of the piecewise constant interpolation, Definition 2.8. For any $0 < s < t < T$, there are two uniquely determined integers m, k such that
$$s \in (m\tau, (m+1)\tau] \quad \text{and} \quad t \in (k\tau, (k+1)\tau].$$

Multiplying (20) by τ and summing from m to $k-1$, we obtain

$$0 = \int_{\Omega}\zeta(x)(\bar\rho_\tau(t, x) - \bar\rho_\tau(s, x))\,\mathrm{d}x + \mathcal{O}(\tau)$$
$$+ \int_s^t\int_{\Omega}\bar\rho_\tau(\bar\tau, x)\partial_x\chi'\big(\bar\rho_\tau(\bar\tau, x) + \bar\eta_\tau(\bar\tau, x)\big)\partial_x\zeta(x)$$
$$- \left(\bar\rho_\tau(\bar\tau, x)F_1\big(\bar\rho_\tau(\bar\tau, x), \bar\eta_\tau(\bar\tau, x)\big) + \bar\eta_\tau(\bar\tau, x)G_1\big(\bar\rho_\tau(\bar\tau, x), \bar\eta_\tau(\bar\tau, x)\big)\right)\zeta(x)\,\mathrm{d}x\mathrm{d}\bar\tau.$$

We are now ready to pass to the limit. The strong convergence obtained in Corollary 3.13 allows us to pass to the limit in the nonlinear reaction terms. Moreover, the cross-diffusion term converges due to the weak-strong $L^2(0, T; L^2(\Omega))$-duality by Lemma 3.10 and Corollary 3.13. Passing to the limit $\tau \to 0$, we obtain the weak formulation for ρ. The same argument for η yields the statement. □

We end the paper with a stunning result which can be seen as a generalization of the result of Bertsch et al.; cf. [8, 7]. In these papers they prove that initially segregated species stay segregated at all times. In these works only the cross-diffusion system is considered, and the authors even allow for vacuum; however, they impose a rather restrictive assumption on the initial datum. We can drop their assumption that

$$\operatorname{supp}(\rho_0) < \operatorname{supp}(\eta_0),$$

i.e., that both species are ordered, and prove the following, more general theorem. In the works [6, 9] the authors study the system including reaction terms for initial data that are not necessarily ordered. However, they impose boundedness away from zero which we are also able to drop.

THEOREM 3.15 (segregation in the case of no cross-reaction). *Assume no cross-reaction terms, i.e., $G_1 \equiv G_2 \equiv 0$, and that both species are initially segregated, i.e.,*

$$\int_\Omega \rho_0(x)\,\eta_0(x)\,\mathrm{d}x = 0.$$

Then, there exists a solution such that both species stay segregated at all times.

Proof. It suffices to show this property at the level of the discrete scheme since, once it is established, we use the strong L^2-convergence of ρ_τ and η_τ to show segregation is also kept in the limit:

$$0 = \lim_{\tau \to 0} \int_\Omega \rho_\tau(t,x)\eta_\tau(t,x)\,\mathrm{d}x = \int_\Omega \rho(t,x)\eta(t,x)\,\mathrm{d}x.$$

Thus, let us now show the property at the level of the approximation. During the reaction step, the segregation of ρ and η is kept as follows. First, recall that in the reaction step we solve (12a), with initial data σ^n, r^n such that

$$r^n\sigma^n \cdot (1 - r^n)\sigma^n = r^n(1 - r^n)\,(\sigma^n)^2 = 0,$$

on the support of σ^n. Thus, we conclude that already

$$r^n(1 - r^n) = 0$$

on $\operatorname{supp}(\sigma^n)$. Using the fact that $\tilde{G}_1 = \tilde{G}_2 \equiv 0$, the equation for r reads

$$\partial_t r = r(1 - r)A_3,$$

with $r(t^n) = r^n$. By the uniqueness of the solution, we have $r(t) = r^n$ on $[t^n, (n+1)\tau]$, and therefore $r^{n+1/2} = r^n$. Next, we notice that $\operatorname{supp}(\sigma^{n+1/2}) \subset \operatorname{supp}(\sigma^n)$. This is due to the fact that the right-hand side of the equation for σ is premultiplied by σ,

$$\partial_t \sigma = \sigma(rA_1 + (1 - r)A_2);$$

i.e., for any $x \in \Omega$ with $\sigma(t^n, x) = 0$ there holds $\sigma(t, x) = 0$ for all $t \in [t^n, (n+1)\tau]$ by the uniqueness of solutions, whence we deduce $\operatorname{supp}(\sigma^{n+1/2}) \subset \operatorname{supp}(\sigma^n)$.

