

**SPLITTING SCHEMES & SEGREGATION IN REACTION-(CROSS-)DIFFUSION SYSTEMS\***

J. A. CARRILLO<sup>†</sup>, S. FAGIOLI<sup>‡</sup>, F. SANTAMBROGIO<sup>§</sup>, AND M. SCHMIDTCHEN<sup>¶</sup>

**Abstract.** One of the most fascinating phenomena observed in reaction-diffusion systems is the emergence of segregated solutions, *i.e.* population densities with disjoint supports. We analyse such a reaction cross-diffusion system. In order to prove existence of weak solutions for a wide class of initial data without restriction about 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

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

$$(1.1) \quad \begin{cases} \partial_t \rho = \partial_x (\rho \partial_x \chi'(\rho + \eta)) + \rho F_1(\rho, \eta) + \eta G_1(\rho, \eta), \\ \partial_t \eta = \partial_x (\eta \partial_x \chi'(\rho + \eta)) + \eta F_2(\rho, \eta) + \rho G_2(\rho, \eta), \end{cases}$$

where  $\chi : [0, \infty) \rightarrow [0, \infty)$  is a  $C^1$  super-linear function modelling 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 (1.1) was originally introduced in the seminal papers [30] and [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 into mathematical biology models, see [42, 20] in the case of the classical Keller-Segel system. They help to avoid blow-up in these aggregation models in a biologically

---

\*JAC was supported by the Engineering and Physical Sciences Research Council (EPSRC) under grant no. EP/P031587/1. Most of the work was performed during a visit of FS to Imperial College thanks to a joint Imperial-CNRS Fellowship, the hospitality and support of both Imperial College and CNRS are warmly acknowledged. SF was partially supported by the local fund of the University of LAquila "DP-LAND" (Deterministic Particles for Local And Nonlocal Dynamics), and by the Erasmus Mundus programme "MathMods", [www.mathmods.eu](http://www.mathmods.eu). SF thanks the Department of Mathematics at Imperial College in London for hosting him for a visiting period during which this work was carried out.

<sup>†</sup>Department of Mathematics, Imperial College London, SW7 2AZ London, UK. [carrillo@imperial.ac.uk](mailto:carrillo@imperial.ac.uk).

<sup>‡</sup>DISIM - Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila, Via Vetoio 1 (Coppito) 67100 L'Aquila (AQ) - Italy [simone.fagioli@univaq.it](mailto:simone.fagioli@univaq.it).

<sup>§</sup>Laboratoire de Mathématiques d'Orsay, Univ. Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay Cedex, FRANCE [filippo.santambrogio@math.u-psud.fr](mailto:filippo.santambrogio@math.u-psud.fr).

<sup>¶</sup>Department of Mathematics, Imperial College London, SW7 2AZ London, UK. [m.schmidtchen15@imperial.ac.uk](mailto:m.schmidtchen15@imperial.ac.uk).

meaningful way and lead generically to asymptotic stabilisation. In the absence of reactions, the system leads to the nonlinear diffusion equation

$$(1.2) \quad \partial_t \sigma = \partial_x (\sigma \partial_x \chi'(\sigma)) = \partial_x^2 \beta(\sigma).$$

with  $\sigma = \rho + \eta$ , in which  $\chi'(\sigma)$  models the resistance to compression of the whole group of individuals  $\sigma$ . The natural assumption on (1.2) in order to be a diffusion equation is  $\beta'(s) > 0$  for  $s > 0$ , possibly degenerating at  $s = 0$ , or equivalently  $\chi''(s) > 0$  for  $s > 0$ . The particular relevant case of  $\chi(s) = s^2/2$  can be understood as the mean-field limit of interacting particles with very localised repulsion, see [40, 16, 12, 35].

Related reaction-diffusion models to (1.1) 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 bricks 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 modelling factors such as extracellular matrix, enzymatic activators and other substances are 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 (1.1), see for instance [34].

The nonlinear diffusion equation (1.2) is well-studied, see [48], and it can be understood as a gradient flow, see [41, 2, 49, 44], of the energy functional

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

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 wellposedness of solutions to the nonlinear diffusion equation (1.2) was obtained in [41] by means of the so called *JKO-scheme*, cf. [33], which is a particular case of the minimising 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 minimisation 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 Eq. (1.2) and a fixed time step  $0 < \tau < 1$  we set  $\sigma_{\tau}^0 = \bar{\sigma}$ , and then recursively define

$$(1.3) \quad \sigma_{\tau}^{n+1} \in \operatorname{argmin}_{\sigma \in X} \left\{ \frac{1}{2\tau} d^2(\sigma, \sigma_{\tau}^n) + \mathcal{E}(\sigma) \right\},$$

for  $n \in \mathbb{N}$ . The seminal work of R. J. 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  $\beta$  is nondecreasing on  $(0, +\infty)$  or equivalently  $\chi$  convex, called the McCann's condition in short. We also refer to [2, Chap. 9], [24, p. 26], or [50, Chap. 17] for this classical notion, and [13] for related issues. Upon choosing a proper time interpolation  $\sigma_{\tau}$ , it can be proven that the sequence  $\{\sigma_{\tau}\}_{\tau}$  converges to a weak solution of Eq. (1.2), 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 (1.1). 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 (1.2), and then we transport both densities  $\rho, \eta$  through the flow generated by the equation for the total population  $\sigma$ . 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 parameterised 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 details in Section 2. The splitting between reaction and diffusion steps is natural from the numerical analysis view point 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 (1.1). 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 characterising the segregation front through this free boundary. This approach can only work in case the support of both populations are at a positive distance to each other initially. Later [6, 9] combined both the nonlinear diffusion and the reaction to obtain a system similar to (1.1) showing similar segregation phenomena by regularisation techniques. However, their approach heavily relies on the absence of vacuum as they assume that  $\sigma_0$  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 become discontinuous at the contact interface immediately. This phenomenon legitimises our functional space choice as bounded functions of bounded variation, see the precise notion of weak solution and assumptions on the initial data in the next section.

