ESAIM: COCV 25 (2019) 8
https://doi.org/10.1051/cocv/2018001

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

AN UNBALANCED OPTIMAL TRANSPORT SPLITTING SCHEME FOR GENERAL ADVECTION-REACTION-DIFFUSION PROBLEMS

Thomas Gallouët[1,2], Maxime Laborde[3]
and Léonard Monsaingeon[4,5,*]

Abstract. In this paper, we show that unbalanced optimal transport provides a convenient framework to handle reaction and diffusion processes in a unified metric setting. We use a constructive method, alternating minimizing movements for the Wasserstein distance and for the Fisher-Rao distance, and prove existence of weak solutions for general scalar reaction-diffusion-advection equations. We extend the approach to systems of multiple interacting species, and also consider an application to a very degenerate diffusion problem involving a Gamma-limit. Moreover, some numerical simulations are included.

Mathematics Subject Classification. 35K15, 35K57, 35K65, 47J30

Received June 27, 2017. Accepted December 21, 2017.

1. Introduction

Since the seminal works of Jordan-Kinderlehrer-Otto [20], it is well known that certain diffusion equations can be interpreted as gradient flows in the space of probability measures, endowed with the quadratic Wasserstein distance $\mathtt{W}$. The well-known JKO scheme (a.k.a. minimizing movement), which is a natural implicit Euler scheme for such gradient flows, naturally leads to constructive proofs of existence for weak solutions to equations or systems with mass conservation such as, for instance, Fokker-Planck equations [20], Porous Media Equations [34], aggregation equation [9], double degenerate diffusion equations [32], general degenerate parabolic equation [1], etc. We refer to the classical textbooks of Ambrosio, Gigli and Savaré [4] and to Villani's books [45, 46] for a detailed account of the theory and extended bibliography. Recently, this theory has been extended to study the evolution of interacting species with mass-conservation, see for examples [8, 15, 21, 24, 47].

Nevertheless in biology, for example for diffusive prey-predator models, the conservation of mass may not hold, and the classical optimal transport theory does not apply. An unbalanced optimal transport theory was recently introduced simultaneously in [11, 12, 22, 26, 27], and the resulting Wasserstein-Fisher-Rao ($\mathtt{WFR}$) metrics (also referred to as the Hellinger-Kantorovich distance $\mathtt{HK}$) allows to compute distances between measures with

Keywords and phrases: Unbalanced optimal transport, Wasserstein-Fisher-Rao, Hellinger-Kantorovich, JKO scheme, reaction-diffusion-advection equations.

[1] Mathematics Department, Université de Liége, Liége, Belgium.
[2] INRIA Paris MOKAPLAN, Paris, France.
[3] McGill University, Montreal, Canada.
[4] IECL Université de Lorraine, Nancy, France.
[5] GFM Universidade de Lisboa, Lisboa, Portugal.

[*] Corresponding author: leonard.monsaingeon@gmail.com

variable masses while retaining a convenient Riemannian structure. See Section 2 for the definition and a short discussions on this WFR metric. We also refer to [16, 39] for earlier attempts to account for mass variations within the framework of optimal transport.

The WFR metrics can be seen as an *inf-convolution* between Wasserstein/transport and Fisher-Rao/reaction processes, and is therefore extremely convenient to control both in a unified metric setting. This allows to deal with non-conservative models of population dynamics, see *e.g.* [22, 23]. In [18], the first and third authors proposed a variant of the JKO scheme for WFR-gradient flows corresponding to some particular class of reaction-diffusion PDEs: roughly speaking, the reaction and diffusion were handled separately in the FR, W metrics, respectively, and then patched together using a particular uncoupling of the inf-convolution, namely $\mathtt{WFR}^2 \approx \mathtt{W}^2 + \mathtt{FR}^2$ in some sense (see [18], Sect. 3 for a thorough discussion). However, the analysis was restricted to very particular structures for the PDE, corresponding to pure WFR gradient-flows.

In this work we aim at extending this splitting scheme in order to handle more general reaction-diffusion problems, not necessarily corresponding to gradient flows. Roughly speaking, the structure of our splitting scheme is the following: the transport./diffusion part of the PDE is treated by a single Wasserstein JKO step

$$\rho^k \xrightarrow[\text{transport}]{\mathtt{W}} \rho^{k+1/2},$$

and the next Fisher-Rao JKO step

$$\rho^{k+1/2} \xrightarrow[\text{reaction}]{\mathtt{FR}} \rho^{k+1}$$

handles the reaction part of the evolution. As already mentioned, the WFR metric will allow to suitably control both steps in a unified metric framework. We will first state a general convergence result for scalar reaction-diffusion equations, and then illustrate on a few particular examples how the general idea can be adapted to treat *e.g.* prey-predator systems or very degenerate Hele-Shaw diffusion problems. In this work we do not focus on optimal results and do not seek full generality, but rather wish to illustrate the efficiency of the general approach.

Another advantage of our splitting scheme is that it is well adapted to existing Monge/Kantorovich/Wasserstein numerical solvers, and the Fisher-Rao step turns out to be a simple pointwise convex problem which can be implemented in a very simple way. See also [10, 13] for a more direct numerical approach by entropic regularization. Throughout the paper we will illustrate the theoretical results with a few numerical tests. All the simulations were implemented with the augmented Lagrangian ALG2-JKO scheme from [6] for the Wasserstein step, and we used a classical Newton algorithm for the Fisher-Rao step.

The paper is organized as follows. In Section 2 we recall the basic definitions and useful properties of the Wasserstein-Fisher-Rao distance WFR. Section 3 contains the precise description of the splitting scheme and a detailed convergence analysis for a broad class of reaction-diffusion equations. In Section 4 we present an extension to some prey-predator multicomponent systems with nonlocal interactions. In Section 5 we extend the general result from Section 3 to a very degenerate tumor growth model studied in [36], corresponding to a pure WFR gradient flow: we show that the splitting scheme captures fine properties of the model, particularly the Γ-convergence of discrete gradient flows as the degenerate diffusion exponent of Porous Medium type $m \to \infty$ (this limit is often referred to as the mesa problem, or the incompressible limit). Section 6 contains an extension to a tumor-growth model coupled with an evolution equation for the nutrients.

2. Preliminaries

Let us first fix some notations. Throughout the whole paper, Ω denotes a possibly unbounded convex subset of $\mathbb{R}^d$, Q_T represents the product space $[0,T] \times \Omega$, for $T > 0$, and we write $\mathcal{M}^+ = \mathcal{M}^+(\Omega)$ for the set of nonnegative finite Radon measures on Ω. We say that a curve of measures $t \mapsto \rho_t \in \mathcal{C}_w([0,1]; \mathcal{M}^+)$ is narrowly continuous if it is continuous with respect to the narrow convergence of measures, namely for the duality with $\mathcal{C}_b(\Omega)$ test-functions.

Definition 2.1. The Fisher-Rao distance between $\rho_0, \rho_1 \in \mathcal{M}^+$ is

$$\mathtt{FR}(\rho_0,\rho_1) := \min_{(\rho_t,r_t)\in\mathcal{A}_{\mathtt{FR}}[\rho_0,\rho_1]} \int_0^1 \int_\Omega |r_t|^2 \,\mathrm{d}\rho_t(x)\mathrm{d}t,$$

where the admissible set $A_{\mathtt{FR}}[\rho_0,\rho_1]$ consists in curves $[0,1] \ni t \mapsto (\rho_t, r_t) \in \mathcal{M}^+ \times \mathcal{M}$ such that $t \mapsto \rho_t$ is narrowly continuous with endpoints ρ_0, ρ_1, and

$$\partial_t \rho_t = \rho_t r_t$$

in the sense of distributions $\mathcal{D}'((0,1)\times\Omega)$.

The Monge-Kantorovich-Wasserstein admits several equivalent definitions and formulations, and we refer *e.g.* to [4, 43, 45, 46] for a complete description. For our purpose we shall only need the dynamical Benamou-Brenier formula:

Theorem 2.2 (Benamou-Brenier formula [4, 5])**.** *There holds*

$$\mathtt{W}^2(\rho_0,\rho_1) = \min_{(\rho,\mathbf{v})\in\mathcal{A}_{\mathtt{W}}[\rho_0,\rho_1]} \int_0^1 \int_\Omega |\mathbf{v}_t|^2 \mathrm{d}\rho_t \mathrm{d}t, \tag{2.1}$$

where the admissible set $\mathcal{A}_{\mathtt{W}}[\rho_0,\rho_1]$ consists in curves $(0,1) \ni t \mapsto (\rho_t, \mathbf{v}_t) \in \mathcal{M}^+ \times \mathcal{M}(\Omega;\mathbb{R}^d)$ such that $t \mapsto \rho_t$ is narrowly continuous with endpoints ρ_0, ρ_1 and solving the continuity equation

$$\partial_t \rho_t + \operatorname{div}(\rho_t \mathbf{v}_t) = 0$$

in the sense of distributions $\mathcal{D}'((0,1)\times\Omega)$.

According to the original definition in [11] we have

Definition 2.3. The Wasserstein-Fisher-Rao distance between $\rho_0, \rho_1 \in \mathcal{M}^+(\Omega)$ is

$$\mathtt{WFR}^2(\rho_0,\rho_1) := \inf_{(\rho,\mathbf{v},r)\in\mathcal{A}_{\mathtt{WFR}}[\rho_0,\rho_1]} \int_0^1 \int_\Omega \left(|\mathbf{v}_t(x)|^2 + |r_t|^2\right) \mathrm{d}\rho_t(x)\mathrm{d}t, \tag{2.2}$$

where the admissible set $\mathcal{A}_{\mathtt{WFR}}[\rho_0,\rho_1]$ is the set of curves $t \in [0,1] \mapsto (\rho_t, v_t, r_t) \in \mathcal{M}^+ \times \mathcal{M}(\Omega;\mathbb{R}^d) \times \mathcal{M}$ such that $t \mapsto \rho_t$ is narrowly continuous with endpoints ρ_0, ρ_1 and solves the continuity equation with source

$$\partial_t \rho_t + \operatorname{div}(\rho_t v_t) = \rho_t r_t.$$

Comparing Definition 2.3 with Definition 2.1 and Theorem 2.2, this dynamical formulation *à la Benamou-Brenier* shows that the WFR distance can be viewed as an inf-convolution of the Wasserstein and Fisher-Rao distances $\mathtt{W}, \mathtt{FR}$. From [11, 12, 22, 26] the infimum in (2.2) is always a minimum, and the corresponding minimizing curves $t \mapsto \rho_t$ are of course constant-speed geodesics $\mathtt{WFR}(\rho_t,\rho_s) = |t-s|\mathtt{WFR}(\rho_0,\rho_1)$. Then $(\mathcal{M}^+, \mathtt{WFR})$ is a complete metric space, and WFR metrizes the narrow convergences of measures (see again [11, 12, 22, 26]). Interestingly, there are other possible formulations of the distance in terms of static unbalanced optimal transportation, primal-dual characterizations with relaxed marginals, lifting to probability measures on a cone over Ω, duality with subsolutions of Hamilton-Jacobi equations, and we refer to [11, 12, 22, 26, 27] for more details.

As a first useful interplay between the $\mathtt{WFR}, \mathtt{W}, \mathtt{FR}$ distances we have

Proposition 2.4 ([18])**.** *Let $\rho_0, \rho_1 \in \mathcal{M}_2^+$ such that $|\rho_0| = |\rho_1|$. Then*

$$\mathtt{WFR}^2(\rho_0, \rho_1) \leqslant \mathtt{W}^2(\rho_0, \rho_1).$$

Similarly for all $\mu_0, \mu_1 \in \mathcal{M}^+$ (with possibly different masses) there holds

$$\mathtt{WFR}^2(\mu_0, \mu_1) \leqslant \mathtt{FR}^2(\mu_0, \mu_1).$$

Finally, for all $\nu_0, \nu_1 \in \mathcal{M}_2^+$ such that $|\nu_0| = |\nu_1|$ and all $\nu \in \mathcal{M}^+$, there holds

$$\mathtt{WFR}^2(\nu_0, \nu) \leqslant 2(\mathtt{W}^2(\nu_0, \nu_1) + \mathtt{FR}^2(\nu_1, \nu)).$$

Moreover, we have the following link between the reaction and the velocity in (2.2), which was the original definition in [22]:

Proposition 2.5 ([18], Prop. 2.2)**.** *Definition (2.3) of the $\mathtt{WFR}$ distance can be restricted to the subclass of admissible paths $(\mathbf{v}_t, r_t) = (\nabla u_t, u_t)$ for potentials $u_t \in H^1(\mathrm{d}\rho_t)$ and continuity equations*

$$\partial_t \rho_t + \mathrm{div}(\rho_t \nabla u_t) = \rho_t u_t.$$

This shows that $(\mathcal{M}^+, \mathtt{WFR})$ can be endowed with the formal Riemannian structure constructed as follow: any two tangent vectors $\xi^1 = \partial_t \rho^1, \xi^2 = \partial_t \rho^2$ can be uniquely identified with potentials u^i by solving the elliptic equations

$$\xi^i = -\,\mathrm{div}(\rho \nabla u^i) + \rho u^i.$$

Then the Riemannian tensor is naturally constructed on the $H^1(\mathrm{d}\rho)$ scalar product, *i.e.*

$$g_\rho(\xi^1, \xi^2) := \langle u^1, u^2 \rangle_{H^1(\mathrm{d}\rho)} = \int_\Omega (\nabla u^1 \cdot \nabla u^2 + u^1 u^2) \mathrm{d}\rho.$$

This is purely formal, and we refer again to [18] for discussions. Given a functional

$$\mathcal{F}(\rho) := \int_\Omega F(\rho) + \int_\Omega \rho V + \frac{1}{2} \int_\Omega (K * \rho)\rho,$$

this Riemannian structure also allows to compute $\mathtt{WFR}$ gradients as

$$\mathrm{grad}_{\mathtt{WFR}}\, \mathcal{F}(\rho) = -\,\mathrm{div}\left(\rho \nabla \frac{\delta \mathcal{F}}{\delta \rho}\right) + \rho \frac{\delta \mathcal{F}}{\delta \rho} = \mathrm{grad}_{\mathtt{W}}\, \mathcal{F}(\rho) + \mathrm{grad}_{\mathtt{FR}}\, \mathcal{F}(\rho),$$

where $\frac{\delta \mathcal{F}}{\delta \rho} = F'(\rho) + V + K * \rho$ denotes the Euclidean first variation of $\mathcal{F}$ with respect to ρ. In other words, the Riemannian tangent vector $\mathrm{grad}_{\mathtt{WFR}}\, \mathcal{F}(\rho)$ is represented in the previous $H^1(\mathrm{d}\rho)$ duality by the scalar potential $u = \frac{\delta \mathcal{F}}{\delta \rho}$.

3. An existence result for general parabolic equations

In this section, we propose to solve scalar parabolic equations of the form

$$\begin{cases} \partial_t \rho = \operatorname{div}(\rho \nabla (F'(\rho) + V)) - \rho(G'(\rho) + U) \\ \rho|_{t=0} = \rho^0 \\ \rho \nabla (F'(\rho) + V)|_{\partial\Omega} \cdot \nu = 0 \end{cases} \tag{3.1}$$

in a bounded domain $\Omega \subset \mathbb{R}^d$ with Neumann boundary condition and suitable initial conditions. Our goal is to extend to the case $F \not\equiv G, V \not\equiv U$ the method initially introduced in [18] for variational WFR-gradient flows, *i.e.* (3.1) with $F \equiv G$ and $V \equiv U$.