Hence, we only need to prove that segregation is kept in the diffusion step. This is done by contradiction. Let us assume there exists an $1 \le n \le N$ such that

$$(21) \qquad \int_\Omega \rho^{n+1/2}(x)\eta^{n+1/2}(x)\,\mathrm{d}x = 0,$$

while

$$\int_\Omega \rho^{n+1}(x)\eta^{n+1}(x)\,\mathrm{d}x > 0.$$

Then there exists $\delta > 0$ and a set $B \subset \Omega$ with $|B| > 0$, such that

$$|\mathcal{T}'(x)| < \infty, \quad \rho^{n+1}(x) > \delta, \quad \text{and} \quad \eta^{n+1}(x) > \delta,$$

for almost all $x \in B$. As both species have in common the same transport map $\mathcal{T}$, there exists a set A such that $A = \mathcal{T}(B)$ and

$$0 < \delta < \rho^{n+1}(x) = \rho^{n+1/2}(\mathcal{T}(x))\mathcal{T}'(x),$$
$$0 < \delta < \eta^{n+1}(x) = \eta^{n+1/2}(\mathcal{T}(x))\mathcal{T}'(x),$$

which is absurd, for we assumed (21). Thus segregation is kept at each iteration, which concludes the proof of the theorem. □

REFERENCES

[1] L. AMBROSIO, *Minimizing movements*, Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5), 19 (1995), pp. 191–246.

[2] L. AMBROSIO, N. GIGLI, AND G. SAVARÉ, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2nd ed., Lectures Math. ETH Zürich, Birkhäuser Verlag, Basel, 2008.

[3] N. ARMSTRONG, K. PAINTER, AND J. SHERRATT, *A continuum approach to modelling cell-cell adhesion*, J. Theoret. Biol., 243 (2006), pp. 98–113.

[4] N. ARMSTRONG, K. PAINTER, AND J. SHERRATT, *Adding adhesion to a chemical signaling model for somite formation*, Bull. Math. Biol., 71 (2009), pp. 1–24.

[5] N. ARMSTRONG, K. PAINTER, AND J. SHERRATT, *The impact of adhesion on cellular invasion processes in cancer and development*, J. Theoret. Biol., 264 (2010), pp. 1057–1067.

[6] M. BERTSCH, R. DAL PASSO, AND M. MIMURA, *A free boundary problem arising in a simplified tumour growth model of contact inhibition*, Interfaces Free Boundaries, 12 (2010), pp. 235–250.

[7] M. BERTSCH, M. GURTIN, AND D. HILHORST, *On a degenerate diffusion equation of the form $c(z)_t = \varphi(z_x)_x$ with application to population dynamics*, J. Differential Equations, 67 (1987), pp. 56–89.

[8] M. BERTSCH, M. GURTIN, D. HILHORST, AND L. PELETIER, *On interacting populations that disperse to avoid crowding: Preservation of segregation*, J. Math. Biol., 23 (1985), pp. 1–13.

[9] M. BERTSCH, D. HILHORST, H. IZUHARA, AND M. MIMURA, *A nonlinear parabolic-hyperbolic system for contact inhibition of cell-growth*, Differ. Equ. Appl., 4 (2012), pp. 137–157.

[10] M. BESSEMOULIN-CHATARD AND F. FILBET, *A finite volume scheme for nonlinear degenerate parabolic equations*, SIAM J. Sci. Comput., 34 (2012), pp. B559–B583, https://doi.org/10.1137/110853807.