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 their result in [8, 6] about segregation fronts, even in the case of initial vacuum, by showing that initial data which are initially segregated remain segregated for all times. An important technical point in our proof relies on displacement convexity of an auxiliary functional that allows 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 organised 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 the existence theorem as well as 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:

$$(2.1) \quad \begin{cases} \partial_t \rho = \partial_x (\rho \partial_x \chi'(\rho + \eta)) + \rho F_1(\rho, \eta) + \eta G_1(\rho, \eta), & \text{in } [0, T] \times \Omega, \\ \partial_t \eta = \partial_x (\eta \partial_x \chi'(\rho + \eta)) + \eta F_2(\rho, \eta) + \rho G_2(\rho, \eta), & \text{in } [0, T] \times \Omega, \\ \rho(t, x) \partial_x \chi'(\rho + \eta)(t, x) = 0, & \text{on } [0, T] \times \partial\Omega, \\ \eta(t, x) \partial_x \chi'(\rho + \eta)(t, x) = 0, & \text{on } [0, T] \times \partial\Omega, \\ \rho(\cdot, 0) = \rho_0, \quad \eta(\cdot, 0) = \eta_0, & \text{in } \Omega, \end{cases}$$

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

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

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

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 generalisations.

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

$$\mathcal{P}^m(\Omega) := \{\mu \in \mathcal{M}_+(\Omega) \mid \mu(\Omega) = m\},$$

that is, the set of positive measures with mass  $m > 0$ . Consider a measure  $\mu \in \mathcal{P}^m(\Omega)$  and a Borel map  $\mathcal{T} : \mathbb{R} \rightarrow \mathbb{R}$ . We denote by  $\nu = \mathcal{T}_\# \mu \in \mathcal{P}^m(\Omega)$  the push-forward measure of  $\mu$  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  $\Omega$ . We call  $\mathcal{T}$  a *transport map* pushing  $\mu$  to  $\nu$ . 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,  $\gamma$ , between  $\mu_1$  and  $\mu_2$ , 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  $\pi^i$  denotes the projection operator on the  $i$ -th component of the product space. If  $\mu_1$  is absolutely continuous with respect to the Lebesgue measure the optimal transport plan,  $\gamma$ , is unique and can be written as  $\gamma = (\text{id}, \mathcal{T})_{\#} \mu$ . In addition there exists a *Kantorovich potential*,  $\varphi$ , that is linked to the transport map,  $\mathcal{T}$ , in the following way

$$(2.2) \quad \mathcal{T}(x) = (\text{id} - \partial_x \varphi)(x).$$

We refer 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 owing to the fact that it is only defined for two measures of the same mass.

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

$$(2.3) \quad d_{\text{BL}}(\mu, \nu) := \sup \left\{ \int_{\Omega} f d(\mu - \nu) \mid \|f\|_{L^\infty(\Omega)}, \|f'\|_{L^\infty(\Omega)} \leq 1 \right\} = \|\mu - \nu\|_{(W^{1,\infty}(\Omega))^*}.$$

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

$$(2.4) \quad d(U, \tilde{U}) := d(\rho, \tilde{\rho}) + d(\eta, \tilde{\eta}),$$

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

**PROPOSITION 2.3** (Properties of  $d_{\text{BL}}$ ). *Let  $\mu, \nu \in L_+^1(\Omega)$  be two densities. Then the following properties hold true:*

- (i.)  $d_{\text{BL}}(\mu, \nu) \leq \|\mu - \nu\|_{L^1(\Omega)}$ ,
- (ii.)  $d_{\text{BL}}(\mu, \nu) \leq d_1(\mu, \nu)$ , whenever  $\mu(\Omega) = \nu(\Omega)$ .

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

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

Taking the supremum over all such functions  $f$  we get the first statement by using definition (2.3). For the second statement, we additionally assume that  $\mu(\Omega) = \nu(\Omega)$ .

We recall the dual definition of  $d_1$ ,

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

where  $\text{Lip}(\Omega)$  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  $i = 1, 2$  as well as  $\mu_2(\Omega) = \mu_3(\Omega)$  and  $\nu_2(\Omega) = \nu_3(\Omega)$ . Then*

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

Throughout two quantities are crucial for our analysis – the sum of the two densities,  $\sigma$ , and the ratio,  $r$ , between one density, say  $\rho$ , 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 these functions formally satisfy the following system of PDEs

$$(2.5) \quad \begin{cases} \partial_t \sigma = \partial_x (\sigma \partial_x \chi'(\sigma)) + \sigma \left( r(\tilde{F}_1 + \tilde{G}_2) + (1-r)(\tilde{G}_1 + \tilde{F}_2) \right), \\ \partial_t r = \partial_x r \partial_x \chi'(\sigma) + r(1-r) \left( \tilde{F}_1 - \tilde{F}_2 \right) + (1-r)^2 \tilde{G}_1 - r^2 \tilde{G}_2, \end{cases}$$

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 (2.5) can be rewritten in the more compact form

$$(2.6) \quad \begin{cases} \partial_t \sigma = \partial_x (\sigma \partial_x \chi'(\sigma)) + \sigma (rA_1 + (1-r)A_2), \\ \partial_t r = \partial_x r \partial_x \chi'(\sigma) + r(1-r)A_3 + (1-r)^2 \tilde{G}_1 - r^2 \tilde{G}_2. \end{cases}$$

Let us note here, that taking  $F_i = G_i = 0$  in (1.1), yields the following nonlinear cross-diffusion system

$$(2.7) \quad \begin{cases} \partial_t \rho = \partial_x (\rho \partial_x \chi'(\sigma)), \\ \partial_t \eta = \partial_x (\eta \partial_x \chi'(\sigma)), \end{cases}$$

where the sum  $\sigma$  satisfies the nonlinear diffusion equation (1.2).