We assume for simplicity that $F : \mathbb{R} \to \mathbb{R}$ is given by

$$F(z) = \begin{cases} z \log z - z & \text{(linear diffusion)} \\ \text{or} \\ \frac{1}{m_1 - 1} z^{m_1} & \text{(Porous Media diffusion)} \end{cases} \tag{3.2}$$

and $G : \mathbb{R} \to \mathbb{R}$ is given by

$$G(z) = \frac{1}{m_2 - 1} z^{m_2}, \quad \text{for some } m_2 > 1. \tag{3.3}$$

Note that we cannot take $G(z) = z \log z - z$ because the Boltzmann entropy is not well behaved (neither regular nor displacement convex) with respect to the Fisher-Rao metric in the reaction step, see [18, 26, 27] for discussions. In addition, we assume that

$$V \in W^{1,\infty}(\Omega) \quad \text{and} \quad U \in L^\infty(\Omega).$$

We denote by $\mathcal{E}_F, \mathcal{E}_G : \mathcal{M}^+ \to \mathbb{R}$ the energy functionals

$$\mathcal{E}_F(\rho) := \mathcal{F}(\rho) + \mathcal{V}(\rho),$$

$$\mathcal{E}_G(\rho) := \mathcal{G}(\rho) + \mathcal{U}(\rho),$$

where

$$\mathcal{F}(\rho) := \begin{cases} \int_\Omega F(\rho) & \text{if } \rho \ll \mathcal{L}_{|\Omega} \\ +\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad \mathcal{V}(\rho) := \int_\Omega V\rho,$$

similarly

$$\mathcal{G}(\rho) := \begin{cases} \int_\Omega G(\rho) & \text{if } \rho \ll \mathcal{L}_{|\Omega} \\ +\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad \mathcal{U}(\rho) := \int_\Omega U\rho.$$

Although more general statements with suitable structural assumptions could certainly be proved, we do not seek full generality here and choose to restrict from the beginning to the above simple (but nontrivial) setting for the sake of exposition.

Definition 3.1. A weak solution of (3.1) with initial datum ρ^0 is a curve $t \mapsto \rho(t,\cdot) \in L^1_+ \cap L^\infty(\Omega)$ defined for all $t \geq 0$ such, that for all $T < \infty$, the pressure $P_F(\rho) := \rho F'(\rho) - F(\rho)$ satisfies $\nabla P_F(\rho) \in L^2([0,T]\times\Omega)$, and

$$\int_0^{+\infty} \left(\int_\Omega (\rho\partial_t\phi - \rho\nabla V\cdot\nabla\phi - \nabla P_F(\rho)\cdot\nabla\phi - \rho(G'(\rho)+U)\phi)\,\mathrm{d}x \right)\mathrm{d}t = -\int_\Omega \phi(0,x)\rho^0(x)\,\mathrm{d}x$$

for every $\phi \in \mathcal{C}^\infty_c([0,+\infty)\times\mathbb{R}^d)$.

Note that the pressure P_F is defined so that the chain rule $\mathrm{div}(\rho\nabla F'(\rho)) = \Delta P_F(\rho)$ holds at least for smooth ρ's.

The starting point of our analysis is that (3.1) can be written, at least formally as,

$$\partial_t\rho = \mathrm{div}(\rho\nabla(F'(\rho)+V)) - \rho(G'(\rho)+U) \quad \leftrightarrow \quad \partial_t\rho = -\operatorname{grad}_{\mathtt{W}}\mathcal{E}_F(\rho) - \operatorname{grad}_{\mathtt{FR}}\mathcal{E}_G(\rho),$$

see again Section 2 from [18] for a discussion on the $\mathtt{FR},\mathtt{W},\mathtt{WFR}$ Riemannian structure and how to compute the respective gradients. Our splitting scheme is a variant of that originally introduced in [18], and can be viewed as an operator splitting method: each part of the PDE above is discretized (in time) in its own $\mathtt{W},\mathtt{FR}$ metric, and corresponds respectively to a $\mathtt{W}$/transport./diffusion step and to a $\mathtt{FR}$/reaction step. More precisely, let $h>0$ be a small time step. Starting from the initial datum $\rho^0_h := \rho^0$, we construct two recursive sequences $(\rho^k_h)_k$ and $(\rho^{k+1/2}_h)_k$ such that

$$\begin{cases} \rho^{k+1/2}_h \in \underset{\rho\in\mathcal{M}^+,|\rho|=|\rho^k_h|}{\operatorname{argmin}} \left\{ \frac{1}{2h}\mathtt{W}^2(\rho,\rho^k_h) + \mathcal{E}_F(\rho) \right\}, \\ \\ \rho^{k+1}_h \in \underset{\rho\in\mathcal{M}^+}{\operatorname{argmin}} \left\{ \frac{1}{2h}\mathtt{FR}^2(\rho,\rho^{k+1/2}_h) + \mathcal{E}_G(\rho) \right\}. \end{cases} \tag{3.4}$$

With our structural assumptions on the various functionals involved and arguing as in [18], the direct method shows that this scheme is well-posed, *i.e.* that each minimizing problem in (3.4) admits a unique minimizer. We construct next two piecewise-constant interpolating curves

$$\begin{cases} \rho_h(t) = \rho^{k+1}_h, \\ \tilde\rho_h(t) = \rho^{k+1/2}_h, \end{cases} \qquad \text{for all } t\in(kh,(k+1)h]. \tag{3.5}$$

Our main result in this section is the constructive existence of weak solutions to (3.1):

Theorem 3.2. *Assume that $\rho^0 \in L^1_+\cap L^\infty(\Omega)$. Then, up to a discrete subsequence (still denoted $h\to 0$ and not relabeled here), ρ_h and $\tilde\rho_h$ converge strongly in $L^1((0,T)\times\Omega)$ to a weak solution ρ of (3.1).*

Note that any uniqueness for (3.1) would imply convergence of the whole (continuous) sequences $\rho_h,\tilde\rho_h \to \rho$ as $h\to 0$, but for the sake of simplicity we shall not address this issue here.

The main technical obstacle in the proof of Theorem 3.2 is to retrieve compactness in time. For the classical minimizing scheme of any energy $\mathcal{E}$ on any metric space (X,d), suitable time compactness is usually retrieved in the form of the *total-square distance estimate*

$$\frac{1}{2h}\sum_{k\geq 0} d^2(x^k,x^{k+1}) \leqslant \mathcal{E}(x^0) - \inf\mathcal{E}.$$

This usually works because only one functional is involved, and $\mathcal{E}(x_0) - \inf\mathcal{E}$ is obtained as a telescopic sum of one-step energy dissipations $\mathcal{E}(x^{k+1}) - \mathcal{E}(x^k)$. Here each of our elementary step in (3.1) involves one of the

W, FR metrics, and we will use the WFR distance to control both simultaneously: this strongly leverages the inf-convolution structure, the WFR distance being precisely built on a compromise between W/transport and FR/reaction. On the other hand we also have two different functionals $\mathcal{E}_F, \mathcal{E}_G$, and we will have to carefully estimate the dissipation of $\mathcal{E}_F$ during the FR reaction step (driven by $\mathcal{E}_G$) as well as the dissipation of $\mathcal{E}_G$ during the W transport./diffusion step (driven by $\mathcal{E}_F$).

We start by collecting one-step estimates, exploiting the optimality conditions for each elementary minimization procedure, and postpone the proof of Theorem 3.2 to the end of the section.

3.1. Optimality conditions and pointwise L^∞ estimates

The optimality conditions for the first Wasserstein step $\rho^k \to \rho^{k+1/2}$ in (3.4) are by now classical [43], and can be written for example

$$\frac{-\nabla\varphi_h^{k+1/2}}{h}\rho_h^{k+1/2} = \nabla P_F(\rho_h^{k+1/2}) + \rho_h^{k+1/2}\nabla V \quad \text{a.e.} \tag{3.6}$$

Here $\varphi_h^{k+1/2}$ is an optimal (backward) Kantorovich potential from $\rho_h^{k+1/2}$ to ρ_h^k.

Lemma 3.3. *For all $k \geqslant 0$,*

$$\|\rho_h^{k+1/2}\|_{L^1} = \|\rho_h^k\|_{L^1} \tag{3.7}$$

and for all constant C such that $V(x) \leqslant C$ a.e,

$$\rho_h^k(x) \leqslant (F')^{-1}(C - V(x)) \text{ a.e.} \quad \Rightarrow \quad \rho_h^{k+1/2}(x) \leqslant (F')^{-1}(C - V(x)) \text{ a.e.} \tag{3.8}$$

Proof. The Wasserstein step in (3.4) is mass conservative by construction, so the first part is obvious.

As for the second part, F. Otto established in [33] a weaker version of (3.8) in the form (with the same assumptions on F):

$$\text{any minimizer} \quad \rho^* = \underset{\rho}{\operatorname{Argmin}} \left\{ \frac{1}{2h}\mathtt{W}^2(\rho,\nu) + \int_\Omega F(\rho) \right\} \quad \text{satisfies} \quad \|\rho^*\|_{L^\infty} \leq \|\nu\|_{L^\infty}, \tag{3.9}$$

which we shall refer to as Otto's maximum principle in the sequel. This was later generalized in [38] to include an external potential $V(x) \not\equiv 0$ exactly as in our statement. The proof goes by contradiction: assuming that $\{\rho^{k+1/2}(x) > (F')^{-1}(C - V(x))\}$ has positive Lebesgue measure, a careful construction based on the convexity of F (and the optimal transport plan γ^k between ρ_h^k and $\rho_h^{k+1/2}$) allows to construct a strictly better competitor $\tilde{\rho}_h^{k+1/2}$ in (3.4). For the sake of brevity we omit the details and refer to proof of Lemma 2 from [38]. Let us however briefly mention that the the convexity $F'' > 0$ precisely makes the continuous PDE $\partial_t\rho = \operatorname{div}(\rho\nabla(F'(\rho) + V)) = \operatorname{div}(\rho F''(\rho)\nabla\rho) + (\dots)$ parabolic. In other words, (3.8) is a discrete equivalent of taking $\overline{\rho}(x) := (F')^{-1}(C - V(x))$ as a stationary Barenblatt supersolution for $\partial_t\rho = \operatorname{div}(\rho\nabla(F'(\rho) + V))$. □

For the second Fisher-Rao reaction step, the optimality condition can be written Section 4.2 from [18] in the form

$$\left(\sqrt{\rho_h^{k+1}} - \sqrt{\rho_h^{k+1/2}}\right)\sqrt{\rho_h^{k+1}} = -\frac{h}{2}\rho_h^{k+1}\left(G'(\rho_h^{k+1}) + U\right) \quad \text{a.e.} \tag{3.10}$$

As a consequence we have

Lemma 3.4. *There is $C \equiv C(U) > 0$ such that for $h \leqslant h_0(U)$ small enough we have*

$$\rho_h^{k+1}(x) \leqslant (1 + Ch)\rho_h^{k+1/2}(x) \quad a.e., \tag{3.11}$$

and for all $M > 0$ there is $c \equiv c(M, U)$ such that if $\|\rho_h^{k+1/2}\|_\infty \leqslant M$ then

$$(1 - ch)\rho_h^{k+1/2}(x) \leqslant \rho_h^{k+1}(x) \quad a.e. \tag{3.12}$$

Note in particular that this immediately implies

$$\text{supp } \rho_h^{k+1} = \text{supp } \rho_h^{k+1/2}, \tag{3.13}$$

which was to be expected since the reaction part $\partial_t\rho = -\rho(G'(\rho) + U)$ of the PDE (3.1) preserves strict positivity.

Proof. We start with the upper bound: inside supp ρ_h^{k+1}, (3.10) and $G' \geqslant 0$ give

$$\begin{aligned}\sqrt{\rho_h^{k+1}(x)} - \sqrt{\rho_h^{k+1/2}(x)} &= -h\sqrt{\rho_h^{k+1}(x)}(G'(\rho_h^{k+1}(x)) + U(x)) \\ &\leqslant -hU(x)\sqrt{\rho_h^{k+1}(x)} \leqslant h\|U\|_\infty\sqrt{\rho_h^{k+1}(x)}\end{aligned}$$

whence

$$\sqrt{\rho_h^{k+1}(x)} \leqslant \frac{1}{1 - h\|U\|_\infty}\sqrt{\rho_h^{k+1/2}(x)}.$$

Taking squares and using

$$\frac{1}{(1 - h\|U\|_\infty)^2} = 1 + 2\|U\|_{L^\infty}h + \mathcal{O}(h^2) \leqslant 1 + 3\|U\|_{L^\infty}h$$

for small h gives the desired inequality.

For the lower bound (3.12), we first observe that since $G'' \geqslant 0$ and from (3.11) we have $G'(\rho_h^{k+1}) \leqslant G'((1 + Ch)\rho_h^{k+1/2}) \leqslant G'(2M)$ if h is small enough. Then (3.10) gives inside supp ρ^{k+1}

$$\begin{aligned}\sqrt{\rho_h^{k+1}(x)} - \sqrt{\rho_h^{k+1/2}(x))} &= -h\sqrt{\rho_h^{k+1}(x)}(G'(\rho_h^{k+1}(x)) + U(x)) \\ &\geqslant -h(G'(2M) + \|U\|_\infty)\sqrt{\rho_h^{k+1}(x)},\end{aligned}$$

hence

$$\rho_h^{k+1}(x) \geqslant \frac{1}{(1 + h(G'(2M) + \|U\|_\infty))^2}\rho_h^{k+1/2}(x) \geqslant (1 - ch)\rho_h^{k+1/2}(x)$$

for small h. □

Combining Lemmas 3.3 and 3.4, we obtain at the continuous level

Proposition 3.5. *For all $T > 0$ there exist constants M_T, M'_T such that for all $t \in [0, T]$,*

$$\|\rho_h(t)\|_{L^1 \cap L^\infty}, \|\tilde\rho_h(t)\|_{L^1 \cap L^\infty} \leqslant M_T$$

and

$$\|\rho_h(t) - \tilde{\rho}_h(t)\|_{L^1} \leqslant hM'_T$$

uniformly in $h \geqslant 0$.

Note from the second estimate that strong $L^1((0,T)\times\Omega)$ convergence of ρ_h will immediately imply convergence of $\tilde{\rho}_h$ to the same limit.

Proof. Combining (3.8) and (3.11) and arguing by induction, we obtain, for all $t \in [0,T]$,

$$\|\rho_h(t)\|_{L^\infty}, \|\tilde{\rho}_h(t)\|_{L^\infty} \leqslant C_T,$$

where C_T is a constant depending on $\|V\|_{L^\infty}$, see Lemma 2 from [38]. The L^1 bound is even easier: since the Wasserstein step is mass preserving, we can integrate (3.11) in space to get

$$\|\rho_h^{k+1}\|_{L^1} \leqslant (1+Ch)\|\rho_h^{k+1/2}\|_{L^1} = (1+Ch)\|\rho_h^k\|_{L^1}.$$

This immediately gives $\|\rho_h^k\|_{L^1} \leq (1+Ch)^k\|\rho^0\|_{L^1}$, and for $t \leqslant T \Leftrightarrow k \leqslant \lfloor T/h \rfloor$ we conclude exploiting $(1+Ch)^{\lfloor T/h \rfloor} \lesssim e^{CT}$.

In order to compare now ρ_h and $\tilde{\rho}_h$, we take advantage of the upper bound (3.11) to write $\rho_h^{k+1/2} \leqslant M_T$ as long as $kh \leqslant T$. Taking $c = c(M_T)$ in (3.12) and combining with (3.11), we have

$$-ch\rho_h^{k+1/2} \leqslant \rho_h^{k+1/2} - \rho_h^{k+1} \leqslant Ch\rho_h^{k+1/2} \quad \text{a.e.}$$

Integrating in Ω we conclude that

$$\|\rho_h(t) - \tilde{\rho}_h(t)\|_{L^1} = \|\rho_h^{k+1} - \rho_h^{k+1/2}\|_{L^1} \leqslant h\max\{c,C\}\|\rho_h^{k+1/2}\|_{L^1} \leqslant h\max\{c,C\}M_T = hM'_T$$

and the proof is complete. □

3.2. Total square distance and energy dissipation

Our goal is here to retrieve a total square distance estimate of the form $\sum_k \mathtt{WFR}^2(\rho_h^k, \rho_h^{k+1}) \leq \ldots$, see Proposition 3.7 below for a precise statement. According to Proposition 2.4 it is enough to control $\mathtt{W}^2(\rho_h^k, \rho_h^{k+1/2})$ and $\mathtt{FR}^2(\rho_h^{k+1/2}, \rho_h^{k+1})$ separately, each step being driven by its own $\mathcal{F}+\mathcal{V}$ and $\mathcal{G}+\mathcal{U}$ functionals, respectively. Along the second Fisher-Rao step $k+1/2 \to k+1$, we will first control the $\mathcal{F}+\mathcal{V}$ increment (which is a crossed dissipation estimate), and then directly show that the $\mathcal{G}+\mathcal{U}$ increment is small enough. This will allow in the end to recover a telescopic sum as in the classical approach, plus a summable term.