[11] A. BLANCHET, V. CALVEZ, AND J. A. CARRILLO, *Convergence of the mass-transport steepest descent scheme for the subcritical Patlak–Keller–Segel model*, SIAM J. Numer. Anal., 46 (2008), pp. 691–721, https://doi.org/10.1137/070683337.

[12] M. BODNAR AND J. J. L. VELÁZQUEZ, *Friction dominated dynamics of interacting particles locally close to a crystallographic lattice*, Math. Methods Appl. Sci., 36 (2013), pp. 1206–1228.

[13] F. BOLLEY AND J. A. CARRILLO, *Nonlinear diffusion: Geodesic convexity is equivalent to Wasserstein contraction*, Comm. Partial Differential Equations, 39 (2014), pp. 1860–1869.

[14] Y. Brenier, *Décomposition polaire et réarrangement monotone des champs de vecteurs*, C.R. Acad. Sci. Paris Sér. I Math., 305 (1987), pp. 805–808.

[15] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Commun. Pure Appl. Math., 44 (1991), pp. 375–417.

[16] M. Burger, V. Capasso, and D. Morale, *On an aggregation model with long and short range interactions*, Nonlinear Anal. Real World Appl., 8 (2007), pp. 939–958.

[17] M. Burger, M. Di Francesco, S. Fagioli, and A. Stevens, *Sorting phenomena in a mathematical model for two mutually attracting/repelling species*, SIAM J. Math. Anal., 50 (2018), pp. 3210–3250, https://doi.org/10.1137/17M1125716.

[18] J. A. Cañizo, J. A. Carrillo, and S. Cuadrado, *Measure solutions for some models in population dynamics*, Acta Appl. Math., 123 (2013), pp. 141–156.

[19] J. A. Cañizo, J. A. Carrillo, and J. Rosado, *A well-posedness theory in measures for some kinetic models of collective motion*, Math. Models Methods Appl. Sci., 21 (2011), pp. 515–539.

[20] V. Calvez and J. A. Carrillo, *Volume effects in the Keller-Segel model: Energy estimates preventing blow-up*, J. Math. Pures Appl. (9), 86 (2006), pp. 155–175.

[21] J. A. Carrillo, A. Chertock, and Y. Huang, *A finite-volume method for nonlinear nonlocal equations with a gradient flow structure*, Commun. Comput. Phys., 17 (2015), pp. 233–258.

[22] J. A. Carrillo, A. Colombi, and M. Scianna, *Adhesion and Volume Constraints via Nonlocal Interactions Lead to Cell Sorting*, preprint, https://arxiv.org/abs/1706.08969, 2017.

[23] J. A. Carrillo, Y. Huang, and M. Schmidtchen, *Zoology of a nonlocal cross-diffusion model for two species*, SIAM J. Appl. Math., 78 (2018), pp. 1078-1104, https://doi.org/10.1137/17M1128782.

[24] J. A. Carrillo, A. Jüngel, P. A. Markowich, G. Toscani, and A. Unterreiter, *Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities*, Monatsh. Math., 133 (2001), pp. 1–82.

[25] J. A. Carrillo, N. Kolbe, and M. Lukáčová-Medvid'ová, *A Hybrid Mass Transport Finite Element Method for Keller-Segel Type Systems*, https://arxiv.org/abs/1709.07394, 2017.

[26] M. A. J. Chaplain and G. Lolas, *Mathematical modelling of cancer cell invasion of tissue. the role of the urokinase plasminogen activation system*, Math. Model. Methods Appl. Sci., 15 (2005), pp. 1685–1734.

[27] G. De Philippis, A. R. Mészáros, F. Santambrogio, and B. Velichkov, *BV estimates in optimal transportation and applications*, Arch. Ration. Mech. Anal., 219 (2016), pp. 829–860.

[28] P. Domschke, D. Trucu, A. Gerisch, and M. Chaplain, *Mathematical modelling of cancer invasion: Implications of cell adhesion variability for tumour infiltrative growth patterns*, J. Theoret. Biol., 361 (2014), pp. 41–60.