In order to be useful for our purposes, we need some properties on the internal energy density. The role of these properties will be clearer after the statement of our main result.

**DEFINITION 2.5 (Internal energy density).** *A function  $\chi : [0, \infty] \rightarrow \mathbb{R}$  is an internal energy density if*

$$(NL-i) \quad \chi \in C^0([0, \infty], \mathbb{R}) \cap C^2((0, \infty), \mathbb{R}) \text{ with } \chi'' > 0,$$

(NL-ii)  $\lim_{h \downarrow 0} \chi'(h) = 0$ .

(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 a metric space. In this case the minimiser in Eq. (1.3) satisfies the optimality condition

$$(2.8) \quad \frac{\varphi}{\tau} + \chi'(\sigma^{n+1}) = \text{const.},$$

cf. [44, Proposition 7.20], for instance. Here  $\varphi$  denotes the associated Kantorovich potential transporting  $\sigma^{n+1}$  to  $\sigma^n$  in  $\mathcal{P}^{m_1+m_2}$ , with  $m_1$  and  $m_2$  the masses of  $\rho$  and  $\eta$ , [44, Theorem 1.17]. Notice that this optimality is only obtained in those references for non-degenerate problems, where  $\beta'(0^+) > 0$ . However, similar approximation techniques as those developed in [41, p. 156] and [24, p. 27] allow to overcome the difficulties associated to degenerate diffusions, where  $\beta'(0^+) = 0$ . In the rest of the paper, we will proceed as if we were dealing with nonlinearities leading to non-degenerate 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  $\rho, \eta \in C(0, T; BV(\Omega)) \cap L^\infty(\Omega)$  is a weak solution to system (2.1) if  $\sigma = \rho + \eta \in L^2(0, T; H^1(\Omega))$  and there holds*

$$\begin{aligned} \int_{\Omega} (\rho(t) - \rho(s)) \zeta dx &= \int_s^t \int_{\Omega} -\rho \partial_x \chi'(\rho + \eta) \partial_x \zeta + (\rho F_1(\rho, \eta) + \eta G_1(\rho, \eta)) \zeta dx d\bar{\tau}, \\ \int_{\Omega} (\eta(t) - \eta(s)) \xi dx &= \int_s^t \int_{\Omega} -\eta \partial_x \chi'(\rho + \eta) \partial_x \xi + (\eta F_2(\rho, \eta) + \rho G_2(\rho, \eta)) \xi dx d\bar{\tau}, \end{aligned}$$

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

**REMARK 2.7.**

1. Notice that the functional

$$\mathcal{K}(\sigma) := \int_{\Omega} K(\sigma) dx,$$

associated to  $\chi$ , as in Definition 2.5, satisfies the McCann condition since  $K''(s) = \frac{\chi''(s)}{s} > 0$ , and is therefore displacement convex. This fact will be crucial in Section 3.3 in order to prove Lemma 3.10. Let us just 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  $G_i$ , for  $i = 1, 2$ . The only assumption we need to impose on the kernels is that they are smooth and integrable. These models are found in modelling pattern formation, as 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 (2.7) as a gradient flow of the functional

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

in the product Wasserstein space.

**2.2. Splitting scheme.** We are now ready to introduce our splitting scheme for equation (2.1). 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

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

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  $\rho$  and  $\eta$  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 ordinary differential equations

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

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

$$(2.9a) \quad \rho^{n+1/2} := r\sigma|_{t=(n+1)\tau}, \quad \text{and} \quad \eta^{n+1/2} := (1-r)\sigma|_{t=(n+1)\tau}.$$

A straightforward computation reveals

$$(2.9b) \quad \begin{aligned} \partial_t(r\sigma) &= r\sigma\tilde{F}_1 + (1-r)\sigma\tilde{G}_1, \\ \partial_t((1-r)\sigma) &= (1-r)\sigma\tilde{F}_2 + r\sigma\tilde{G}_2, \end{aligned}$$

formally verifying that  $r\sigma$  and  $(1-r)\sigma$  solve the reaction part of Eq. (2.1). Note that we shall solve the system rigorously in the subsequent section.

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

$$(2.9c) \quad U^{n+1} \in \operatorname{argmin}_{U \in \mathcal{P}^{m_1}(\Omega) \times \mathcal{P}^{m_2}(\Omega)} \left\{ \frac{1}{2\tau} d_2^2(U, U^{n+1/2}) + \mathcal{E}(U) \right\},$$

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

$$(2.9d) \quad \sigma^{n+1} := \rho^{n+1} + \eta^{n+1}, \quad \text{and} \quad r^{n+1} := r^{n+1/2} \circ \mathcal{T}.$$

While the definition of  $\sigma^{n+1}$  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 to Section 3.2, Eq. (3.3).



**2.2.3. Combination of both steps – construction of a solution.** Throughout this paper we refer to Eq. (2.9a) as *reaction step* and to Eq. (2.9c) as *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

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

as well as

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

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  $0 \leq r_0 \leq 1$ . Furthermore, we assume that the nonlinearity  $\chi$  satisfies the assumptions in Definition 2.5 and the reaction terms  $F_i$  and  $G_i$ ,  $i = 1, 2$  are bounded and Lipschitz with respect to  $\rho$  and  $\eta$  with  $G_1(0, \cdot) \geq 0$  and  $G_2(\cdot, 0) \geq 0$ . 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 (2.1) in the sense of Definition 2.6. Moreover, if initially the two densities  $\rho_0, \eta_0$  are segregated, then, in the absence of cross-reactions, the limit densities  $\rho(t, \cdot), \eta(t, \cdot)$  remain segregated for 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 organised as follows: in Section 3.1 and 3.2 we establish the crucial BV-estimates and  $L^\infty$ -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 (2.1).