Testing $\rho = \rho_h^k$ in the first Wasserstein step in (3.4), we get as usual

$$\frac{1}{2h}\mathtt{W}^2(\rho_h^{k+1/2}, \rho_h^k) \leqslant \mathcal{F}(\rho_h^k) - \mathcal{F}(\rho_h^{k+1/2}) + \int_\Omega V(\rho_h^k - \rho_h^{k+1/2}). \tag{3.14}$$

Let $\mathbf{t}$ be the optimal map from ρ_h^k to $\rho_h^{k+1/2}=\mathbf{t}_{\#}\rho_h^k$, given by the Gangbo-McCann theorem [19]. Since V is Globally Lipschitz we get by Young's inequality with ε

$$\left|\int_\Omega V(\rho_h^k-\rho_h^{k+1/2})\right| \le \int_\Omega |V(x)-V(\mathbf{t}(x))|\rho_h^k(x) \le \|\nabla V\|_\infty \int |\mathbf{t}(x)-x|\rho_h^k(x)$$
$$\le \|\nabla V\|_\infty \int \left\{\frac{\varepsilon}{h}|\mathbf{t}(x)-x|^2+\frac{h}{\varepsilon}\right\}\rho_h^k(x) = \|\nabla V\|_\infty\left(\frac{\varepsilon}{h}\mathtt{W}^2(\rho_h^k,\rho_h^{k+1/2})+\frac{h}{\varepsilon}\|\rho_h^k\|_{L^1}\right).$$

Choosing $\varepsilon=\frac{1}{4\|\nabla V\|_\infty}$ and recalling from Proposition 3.5 that the mass $\|\rho_h^k\|_{L^1}\le M_T$ is bounded uniformly, we can thus reabsorb the $\mathtt{W}^2$ term in the left-hand side of (3.14) to obtain

$$\frac{1}{4h}\mathtt{W}^2(\rho_h^{k+1/2},\rho_h^k) \leqslant \mathcal{F}(\rho_h^k)-\mathcal{F}(\rho_h^{k+1/2})+C_Th. \tag{3.15}$$

The next proposition allows to replace $\mathcal{F}(\rho_h^{k+1/2})$ by $\mathcal{F}(\rho_h^k)$ above in order to retrieve a telescopic sum, up to a $\mathcal{O}(h)$ term:

Proposition 3.6. *For all $T>0$ there exists a constant $C_T>0$ such that, for all $k\geqslant 0$ and $k\le\lfloor T/h\rfloor$,*

$$\mathcal{F}(\rho_h^{k+1}) \leqslant \mathcal{F}(\rho_h^{k+1/2})+C_Th. \tag{3.16}$$

Proof. We first treat the case of $F(z)=\frac{1}{m_1-1}z^{m_1}$ with $m_1>1$. Since F is increasing, we use (3.11) to obtain

$$\mathcal{F}(\rho_h^{k+1})-\mathcal{F}(\rho_h^{k+1/2}) \leqslant \frac{((1+Ch)^{m_1}-1)}{m_1-1}\int_\Omega (\rho_h^{k+1/2})^{m_1}$$
$$\leqslant Ch\|\rho^{k+1/2}\|_{L^\infty}^{m_1-1}\|\rho^{k+1/2}\|_{L^1},$$

and we conclude from Proposition 3.5.

In the second case $F(z)=z\log(z)-z$, we have

$$\mathcal{F}(\rho_h^{k+1})=\int_{\{\rho_h^{k+1}\leqslant e^{-1}\}}\rho_h^{k+1}\log(\rho_h^{k+1})+\int_{\{\rho_h^{k+1}\geqslant e^{-1}\}}\rho_h^{k+1}\log(\rho_h^{k+1})-\int_\Omega\rho_h^{k+1}.$$

Note from Proposition 3.5 that the z contribution in $F(z)=z\log z-z$ is immediately controlled by $|\int\rho_h^{k+1}-\int\rho_h^{k+1/2}|\leqslant\|\rho_h^{k+1}-\rho_h^{k+1/2}\|_{L^1}\leqslant hM_T'$, so we only have to estimate the $z\log z$ contribution. Since $z\mapsto z\log z$ is increasing on $\{z\geqslant e^{-1}\}$ and using (3.11), the second term in the right hand side becomes

$$\int_{\{\rho_h^{k+1}\geqslant e^{-1}\}}\rho_h^{k+1}\log(\rho_h^{k+1}) \leqslant \int_{\{\rho_h^{k+1}\geqslant e^{-1}\}}(1+Ch)\rho_h^{k+1/2}\log((1+Ch)\rho_h^{k+1/2})$$
$$\leqslant \int_{\{\rho_h^{k+1}\geqslant e^{-1}\}}\rho_h^{k+1/2}\log(\rho_h^{k+1/2})+Ch\int_{\{\rho_h^{k+1}\geqslant e^{-1}\}}\rho_h^{k+1/2}\log(\rho_h^{k+1/2})$$
$$+(1+Ch)\int_{\{\rho_h^{k+1}\geqslant e^{-1}\}}\rho_h^{k+1/2}\log(1+Ch)$$
$$\leqslant \int_{\{\rho_h^{k+1}\geqslant e^{-1}\}}\rho_h^{k+1/2}\log(\rho_h^{k+1/2})+C_Th,$$

where we used $\|\rho_h^{k+1/2}\|_{L^1} \leqslant M_T$ from Proposition 3.5 as well as $\log(1+Ch) \leqslant Ch$ in the last inequality. Using the same method with the bound from below (3.12) on $\{\rho_h^{k+1} \leqslant e^{-1}\}$ (where $z \mapsto z\log z$ is now decreasing), we obtain similarly

$$\int_{\{\rho_h^{k+1} \leqslant e^{-1}\}} \rho_h^{k+1}\log(\rho_h^{k+1}) \leqslant \int_{\{\rho_h^{k+1} \leqslant e^{-1}\}} \rho_h^{k+1/2}\log(\rho_h^{k+1/2}) + C_T h.$$

Combining both inequalities gives

$$\int_\Omega \rho_h^{k+1}\log(\rho_h^{k+1}) \leqslant \int_\Omega \rho_h^{k+1/2}\log(\rho_h^{k+1/2}) + C_T h$$

and the proof is complete. □

Summing (3.15) and (3.16) over k we obtain

$$\frac{1}{4h}\sum_{k=0}^{N-1} \mathtt{W}^2(\rho_h^{k+1/2}, \rho_h^k) \leqslant \mathcal{F}(\rho^0) - \mathcal{F}(\rho_h^N) + C_T, \tag{3.17}$$

where $N = \lfloor \frac{T}{h} \rfloor$.

The goal is now to estimate the $\mathtt{FR}^2$ increment along the second reaction step. Testing $\rho = \rho_h^{k+1/2}$ in the Fisher-Rao minimization in (3.4), we obtain

$$\frac{1}{2h}\mathtt{FR}^2(\rho_h^{k+1}, \rho_h^{k+1/2}) \leqslant \mathcal{G}(\rho_h^{k+1/2}) - \mathcal{G}(\rho_h^{k+1}) + \int_\Omega U(\rho_h^{k+1/2} - \rho_h^{k+1}). \tag{3.18}$$

Since we assumed $U \in L^\infty(\Omega)$ and because $\rho_h(t) = \rho_h^{k+1}$ remains close to $\tilde{\rho}_h(t) = \rho_h^{k+1/2}$ in L^1 uniformly in t, h by Proposition 3.5, we immediately control the potential part as

$$\int_\Omega U(\rho_h^{k+1/2} - \rho_h^{k+1}) \leqslant \|U\|_\infty C_T h. \tag{3.19}$$

For the internal energy we argue exactly as in the proof Proposition 3.6 (for the Porous Media part, since we chose here $G(z) = \frac{1}{m_2-1}z^{m_2}$), and obtain

$$\mathcal{G}(\rho_h^{k+1/2}) - \mathcal{G}(\rho_h^{k+1}) \leqslant C_T h. \tag{3.20}$$

Combining (3.18), (3.19) and (3.20), we immediately deduce that

$$\frac{1}{2h}\sum_{k=0}^{N-1} \mathtt{FR}^2(\rho_h^{k+1/2}, \rho_h^{k+1}) \leqslant C_T, \tag{3.21}$$

where $N = \lfloor \frac{T}{h} \rfloor$ as before.

Finally, we recover an approximate compactness in time in the form

Proposition 3.7. *There exists a constant $C_T > 0$ such that for all h small enough and $k \leqslant N = \lfloor T/h \rfloor$,*

$$\frac{1}{h}\sum_{k=0}^{N-1} \mathtt{WFR}^2(\rho_h^k, \rho_h^{k+1}) \leqslant 8\mathcal{F}(\rho^0) + C_T. \tag{3.22}$$

Proof. Adding (3.17) and (3.21) gives

$$\frac{1}{h}\sum_{k=0}^{N-1} \mathtt{W}^2(\rho_h^k, \rho_h^{k+1/2}) + \mathtt{FR}^2(\rho_h^{k+1/2}, \rho_h^{k+1}) \leqslant 4\left(\mathcal{F}(\rho^0) - \mathcal{F}(\rho_h^N) + C_T\right) + 2C_T \leqslant 4\mathcal{F}(\rho^0) + C_T,$$

since in any case $F(z) = \frac{1}{m_1-1} z^{m_1} \geqslant 0$ and $F(z) = z\log z - z \geqslant -1$ is bounded from below on the bounded domain Ω, hence $\mathcal{F}(\rho_h^N) \geqslant -C_\Omega$ uniformly. It then follows from Proposition 2.4 that $\mathtt{W}^2(\rho_h^k, \rho_h^{k+1/2}) + \mathtt{FR}^2(\rho_h^{k+1/2}, \rho_h^{k+1}) \geqslant \frac{1}{2}\mathtt{WFR}^2(\rho_h^k, \rho_h^{k+1})$ in the left-hand side, and the result immediately follows.

□

3.3. Estimates and convergences

From the total-square distance estimate (3.22) we recover as usual the approximate $\frac{1}{2}$-Hölder estimate

$$\mathtt{WFR}(\rho_h(t), \rho_h(s)) + \mathtt{WFR}(\tilde\rho_h(t), \tilde\rho_h(s)) \leqslant C_T|t - s + h|^{1/2} \tag{3.23}$$

for all fixed $T > 0$ and $t, s \in [0, T]$. From (3.21) and Proposition 2.4 we have moreover

$$\mathtt{WFR}(\rho_h(t), \tilde\rho_h(t)) \leqslant \mathtt{FR}(\rho_h(t), \tilde\rho_h(t)) \leqslant C\sqrt{h}. \tag{3.24}$$

Using a refined version of Ascoli-Arzelà theorem, Proposition 3.3.1 from [4] and arguing exactly as in Proposition 4.1 from [18], we see that for all $T > 0$ and up to extraction of a discrete subsequence, ρ_h and $\tilde\rho_h$ converge uniformly to the same WFR-continuous curve $\rho \in \mathcal{C}^{1/2}([0,T], \mathcal{M}^+_{\mathtt{WFR}})$ as

$$\sup_{t\in[0,T]} \left(\mathtt{WFR}(\rho_h(t), \rho(t)) + \mathtt{WFR}(\tilde\rho_h(t), \rho(t))\right) \to 0.$$

In order to pass to the limit in the nonlinear terms, we first strengthen this WFR-convergence into a more tractable strong L^1 convergence. The first step is to retrieve compactness in space:

Proposition 3.8. *For all $T > 0$, ρ_h and $\tilde\rho_h$ satisfies*

$$\|P_F(\tilde\rho_h)\|_{L^2([0,T];H^1(\Omega))} \leqslant C_T. \tag{3.25}$$

Proof. From (3.6) and the $L^1 \cap L^\infty$ bounds from Proposition 3.5 we see that

$$\begin{aligned}
\int_\Omega |\nabla P_F(\rho_h^{k+1/2})|^2 &\leqslant \frac{2}{h^2}\int_\Omega |\nabla\varphi_h^{k+1/2}|^2(\rho_h^{k+1/2})^2 + 2\int_\Omega |\nabla V|^2(\rho_h^{k+1/2})^2 \\
&\leqslant \frac{2C_T}{h^2}\int_\Omega |\nabla\varphi_h^{k+1/2}|^2\rho_h^{k+1/2} + 2\|\nabla V\|_\infty^2 \int_\Omega (\rho_h^{k+1/2})^2 \\
&\leqslant C_T\left(\frac{\mathtt{W}^2(\rho_h^{k+1/2}, \rho_h^k)}{h^2} + 1\right)
\end{aligned}$$

since $\varphi_h^{k+1/2}$ is the optimal (backward) Kantorovich potential from $\rho_h^{k+1/2}$ to ρ_h^k. Multiplying by $h>0$, summing over k, and exploiting (3.17) gives

$$\|P_F(\tilde\rho_h)\|^2_{L^2([0,T];H^1(\Omega))} \leqslant \sum_{k=0}^{N-1} h\|P_F(\rho_h^{k+1/2})\|^2_{H^1} \leqslant C_T(\mathcal{F}(\rho^0)-\mathcal{F}(\rho_h^N)+1)\leqslant C_T,$$

where we used as before $\mathcal{F}(\rho_h^N)\geqslant -C_\Omega$ in the last inequality. □

We are now finally in position of proving our main result:

Proof of Theorem 3.2. Exploiting (3.22) and (3.25), we can apply a generalized Aubin-Lions lemma Theorem 2 from [41], to obtain that $\tilde\rho_h$ converges to ρ strongly in $L^1(Q_T)$ (see [24]). By diagonal extraction if needed, we can assume that the convergence holds in $L^1(Q_T)$ for all fixed $T>0$. Then by Proposition 3.5 we have

$$\|\rho_h-\rho\|_{L^1(Q_T)} \leqslant \|\rho_h-\tilde\rho_h\|_{L^1(Q_T)} + \|\tilde\rho_h-\rho\|_{L^1(Q_T)} \leqslant C_T h + \|\tilde\rho_h-\rho\|_{L^1(Q_T)} \to 0$$

hence $\rho_h\to\rho$ strongly in $L^1(Q_T)$ as well. Moreover, since $P_F(\tilde\rho_h)$ is bounded in $L^2((0,T),H^1(\Omega))$ we can assume that $\nabla P_F(\tilde\rho_h)\rightharpoonup\nabla P_F(\rho)$ in $L^2((0,T),L^2(\Omega))$ for all $T>0$. Following [18], one easily exploits the Euler-Lagrange equations (3.6) and (3.10) to check that

$$\begin{aligned}\int_\Omega(\rho_h^{k+1}-\rho_h^k)\varphi = &-h\int_\Omega\left\{\nabla P_F(\rho_h^{k+1/2})+\rho_h^{k+1/2}\nabla V\right\}\cdot\nabla\varphi + \mathcal{O}\left(\|D^2\varphi\|_{L^\infty}\mathtt{W}^2(\rho_h^{k+1/2},\rho_h^k)\right)\\ &-\frac{h}{2}\int_\Omega\sqrt{\rho_h^{k+1}}\left(\sqrt{\rho_h^{k+1}}+\sqrt{\rho_h^{k+1/2}}\right)\left(G'(\rho_h^{k+1})+U\right)\varphi\end{aligned}$$

for all smooth $\varphi\in C^2(\bar\Omega)$. Summing from $k_1=\lfloor t_1/h\rfloor$ to $k_2=\lfloor t_2/h\rfloor$ and leveraging the strong convergence $\rho_h,\tilde\rho_h\to\rho$ one easily gets, in the limit $h\to0$,

$$\int_\Omega\rho(t_2)\varphi-\rho(t_1)\varphi = -\int_{t_1}^{t_2}\int_\Omega\left\{\nabla P_F(\rho)+\rho\nabla V\right\}\cdot\nabla\varphi+\rho(G'(\rho)+U)\varphi$$

for all $0<t_1<t_2$ and $\varphi\in\mathcal{C}_b^1(\Omega)$. Here we omit the details and refer to the proof of Theorem 4 from [18]. Since $\rho\in\mathcal{C}([0,T];\mathcal{M}^+_{\mathtt{WFR}})$ takes the initial datum $\rho(0)=\rho^0$ and WFR metrizes the narrow convergence of measures, this is well-known to be equivalent to our weak formulation in Definition 3.1, and the proof is complete. □

Remark 3.9. In the above proofs one can check that Theorem 3.2 extends in fact to all $\mathcal{C}^1$ nonlinearities G such that $G'\geqslant -C$ is bounded from below (so that the reaction term $\partial_t\rho=(\dots)-\rho G'(\rho)$ does not allow blow-up). Likewise, we stated and proved our main result in bounded domains for convenience: all the above arguments immediately extend to $\Omega=\mathbb{R}^d$ at least for $F(z)=\frac{1}{m_1-1}z^{m_1}\geqslant0$. The only argument actually exploiting the boundedness of Ω was the proof of Proposition 3.7, when we bounded from below $\mathcal{F}(\rho_h^N)\geqslant -C_\Omega$ in order to retrieve the total-square distance estimate. In the case $\Omega=\mathbb{R}^d$ and $F(z)=z\log z-z$ a lower bound $\mathcal{F}(\rho_h^N)\geqslant -C_T$ still holds, but the proof involves Carleman inequalities and requires a tedious control of the second moments $\mathfrak{m}_2(\rho)=\int_{\mathbb{R}^d}|x|^2\rho$ and we chose not to address this technical issue for the ease of exposition.

