Networks and Heterogeneous Media
Vol. 17, No. 5, October 2022, pp. 783–801
doi:10.3934/nhm.2022026

SPLITTING SCHEME FOR A MACROSCOPIC CROWD MOTION MODEL WITH CONGESTION FOR A TWO-TYPED POPULATION.

FÉLICIEN BOURDIN
LMO - Laboratoire de Mathématiques d'Orsay
France

(Communicated by the Pierre-Emmanuel Jabin)

ABSTRACT. We study the extension of the macroscopic crowd motion model with congestion to a population divided into two types. As the set of pairs of density whose sum is bounded is not geodesically convex in the product of Wasserstein spaces, the generic splitting scheme may be ill-posed. We thus analyze precisely the projection operator on the set of admissible densities, and link it to the projection on the set of measures of bounded density in the mono-type case. We then derive a numerical scheme to adapt the one-typed population splitting scheme.

1. **Introduction.** The modeling of the collective motion of entities - as a human crowd, a population of cells in biology or self-propelled particles in physics - involves a large range of approaches based on the level of description of the particle and its interactions. On the one hand, microscopic models rely on the description of each particle's motion. The population is thus described as a collection of n points moving and interacting in the space. Formally, the system evolves according to a system of coupled ordinary differential equations

$$\frac{dX_i}{dt} = v_i(t, X_1, \dots, X_n) \quad i = 1, \dots, n. \tag{1}$$

It is said as "Lagrangian", since we follow each trajectory in time. It is tractable for a reasonable number of particles, but it becomes unusable when the size of the system increases since it involves the computation of many interactions (roughly increasing as the square of the total population).
On the other hand, in macroscopic models the population is described by its density, that is a number of particles by surface unit. The system is thus a partial differential equation on this very density ρ, of the type

$$\partial_t \rho + \nabla \cdot (\rho u) = 0 \tag{2}$$

where u is the "Eulerian" velocity of the crowd in the space. It allows the description of a system of arbitrarily large number of particles, but does not provide the information of every trajectory.

2020 *Mathematics Subject Classification.* 35Q92, 49Q22, 92C17.
Key words and phrases. Crowd motion, optimal transport, Wasserstein spaces, splitting scheme.
The work of Félicien Bourdin is supported by the ERC grant NORIA. .

Congestion. When modeling the collecting motion of particles, one shall often introduce a congestion constraint in the description of the motion. This congestion encodes the impenetrability of the entities at a microscopic scale. Generically, one distinguishes between a desired velocity (or self-propulsion in physics) and an effective velocity, which is the resulting velocity due to the interactions ruled by congestion. There are two main approaches to introduce congestion on a crowd motion model. On the one hand, **soft** models rely on a relaxation of the constraint that affects the behaviour of the system as some configurations get close of the saturation. In microscopic space-continuous models, a highly repulsive potential of interaction at short-range is usually introduced in the model to prevent any overlapping between particles for instance in the context of pedestrians [11], or collision of solids, see [2], [23]. For discrete-space microscopic models of motion of pedestrians, cellular automaton based models place people on a given grid [3], [5]. In macroscopic models, the porous media equation [8], [13], [19] can be seen as a gradient flow for the functional

$$\rho \mapsto \int \rho^m(x)dx, \tag{3}$$

hence penalizing the high values of the density. A kernel can also be considered [17], by introducing a term of the form

$$\nabla(G * \rho), \tag{4}$$

where $G * \rho$ is the convolution between the density ρ and a repulsive kernel $G(x, y)$ generically of the form $V(\|x - y\|)$, for instance $V(r) = e^{-\frac{r^2}{2\sigma^2}}$ (see [16] for various choices of V).

In this article, we will be focused on **hard** models of congestion, that are models that do not allow any overlapping or violation of the constraint. A microscopic hard model of crowd motion is introduced in [15], in which the desired velocities are projected at every time on the set of admissible velocities, that are the velocities that lead to non-overlapping configurations. This complex model (the particles are driven by a velocity that is the projection on a set that evolves in time) is shown to fit into the framework of the sweeping process, introduced by Moreau [18]. The well-posedness of the model is based on a catching-up, or splitting algorithm which has generically two steps:

- Step 1 (prediction): apply the unconstrained velocities during a fixed time step,
- Step 2 (correction): project the obtained positions on the set of admissible configurations.