**3.1. Estimates for 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 ( $L^\infty$  estimates of the reaction step). Let  $(r^n, \sigma^n)$  be given by our splitting scheme. Then there holds

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

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

*Proof.* We show first that there holds  $r^{n+1/2} \in [0, 1]$ . 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  $t^*(x) \in (t^n, t^{n+1})$  such that  $r(t^*, x) = 0$  and  $\partial_t r(t^*, x) < 0$ . However, this is absurd as

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

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

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

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

Since we control the  $L^\infty$ -norm of both  $r$  and  $\sigma$  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  $(r^n, \sigma^n)$  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  $L^\infty$ -bounds on  $F_i, G_i$ , for  $i = 1, 2$ .

*Proof.* Using the transformed system, Eqs.(2.6),  $r$  and  $\sigma$  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

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

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

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

whence

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

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

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

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

**3.2. Estimates for diffusive step.** This section is devoted to establishing BV-estimates and  $L^\infty$ -bounds in the diffusive step. To this end we will make use of the following lemma.

LEMMA 3.3 (Same optimality conditions).  $(\rho^{n+1}, \eta^{n+1})$  be given by the JKO step for  $\mathcal{E}$ , cf. (2.9c). 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  $\mathcal{T}$  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  $\operatorname{supp}(\rho^{n+1}) \cup \operatorname{supp}(\eta^{n+1})$  for the associated optimal transport maps. Moreover the optimality conditions read

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

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

*Proof.* Let

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

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 following form

$$\begin{aligned} \chi'(\rho^{n+1} + \eta^{n+1}) + \frac{\varphi_\rho}{\tau} &= c_\rho, & \text{on } \operatorname{supp}(\rho^{n+1}), \\ &\geq c_\rho, & \text{elsewhere,} \end{aligned}$$

for  $\rho$  and

$$\begin{aligned} \chi'(\rho^{n+1} + \eta^{n+1}) + \frac{\varphi_\eta}{\tau} &= c_\eta, & \text{on } \operatorname{supp}(\eta^{n+1}), \\ &\geq c_\eta, & \text{elsewhere,} \end{aligned}$$

for  $\eta$ , respectively. Thus the optimal maps are given by

$$\begin{aligned} \mathcal{T}_\rho &= \operatorname{id} + \tau \partial_x \chi'(\rho^{n+1} + \eta^{n+1}), & \text{on } \operatorname{supp}(\rho^{n+1}), \\ \mathcal{T}_\eta &= \operatorname{id} + \tau \partial_x \chi'(\rho^{n+1} + \eta^{n+1}), & \text{on } \operatorname{supp}(\eta^{n+1}). \end{aligned}$$

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

$$(3.1) \quad \min \left\{ \frac{1}{2\tau} d_2^2(\sigma, \sigma^{n+1/2}) + \mathcal{E}(\sigma) \right\}.$$

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

$$\begin{aligned}
& \frac{1}{2\tau} d_2^2(U^{n+1}, U^{n+1/2}) + \mathcal{E}(U^{n+1}) \\
&= \frac{1}{2\tau} \left( \int_{\Omega} |x - \mathcal{T}_{\rho}(x)|^2 \rho^{n+1}(x) \, dx + \int_{\Omega} |x - \mathcal{T}_{\eta}(x)|^2 \eta^{n+1}(x) \, dx \right) + \mathcal{E}(U^{n+1}) \\
&= \frac{1}{2\tau} \left( \int_{\text{supp}(\rho^{n+1})} |x - \mathcal{T}_{\rho}(x)|^2 \rho^{n+1}(x) \, dx + \int_{\text{supp}(\eta^{n+1})} |x - \mathcal{T}_{\eta}(x)|^2 \eta^{n+1}(x) \, dx \right) + \mathcal{E}(U^{n+1}) \\
&= \frac{1}{2\tau} \left( \int_{\text{supp}(\rho^{n+1}) \setminus (\text{supp}(\rho^{n+1}) \cap \text{supp}(\eta^{n+1}))} |x - \mathcal{T}_{\rho}(x)|^2 \rho^{n+1}(x) \, dx \right. \\
&\quad + \int_{\text{supp}(\eta^{n+1}) \setminus (\text{supp}(\rho^{n+1}) \cap \text{supp}(\eta^{n+1}))} |x - \mathcal{T}_{\eta}(x)|^2 \eta^{n+1}(x) \, dx \\
&\quad \left. + \int_{(\text{supp}(\rho^{n+1}) \cap \text{supp}(\eta^{n+1}))} |x - \mathcal{T}(x)|^2 \sigma^{n+1}(x) \, dx \right) + \mathcal{E}(U^{n+1}) \\
&= \frac{1}{2\tau} \left( \int_{\text{supp}(\rho^{n+1}) \setminus (\text{supp}(\rho^{n+1}) \cap \text{supp}(\eta^{n+1}))} |x - \mathcal{T}_{\rho}(x)|^2 \sigma^{n+1}(x) \, dx \right. \\
&\quad + \int_{\text{supp}(\eta^{n+1}) \setminus (\text{supp}(\rho^{n+1}) \cap \text{supp}(\eta^{n+1}))} |x - \mathcal{T}_{\eta}(x)|^2 \sigma^{n+1}(x) \, dx \\
&\quad \left. + \int_{(\text{supp}(\rho^{n+1}) \cap \text{supp}(\eta^{n+1}))} |x - \mathcal{T}(x)|^2 \sigma^{n+1}(x) \, dx \right) + \mathcal{E}(U^{n+1}) \\
&\geq \frac{1}{2\tau} d_2^2(\sigma^{n+1}, \sigma^{n+1/2}) + \mathcal{E}(\sigma^{n+1}) \\
&> \frac{1}{2\tau} d_2^2(\sigma^*, \sigma^{n+1/2}) + \mathcal{E}(\sigma^*) \\
&= \frac{1}{2\tau} \int_{\Omega} |x - \mathcal{S}(x)|^2 \sigma^*(x) \, dx + \mathcal{E}(\sigma^*),
\end{aligned}$$