4. Application to systems

In this section we wish to illustrate the tractability of our splitting approach, and as an example we will solve systems of the form

$$\begin{cases} \partial_t \rho_1 = \operatorname{div}(\rho_1 \nabla(F_1'(\rho_1) + V_1[\rho_1, \rho_2])) - \rho_1(G_1'(\rho_1) + U_1[\rho_1, \rho_2]), \\ \partial_t \rho_2 = \operatorname{div}(\rho_2 \nabla(F_2'(\rho_2) + V_2[\rho_1, \rho_2])) - \rho_2(G_2'(\rho_2) + U_2[\rho_1, \rho_2]), \\ \rho_{1|t=0} = \rho_1^0, \ \rho_{2|t=0} = \rho_2^0. \end{cases} \tag{4.1}$$

For simplicity we assume again that Ω is a smooth, bounded subset of $\mathbb{R}^d$. The system (4.1) is naturally supplemented with Neumann boundary conditions,

$$\rho_1 \nabla(F_1'(\rho_1) + V_1[\rho_1, \rho_2]) \cdot \nu = 0 \text{ and } \rho_2 \nabla(F_2'(\rho_2) + V_2[\rho_1, \rho_2]) \cdot \nu = 0 \quad \text{on } \mathbb{R}^+ \times \partial\Omega,$$

where ν is the outward unit normal to $\partial\Omega$. Here we allow interactions between densities in the potential terms $V_i[\rho_1, \rho_2]$ and $U_i[\rho_1, \rho_2]$. In the mass-conservative case (without reaction terms), this system has already been studied in [8, 15, 24], using a semi-implicit JKO scheme introduced by Di Francesco and Fagioli [15]. This section combines the splitting scheme introduced in the previous section and semi-implicit schemes both for the Wasserstein JKO step and for the Fisher-Rao JKO step.

For the ease of exposition we keep the same assumptions for F_i and G_i as in the previous section, *i.e.* the diffusion terms F_i satisfy (3.2) and the reaction terms G_i satisfy (3.3). Moreover, since the potentials depend now on the densities ρ_1 and ρ_2, we need stronger hypotheses: we assume that $V_i : L^1(\Omega; \mathbb{R}^+)^2 \to W^{1,\infty}(\Omega)$ are continuous and verify, uniformly in $\rho_1, \rho_2 \in L^1(\Omega; \mathbb{R}^+)$,

$$\begin{aligned} \Big\| V_i[\rho_1, \rho_2] \Big\|_{W^{1,\infty}(\Omega)} &\leqslant K(1 + \|\rho_1\|_{L^1(\Omega)} + \|\rho_2\|_{L^1(\Omega)}), \\ \Big\| \nabla(V_i[\rho_1, \rho_2]) - \nabla(V_i[\mu_1, \mu_2]) \Big\|_{L^\infty(\Omega)} &\leqslant K(\|\rho_1 - \mu_1\|_{L^1(\Omega)} + \|\rho_2 - \mu_2\|_{L^1(\Omega)}). \end{aligned} \tag{4.2}$$

The interacting potentials we have in mind are of the form $V_i[\rho_1, \rho_2] = K_{i,1} * \rho_1 + K_{i,2} * \rho_2$, where $K_{i,1}, K_{i,2} \in W^{1,\infty}(\Omega)$ and then V_i satisfies (4.2).

$$U_i[\rho_1, \rho_2] \geqslant -K, \quad \forall\, \rho_1, \rho_2 \in L^1(\Omega; \mathbb{R}^+) \tag{4.3}$$

for some $K \in \mathbb{R}$, and

$$\|U_i[\rho_1, \rho_2]\|_{L^\infty(\Omega)} \leqslant K_M, \quad \forall \|\rho_1\|_{L^1(\Omega)}, \|\rho_2\|_{L^1(\Omega)} \leqslant M \tag{4.4}$$

for some nondecreasimg function $K_M \geqslant 0$ of M. The examples we have in mind are of the form

$$U_1[\rho_1, \rho_2] = C_1 \frac{\rho_2}{1 + \rho_1}, \quad U_2[\rho_1, \rho_2] = -C_2 \frac{\rho_1}{1 + \rho_1}$$

for some constants $C_i \geq 0$, or nonlocal reactions

$$U_i[\rho_1, \rho_2](x) = \int_\Omega K_{i,1}(x, y) \rho_1(y) \, \mathrm{d}y + \int_\Omega K_{i,2}(x, y) \rho_2(y) \, \mathrm{d}y$$

for some nonnegative kernels $K_{i,j} \in L^1 \cap L^\infty$. Such reaction models appear for example in biological adaptive dynamics [35].

Definition 4.1. We say that $(\rho_1, \rho_2) : \mathbb{R}^+ \to L^1_+ \cap L^\infty_+(\Omega)$ is a weak solution of (4.1) with initial datum (ρ_1^0, ρ_2^0) if, for $i \in \{1,2\}$ and all $T < +\infty$, the pressure $P_i(\rho_i) := \rho_i F_i'(\rho_i) - F_i(\rho_i)$ satisfies $\nabla P_i(\rho_i) \in L^2([0,T]\times\Omega)$, and

$$\int_0^{+\infty} \left(\int_\Omega (\rho \partial_t \phi_i - \rho_i \nabla V_i[\rho_1,\rho_2]\cdot\nabla\phi_i - \nabla P_i(\rho_i)\cdot\nabla\phi_i - \rho_i(G_i'(\rho_i) + U_i[\rho_1,\rho_2])\phi_i)\mathrm{d}x \right)\mathrm{d}t = -\int_\Omega \phi_i(0,x)\rho_i^0(x)\mathrm{d}x, \tag{4.5}$$

for all $\phi_i \in \mathcal{C}_c^\infty([0,+\infty)\times\mathbb{R}^d)$.

Then, the following result holds,

Theorem 4.2. *Assume that $\rho_1^0, \rho_2^0 \in L^1 \cap L^\infty_+(\Omega)$ and that V_i, U_i satisfy (4.2)–(4.4). Then (4.1) admits at least one weak solution.*

Note that this result can be easily adapted to systems with an arbitrary number of species $N \geqslant 2$, coupled by nonlocal terms $V_i[\rho_1,\dots,\rho_N]$ and $U_i[\rho_1,\dots,\rho_N]$.

Remark 4.3. A refined analysis shows that our approach would allow to handle systems of the form

$$\begin{cases} \partial_t\rho_1 - \operatorname{div}(\rho_1\nabla(F_1'(\rho_1) + V_1)) = -\rho_1 H_1(\rho_1,\rho_2), \\ \partial_t\rho_2 - \operatorname{div}(\rho_2\nabla(F_2'(\rho_2) + V_2)) = +\rho_2 H_2(\rho_1), \end{cases}$$

where H_1 is a nonnegative continuous function and H_2 is a continuous functions. Indeed since $H_1 \geq 0$ the reaction term is the first equation is nonpositive, hence $\|\rho_1(t)\|_{L^\infty(\Omega)} \leqslant C_T$. Then it follows that $-H_2(\rho_1)$ satisfies assumptions (4.3) and (4.4). A classical example is $H_2(\rho_1) = \rho_1^\alpha$ and $H_1(\rho_1,\rho_2) = \rho_1^{\alpha-1}\rho_2$, where $\alpha \geqslant 1$, see for example [40] for more discussions.

As already mentioned, the proof of Theorem 4.2 is based on a semi-implicit splitting scheme. More precisely, we construct four sequences $\rho_{1,h}^{k+1/2}, \rho_{1,h}^{k+1}, \rho_{2,h}^{k+1/2}, \rho_{2,h}^{k+1}$ defined recursively as

$$\begin{cases} \rho_{i,h}^{k+1/2} \in \underset{\rho\in\mathcal{M}^+, |\rho| = |\rho_{i,h}^k|}{\operatorname{argmin}} \left\{ \frac{1}{2h}\mathtt{W}^2(\rho,\rho_{i,h}^k) + \mathcal{F}_i(\rho) + \mathcal{V}_i(\rho|\rho_{1,h}^k,\rho_{2,h}^k) \right\} \\ \\ \rho_{i,h}^{k+1} \in \underset{\rho\in\mathcal{M}^+}{\operatorname{argmin}} \left\{ \frac{1}{2h}\mathtt{FR}^2(\rho,\rho_{i,h}^{k+1/2}) + \mathcal{G}_i(\rho) + \mathcal{U}_i(\rho|\rho_{1,h}^k,\rho_{2,h}^k) \right\}, \end{cases} \tag{4.6}$$

where the fully implicit terms

$$\mathcal{F}_i(\rho) := \begin{cases} \int_\Omega F_i(\rho) & \text{if } \rho \ll \mathcal{L}_{|\Omega} \\ +\infty & \text{otherwise} \end{cases} \quad \text{and} \quad \mathcal{G}_i(\rho) := \begin{cases} \int_\Omega G_i(\rho) & \text{if } \rho \ll \mathcal{L}_{|\Omega} \\ +\infty & \text{otherwise} \end{cases},$$

and the semi-implicit terms

$$\mathcal{V}_i(\rho|\mu_1,\mu_2) := \int_\Omega V_i[\mu_1,\mu_2]\rho \quad \text{and} \quad \mathcal{U}_i(\rho|\mu_1,\mu_2) := \int_\Omega U_i[\mu_1,\mu_2]\rho.$$

In the previous section, the proof of Theorem 3.2 for scalar equations strongly leveraged the uniform $L^\infty(\Omega)$-bounds on the discrete solutions. Here an additional difficulty arises due to the nonlocal terms $\nabla V_i[\rho_1,\rho_2]$ and $U_i[\rho_1,\rho_2]$, which are *a priori* not uniformly bounded in $L^\infty(\Omega)$. Using assumption (4.3) we will first obtain a uniform $L^1(\Omega)$-bound on ρ_1,ρ_2, and then extend Proposition 3.5 to the system (4.1). This in turn will give a uniform $W^{1,\infty}$ control on $V_i[\rho_1,\rho_2]$ and L^∞ control on $U_i[\rho_1,\rho_2]$ through our assumptions (4.2)–(4.4), which will finally allow to argue as in the previous section and give L^∞ control on ρ_1,ρ_2. Numerical simulations for a diffusive prey-predator system are presented at the end of this section.

4.1. Properties of discrete solutions

Arguing as in the case of one equation, the optimality conditions for the Wasserstein step and for the Fisher-Rao step first give

Lemma 4.4. *For all $k \geqslant 0$ and $i \in \{1,2\}$, we have*

$$\|\rho_{i,h}^{k+1/2}\|_{L^1} = \|\rho_{i,h}^{k}\|_{L^1}. \tag{4.7}$$

Moreover, there exists $C_i \equiv C(U_i) > 0$ (uniform in k) such that

$$\rho_{i,h}^{k+1}(x) \leqslant (1 + C_i h)\rho_{i,h}^{k+1/2}(x) \quad a.e. \tag{4.8}$$

Proof. The first part is simply the mass conservation in the Wasserstein step, and the second part follows the lines of the proof of (3.11) in Lemma 3.4 using assumption (4.3). □

As a direct consequence we have uniform control on the L^1-norms:

Lemma 4.5. *For all $T > 0$ there exist constants $C_T, C'_T > 0$ such that, for all $t \in [0,T]$,*

$$\|\rho_{i,h}(t)\|_{L^1}, \|\tilde{\rho}_{i,h}(t)\|_{L^1} \leqslant C_T$$

and

$$\Big\|V_i[\rho_{1,h}(t),\rho_{2,h}(t)]\Big\|_{W^{1,\infty}}, \Big\|V_i[\tilde{\rho}_{1,h}(t),\tilde{\rho}_{2,h}(t)]\Big\|_{W^{1,\infty}} \leqslant C'_T. \tag{4.9}$$

Proof. Integrating (4.8) and iterating with (4.7), we obtain for all $t \leqslant T$ and $k \leqslant \lfloor T/h \rfloor$

$$\|\rho_{i,h}^{k+1}\|_{L^1} \leqslant (1+C_i h)\|\rho_{i,h}^{k}\|_{L^1} \leqslant (1+C_i h)^k \|\rho_i^0\|_{L^1} \leqslant e^{C_i T}\|\rho_i^0\|_{L^1}.$$

Then (4.9) follows from our assumption (4.2) on the interactions. □

Combining (4.8) and (4.9), we deduce

Proposition 4.6. *For all $T > 0$, there exists M_T such that for all $t \in [0,T]$,*

$$\|\rho_{i,h}(t)\|_{L^\infty}, \|\tilde{\rho}_{i,h}(t)\|_{L^\infty} \leqslant M_T.$$

Then, there exists $c_i \equiv c(M_T, U_i) \geq 0$, such that, for all $k \leqslant \lfloor T/h \rfloor$ and $h \leqslant h_0(U_1,U_2)$,

$$(1 - c_i h)\rho_{i,h}^{k+1/2} \leqslant \rho_{i,h}^{k+1}.$$

In particular, there exist $M'_T > 0$ such that for all $t \in [0,T]$,

$$\|\rho_{i,h}(t) - \tilde{\rho}_{i,h}(t)\|_{L^1} \leqslant h M'_T.$$

Proof. The first L^∞ estimate can be found in Lemma 2 from [38], and the rest of our statement can be proved exactly as in Lemma 3.4 and Proposition 3.5. □

4.2. Estimates and convergences

Since we proved that $V_1[\rho_{1,h}, \rho_{2,h}]$ and $V_2[\rho_{1,h}, \rho_{2,h}]$ are bounded in $L^\infty([0,T], W^{1,\infty}(\Omega))$, we can argue exactly as in Section 3.2 for the Wasserstein step and obtain

$$\frac{1}{4h}\mathtt{W}^2(\rho_{i,h}^{k+1/2}, \rho_{i,h}^k) \leqslant \mathcal{F}_i(\rho_{i,h}^k) - \mathcal{F}_i(\rho_{i,h}^{k+1/2}) + C_T h, \tag{4.10}$$

see (3.14) and (3.15) for details. Since $\tilde{\rho}_{1,h}$ and $\tilde{\rho}_{2,h}$ are uniformly bounded in $L^1(\Omega)$ (Lem. 4.5), our assumption (4.4) ensures that $U_1[\rho_{1,h}^{k+1/2}, \rho_{2,h}^{k+1/2}]$ and $U_2[\rho_{1,h}^{k+1/2}, \rho_{2,h}^{k+1/2}]$ are uniformly bounded in $L^\infty(\Omega)$. Proposition 4.6 then allows to argue exactly as in (3.18) and (3.19) for the Fisher-Rao step, and we get

$$\frac{1}{2h}\mathtt{FR}^2(\rho_{i,h}^{k+1}, \rho_{i,h}^{k+1/2}) \leqslant \mathcal{G}_i(\rho_{i,h}^{k+1/2}) - \mathcal{G}_i(\rho_{i,h}^{k+1}) + C_T h. \tag{4.11}$$

In order to retrieve the total square distance estimate we need to control the energy increments as before:

Proposition 4.7. *For all $T > 0$ and $i \in \{1, 2\}$, there exist constants $C_T, C'_T > 0$ such that, for all $k \geqslant 0$ with $hk \leqslant T$,*

$$\mathcal{F}_i(\rho_{i,h}^{k+1}) \leqslant \mathcal{F}_i(\rho_{i,h}^{k+1/2}) + C_T h,$$
$$\mathcal{G}_i(\rho_{i,h}^{k+1/2}) \leqslant \mathcal{G}_i(\rho_{i,h}^{k+1}) + C'_T h.$$

Proof. The dissipation of $\mathcal{F}_i$ along the Fisher-Rao step is obtained in the same way as Proposition 3.6, the increment of $\mathcal{G}_i$ along the Fisher-Rao step is obtained arguing as in (3.20) and we omit the details. □

From (4.10) and (4.11) this immediately gives a telescopic sum

$$\frac{1}{2h}\left(\mathtt{W}^2(\rho_{i,h}^k, \rho_{i,h}^{k+1/2}) + \mathtt{FR}^2(\rho_{i,h}^{k+1/2}, \rho_{i,h}^k)\right) \leqslant 2[\mathcal{F}_i(\rho_{i,h}^k) - \mathcal{F}_i(\rho_{i,h}^{k+1})] + C_T h$$

which in turn yields an approximate $\frac{1}{2}$-Hölder estimate as in Proposition 3.7:

$$\frac{1}{h}\sum_{k=0}^{N-1} \mathtt{WFR}^2(\rho_{i,h}^k, \rho_{i,h}^{k+1}) \leqslant 8\mathcal{F}_i(\rho_i^0) + C_T.$$

The rest of the proof of Theorem 4.2 is then identical to Section 3 and we omit the details.