The macroscopic model of crowd motion introduced in [14] is built on the same approach of projection of the velocities. Given a density on a domain $\Omega \subset \mathbb{R}^2$ under maximal congestion $\rho \leq 1$, the set of admissible velocities writes

$$C_\rho = \left\{ v \in \mathbb{L}^2(\Omega),\ \nabla \cdot v \geq 0 \text{ where } \rho = 1 \right\}. \tag{5}$$

The models then reads

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho u) = 0 \\ u \in \underset{v \in C_\rho}{\operatorname{argmin}} \|v - U\|_2^2 \end{cases} \tag{6}$$

U being the desired velocity. The well-posedness of the model is established in [14], and it is shown in [15] that the Eq. (6) fits in a sweeping framework.

Macroscopic multi-type models. As the microscopic models allow natively to consider any heterogeneity in the behaviours of the population, there are few extensions of macroscopic congestion models for a population composed by different types. Multiphase flows have been widely studied in physics, see [4] and [9], even under a total density constraint [6], but may not be applied to crowd (or cell) motion since they involve a phase separation and not a total mixture of the types. A separation of phase has been also considered in order to study the motion of cars structured in a multilane traffic [24]. Conversely, in the model we shall consider we assume a mixture between two densities ρ_1, ρ_2 with a constraint on the total density $\rho_1 + \rho_2 \leq 1$, as considered in [13] and [15]. It can be seen as the formal limit as $m \to \infty$ of a cross-diffusion model with a common pressure term

$$p = \frac{m}{m-1}(\rho_1 + \rho_2)^{m-1}, \tag{7}$$

that reduces to the classical porous media equation

$$\partial_t \rho - \Delta \rho^m = 0 \tag{8}$$

in the absence of one of the two types.

In what follows, we study the two-species adaptation of the crowd motion model (6), that writes

$$\begin{cases} \partial_t \rho_1 + \nabla \cdot (\rho_1 u_1) &= 0 \\ \partial_t \rho_2 + \nabla \cdot (\rho_2 u_2) &= 0 \\ (u_1, u_2) &= P_{C_{\rho_1,\rho_2}}(U_1, U_2) \end{cases} \tag{9}$$

where U_1, U_2 are the spontaneous velocities, C_{ρ_1,ρ_2} is the set of admissible velocities that transport the densities such that $\rho_1 + \rho_2$ remains smaller than 1 almost everywhere and $P_{C_{\rho_1,\rho_2}}$ the orthogonal projection operator on C_{ρ_1,ρ_2}.

Optimal transport, Wasserstein spaces and admissible set of densities. In order to define a splitting scheme that shall be the macroscopic counterpart of the microscopic splitting scheme defined above, we need to introduce the formalism of Wasserstein spaces. Being given a convex compact domain Ω of $\mathbb{R}^d$, and two probability measures μ, ν on Ω, one denotes $\Pi(\mu, \nu)$ the set of transport plans from μ to ν. The 2-Wasserstein distance between μ and ν is then defined as

$$W_2^2(\mu, \nu) = \inf_{\gamma \in \Pi(\mu,\nu)} \int_\Omega \|x - y\|^2 d\gamma(x, y). \tag{10}$$

It defines a distance on $\mathbb{P}(\Omega)$ that metrizes the weak convergence of probability measures, at makes $\mathbb{P}(\Omega)$ a geodesic space, denoted by $W_2(\Omega)$. We refer to [22] for a general introduction to Optimal Transport and Wasserstein spaces. Under this setting, the splitting scheme introduced in [15] for (6) writes, given a time step $\tau > 0$ and ρ_t a density at time t:

- Step 1 (prediction): Transport the density by the desired velocity during τ, with
$$\mu_t = (\mathrm{Id} + \tau U)_\# \rho_t, \tag{11}$$
where $\#$ is the pushforward operator.
- Step 2 (correction): Project the intermediate density on the set of admissible densities
$$K_1 = \{\rho \in W_2(\Omega), \rho \leq 1\}. \tag{12}$$