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

$$\begin{aligned}
& \frac{1}{2\tau} d_2^2(U^{n+1}, U^{n+1/2}) + \mathcal{E}(U^{n+1}) \\
&> \frac{1}{2\tau} \int_{\Omega} |x - \mathcal{S}(x)|^2 \rho^*(x) \, dx + \frac{1}{2\tau} \int_{\Omega} |x - \mathcal{S}(x)|^2 \eta^*(x) \, dx + \mathcal{E}(\rho^* + \eta^*) \\
&\geq \frac{1}{2\tau} d_2^2(U^*, U^{n+1/2}) + \mathcal{E}(U^*),
\end{aligned}$$

which is absurd for it implies  $U^{n+1}$  is not a minimiser. Hence we conclude that  $\sigma^{n+1}$  is a minimiser of problem (3.1) and, as a byproduct,  $\mathcal{T}_{\rho} \equiv \mathcal{T}_{\eta}$  on  $\text{supp}(\rho^{n+1}) \cup \text{supp}(\eta^{n+1}) = \text{supp}(\sigma^{n+1})$ .

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

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

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  $\sigma^*$ , which concludes the proof.  $\square$

**PROPOSITION 3.4** ( $L^\infty$  stability of the diffusive step). *Let  $(r^{n+1/2}, \sigma^{n+1/2})$  be given by the splitting scheme (2.9a). Then these quantities satisfy*

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

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

*Proof.* Choose  $x_0 \in \text{argmax}(\sigma^{n+1})$ . Then, by the optimality condition, Eq. (2.8),

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

we have  $x_0 \in \text{argmin}(\varphi)$  where we used the fact that  $\chi'' \geq 0$ , cf. Definition 2.5, (NL – iii). Hence  $\varphi''(x_0) \geq 0$  and consequently, by passing to the derivative in Eq. (2.2) we get

$$(3.2) \quad \mathcal{T}'(x_0) = 1 - \partial_{xx}\varphi(x_0) \leq 1.$$

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

$$\begin{aligned} \sigma^{n+1}(x) &\leq \|\sigma^{n+1}\|_{L^\infty} = \sigma^{n+1}(x_0) = \mathcal{T}'(x_0)\sigma^{n+1/2}(\mathcal{T}(x_0)) \stackrel{(3.2)}{\leq} \sigma^{n+1/2}(\mathcal{T}(x_0)) \\ &\leq \|\sigma^{n+1/2}\|_{L^\infty}, \end{aligned}$$

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

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

Finally the bounds for  $r^{n+1}$  follow from its definition, cf. Eq. (2.9d), 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  $\sigma^{n+1}$ ). *Let  $(r^{n+1/2}, \sigma^{n+1/2})$  be given. After the diffusive step they satisfy the following estimate.*

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

*Proof.* The result for the BV-norm of the minimiser,  $\sigma^{n+1}$ , 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. Eq. (2.9d), 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  $\Omega$ . 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  $\text{supp}(\sigma^{n+1})$  as we shall see now. According to Lemma 3.3, the same transport map  $\mathcal{T}$  pushes  $\rho^{n+1}$  onto  $\rho^{n+1/2}$  and  $\sigma^{n+1}$  onto  $\sigma^{n+1/2}$ . As a consequence the densities satisfy

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

whence

$$(3.3) \quad r^{n+1}(x) := r^{n+1/2} \circ \mathcal{T}(x) = \frac{\rho^{n+1/2}}{\sigma^{n+1/2}} \circ \mathcal{T}(x) = \frac{\rho^{n+1}}{\sigma^{n+1}}(x). \quad \square$$

**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  $L^\infty$ -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*

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

and

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

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

*Proof.* The uniform  $L^\infty$ -bound is a consequence of combining Propositions 3.1 and 3.4. We use these uniform  $L^\infty$ -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 Lemma 3.6 and Lemma 2.2 we obtain the following corollary.

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

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

where the constant  $C > 0$  is independent of the parameter  $\tau$

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.* that it is decayed along solutions. Since we are interested in a uniform estimate we shall 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 straight forward computation yields

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

using the uniform  $L^\infty$ -bound on  $\sigma$  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  $n$ . □

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

LEMMA 3.9 ( $H^1$ -bound for  $\sigma^{n+1}$ ). *The minimiser of the JKO step satisfies the following 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

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

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

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

and

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

for the associated Kantorovich potential,  $\varphi$ , *cf.* Eq. (2.2). As a consequence the velocity field is given by

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

satisfying the following continuity equation,

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

We differentiate the entropy along the geodesic and obtain

$$\begin{aligned} \frac{d}{ds} \int_{\Omega} K(\sigma_s) dx &= \int_{\Omega} \kappa(\sigma_s) \partial_s \sigma_s dx = - \int_{\Omega} \kappa'(\sigma_s) \partial_x \sigma_s \sigma_s v_s dx \\ &= - \int_{\Omega} \chi''(\sigma_s) \partial_x \sigma_s v_s dx = - \int_{\Omega} \partial_x \chi'(\sigma_s) v_s dx. \end{aligned}$$

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

$$\frac{d}{ds} \int_{\Omega} K(\sigma_s) dx \Big|_{s=0} = \int_{\Omega} \partial_x \chi'(\sigma^{n+1}) \partial_x \varphi dx.$$

Using the optimality condition, Eq. (2.8), we obtain

$$\begin{aligned} \tau \int_{\Omega} |\partial_x \chi'(\sigma^{n+1})|^2 dx &= - \int_{\Omega} \partial_x \chi'(\sigma^{n+1}) \partial_x \varphi dx \\ &= \frac{d}{ds} \int_{\Omega} K(\sigma_s) dx \Big|_{s=0} \\ &\leq \mathcal{K}(\sigma^{n+1/2}) - \mathcal{K}(\sigma^{n+1}), \end{aligned}$$

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

We combine the previous lemmas to obtain the desired estimate for a full iteration and finally for the piecewise constant interpolation,  $\sigma_{\tau}$ .