4.3. Numerical application: prey-predator systems

Our constructive scheme can be implemented numerically, by simply discretizing (4.6) in space. We use the augmented Lagrangian method ALG-JKO from [6] to solve the Wasserstein step, and the Fisher-Rao step is just a convex pointwise minimization problem. Indeed, it is known [18, 28] that $\mathtt{FR}^2(\rho, \mu) = 4\|\sqrt{\rho} - \sqrt{\mu}\|_{L^2}^2$, hence the Fisher-Rao step in (4.6) is a mere convex pointwise minimization problem of the form: for all $x \in \Omega$ (and omitting all indexes $\rho_{i,h}$),

$$\rho^{k+1}(x) = \operatorname*{argmin}_{\rho \in \mathbb{R}^+} \left\{ 4\left|\sqrt{\rho} - \sqrt{\rho^{k+1/2}(x)}\right|^2 + 2hF(\rho) \right\}.$$

This is easily solved using any simple Newton procedure.

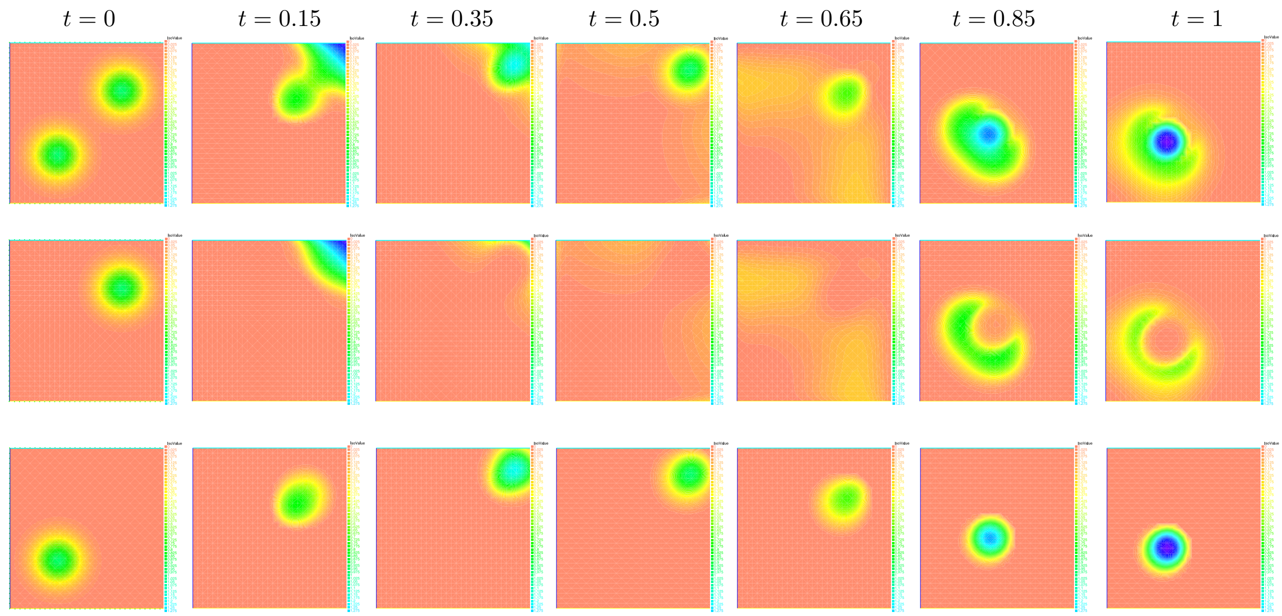


FIGURE 1. Evolution of two species with prey-predator interactions. *First row:* total density $\rho_1 + \rho_2$. *Second row:* prey density ρ_1. *Third row:* predator density ρ_1.

Figure 1 shows the numerical solution of the following diffusive prey-predator system

$$\begin{cases} \partial_t \rho_1 - \Delta \rho_1 - \operatorname{div}(\rho_1 \nabla V_1[\rho_1, \rho_2]) = A\rho_1 (1 - \rho_1) - B\frac{\rho_1 \rho_2}{1+\rho_1}, \\ \partial_t \rho_2 - \Delta \rho_2 - \operatorname{div}(\rho_2 \nabla V_2[\rho_1, \rho_2]) = \frac{B\rho_1 \rho_2}{1+\rho_1} - C\rho_2. \end{cases}$$

Here the ρ_1 species are preys and ρ_2 are predators, see for example [31], the parameters $A = 10, C = 5, B = 70$, and the interactions are chosen as

$$V_1[\rho_1, \rho_2] = |x|^2 * \rho_1 - |x|^2 * \rho_2, \quad V_2[\rho_1, \rho_2] = |x|^2 * \rho_1 + |x|^2 * \rho_2.$$

In (4.1) this corresponds to

$$G_1(\rho_1) = A\frac{\rho_1^2}{2}, \quad G_2(\rho_2) = 0, \quad U_1[\rho_1, \rho_2] = \frac{B\rho_2}{1+\rho_1} - A, \quad U_2[\rho_1, \rho_2] = -\frac{B\rho_1}{1+\rho_1} + C.$$

Of course, U_1 and U_2 satisfy assumptions (4.3) and (4.4), and then Theorem 4.2 gives a solution of the prey-predator system. As before, we shall disregard the uniqueness issue for the sake of simplicity. Figure 2 depicts the time evolution of the L^1 mass for each prey/predator species: we observe the usual oscillations in time with phase opposition, a characteristic behaviour for Lotka-Volterra types of systems.

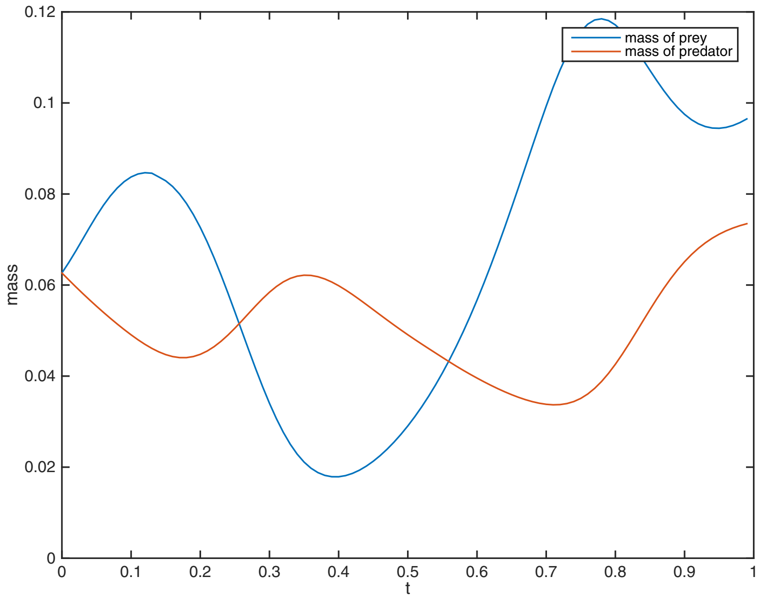


FIGURE 2. Mass evolution for two-species prey-predator interactions.

5. Application to a tumor growth model with very degenerate diffusion and singular energy

In this section we consider the problem

$$\begin{cases} \partial_t \rho = \operatorname{div}(\rho \nabla p) + \rho(1-p), \\ p \geqslant 0 \quad \text{and} \quad p(1-\rho) = 0 \\ 0 \leqslant \rho \leqslant 1, \\ \rho_{|t=0} = \rho^0. \end{cases} \tag{5.1}$$

This is a particular case of the tumor growth models studied in [36, 37], and solutions exhibit a Hele-Shaw patch dynamics: if $\rho^0 = \chi_{\Omega_0}$ then the solution remains an indicator $\rho(t) = \chi_{\Omega(t)}$ and the boundary moves with normal velocity $V = -\nabla p|_{\partial\Omega(t)}$, see [2] for a rigorous analysis in the framework of viscosity solutions. At least formally, we remark that (5.1) is the Wasserstein-Fisher-Rao gradient flow of the singular functional

$$\mathcal{F}(\rho) := \mathcal{F}_\infty(\rho) - \int_\Omega \rho,$$

where

$$\mathcal{F}_\infty(\rho) := \begin{cases} 0 & \text{if } \rho \leqslant 1 \text{ a.e}, \\ +\infty & \text{otherwise}. \end{cases}$$

Indeed, the compatibility conditions $p \geqslant 0$ and $p(1-\rho)=0$ in (5.1) really mean that the pressure p belongs to the subdifferential $\partial\mathcal{F}_\infty(\rho)$, and (5.1) thus reads as the gradient flow

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\operatorname{grad}_{\mathtt{WFR}}\mathcal{F}(\rho) \quad \leftrightarrow \quad \begin{cases}\partial_t\rho = \operatorname{div}(\rho\nabla u) - \rho u\\ u = p-1 \in -\partial\mathcal{F}(\rho).\end{cases}$$

However, this functional is too singular for the previous splitting scheme to correctly capture the very degenerate diffusion. Indeed, the naive and direct approach from Section 3 would lead to

$$\begin{cases}\rho_h^{k+1/2} \in \underset{\rho\leqslant 1,\,|\rho|=|\rho_h^k|}{\operatorname{argmin}} \left\{\frac{1}{2h}\mathtt{W}^2(\rho,\rho_h^k) - \int_\Omega \rho\right\},\\[2ex] \rho_h^{k+1} \in \underset{\rho\leqslant 1}{\operatorname{argmin}}\left\{\frac{1}{2h}\mathtt{FR}^2(\rho,\rho_h^{k+1/2}) - \int_\Omega\rho\right\}.\end{cases}$$

Since the Wasserstein step is mass-conservative by definition, the $\int\rho$ term has no effect in the first step and the latter reads as "project ρ_h^k on $\{\rho\leqslant 1\}$ w.r.t to the $\mathtt{W}$ distance". Since the output of the reaction step $\rho_h^{k+1}\leqslant 1$, the Wasserstein step will never actually project anything, and the diffusion is completely shut down. As an example, it is easy to see that if the initial datum is an indicator $\rho^0=\chi_{\Omega_0}$ then the above naive scheme leads to a stationary solution $\rho_h^{k+1}=\rho_h^{k+1/2}=\rho^0$ for all $k\geqslant 0$, while the real solution should evolve according to the aforementioned Hele-Shaw dynamics $\rho(t)=\chi_{\Omega(t)}$ [2, 36]. One could otherwise try to write a semi-implicit scheme as follows: (1) keep the projection on $\{\rho\leqslant 1\}$ in the first Wasserstein step. As in [30] a pressure term $p_h^{k+1/2}$ appears as a Lagrange multiplier in the Wasserstein projection; (2) in the $\mathtt{FR}$/reaction step, relax the constraint $\rho\leqslant 1$ and minimize instead $\rho^{k+1}\in\operatorname{argmin}\left\{\frac{1}{2h}\mathtt{FR}^2(\rho)+\int\rho p^{k+1/2}-\int\rho\right\}$, and keep iterating. This seems to correctly capture the diffusion at least numerically speaking, but raises technical issues in the rigorous proof of convergence and most importantly destroys the variational structure at the discrete level (due to the fact that the reaction step becomes semi-explicit).

We shall use instead an approximation procedure, which preserves the variational structure at the discrete level: it is well-known that the Porous-Medium functional

$$\mathcal{F}_m(\rho) := \begin{cases}\int_\Omega \frac{\rho^m}{m-1} & \text{if } \rho^m\in L^1(\Omega)\\ +\infty & \text{otherwise}\end{cases}$$

Γ-converges to $\mathcal{F}_\infty$ in the incompressible limit $m\to\infty$, see [7]. In the spirit of [42], one should therefore expect that the gradient flow ρ_m of $\mathcal{F}_m(\rho)-\int\rho$ converges to the gradient flow ρ_∞ of the limiting functional $\mathcal{F}(\rho)=\mathcal{F}_\infty(\rho)-\int\rho$. Implementing the splitting scheme for the regular energy functional $\mathcal{F}_m(\rho)-\int\rho$ gives a sequence $\rho_{h,m}$, and we shall prove below that $\rho_{h,m}$ converges to a solution of the limiting gradient flow as $m\to\infty$ and $h\to 0$. However, it is known [17] that the limit depends in general on the interplay between the time-step h and the regularization parameter ($m\to\infty$ here), and for technical reasons we shall enforce the CFL-like condition

$$mh\to 0 \quad \text{as} \quad m\to\infty \text{ and } h\to 0.$$

Note that [36] already contained a similar approximation $m\to\infty$ but without exploiting the variational structure of the m- gradient flow, and our approach is thus different. The above gradient-flow structure was already noticed and fully exploited in the independent work [10], where existence and uniqueness of weak solutions is proved and numerical simulations are performed needless of any splitting an using directly the $\mathtt{WFR}$ structure (with an additional entropic approximation). Here we rather emphasize the fact that the splitting does capture delicate Γ-convergence phenomena.

In order to make the above discussion more rigorous, we fix a time step $h > 0$ and construct two sequences $(\rho_{h,m}^{k+1/2})_k$ and $(\rho_{h,m}^k)_k$, with $\rho_{h,m}^0 = \rho^0$, defined recursively as

$$\begin{cases} \rho_h^{k+1/2} \in \underset{\rho\in\mathcal{M}^+,\,|\rho|=|\rho_h^k|}{\operatorname{argmin}} \left\{ \frac{1}{2h}\mathtt{W}^2(\rho,\rho_{h,m}^k) + \mathcal{F}_m(\rho) - \int_\Omega \rho \right\}, \\ \rho_h^{k+1} \in \underset{\rho\in\mathcal{M}^+}{\operatorname{argmin}} \left\{ \frac{1}{2h}\mathtt{FR}^2(\rho,\rho_h^{k+1/2}) + \mathcal{F}_m(\rho) - \int_\Omega \rho \right\}. \end{cases} \tag{5.2}$$

As is common in the classical theory of Porous Media equations [44], we define the pressure as the first variation

$$p_m := F_m'(\rho) = \frac{m}{m-1}\rho^{m-1},$$

and we write accordingly

$$p_{h,m}^{k+1/2} := \frac{m}{m-1}(\rho_{h,m}^{k+1/2})^{m-1} \quad \text{and} \quad p_{h,m}^{k+1} := \frac{m}{m-1}(\rho_{h,m}^{k+1})^{m-1}$$

for the discrete pressures. As in Section 3 we denote by $\rho_{h,m}(t), p_{h,m}(t)$ and $\tilde\rho_{h,m}(t), \tilde p_{h,m}(t)$ the piecewise constant interpolations of $\rho_{h,m}^{k+1}, p_{h,m}^{k+1}$ and $\rho_{h,m}^{k+1/2}, p_{h,m}^{k+1/2}$, respectively.

Our main result is

Theorem 5.1. *Assume that $\rho^0 \in BV(\Omega)$, $\rho^0 \leqslant 1$, and $mh \to 0$ as $h \to 0$ and $m \to \infty$. Then for all $T > 0$, $\rho_{h,m}, \tilde\rho_{h,m}$ both converge to some ρ strongly in $L^1((0,T)\times\Omega)$, the pressures $p_{h,m}, \tilde p_{h,m}$ both converge to some p weakly in $L^2((0,T),H^1(\Omega))$, and (ρ,p) is the unique weak solution of (5.1).*

Since we have a WFR gradient-flow structure, uniqueness should formally follows from the -1 geodesic convexity of the driving functional $\mathcal{E}_\infty(\rho) - \int_\Omega \rho$ with respect to the WFR distance [25, 27] and the resulting contractivity estimate $\mathtt{WFR}(\rho^1(t),\rho^2(t)) \le e^t \mathtt{WFR}(\rho^{0,1},\rho^{0,2})$ between any two solutions ρ^1,ρ^2. This is proved rigorously in [10], and therefore we retrieve convergence of the whole sequence $\rho_{h,m} \to \rho$ in Theorem 5.1 (and not only for subsequences). Given this uniqueness, it is clearly enough to prove convergence along any discrete (sub)sequence, and this is exactly what we show below.