Numerical schemes for hard congestion macroscopic models. We shall end this introduction by saying a few words on schemes developed to approximate the macroscopic crowd motion model (6). In [14], existence of solutions to (6) is proven by considering the limit of a JKO-scheme. This scheme originally introduced in [12] approximates the model by the iterations

$$\rho^{k+1} = \underset{\rho\in W_2(\Omega)}{\operatorname{argmin}} \int_\Omega D d\rho + \mathbf{1}_{K_1}(\rho) + \frac{1}{2\tau} W_2^2(\rho, \rho^k), \tag{13}$$

when the desired velocity is of the form $-\nabla D$. As highlighted in [21], there are two main issues when using this algorithm as a numerical scheme. The first one is that it involves the computation of a Wasserstein distance, which is a linear problem under constraint and thus can be long. The use of regularized optimal transport and Sinkhorn algorithm can overcome this difficulty [10], [20]. The second problem is a freezing phenomenon: when implementing the algorithm on a given mesh, for small timesteps the Wasserstein part of (13) becomes large and forbids any mass displacement. One should then work under the inverse of a CFL condition of the type $dx < C\tau$. This condition forces to use small meshes, resulting on a high increase of the time of computation. The last constraint is that the use of a JKO scheme is possible only if the desired velocity has a specific form that defines a model that has a gradient flow structure, see [22].

We shall consider in this paper an implementation of the splitting scheme (11) (12) which extends the implementation presented in [15] for a mono-type population. The prediction step (11) is computed by transportation of the cells of the mesh by the desired velocity, in a Lagrangian approach. The correction step (12) is computed by a stochastic algorithm: from each oversaturated cell a random walk starts and dispatches the excess of mass when it reaches a free cell where the density is lower than 1. The analysis of the convergence of this scheme is far beyond the scope of this paper, but it is easy to compute and allows to choose large time steps, providing a fast computation of the scheme for large total time.

The rest of the paper is organized as follows. In section 2, the model is introduced and a splitting algorithm is derived. In section 3, we analyze the projection operator on the set of admissible pairs of densities and state a well-posedness theorem on the scheme. In section 4, we present an implementation of the numerical algorithm and show some simulations.

2. Macroscopic crowd model for a two-type population, splitting algorithm.

Let Ω be a convex bounded domain of $\mathbb{R}^2$. In what follows, we identify a density measure to a function. We want to model the evolution of $\rho_1, \rho_2 \in W_2(\Omega)$ by integrating a total density constraint $\rho_1 + \rho_2 \leq 1$ as it is done for the mono-type model in [14]. We denote by $U_j \in \mathbb{L}^2(\Omega, \mathbb{R}^2)$ the desired velocity of the type j. Being given two admissible densities, the cone of the admissible velocities writes

$$C_{\rho_1,\rho_2} = \left\{ u_1, u_2 \in \mathbb{L}^2\left(\Omega, \mathbb{R}^2\right), \nabla \cdot (\rho_1 u_1 + \rho_2 u_2) \geq 0 \text{ where } \rho_1 + \rho_2 = 1 \right\} \tag{14}$$

where "$\phi \geq 0$ where $\rho_1 + \rho_2 = 1$" for $\phi \in \left(H^1(\Omega)\right)'$ has a weak formulation: for every $q \in H^1_{\rho_1+\rho_2}(\Omega)$,

$$\langle \phi, q \rangle \geq 0, \tag{15}$$

where

$$H^1_{\rho_1+\rho_2}(\Omega) = \left\{ q \in H^1(\Omega), q \geq 0 \text{ a.e.}, q(1-\rho_1-\rho_2) = 0 \right\}. \tag{16}$$

We shall thus study the following model:

Problem 1. *For $\rho_1^0, \rho_2^0 \in W_2(\Omega)$ such that $\rho_1^0 + \rho_2^0 \leq 1$ and for $U_1, U_2 \in \mathbb{L}^2\left(\Omega, \mathbb{R}^2\right)$, we look for $\rho_1, \rho_2 \in \mathbb{L}^2\left([0,T], \mathbb{L}^2(\Omega)\right)$ that satisfy (in a weak sense)*

$$\begin{cases} \partial_t \rho_1 + \nabla \cdot (\rho_1 u_1) &= 0 \\ \partial_t \rho_2 + \nabla \cdot (\rho_2 u_2) &= 0 \\ (u_1, u_2) &= P_{C_{\rho_1,\rho_2}}(U_1, U_2) \\ \rho_1(0,\cdot), \rho_2(0,\cdot) &= \rho_1^0, \rho_2^0, \end{cases} \tag{17}$$

where $P_{C_{\rho_1,\rho_2}}$ is the projection for the $\mathbb{L}^2$ norm on C_{ρ_1,ρ_2}.