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 Lemma 3.8 and Lemma 3.9 to get

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

Summing over  $n = 0 \dots N - 1$  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.  $\square$

LEMMA 3.11 (Total-square 2-Wasserstein distance estimates). *For every  $n \in \{0, \dots, N\}$  consider two consecutive steps for (2.9c),  $U^{n+1/2} = (\rho^{n+1/2}, \eta^{n+1/2})$  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 minimising 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, and considering that

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

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

$$(3.4) \quad \mathcal{E}(U^{n+1/2}) - \mathcal{E}(U^n) \leq \|\chi'\|_{L^\infty} \int_{\Omega} \int_{t^n}^{t^{n+1}} |\sigma(rA_1 + (1-r)A_2)| \, ds \, dx \leq C\tau,$$

where, in the last inequality, we used the  $L^\infty$ -bounds on  $\sigma, r$  as well as on  $F_i, G_i$ . Hence, we may summarise

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

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

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

### 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  $d_{\text{BL}}$  with values in  $\mathcal{M}_+(\Omega)$ . Moreover,  $\bar{\rho}$  and  $\bar{\eta}$  are  $d_{\text{BL}}$ -continuous functions on  $[0, T]$ .*

*Proof.* The proof is based on the application of a generalised version of the Ascoli-Arzelà theorem, cf. Ref. [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

$$(3.5) \quad d_{\text{BL}}(U_\tau(s), U_\tau(t)) \leq \sum_{n=m}^{k-1} d_{\text{BL}}(U^n, U^{n+1}) \leq \left( \sum_{n=m}^{k-1} d_{\text{BL}}^2(U^n, U^{n+1}) \right)^{\frac{1}{2}} |k-m|^{\frac{1}{2}},$$

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

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

For the reaction step an argument similar to Eq. (3.4) yields

$$\begin{aligned} \|U^n - U^{n+1/2}\|_{L^1(\Omega)}^2 &= \left| \int_{\Omega} |\rho^{n+1/2} - \rho^n| + |\eta^{n+1/2} - \eta^n| \right|^2 \\ &= \left| \int_{\Omega} \left| \int_{t^n}^{(n+1)\tau} r\sigma\tilde{F}_1 + (1-r)\sigma\tilde{G}_1 \, ds \right| \right. \\ &\quad \left. + \left| \int_{t^n}^{(n+1)\tau} (1-r)\sigma\tilde{F}_2 + r\sigma\tilde{G}_2 \, ds \right| \, dx \right|^2 \\ &\leq C\tau^2, \end{aligned}$$

by Eq. (2.9b), where  $C$  depends on the  $L^\infty$ -bounds of  $r, \sigma$  as well as  $F_i, G_i$ , for  $i = 1, 2$ . Using that the  $p$ -Wasserstein distances are ordered, we even have

$$\begin{aligned} \sum_{n=m}^{k-1} d_{\text{BL}}^2(U^n, U^{n+1}) &\leq 2 \sum_{n=0}^N \|U^n - U^{n+1/2}\|_{L^1(\Omega)}^2 + 2 \sum_{n=0}^N d_2^2(U^{n+1/2}, U^{n+1}) \\ &\leq C\tau, \end{aligned}$$

where we used the total-square estimate, Lemma 3.11. Therefore Eq. (3.5) becomes

$$\begin{aligned} d_{\text{BL}}(U_\tau(s), U_\tau(t)) &\leq \left( \sum_{n=m}^{k-1} d_{\text{BL}}^2(U^n, U^{n+1}) \right)^{\frac{1}{2}} |k-m|^{\frac{1}{2}} \\ &\leq C\sqrt{\tau} \left( \frac{|t-s|}{\tau} + 1 \right)^{\frac{1}{2}} \\ &\leq C(\sqrt{|t-s|} + \sqrt{\tau}). \end{aligned}$$

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 Ascoli-Arzelà’s theorem, cf. [2], Section 3.  $\square$

**COROLLARY 3.13** (Strong convergence in  $L^p(0, T; L^q(\Omega))$ ). *Let  $1 \leq p, q < \infty$  and  $(\rho_\tau)_{\tau>0}$  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*

$$\rho_\tau \rightarrow \rho, \quad \text{and} \quad \eta_\tau \rightarrow \eta,$$

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

$$\rho_\tau(t) \rightarrow \rho(t), \quad \text{and} \quad \eta_\tau(t) \rightarrow \eta(t),$$

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  $L^\infty$ -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

$$\rho_\tau \rightarrow \rho,$$

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

LEMMA 3.14 (Identification of the limit). *The sequence constructed in (2.9a), (2.9c) converges to a weak solution of Eqs. (1.1).*

*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  $\rho$  as the corresponding result for  $\eta$  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 (2.9c),

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

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

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

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

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

Moreover, using the fact that  $\mathcal{T}(x) = x - \partial_x \varphi(x)$ , where  $\varphi$  is the Kantorovich potential

associated to the optimal map  $\mathcal{T}$ , the integral above can be rewritten as

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

Thanks to the optimality condition, Lemma 3.3, we get

$$(3.6) \quad \begin{aligned} & \frac{1}{\tau} \int_{\Omega} \left( \rho^{n+1} - \rho^{n+1/2} \right) \zeta dx \\ &= - \int_{\Omega} \rho^{n+1} \partial_x \chi'(\sigma^{n+1}) \partial_x \zeta dx + \mathcal{O} \left( d_2^2(\rho^{n+1}, \rho^{n+1/2}) \right). \end{aligned}$$

**Reaction part.** Note that

$$(3.7) \quad \begin{aligned} \int_{\Omega} \frac{\rho^{n+1/2} - \rho^n}{\tau} \zeta dx &= \int_{\Omega} \frac{r^{n+1/2} \sigma^{n+1/2} - r^n \sigma^n}{\tau} \zeta dx \\ &= \int_{\Omega} \int_{t^n}^{t^{n+1}} \frac{1}{\tau} \partial_t (r(\bar{\tau}) \sigma(\bar{\tau})) \zeta d\bar{\tau} dx, \\ &= \int_{\Omega} \left( r^{n+1/2} \sigma^{n+1/2} \tilde{F}_1 + (1 - r^{n+1/2}) \sigma^{n+1/2} \tilde{G}_1 \right) \zeta dx + \mathcal{O}(\tau) \\ &= \int_{\Omega} \left( \rho^{n+1/2} F_1^{n+1/2} + \eta^{n+1/2} G_1^{n+1/2} \right) \zeta dx + \mathcal{O}(\tau), \end{aligned}$$

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 Eq. (2.9b).