The strategy of proof for Theorem 5.1 is exactly as in Section 3, except that we need now the estimates to be uniform in both $h \to 0$ and $m \to \infty$.

5.1. Estimates and convergences

In this section, we improve the previous estimates from Section 3. We start with a universal L^∞-bound:

Lemma 5.2. *Assume that $\rho^0 \leqslant 1$, then for all $t \ge 0$,*

$$\|\rho_{h,m}(t,\cdot)\|_\infty, \|\tilde\rho_{h,m}(t,\cdot)\|_\infty \leqslant 1.$$

Proof. We argue by induction at the discrete level, starting from $\rho^0 = \rho_{h,m}^0 \leqslant 1$ by assumption. Since $\|\rho_{h,m}^k\|_\infty \leqslant 1$, Otto's maximum principle (3.9) implies that $\|\rho_{h,m}^{k+1/2}\|_\infty \leqslant \|\rho_{h,m}^k\|_\infty \leqslant 1$ in the Wasserstein step.

We establish now $\|\rho_{h,m}^{k+1}\|_\infty \leqslant \|\rho_{h,m}^{k+1/2}\|_\infty$. Assume by contradiction that $E := \{\rho_{h,m}^{k+1} > 1\}$ has positive Lebesgue measure. The optimality condition (3.10) for the Fisher-Rao minimization step gives, dividing by

$\sqrt{\rho_{h,m}^{k+1}} > 0$ in E,

$$\sqrt{\rho_{h,m}^{k+1}} - \sqrt{\rho_{h,m}^{k+1/2}} = \frac{h}{2}\sqrt{\rho_{h,m}^{k+1}}\left(1 - \frac{m}{m-1}(\rho_{h,m}^{k+1})^{m-1}\right)$$

Then $1 - \frac{m}{m-1}(\rho_{h,m}^{k+1})^{m-1} \leqslant 1 - \frac{m}{m-1} < 0$ in the right-hand side, hence $\sqrt{\rho_{h,m}^{k+1}} - \sqrt{\rho_{h,m}^{k+1/2}} \le 0$ in E. This is the desired contradiction since $\rho_{h,m}^{k+1} > 1$ in E but $\rho_{h,m}^{k+1/2} \leqslant 1$ everywhere. □

Noticing that the functional $\frac{1}{m-1}\int \rho^m - \int \rho$ corresponds to taking explicitly $F_2(z) = z^m/m - 1$ and $V_2(x) \equiv -1$ in Section 3, it is easy to reproduce the computations from the proof of Lemma 3.4 and carefully track the dependence of the constants w.r.t $m > 1$ to obtain

Lemma 5.3. *There exists $c > 0$ such that, for all $m > m_0$ large enough and all $h \le h_0$ small enough,*

$$(1 - ch)\rho_{h,m}^{k+1/2}(x) \leqslant \rho_{h,m}^{k+1}(x) \leqslant (1 + h)\rho_{h,m}^{k+1/2}(x) \quad a.e. \tag{5.3}$$

Note that this holds regardless of any compatibility such as $hm \to 0$. The key point is here that the lower bound c previously depended on an upper bound $\rho^{k+1/2} \le M$ in Lemma 3.4, but since we just obtained in Lemma 5.2 the universal upper bound $\rho^{k+1/2} \leqslant 1$ we end up with a lower bound which is also uniform in h, m. The proof is identical to that of Lemma 3.4 and we omit the details for simplicity.

Recalling that the Wasserstein step is mass-preserving, we obtain by immediate induction and for all $0 \le t \le T$

$$\|\rho_{h,m}(t)\|_{L^1},\ \|\tilde\rho_{h,m}(t)\|_{L^1} \leqslant e^T\|\rho^0\|_{L^1}$$

as well as

$$\|\rho_{h,m}(t) - \tilde\rho_{h,m}(t)\|_{L^1} \leqslant C_T h. \tag{5.4}$$

Testing successively $\rho = \rho_{h,m}^k$ and $\rho = \rho_{h,m}^{k+1/2}$ in (5.2), we get

$$\frac{1}{2h}\left(\mathtt{W}^2(\rho_{h,m}^k, \rho_{h,m}^{k+1/2}) + \mathtt{FR}^2(\rho_{h,m}^{k+1/2}, \rho_{h,m}^{k+1})\right) \leqslant \mathcal{F}_m(\rho_{h,m}^k) - \mathcal{F}_m(\rho_{h,m}^{k+1}) + \int_\Omega (\rho_{h,m}^{k+1/2} - \rho_{h,m}^{k+1}).$$

Using Proposition 2.4 to control $\mathtt{WFR}^2 \lesssim 2(\mathtt{W}^2 + \mathtt{FR}^2)$ and the lower bound in (5.3) yields

$$\begin{aligned}\frac{1}{4h}\mathtt{WFR}^2(\rho_{h,m}^{k+1}, \rho_{h,m}^k) &\leqslant \frac{1}{2h}\left(\mathtt{W}^2(\rho_{h,m}^k, \rho_{h,m}^{k+1/2}) + \mathtt{FR}^2(\rho_{h,m}^{k+1/2}, \rho_{h,m}^{k+1})\right) \leqslant \mathcal{F}_m(\rho_{h,m}^k) - \mathcal{F}_m(\rho_{h,m}^{k+1}) + \int_\Omega (\rho_{h,m}^{k+1/2} - \rho_{h,m}^{k+1}) \\ &\leqslant \mathcal{F}_m(\rho_{h,m}^k) - \mathcal{F}_m(\rho_{h,m}^{k+1}) + ch\int_\Omega \rho_{h,m}^{k+1/2} \leqslant \mathcal{F}_m(\rho_{h,m}^k) - \mathcal{F}_m(\rho_{h,m}^{k+1}) + che^T\end{aligned}$$

for all $k \leqslant N := \lfloor T/h \rfloor$. Summing over k we get

$$\begin{aligned}\frac{1}{4h}\sum_{k=0}^{N-1}\mathtt{WFR}^2(\rho_{h,m}^k, \rho_{h,m}^{k+1}) &\leqslant \mathcal{F}_m(\rho^0) - \mathcal{F}_m(\rho_{h,m}^N) + C_T \\ &\leqslant \frac{1}{m-1}\int_\Omega (\rho^0)^m + C_T \leqslant \frac{1}{m-1}\int_\Omega \rho^0 + C_T \leqslant C_T,\end{aligned}$$

where we used successively $F_m \geq 0$ to get rid of $\mathcal{F}_m(\rho^N_{h,m})$, and $(\rho^0)^m \leq \rho^0$ for $\rho^0 \leq 1$ and $m > 1$. Consequently, for all fixed $T > 0$ and any $t, s \in [0, T]$ we obtain the classical $\frac{1}{2}$-Hölder estimate

$$\begin{cases} \mathtt{WFR}(\rho_{h,m}(t), \rho_{h,m}(s)) \leqslant C_T|t-s+h|^{1/2}, \\ \mathtt{WFR}(\tilde{\rho}_{h,m}(t), \tilde{\rho}_{h,m}(s)) \leqslant C_T|t-s+h|^{1/2}. \end{cases} \tag{5.5}$$

Exploiting the explicit algebraic structure of $F_m(z) = \frac{1}{m-1}z^m$, compactness in space will be given here by

Lemma 5.4. *If $\rho^0 \in BV(\Omega)$ then*

$$\sup_{t\in[0,T]} \left\{ \|\rho_{h,m}(t,\cdot)\|_{BV(\Omega)}, \|\tilde{\rho}_{h,m}(t,\cdot)\|_{BV(\Omega)} \right\} \leqslant e^T \|\rho^0\|_{BV(\Omega)}.$$

Proof. The argument closely follows the lines of Proposition 5.1 from [18], and strongly exploits the following result:

Theorem 5.5 (BV estimates for JKO minimizers, Theorem 1.1 from [14])**.** *Let $F : \mathbb{R}^+ \to \mathbb{R} \cup \{+\infty\}$ be convex and l.s.c, and $g \in \mathcal{P}_2 \cap BV(\Omega)$. If ρ^* is a minimizer of the variational problem*

$$\min_\rho \left\{ \frac{1}{2h} \mathtt{W}^2(\rho, g) + \int F(\rho(x))\mathrm{d}x \right\},$$

then $\rho^* \in BV(\Omega)$ as well and

$$\int_\Omega |\nabla \rho^*| \leq \int_\Omega |\nabla g|$$

in the sense of measures.

This type of BV estimates is highly non trivial and cannot be simply sketched here, but one can think of this as a quantified smoothing effect for the (formally) parabolic PDE $\partial_t \rho = \mathrm{div}(\rho \nabla F'(\rho))$, of which the first Wasserstein step in (5.2) is a time-discretization.

Applied to the minimizer $\rho^{k+1/2}_{h,m}$ with the convex functional $F_m(\rho) = \frac{\rho^m}{m-1}$, this estimate immediately gives, together with the mass conservation $\|\rho^{k+1/2}_{h,m}\|_{L^1} = \|\rho^k_{h,m}\|_{L^1}$, the BV bound

$$\|\rho^{k+1/2}_{h,m}\|_{BV(\Omega)} \leqslant \|\rho^k_{h,m}\|_{BV(\Omega)}.$$

Using as before the implicit function theorem, we show below that $\rho^{k+1}_{h,m} = R(\rho^{k+1/2}_{h,m})$ for some suitable $(1+h)$-Lispchitz function R. By standard $Lip \circ BV$ composition [3] this will prove that

$$\|\rho^{k+1}_{h,m}\|_{BV(\Omega)} \leqslant (1+h)\|\rho^{k+1/2}_{h,m}\|_{BV(\Omega)}$$

and will conclude the proof by immediate induction.

Indeed, we already know from (5.3) that $\rho^{k+1/2}_{h,m}$ and $\rho^{k+1}_{h,m}$ share the same support. In this support and from (3.10) it is easy to see that $\rho = \rho^{k+1}_{h,m}(x)$ is the unique positive solution of $f(\rho, \rho^{k+1/2}_{h,m}(x)) = 0$ with

$$f(\rho, \mu) = \sqrt{\rho}\left(1 - \frac{h}{2}\left(1 - \frac{m}{m-1}\rho^{m-1}\right)\right) - \sqrt{\mu}.$$

For $\mu > 0$, the implicit function theorem gives the existence of a $\mathcal{C}^1$ map R such that $f(\rho,\mu) = 0 \Leftrightarrow \rho = R(\mu)$, with $R(0) = 0$. An algebraic computation shows moreover that $0 < \frac{dR}{d\mu} = -\frac{\partial_\mu f}{\partial_\rho f}_{|\rho=R(\mu)} \leqslant (1+h)$ uniformly in $m > 1$, hence R is $(1+h)$-Lipschitz as claimed and the proof is complete. □

Proposition 5.6. *Up to extraction of a discrete sequence $h \to 0, m \to \infty$, there holds*

$$\rho_{h,m},\, \tilde{\rho}_{h,m} \to \rho \quad \textit{strongly in } L^1(Q_T)$$

$$p_{h,m} \rightharpoonup p \quad \textit{and} \quad \tilde{p}_{h,m} \rightharpoonup \tilde{p} \quad \textit{weakly in all } L^q(Q_T)$$

for all $T > 0$. If in addition $mh \to 0$ then $p = \tilde{p}$.

Proof. The first part of the statement follows exactly as in Section 3, exploiting the $\frac{1}{2}$-Hölder estimates (5.5) and the space compactness from Proposition 5.4 in order to apply the Rossi-Savaré theorem [41]. The fact that $\rho_{h,m}, \tilde{\rho}_{h,m}$ have the same limit comes from (5.4).

For the pressures, we simply note from $\rho_{h,m} \leqslant 1$ and $m \gg 1$ that $p_{h,m} = \frac{m}{m-1}\rho_{h,m}^{m-1} \leqslant 2\rho_{h,m}$ is bounded in $L^1 \cap L^\infty(Q_T)$ uniformly in h, m in any finite time interval $[0,T]$. Thus up to extraction of a further sequence we have $p_{h,m} \rightharpoonup p$ in all $L^q(Q_T)$, and likewise for $\tilde{p}_{h,m} \rightharpoonup \tilde{p}$.

Finally, we only have to check that $p = \tilde{p}$, and this is precisely why we need $hm \to 0$. Because $\rho_{h,m}, \tilde{\rho}_{h,m} \leqslant 1$ and $z \mapsto z^{m-1}$ is $(m-1)$-Lipschitz on $[0,1]$ we have for all fixed $t \geqslant 0$ that

$$\begin{aligned}\int_\Omega |p_{m,h}(t,\cdot) - \tilde{p}_{m,h}(t,\cdot)| &= \int_\Omega \frac{m}{m-1}|\rho_{h,m}^{m-1}(t,\cdot) - \tilde{\rho}_{h,m}^{m-1}(t,\cdot)| \\ &\leqslant m\int_\Omega |\rho_{h,m}(t) - \tilde{\rho}_{h,m}(t)| \leqslant C_T hm \longrightarrow 0,\end{aligned}$$

where we used (5.4) to control $\|\rho_{h,m}(t) - \tilde{\rho}_{h,m}(t)\|_{L^1}$ in the last inequality. Hence $p = \tilde{p}$ and the proof is complete. □

In order to pass to the limit in the diffusion term $\operatorname{div}(\rho\nabla p)$ we first improve the convergence of $\tilde{p}_{h,m}$:

Lemma 5.7. *There exists a constant C_T, independent of h and m, such that*

$$\|\tilde{p}_{h,m}\|_{L^2((0,T),H^1(\Omega))} \leqslant C_T$$

for all $T > 0$. Consequently, up to a subsequence, $\tilde{p}_{h,m}$ converges weakly in $L^2((0,T),H^1(\Omega))$ to p.

Proof. The proof is based on the flow interchange technique developed by Matthes et al. in [29]. For fixed small $\varepsilon > 0$ let $\eta(s,x)$ be the (smooth) solution of

$$\begin{cases}\partial_s\eta = \Delta\eta^{m-1} + \varepsilon\Delta\eta, \\ \eta|_{s=0} = \rho_{h,m}^{k+1/2}\end{cases}$$

with Neumann boundary conditions. It is well known [4] that η is the Wasserstein gradient flow of

$$\mathcal{G}(\rho) := \int_\Omega \frac{\rho^{m-1}}{m-2} + \varepsilon\int_\Omega \rho\log(\rho).$$

Since $\mathcal{G}$ is geodesically 0-convex (in the whole space $\Omega = \mathbb{R}^d$), η satisfies the Evolution Variational Inequality (EVI)

$$\frac{1}{2}\frac{d^+}{ds}\mathtt{W}^2(\eta(s),\rho) \leqslant \mathcal{G}(\rho) - \mathcal{G}(\eta(s)), \tag{5.6}$$

for all $s > 0$ and for all $\rho \in \mathcal{P}^{\mathrm{ac}}(\Omega)$, where the upper derivative $\frac{d^+}{ds}f(s) := \limsup_{\tau\to 0^+} \frac{f(s+\tau)-f(s)}{\tau}$. By optimality of $\rho_{h,m}^{k+1/2} = \eta_{|s=0}$ in (5.2), we obtain that

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}s}\bigg|_{s=0} \mathtt{W}^2(\eta(s),\rho_{h,m}^k) \geqslant -h\,\frac{\mathrm{d}^+}{\mathrm{d}s}\bigg|_{s=0} \mathcal{F}_m(\eta(s)). \tag{5.7}$$

Since η is smooth due to the regularizing $\varepsilon\Delta$ term, we can legitimately integrate by parts for all $s > 0$

$$\begin{aligned}
\frac{d}{ds}\mathcal{F}_m(\eta(s)) &= \int_\Omega \frac{m}{m-1}\eta(s)^{m-1}(\Delta\eta(s)^{m-1} + \varepsilon\Delta\eta(s)) \\
&= -\int_\Omega \frac{m}{m-1}|\nabla\eta(s)^{m-1}|^2 - \varepsilon\int_\Omega m\eta(s)^{m-2}|\nabla\eta(s)|^2 \\
&\leqslant -\int_\Omega \frac{m}{m-1}|\nabla\eta(s)^{m-1}|^2 = -\frac{m-1}{m}\int_\Omega \left|\nabla\left(\frac{m}{m-1}\eta(s)^{m-1}\right)\right|^2
\end{aligned}$$