Let us first recall how the existence is proven in [14] in the case of a single type when the desired velocity is of the form $U = -\nabla D$. A JKO scheme is introduced, that writes

$$\rho^{k+1} = \underset{\rho \in W_2(\Omega)}{\operatorname{argmin}} \int_\Omega D d\rho + \mathbf{1}_{K_1}(\rho) + \frac{1}{2\tau} W_2^2(\rho, \rho^k), \tag{18}$$

with

$$K_1 = \{\rho \in W_2(\Omega),\ \rho \leq 1 \text{ a.e.}\} \tag{19}$$

the set of admissible densities. The existence of a unique minimizer is obtained by the convexity of K_1 for generalized geodesics of base ρ^k joining two minimizers μ_0 and μ_1, of the form $\mu_t = (tr_0 + (1-t)r_1)_{\#}\rho^k$ (where r_i is a transport plan from ρ^k to μ_i) and from the strict convexity of the functional defined by (13) along these generalized geodesics. In the 2-dimensional case, the set of admissible densities

$$K_2 = \{\rho_1, \rho_2 \in W_2(\Omega), \rho_1 + \rho_2 \leq 1\} \tag{20}$$

is no more convex along a generalized geodesic.

Example 1. Consider for instance ρ_1 and ρ_2 being the indicator function of the balls $B(0,\epsilon)$, $B(1,\epsilon)$ for a small $\epsilon > 0$ (we can normalize in order to have total masses equal to 1). Let $\mu_1 = \nu_2 = \rho_1$ and $\mu_2 = \nu_1 = \rho_2$. The transports plans from (ρ_1, ρ_2) to (μ_1, μ_2) are

$$r_1(x,y) = r_2(x,y) = (x,y) \tag{21}$$

and the transport plans from (ρ_1, ρ_2) to (ν_1, ν_2) are

$$\begin{aligned} s_1(x,y) &= (x+1, y) \\ s_2(x,y) &= (x-1, y). \end{aligned} \tag{22}$$

The generalized geodesic between (μ_1, μ_2) and (ν_1, ν_2) with respect to (ρ_1, ρ_2) is thus

$$(\alpha_1^t, \alpha_2^t) = \left(\mathbf{1}_{B(t,\epsilon)}, \mathbf{1}_{B(1-t,\epsilon)}\right), \tag{23}$$

which is not admissible for $t = 1/2$, as one can see on figure 1.

Remark 1. As a matter of fact, K_2 is not even geodesically convex in the product space $W_2(\Omega) \times W_2(\Omega)$ endowed with

$$d((\mu_1, \mu_2), (\nu_1, \nu_2))^2 = d(\mu_1, \nu_1)^2 + d(\mu_2, \nu_2)^2. \tag{24}$$

Indeed, let (μ_1, μ_2) and $(\nu_1, \nu_2) \in K_2$, and consider two geodesics of constant speed $x_t : [0,1] \to W_2(\Omega)$, $y_t : [0,1] \to W_2(\Omega)$ respectively from μ_1 to ν_1 and from μ_2 to ν_2. Then $z_t = (x_t, y_t)$ is a constant speed geodesic from (μ_1, μ_2) to (ν_1, ν_2) in $W_2(\Omega) \times W_2(\Omega)$. In the previous example, the generalized geodesics are thus actual geodesics and do not remain into K_2.

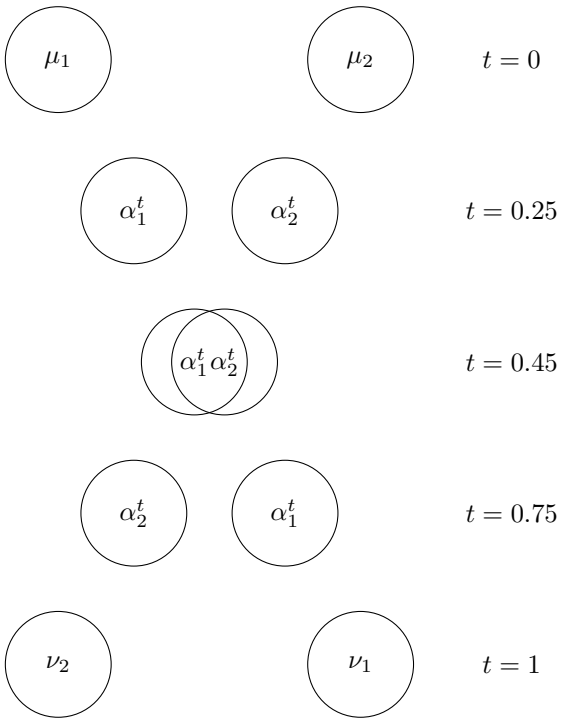

FIGURE 1. The interpolation between two opposite configurations of spheres along generalized geodesics. In particular, for $t = 0.45$ the pair (α_1^t, α_2^t) is not in K_2.