**Combination of both steps.** Let us combine the reaction step and the diffusion step of the splitting scheme. Upon summing up Eqs.(3.6, 3.7) we obtain

$$(3.8) \quad \begin{aligned} 0 &= \frac{1}{\tau} \int_{\Omega} \zeta(x) (\rho^{n+1}(x) - \rho^n(x)) dx + \mathcal{O}(d_2^2(\rho^{n+1}, \rho^{n+1/2})) \\ &+ \int_{\Omega} \rho^{n+1} \partial_x \chi'(\rho^{n+1}(x) + \eta^{n+1}(x)) \partial_x \zeta(x) dx \\ &- \int_{\Omega} \left( \rho^{n+1/2} F_1(\rho^{n+1/2}, \eta^{n+1/2}) + \eta^{n+1/2} G_1(\rho^{n+1/2}, \eta^{n+1/2}) \right) \zeta dx \\ &+ \mathcal{O}(\tau). \end{aligned}$$

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 Eq. (3.8) by  $\tau$  and summing from  $m$  to  $k-1$ , we obtain

$$\begin{aligned} 0 &= \int_{\Omega} \zeta(x)(\bar{\rho}_{\tau}(t, x) - \bar{\rho}_{\tau}(s, x)) dx + \mathcal{O}(\tau) \\ &\quad + \int_s^t \int_{\Omega} \bar{\rho}_{\tau}(\bar{\tau}, x) \partial_x \chi'(\bar{\rho}_{\tau}(\bar{\tau}, x) + \bar{\eta}_{\tau}(\bar{\tau}, x)) \partial_x \zeta(x) \\ &\quad - \left( \bar{\rho}_{\tau}(\bar{\tau}, x) F_1(\bar{\rho}_{\tau}(\bar{\tau}, x), \bar{\eta}_{\tau}(\bar{\tau}, x)) + \bar{\eta}_{\tau}(\bar{\tau}, x) G_1(\bar{\rho}_{\tau}(\bar{\tau}, x), \bar{\eta}_{\tau}(\bar{\tau}, x)) \right) \zeta(x) dx d\bar{\tau}. \end{aligned}$$

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  $\tau \rightarrow 0$  we obtain the weak formulation for  $\rho$ . The same argument for  $\eta$  yields the statement.  $\square$

We end the paper with a stunning result which can be seen as a generalisation of the result of Bertsch *et al.*, cf. [8, 7]. In their 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 imposing a rather restrictive assumption on the initial datum. We can drop their assumption that

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

*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.*

$$\int_{\Omega} \rho_0(x) \eta_0(x) dx = 0.$$

*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  $\rho_{\tau}$  and  $\eta_{\tau}$  to show segregation is also kept in the limit:

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

Thus let us now show the property at the level of the approximation. During the reaction step the segregation of  $\rho$  and  $\eta$  is kept as can be seen as follows. First recall that in the reaction step we solve Eq. (2.9a), with initial data  $\sigma^n, r^n$  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  $\sigma^n$ . Thus we conclude that already

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

on  $\text{supp}(\sigma^n)$ . 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  $\text{supp}(\sigma^{n+1/2}) \subset \text{supp}(\sigma^n)$ . This is due to the fact that the right-hand side of the equation for  $\sigma$  is premultiplied by  $\sigma$

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

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  $\text{supp}(\sigma^{n+1/2}) \subset \text{supp}(\sigma^n)$ .

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  $1 \leq n \leq N$  such that

$$(3.9) \quad \int_{\Omega} \rho^{n+1/2}(x) \eta^{n+1/2}(x) dx = 0,$$

while

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

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 the same transport map  $\mathcal{T}$  in common there exists a set  $A$  such that  $A = \mathcal{T}(B)$  and

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

which is absurd, for we assumed (3.9). Thus segregation is kept at each iteration which concludes the proof of the theorem.  $\square$

#### REFERENCES

- [1] L. Ambrosio. Minimizing movements. *Rend. Accad. Naz. Sci. XL Mem. Mat. Appl.* (5), 19:191–246, 1995.
- [2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [3] N. Armstrong, K. Painter, and J. Sherratt. A continuum approach to modelling cellcell adhesion. *Journal of Theoretical Biology*, 243(1):98–113, 2006.
- [4] N. Armstrong, K. Painter, and J. Sherratt. Adding adhesion to a chemical signaling model for somite formation. *Bulletin of Mathematical Biology*, 71(1):1–24, 2009.
- [5] N. Armstrong, K. Painter, and J. Sherratt. The impact of adhesion on cellular invasion processes in cancer and development. *Journal of Theoretical Biology*, 264(3):1057–1067, 2010.
- [6] M. Bertsch, R. Dal Passo, and M. Mimura. A free boundary problem arising in a simplified tumour growth model of contact inhibition. *Interfaces and Free Boundaries*, 12(2):235–250, 2010.
- [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. *Journal of differential equations*, 67(1):56–89, 1987.
- [8] M. Bertsch, M. Gurtin, D. Hilhorst, and L. Peletier. On interacting populations that disperse to avoid crowding: preservation of segregation. *Journal of mathematical biology*, 23(1):1–13, 1985.
- [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(1):137–157, 2012.