[29] A. Gerisch and M. Chaplain, *Mathematical modelling of cancer cell invasion of tissue: Local and non-local models and the effect of adhesion*, J. Theoret. Biol., 250 (2008), pp. 684–704.

[30] M. E. Gurtin and A. C. Pipkin, *A note on interacting populations that disperse to avoid crowding*, Quart. Appl. Math., 42 (1984), pp. 87–94.

[31] P. Gwiazda and A. Marciniak-Czochra, *Structured population equations in metric spaces*, J. Hyperbolic Differential Equation, 7 (2010), pp. 733–773.

[32] M. D. Johnston, P. K. Maini, S. Jonathan-Chapman, C. M. Edwards, and W. F. Bodmer, *On the proportion of cancer stem cells in a tumour*, J. Theoret. Biol., 266 (2010), pp. 708–711.

[33] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formulation of the Fokker–Planck equation*, SIAM J. Math. Anal., 29 (1998), pp. 1–17, https://doi.org/10.1137/S0036141096303359.

[34] I. Kim and A. R. Mészáros, *On nonlinear cross-diffusion systems: An optimal transport approach*, Calc. Var. Partial Differential Equations, 57 (2018), 79.

[35] A. Klar and S. Tiwari, *A multiscale meshfree method for macroscopic approximations of interacting particle systems*, Multiscale Model. Simul., 12 (2014), pp. 1167–1192, https://doi.org/10.1137/130945788.

[36] S. Kondo, *An updated kernel-based Turing model for studying the mechanisms of biological pattern formation*, J. Theoret. Biol., 414 (2017), pp. 120–127.

[37] R. J. McCann, *A convexity principle for interacting gases*, Adv. Math., 128 (1997), pp. 153–179.

[38] H. Murakawa and H. Togashi, *Continuous models for cell–cell adhesion*, J. Theoret. Biol., 374 (2015), pp. 1–12.

[39] H. Neunzert, *An introduction to the nonlinear Boltzmann-Vlasov equation*, in Kinetic Theories and the Boltzmann Equation (Montecatini, 1981), Lecture Notes in Math. 1048, Springer, Berlin, 1984, pp. 60–110.

[40] K. Oelschläger, *Large systems of interacting particles and the porous medium equation*, J. Differential Equations, 88 (1990), pp. 294–346.

[41] F. Otto, *The geometry of dissipative evolution equations: The porous medium equation*, Comm. Partial Differential Equations, 26 (2001), pp. 101–174.

[42] K. J. Painter and T. Hillen, *Volume-filling and quorum-sensing in models for chemosensitive movement*, Can. Appl. Math. Q., 10 (2002), pp. 501–543.

[43] L. Preziosi, *Cancer Modelling and Simulation*, CRC Press, Boca Raton, FL, 2003.

[44] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, Progr. Nonlinear Differential Equations Appl. 86, Birkhäuser Verlag, Basel, 2015.

[45] N. Sfakianakis, N. Kolbe, N. Hellmann, and M. Lukáčová-Medvid'ová, *A multiscale approach to the migration of cancer stem cells: Mathematical modelling and simulations*, Bull. Math. Biol., 79 (2017), pp. 209–235.

[46] H. Spohn, *Large Scale Dynamics of Interacting Particles*, Theoret. Math. Phys., Springer, Berlin, 1991.

[47] C. Stinner, C. Surulescu, and A. Uatay, *Global existence for a go-or-grow multiscale model for tumor invasion with therapy*, Math. Models Methods Appl. Sci., 26 (2016), pp. 2163–2201.

[48] J. L. Vázquez, *The Porous Medium Equation. Mathematical Theory*, Oxford Math. Monogr., Oxford University Press, Oxford, 2007.

[49] C. Villani, *Topics in Optimal Transportation*, Grad. Stud. Math. 58, American Mathematical Society, Providence, RI, 2003.

[50] C. Villani, *Optimal Transport: Old and New*, Grundlehren Math. Wiss. 338, Springer, Berlin, 2009.

[51] A. Volkening and B. Sandstede, *Modelling stripe formation in zebrafish: An agent-based approach*, J. Roy. Soc. Interface, 12 (112), 2015.