Remarking that $\frac{m}{m-1}\eta(s)^{m-1} \to \frac{m}{m-1}\left(\rho_{h,m}^{k+1/2}\right)^{m-1} = p_{h,m}^{k+1/2}$ in the right-hand side above as $s \to 0$, a classical lower semi-continuity argument gives that

$$\int_\Omega \frac{m-1}{m}|\nabla p_{h,m}^{k+1/2}|^2 = \int_\Omega \frac{m}{m-1}|\nabla(\rho_{h,m}^{k+1/2})^{m-1}|^2 \leqslant -\liminf_{s\searrow 0}\frac{d^+}{ds}\mathcal{F}_m(\eta(s)).$$

Then from the chain of inequalities (5.6) (and strong continuity $\mathcal{G}(\eta(s)) \to \mathcal{G}(\eta(0))$) we have

$$\begin{aligned}
h\int_\Omega \frac{m-1}{m}|\nabla p_{h,m}^{k+1/2}|^2 &\leqslant \mathcal{F}_{m-1}(\rho_{h,m}^k) - \mathcal{F}_{m-1}(\rho_{h,m}^{k+1/2}) \\
&\quad + \varepsilon\left(\int_\Omega \rho_{h,m}^k\log(\rho_{h,m}^k) - \int_\Omega \rho_{h,m}^{k+1/2}\log(\rho_{h,m}^{k+1/2})\right).
\end{aligned}$$

First arguing as in Proposition 3.6 to control

$$\mathcal{F}_{m-1}(\rho_{h,m}^{k+1}) \leqslant \mathcal{F}_{m-1}(\rho_{h,m}^{k+1/2}) + C_T h,$$

and then passing to the limit $\varepsilon \searrow 0$, we obtain

$$h\int_\Omega \frac{m-1}{m}|\nabla p_{h,m}^{k+1/2}|^2 \leqslant \mathcal{F}_{m-1}(\rho_{h,m}^k) - \mathcal{F}_{m-1}(\rho_{h,m}^{k+1}) + C_T h.$$

Summing over k gives

$$\int_0^T\int_\Omega |\nabla\tilde{p}_{h,m}(t,x)|^2 \leqslant \frac{m}{m-1}\left(\mathcal{F}_{m-1}(\rho^0) - \mathcal{F}_{m-1}(\rho_{h,m}^N) + C_T\right) \leqslant 2\mathcal{F}_{m-1}(\rho^0) + C_T$$

for all $T < +\infty$. Due to $\rho^0 \leqslant 1$ and $m \gg 1$ we can bound $\mathcal{F}_{m-1}(\rho^0) = \frac{1}{m-2}\int(\rho^0)^{m-1} \leqslant \frac{1}{m-2}\int \rho^0 \leqslant \|\rho^0\|_{L^1(\Omega)}$ and the result finally follows. □

5.2. Properties of the pressure p and conclusion

We start by showing that the limits ρ, p satisfy the compatibility conditions in (5.1).

Lemma 5.8. *There holds*

$$0 \leqslant \rho, p \leqslant 1 \quad \textit{and} \quad p(1-\rho) = 0 \ \textit{ a.e. in } Q_T.$$

Proof. By Lemma 5.2 it is obvious that $0 \leqslant \rho \leqslant 1$ and $0 \leqslant p \leqslant 1$ are inherited from $0 \leqslant \rho_{h,m} \leqslant 1$ and $0 \leqslant p_{h,m} = \frac{m}{m-1}\rho_{h,m}^{m-1} \leqslant \frac{m}{m-1}$. In order to prove that $p(1-\rho) = 0$, we first claim that

$$p_{h,m}(1-\rho_{h,m}) \to 0 \quad \text{a.e. in } Q_T.$$

Indeed, since $\rho_{h,m} \to \rho$ strongly in $L^1(Q_T)$ we can assume $\rho_{h,m}(t,x) \to \rho(t,x)$ a.e. If the limit $\rho(t,x) < 1$ then $\rho_{h,m}(t,x) \leqslant (1-\varepsilon)$ for all small h and large m, and some $\varepsilon = \varepsilon(t,x) > 0$. Hence $p_{h,m}(t,x) = \frac{m}{m-1}\rho_{h,m}^{m-1} \leqslant \frac{m}{m-1}(1-\varepsilon)^{m-1} \to 0$ while $1-\rho_{h,m}$ remains bounded, and therefore the product $p_{h,m}(1-\rho_{h,m}) \to 0$. Now if the limit $\rho(t,x) = 1$ then the pressure $p_{h,m} = \frac{m}{m-1}\rho_{h,m}^{m-1} \leqslant \frac{m}{m-1}$ remains bounded, while $1-\rho_{h,m}(t,x) \to 0$ hence the product goes to zero in this case too.

Thanks to the uniform L^∞ bounds $\rho_{h,m} \leqslant 1$ and $p_{h,m} \leqslant \frac{m}{m-1} \leqslant 2$ we can apply Lebesgue's convergence theorem to deduce from this pointwise a.e. convergence that, for all fixed nonnegative $\varphi \in C_c^\infty(Q_T)$, there holds

$$\lim_{h,m} \int_{Q_T} p_{h,m}(1-\rho_{h,m})\varphi = 0$$

(the test function φ is only needed to localize). On the other hand since $\rho_{h,m} \to \rho$ strongly in $L^1(Q_T)$ hence a.e, and because $0 \leqslant \rho_{h,m} \leqslant 1$, we see that $(1-\rho_{h,m})\varphi \to (1-\rho)\varphi$ in all $L^q(Q_T)$. From Proposition 5.6 we also had that $p_{h,m} \rightharpoonup p$ in all $L^q(Q_T)$, hence by strong-weak convergence we have that

$$\int_{Q_T} p(1-\rho)\varphi = \lim \int_{Q_T} p_{h,m}(1-\rho_{h,m})\varphi = 0$$

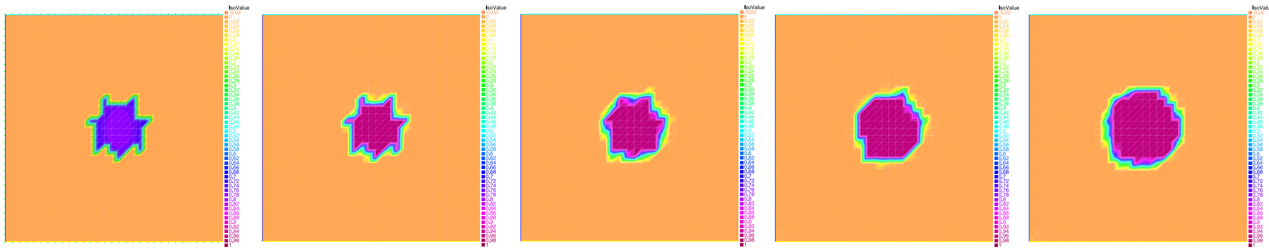
for all $\varphi \geqslant 0$. Because $p(1-\rho) \geqslant 0$ we conclude that $p(1-\rho) = 0$ a.e. in Q_T and the proof is achieved. □

We end this section with

Proof of Theorem 5.1. We only sketch the argument and refer to [18] for the details. Fix any $0 < t_1 < t_2$ and $\varphi \in \mathcal{C}_c^2(\mathbb{R}^d)$. Exploiting the Euler-Lagrange equations (3.6) and (3.10) and summing from $k = k_1 = \lfloor t_1/h \rfloor$ to $k = k_2 - 1 = \lfloor t_2/h \rfloor - 1$, we first obtain

$$\int_{\mathbb{R}^d} \rho_{h,m}(t_2)\varphi - \rho_{h,m}(t_1)\varphi + \int_{k_1 h}^{k_2 h}\int_{\mathbb{R}^d} \tilde{\rho}_{h,m}\nabla\tilde{p}_{h,m}\cdot\nabla\varphi = -\int_{k_1 h}^{k_2 h}\int_{\mathbb{R}^d} \rho_{h,m}(1-p_{h,m})\varphi + R(h,m),$$

where the remainder $R(h,m) \to 0$ for fixed φ. The strong convergence $\rho_{h,m}, \tilde{\rho}_{h,m} \to \rho$, the weak convergences $\nabla\tilde{p}_{h,m} \rightharpoonup \nabla\tilde{p} = \nabla p$ and $p_{h,m} \rightharpoonup p$, and $k_i h \to t_i$ are then enough pass to the limit to get the corresponding weak formulation for all $0 < t_1 < t_2$. Moreover since the limit $\rho \in \mathcal{C}([0,T];\mathcal{M}^+_{\mathtt{WFR}})$ the initial datum $\rho(0) = \rho^0$ is taken at least in the sense of measures. This gives an admissible weak formulation of (5.1), and the proof is complete. □



$t = 0$ $t = 0.3$ $t = 0.5$ $t = 0.7$ $t = 1$

FIGURE 3. Snapshot of the approximate solution $\rho_{h,m}(t,.)$ to (5.1), with $m = 100$, $h = 0.005$.

5.3. Numerical simulation

The constructive scheme (5.2) naturally leads to a fully discrete algorithm, simply discretizing the minimization problem in space for each W, FR step. We use again the ALG2-JKO scheme [6] for the Wasserstein steps. As already mentioned the Fisher-Rao step is a mere convex pointwise minimization problem, here explicitly given by: for all $x \in \Omega$,

$$\rho_{h,m}^{k+1}(x) = \operatorname*{argmin}_{\rho \in \mathbb{R}^+} \left\{ 4 \left| \sqrt{\rho} - \sqrt{\rho_{h,m}^{k+1/2}(x)} \right|^2 + 2h \left(\frac{\rho^m}{m-1} - 1 \right) \right\}$$

and poses no difficulty in the practical implementation using a standard Newton method.

Figure 3 depicts the evolution of the numerical solution $\rho_{h,m}$ for $m = 100$ and with a time step $h = 0.005$. We remark that the tumor first saturates the constraint ($\rho \nearrow 1$) in its initial support, and then starts diffusing outwards. This is consistent with the qualitative behaviour described in [36].

6. A TUMOR GROWTH MODEL WITH NUTRIENT

In this section we use the same approach for the following tumor growth model with nutrients, appearing *e.g.* in [36]

$$\begin{cases} \partial_t \rho - \operatorname{div}(\rho \nabla p) = \rho \left((1-p)(c + c_1) - c_2 \right), \\ \partial_t c - \Delta c = -\rho c, \\ 0 \leqslant \rho \leqslant 1, \\ p \geqslant 0 \text{ and } p(1-\rho) = 0, \\ \rho_{|t=0} = \rho^0, \; c_{|t=0} = c^0. \end{cases} \tag{6.1}$$

Here c_1 and c_2 are two positive constants, and the nutrient c diffuses in Ω while being consumed by the tumor ρ. For technical convenience we work here on a convex bounded domain $\Omega \subset \mathbb{R}^d$, endowed with natural Neumann boundary conditions for both ρ and c.

Contrarily to Section 5 this is not a WFR gradient flow anymore, and we therefore introduce a semi-implicit splitting scheme. We approximate again the singular diffusion by the incompressible limit $m \to \infty$. Starting from the initial data $\rho_{h,m}^0 := \rho^0, c_{h,m}^0 := c^0$ we construct four sequences $\rho_{h,m}^{k+1/2}, \rho_{h,m}^k, c_{h,m}^{k+1/2}, c_{h,m}^k$, defined recursively as

$$\begin{cases} \rho_{h,m}^{k+1/2} \in \underset{\rho\in\mathcal{M}^+,|\rho|=|\rho_{h,m}^k|}{\operatorname{argmin}} \left\{ \frac{1}{2h}\mathtt{W}^2(\rho,\rho_{h,m}^k) + \mathcal{F}_m(\rho) \right\}, \\ \\ c_{h,m}^{k+1/2} \in \underset{c\in\mathcal{M}^+,|c|=|c_{h,m}^k|}{\operatorname{argmin}} \left\{ \frac{1}{2h}\mathtt{W}^2(c,c_{h,m}^k) + \mathcal{H}(\rho) \right\}, \end{cases} \tag{6.2}$$

and

$$\begin{cases} \rho_{h,m}^{k+1} \in \underset{\rho\in\mathcal{M}^+}{\operatorname{argmin}} \left\{ \frac{1}{2h}\mathtt{FR}^2(\rho,\rho_{h,m}^{k+1/2}) + \mathcal{E}_{1,m}(\rho|c_{h,m}^{k+1/2}) \right\}, \\ \\ c_{h,m}^{k+1} \in \underset{c\in\mathcal{M}^+}{\operatorname{argmin}} \left\{ \frac{1}{2h}\mathtt{FR}^2(c,c_{h,m}^{k+1/2}) + \mathcal{E}_2(c|\rho_{h,m}^{k+1/2}) \right\}, \end{cases} \tag{6.3}$$

where

$$\mathcal{H}(\rho) := \int_\Omega \rho\log(\rho),$$

$$\mathcal{E}_{1,m}(\rho|c) := \int_\Omega (c+c_1)\,\frac{\rho^m}{m-1} + \int_\Omega (c_2 - c - c_1)\rho,$$

and

$$\mathcal{E}_2(c|\rho) := \int_\Omega \rho c.$$

As earlier it is easy to see that these sequences are well-defined (*i.e.* there exists a unique minimizer for each step), and the pressures are defined as before as

$$p_{h,m}^{k+1/2} := \frac{m}{m-1}(\rho_{h,m}^{k+1/2})^{m-1} \quad \text{and} \quad p_{h,m}^{k+1} := \frac{m}{m-1}(\rho_{h,m}^{k+1})^{m-1}.$$

We denote again by $a_{h,m}(t), \tilde{a}_{h,m}(t)$ the piecewise constant interpolation of any discrete quantity $a_{h,m}^{k+1}, a_{h,m}^{k+1/2}$ respectively. Our main result reads:

Theorem 6.1. *Assume $\rho^0 \in BV(\Omega)$ with $\rho^0 \leqslant 1$ and $c^0 \in L^\infty(\Omega) \cap BV(\Omega)$. Then $\rho_{h,m}$ and $\tilde{\rho}_{h,m}$ strongly converge to ρ in $L^1((0,T)\times\Omega)$ and $c_{h,m}$ and $\tilde{c}_{h,m}$ strongly converge to c in $L^1((0,T)\times\Omega)$ when $h \searrow 0$ and $m \nearrow +\infty$. Moreover, if $mh \to 0$, then $p_{h,m}, \tilde{p}_{h,m}$ converge weakly in $L^2((0,T), H^1(\Omega))$ to a unique p, and (ρ, p, c) is a solution of (6.1).*

Note that uniqueness of solutions would result in convergence of the whole sequence. Uniqueness was proved in Theorem 4.2 from [36], for slightly more regular weak solutions, but we did not push in this direction for the sake of simplicity. The method of proof is almost identical to Section 5 so we only sketch the argument and emphasize the main differences.

We start by recalling the optimality conditions for the scheme (6.2) and (6.3). The Euler-Lagrange equations for the tumor densities in the Wasserstein and Fisher-Rao steps are

$$\begin{cases}\rho_{h,m}^{k+1/2}\nabla p_{h,m}^{k+1/2} = \frac{\nabla\varphi}{h}\rho_{h,m}^{k+1/2},\\ \sqrt{\rho_{h,m}^{k+1}} - \sqrt{\rho_{h,m}^{k+1/2}} = \frac{h}{2}\sqrt{\rho_{h,m}^{k+1}}\left((1-p_{h,m}^{k+1})(c_{h,m}^{k+1/2}+c_1)-c_2\right),\end{cases} \tag{6.4}$$

where φ is a (backward) Kantorovich potential for $\mathtt{W}(\rho_{h,m}^{k+1/2},\rho_{h,m}^{k})$. For the nutrient, the Euler-Lagrange equations are

$$\begin{cases}\nabla c_{h,m}^{k+1/2} = \frac{\nabla\psi}{h}c_{h,m}^{k+1/2},\\ \sqrt{c_{h,m}^{k+1}} - \sqrt{c_{h,m}^{k+1/2}} = -\frac{h}{2}\sqrt{c_{h,m}^{k+1}}\rho_{h,m}^{k+1/2},\end{cases} \tag{6.5}$$

with ψ a Kantorovich potential for $\mathtt{W}(c_{h,m}^{k+1/2},c_{h,m}^{k})$.