- [10] M. Bessemoulin-Chatard and F. Filbet. A finite volume scheme for nonlinear degenerate parabolic equations. *SIAM J. Sci. Comput.*, 34(5):B559–B583, 2012.
- [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(2):691–721, 2008.
- [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(10):1206–1228, 2013.
- [13] F. Bolley and J. A. Carrillo. Nonlinear diffusion: geodesic convexity is equivalent to Wasserstein contraction. *Comm. Partial Differential Equations*, 39(10):1860–1869, 2014.
- [14] Y. Brenier. Décomposition polaire et réarrangement monotone des champs de vecteurs. *CR Acad. Sci. Paris Sér. I Math*, 305(19):805–808, 1987.
- [15] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematics*, 44(4):375–417, 1991.
- [16] M. Burger, V. Capasso, and D. Morale. On an aggregation model with long and short range interactions. *Nonlinear Anal. Real World Appl.*, 8(3):939–958, 2007.
- [17] M. Burger, M. Di Francesco, S. Fagioli, and A. Stevens. Sorting phenomena in a mathematical model for two mutually attracting/repelling species. *preprint arXiv:1704.04179*.
- [18] J. A. Cañizo, J. A. Carrillo, and S. Cuadrado. Measure solutions for some models in population dynamics. *Acta Appl. Math.*, 123:141–156, 2013.
- [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(3):515–539, 2011.
- [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(2):155–175, 2006.
- [21] J. A. Carrillo, A. Chertock, and Y. Huang. A finite-volume method for nonlinear nonlocal equations with a gradient flow structure. *Communications in Computational Physics*, 17(01):233–258, 2015.
- [22] J. A. Carrillo, A. Colombi, and M. Scianna. Adhesion and volume constraints via nonlocal interactions lead to cell sorting. *preprint arXiv:1706.08969*.
- [23] J. A. Carrillo, Y. Huang, and M. Schmidtchen. Zoology of a non-local cross-diffusion model for two species. 2017.
- [24] J. A. Carrillo, A. Jiingel, P. A. Markowich, G. Toscani, and A. Unterreiter. Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities. *Monatsh. Math.*, 133(1):1–82, 2001.
- [25] J. A. Carrillo, N. Kolbe, and M. Lukáčová-Medvid’ová. A hybrid mass transport finite element method for keller-segel type systems. *preprint arXiv:1709.07394*.
- [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. Mod. Meth. Appl. S.*, 15(11):1685–1734, 2005.
- [27] G. De Philippis, A. R. Mészáros, F. Santambrogio, and B. Velichkov. BV estimates in optimal transportation and applications. *Archive for Rational Mechanics and Analysis*, 219(2):829–860, 2016.
- [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. *Journal of Theoretical Biology*, 361:41–60, 2014.
- [29] A. Gerisch and M. Chaplain. Mathematical modelling of cancer cell invasion of tissue: Local and non-local models and the effect of adhesion. *Journal of Theoretical Biology*, 250(4):684–704, 2008.
- [30] M. E. Gurtin and A. Pipkin. A note on interacting populations that disperse to avoid crowding. *Quarterly of Applied Mathematics*, pages 87–94, 1984.
- [31] P. Gwiazda and A. Marciniak-Czochra. Structured population equations in metric spaces. *J. Hyperbolic Differ. Equ.*, 7(4):733–773, 2010.
- [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. Theor. Biol.*, 266(4):708–711, 2010.
- [33] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker–Planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.
- [34] I. Kim and A. R. Mészáros. On nonlinear cross-diffusion systems: an optimal transport approach. *preprint arXiv:1705.02457*.
- [35] A. Klar and S. Tiwari. A multiscale meshfree method for macroscopic approximations of interacting particle systems. *Multiscale Model. Simul.*, 12(3):1167–1192, 2014.
- [36] S. Kondo. An updated kernel-based Turing model for studying the mechanisms of biological pattern formation. *Journal of Theoretical Biology*, 414:120–127, Feb. 2017.
- [37] R. J. McCann. A convexity principle for interacting gases. *Advances in mathematics*,

- 128(1):153–179, 1997.
- [38] H. Murakawa and H. Togashi. Continuous models for cell–cell adhesion. *Journal of theoretical biology*, 374:1–12, 2015.
  - [39] H. Neunzert. An introduction to the nonlinear Boltzmann-Vlasov equation. In *Kinetic theories and the Boltzmann equation (Montecatini, 1981)*, volume 1048 of *Lecture Notes in Math.*, pages 60–110. Springer, Berlin, 1984.
  - [40] K. Oelschläger. Large systems of interacting particles and the porous medium equation. *J. Differential Equations*, 88(2):294–346, 1990.
  - [41] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.
  - [42] K. J. Painter and T. Hillen. Volume-filling and quorum-sensing in models for chemosensitive movement. *Can. Appl. Math. Q.*, 10(4):501–543, 2002.
  - [43] L. Preziosi. *Cancer modelling and simulation*. CRC Press, 2003.
  - [44] F. Santambrogio. *Optimal Transport for Applied Mathematicians*, volume 86 of *Progress in Nonlinear Differential Equations and Their Applications*. 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. *Bulletin of Mathematical Biology*, 79(1):209–235, 2017.
  - [46] H. Spohn. Large scale dynamics of interacting particles. Texts and Monographs in Physics. Springer, 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(11):2163–2201, 2016.
  - [48] J. L. Vázquez. *The porous medium equation*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, Oxford, 2007. Mathematical theory.
  - [49] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
  - [50] C. Villani. *Optimal transport : old and new*. Grundlehren der mathematischen Wissenschaften. Springer, Berlin, 2009.
  - [51] A. Volkening and B. Sandstede. Modelling stripe formation in zebrafish: an agent-based approach. *Journal of The Royal Society Interface*, 12(112), 2015.