Using the optimality conditions for the Fischer-Rao steps, we obtain directly the following L^∞ bounds:

Lemma 6.2. *For all $k \geqslant 0$*

$$\|c_{h,m}^{k+1}\|_{L^\infty(\Omega)} \leqslant \|c_{h,m}^{k+1/2}\|_{L^\infty(\Omega)} \leqslant \|c_{h,m}^{k}\|_{L^\infty(\Omega)},$$

and at the continuous level

$$\|c_{h,m}(t,\cdot)\|_{L^\infty(\Omega)}, \|\tilde{c}_{h,m}(t,\cdot)\|_{L^\infty(\Omega)} \leqslant \|c_0\|_{L^\infty(\Omega)} \quad \forall\, t \geq 0.$$

Moreover,

$$\|\rho_{h,m}(t,\cdot)\|_\infty, \|\tilde{\rho}_{h,m}(t,\cdot)\|_\infty \leqslant 1$$

and there exists $c_T \equiv c_T(\|c_0\|_{L^\infty}), C_T \equiv C_T(\|c_0\|_{L^\infty}) > 0$ such that

$$\begin{aligned}&(1-c_T h)\rho_{h,m}^{k+1/2}(x) \leqslant \rho_{h,m}^{k+1}(x) \leqslant (1+C_T h)\rho_{h,m}^{k+1/2}(x) \quad \textit{a.e. in } \Omega.\\ &(1-h)c_{h,m}^{k+1/2}(x) \leqslant c_{h,m}^{k+1}(x) \leqslant c_{h,m}^{k+1/2}(x) \quad \textit{a.e. in } \Omega.\end{aligned} \tag{6.6}$$

Proof. The proof of the estimates on $c_{h,m}$ and $\tilde{c}_{h,m}$ is obvious because one step of Wasserstein gradient flow with the Boltzmann entropy decreases the L^∞-norm in (6.2) (see [1, 34]), and, because the product $\sqrt{c_{h,m}^{k+1}}\rho_{h,m}^{k+1/2}$ is nonnegative in (6.5), the L^∞-norm is also nonincreasing during the Fischer-Rao step. The proof for $\rho_{h,m}$ and $\tilde{\rho}_{h,m}$ is the same as in Lemma 5.2. Using the fact that $\|\tilde{\rho}_{h,m}(t,\cdot)\|_\infty \leqslant 1$, we see that the term $\Phi(p_{h,m}^{k+1},c_{h,m}^{k+1/2}) := (1-p_{h,m}^{k+1})(c_{h,m}^{k+1/2}+c_1)-c_2$ in (6.4) is bounded in L^∞ uniformly in k. This allows to argue exactly as in Lemma 3.4 to retrieve the estimate (6.6) and concludes the proof. ☐

With these bounds it is easy to prove as in Proposition 3.6 that

$$\begin{array}{c}\mathcal{F}_m(\rho_{h,m}^{k+1}) \leqslant \mathcal{F}_m(\rho_{h,m}^{k+1/2}) + C_T h,\\ \mathcal{E}_{1,m}(\rho_{h,m}^{k+1/2}|c_{h,m}^{k+1/2}) - \mathcal{E}_{1,m}(\rho_{h,m}^{k+1}|c_{h,m}^{k+1/2}) \leqslant C_T h,\\ \mathcal{H}(c_{h,m}^{k+1}) \leqslant \mathcal{H}(c_{h,m}^{k+1/2}) + C_T h,\\ \mathcal{E}_2(c_{h,m}^{k+1/2}|\rho_{h,m}^{k+1/2}) - \mathcal{E}_2(c_{h,m}^{k+1}|\rho_{h,m}^{k+1/2}) \leqslant C_T h,\end{array}.$$

for some C_T independent of m. Then we obtain the usual $\frac{1}{2}$-Hölder estimates in time with respect to the WFR distance, which in turn implies that $\rho_{h,m}, \tilde{\rho}_{h,m}$ converge to some $\rho \in L^\infty([0,T], L^1(\Omega))$ and $c_{h,m}, \tilde{c}_{h,m}$ converge to some $c \in L^\infty([0,T], L^1(\Omega))$ pointwise in time with respect to WFR, see (3.21), Proposition 3.7, and (3.23) for details.

As before we need to improve the convergence in order to pass to the limit in the nonlinear terms. For $\rho_{h,m}$ and $\tilde{\rho}_{h,m}$, this follows from

Lemma 6.3. *For all $T > 0$, if $\rho^0, c^0 \in BV(\Omega)$,*

$$\sup_{t\in[0,T]} \left\{ \|\rho_{h,m}(t,\cdot)\|_{BV(\Omega)} + \|c_{h,m}(t,\cdot)\|_{BV(\Omega)} \right\} \leqslant e^{C_T T}(\|\rho^0\|_{BV(\Omega)} + \|c^0\|_{BV(\Omega)})$$
$$\sup_{t\in[0,T]} \left\{ \|\tilde{\rho}_{h,m}(t,\cdot)\|_{BV(\Omega)} + \|\tilde{c}_{h,m}(t,\cdot)\|_{BV(\Omega)} \right\} \leqslant e^{C_T T}(\|\rho^0\|_{BV(\Omega)} + \|c^0\|_{BV(\Omega)}).$$

Proof. The argument is a generalization of Lemma 5.4, see Remark 5.1 from [18]. First, the BV-norm is nonincreasing during the Wasserstein step, Theorem 1.1 from [14],

$$\|\rho^{k+1/2}_{h,m}\|_{BV(\Omega)} \leqslant \|\rho^k_{h,m}\|_{BV(\Omega)} \quad \text{and} \quad \|c^{k+1/2}_{h,m}\|_{BV(\Omega)} \leqslant \|c^k_{h,m}\|_{BV(\Omega)}.$$

Arguing as in Lemma 5.4, we observe that, inside $\operatorname{supp} \rho^{k+1/2}_{h,m} = \operatorname{supp} \rho^{k+1}_{h,m}$, the minimizer $\rho = \rho^{k+1}_{h,m}(x)$ is the unique positive solution of $f(\rho, \rho^{k+1/2}_{h,m}(x), c^{k+1/2}_{h,m}(x)) = 0$, with

$$f(\rho,\mu,c) = \sqrt{\rho}\left(1 - \frac{h}{2}\left(\left(1 - \frac{m}{m-1}\rho^{m-1}\right)(c + c_1) - c_2\right)\right) - \sqrt{\mu}.$$

For $\mu > 0$ the implicit function theorem gives as before a $\mathcal{C}^1$ map R such that $f(\rho,\mu,c) = 0 \Leftrightarrow \rho = R(\mu,c)$. An easy algebraic computation and (6.6) then give $0 < \partial_\mu R(\mu,c) \leqslant (1 + C_T h)$ and $|\partial_c R(\mu,c)| \leqslant C_T h$ for some constant $C_T > 0$ independent of h, m, k. This implies that

$$\begin{aligned}\|\rho^{k+1}_{h,m}\|_{BV(\Omega)} &\leqslant (1 + C_T h)\|\rho^{k+1/2}_{h,m}\|_{BV(\Omega)} + C_T h\|c^{k+1/2}_{h,m}\|_{BV(\Omega)} \\ &\leqslant (1 + C_T h)\|\rho^k_{h,m}\|_{BV(\Omega)} + C_T h\|c^k_{h,m}\|_{BV(\Omega)}.\end{aligned}$$

The same argument shows that

$$\|c^{k+1}_{h,m}\|_{BV(\Omega)} \leqslant (1 + C_T h)\|c^k_{h,m}\|_{BV(\Omega)} + C_T h\|\rho^k_{h,m}\|_{BV(\Omega)},$$

and a simple induction allows to conclude. □

Proposition 6.4. *Up to extraction of a discrete sequence $h \to 0, m \to +\infty$,*

$$\rho_{h,m}, \tilde{\rho}_{h,m} \to \rho \quad \text{strongly in } L^1(Q_T)$$

$$p_{h,m} \rightharpoonup p \quad \text{and} \quad \tilde{p}_{h,m} \rightharpoonup \tilde{p} \quad \text{weakly in all } L^q(Q_T)$$

for all $T > 0$. If in addition $mh \to 0$ then $p = \tilde{p} \in L^2((0,T), H^1(\Omega))$ and (ρ, p) satisfies

$$0 \leqslant \rho, p \leqslant 1 \quad \text{and} \quad p(1-\rho) = 0 \quad \text{a.e. in } Q_T.$$

Proof. The argument is the same as for Proposition 5.6, Lemmas 5.7, and Lemma 5.8. □

In order to conclude the proof of Theorem 6.1 we only need to check that ρ, p, c satisfy the weak formulation of (6.1): the strong convergence of $\rho_{h,m}, c_{h,m}$ and the weak convergence of $p_{h,m}$ are enough to take the limit in the nonlinear terms as in Section 5.2, and we omit the details.

Acknowledgements. We warmly thank G. Carlier for fruitful discussions and suggesting us the problem in Section 3. This work was partially supported by the ANR contract GEOPOR (ANR-13-JS01-0007-01). T.O. Gallouët was supported by the ANR contract ISOTACE (ANR-12-MONU-013) and by the Fonds de la Recherche Scientifique - FNRS under Grant MIS F.4539.16.

References

[1] M. Agueh, Existence of solutions to degenerate parabolic equations via the Monge-Kantorovich theory. *Adv. Differ. Equ.* **10** (2005) 309–360.
[2] D. Alexander, I. Kim and Y. Yao, Quasi-static evolution and congested crowd transport. *Nonlinearity* **27** (2014) 823.
[3] L. Ambrosio, N. Fusco and D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems. *Oxford Mathematical Monographs*. The Clarendon Press, Oxford University Press, New York (2000).
[4] 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 (2005).
[5] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.* **84** 2000 375–393.
[6] J.-D. Benamou, G. Carlier and M. Laborde, An augmented Lagrangian approach to Wasserstein gradient flows and applications. *ESAIM: PROCs.* **54** (2016) 1–17.
[7] A. Braides. *Γ*-Convergence for Beginners. Vol. 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford (2002).
[8] G. Carlier and M. Laborde, A splitting method for nonlinear diffusions with nonlocal, nonpotential drifts. *Nonlinear Anal.: Theory Methods Appl.* **150** (2017) 1–18.
[9] J. A. Carrillo, M. DiFrancesco, A. Figalli, T. Laurent and D. Slepčev, Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. *Duke Math. J.* **156** (2011) 229–271.
[10] L. Chizat and S. Di Marino, A Tumor Growth Model of Hele-Shaw Type as a Gradient Flow. Preprint arXiv:1712.06124 (2017).
[11] L. Chizat, G. Peyré, B. Schmitzer and F.-X. Vialard, An Interpolating Distance Between Optimal Transport and Fischer-Rao. Preprint arXiv:1506.06430 (2015).
[12] L. Chizat, G. Peyré, B. Schmitzer and F.-X. Vialard, Unbalanced Optimal Transport: Geometry and Kantorovich Formulation. Preprint arXiv:1508.05216 (2015).
[13] L. Chizat, G. Peyré, B. Schmitzer and F.-X. Vialard, Scaling Algorithms for Unbalanced Transport Problems. Preprint arXiv:1607.05816 (2016).
[14] G. De Philippis, A.R. Mészáros, F. Santambrogio and B. Velichkov, BV estimates in optimal transportation and applications. *Arch. Ration. Mech. Anal.* **219** (2016) 829–860.
[15] M. Di Francesco and S. Fagioli, Measure solutions for non-local interaction PDEs with two species. *Nonlinearity* **26** (2013) 2777–2808.
[16] A. Figalli and N. Gigli, A new transportation distance between non-negative measures, with applications to gradients flows with dirichlet boundary conditions. *J. Math. Pures Appl.* **94** (2010) 107–130.
[17] F. Fleißner, Gamma-Convergence and Relaxations for Gradient Flows in Metric Spaces: A Minimizing Movement Approach. Preprint arXiv:1603.02822 (2016).
[18] T. Gallouët and L. Monsaingeon, A JKO Splitting Scheme for Kantorovich-Fischer-Rao Gradient Flows. Preprint arXiv:1602.04457 (2016).
[19] W. Gangbo and R.J. McCann, The geometry of optimal transportation. *Acta Math.* **177** (1996) 113–161.
[20] R. Jordan, D. Kinderlehrer and F. Otto, The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **29** (1998) 1–17.
[21] D. Kinderlehrer, L. Monsaingeon and X. Xu, A Wasserstein Gradient Flow Approach to Poisson-Nernst-Planck Equations. Preprint arXiv:1501.04437 (2015).
[22] S. Kondratyev, L. Monsaingeon and D. Vorotnikov, A New Optimal Transport Distance on the Space of Finite Radon Measures. Preprint arXiv:1505.07746 (2015).
[23] S. Kondratyev, L. Monsaingeon and D. Vorotnikov, A fitness-driven cross-diffusion system from population dynamics as a gradient flow. *J. Differ. Equ.* **261** (2016) 2784–2808.
[24] M. Laborde, On Some Non linear Evolution Systems Which Are Perturbations of Wasserstein Gradient Flows. *Radon Ser. Comput. Appl. Math.* (2015).
[25] M. Liero and A. Mielke, Gradient structures and geodesic convexity for reaction–diffusion systems. *Philos. Trans. R. Soc. A* **371** (2013) 20120346.
[26] M. Liero, A. Mielke and G. Savaré, Optimal Entropy-Transport Problems and a New Hellinger-Kantorovich Distance Between Positive Measures. *Invent. Math.* **211** (2018) 969–1117.

[27] M. Liero, A. Mielke and G. Savaré, Optimal Transport in Competition with Reaction: The Hellinger-Kantorovich Distance and Geodesic Curves. *SIAM J. Math. Anal.* **48** (2016) 2869–2911.
[28] S. Lisini, D. Matthes and G. Savaré, Cahn-Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics. *J. Differ. Equ.* **253** (2012) 814–850.
[29] D. Matthes, R.J. McCann and G. Savaré, A family of nonlinear fourth order equations of gradient flow type. *Commun. Partial Diff. Equ.* **34** (2009) 1352–1397.
[30] B. Maury, A. Roudneff-Chupin, F. Santambrogio and J. Venel, Handling congestion in crowd motion modeling. *Netw. Heterog. Media* **6** (2011) 485–519.
[31] J.D. Murray, Mathematical Biology II. Spatial Models and Biomedical Applications, 3rd edn. Vol. 18 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York (2003).
[32] F. Otto, Double Degenerate Diffusion Equations as Steepest Descent (1996).
[33] F. Otto, Dynamics of labyrinthine pattern formation in magnetic fluids: A mean-field theory. *Arch. Ration. Mech. Anal.* **141** (1998) 63–103.
[34] F. Otto, The geometry of dissipative evolution equations: the porous medium equation. *Commun. Partial Differ. Equ.* **26** (2001) 101–174.
[35] B. Perthame, Transport Equations in Biology. Frontiers in Mathematics. Birkhäuser Verlag, Basel (2007).
[36] B. Perthame, F. Quirós and J.L. Vázquez, The Hele-Shaw asymptotics for mechanical models of tumor growth. *Arch. Ration. Mech. Anal.* **212** (2014) 93–127.
[37] B. Perthame, M. Tang and N. Vauchelet, Traveling wave solution of the Hele-Shaw model of tumor growth with nutrient. *Math. Models Methods Appl. Sci.* **24** (2014) 2601–2626.
[38] L. Petrelli and A. Tudorascu, Variational principle for general diffusion problems. *Appl. Math. Optim.* **50** (2004) 229–257.
[39] B. Piccoli and F. Rossi, Generalized Wasserstein distance and its application to transport equations with source. *Arch. Ration. Mech. Anal.* **211** (2014) 335–358.
[40] M. Pierre, Global existence in reaction-diffusion systems with control of mass: a survey. *Milan J. Math.* **78** (2010) 417–455.
[41] R. Rossi and G. Savaré, Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **2** (2003) 395–431.
[42] E. Sandier and S. Serfaty, Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Commun. Pure Appl. Math.* **57** (2004) 1627–1672.
[43] F. Santambrogio, Optimal Transport for Applied Mathematicians. Vol. 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkasauser Verlag, Basel (2015).
[44] J.L. Vázquez, The Porous Medium Equation: Mathematical Theory. Oxford University Press (2007).
[45] C. Villani, Topics in Optimal Transportation. Vol. 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI (2003).
[46] C. Villani, Optimal Transport. Old and new Vol. 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin (2009).
[47] J. Zinsl, Geodesically convex energies and confinement of solutions for a multi-component system of nonlocal interaction equations. Technical report (2014).