Even in the absence of an existence theorem for problem 1, we can write a splitting scheme for Eq. (17).

Definition 2.1. Given $\rho_1^0, \rho_2^0 \in W_2(\Omega)$ such that $\rho_1^0 + \rho_2^0 \leq 1$, $U_1, U_2 \in \mathbb{L}^2\left(\Omega, \mathbb{R}^2\right)$ and a time step $\tau > 0$, we define the splitting scheme as

$$\begin{aligned} \mu_j^{k+1} &= (Id + \tau U_j)_{\#}\rho_j^k && \text{(prediction)} \\ (\rho_1^{k+1}, \rho_2^{k+1}) &\in \operatorname*{argmin}_{(\rho_1,\rho_2)\in K_2} W_2^2(\rho_1, \mu_1^{k+1}) + W_2^2(\rho_2, \mu_2^{k+1}) && \text{(correction)} \end{aligned} \tag{25}$$

From its pushforward structure, the prediction step uniquely defines a pair of intermediate probability measures. It is less straightforward that the projection on K_2 is unique. Indeed, even if the argmin of the correction step of (25) is not empty from the minimization of a continuous functional on a compact set in $W_2(\Omega) \times W_2(\Omega)$, it is unclear that it is a singleton. The next section is dedicated to a special analysis of this correction step in order to get well-posedness of the splitting scheme (25).

3. Properties of the projection on K_2.

The goal of this section is to show that under tractable assumptions, the splitting scheme (25) is uniquely defined.

3.1. Uniqueness of the projection.

Let us start with a lemma that links the projection on K_2 with the projection of the total density on K_1.

Lemma 3.1. *Let $\mu_1, \mu_2 \in W_2(\Omega)$. We denote ρ the projection of $\mu_1 + \mu_2$ on K_1. Let*

$$(\rho_1, \rho_2) \in \underset{(\nu_1,\nu_2)\in K_2}{\operatorname{argmin}} W_2^2(\nu_1, \mu_1) + W_2^2(\nu_2, \mu_2). \tag{26}$$

Then $\rho_1 + \rho_2 = \rho$ and

$$W_2^2(\rho_1 + \rho_2, \mu_1 + \mu_2) = W_2^2(\rho_1, \mu_1) + W_2^2(\rho_2, \mu_2). \tag{27}$$

Proof. From the definition of ρ, one has

$$W_2^2\left(\rho, \mu_1 + \mu_2\right) \leq W_2^2\left(\rho_1 + \rho_2, \mu_1 + \mu_2\right). \tag{28}$$

Let T the optimal transport map from $\rho_1 + \rho_2$ to $\mu_1 + \mu_2$, T_1 the optimal transport map from ρ_1 to μ_1 and T_2 from ρ_2 to μ_2. Let $\gamma = (Id, T_1)_{\#}\,\rho_1 + (Id, T_2)_{\#}\,\rho_2$. The marginals of γ are respectively $\rho_1 + \rho_2$ and $\mu_1 + \mu_2$. The quadratic cost associated to γ is $W_2^2\left(\rho_1, \mu_1\right) + W_2^2\left(\rho_2, \mu_2\right)$. One thus has

$$W_2^2\left(\rho_1 + \rho_2, \mu_1 + \mu_2\right) \leq W_2^2\left(\rho_1, \mu_1\right) + W_2^2\left(\rho_2, \mu_2\right). \tag{29}$$

Let ζ be an optimal transport plan (in the sense of Kantorovich) from $\mu_1 + \mu_2$ to ρ. One has $\mu_1 \ll \mu_1 + \mu_2$, $\mu_2 \ll \mu_1 + \mu_2$: denote f and g the corresponding Radon-Nikodym derivatives, and

$$\begin{aligned} d\zeta_1(x, y) &= f(x) d\zeta(x, y) \\ d\zeta_2(x, y) &= g(x) d\zeta(x, y). \end{aligned} \tag{30}$$

One has $\zeta_1 + \zeta_2 = \zeta$, $\pi^x_{\#}\zeta_1 = \mu_1$ and $\pi^x_{\#}\zeta_2 = \mu_2$ (π^x and π^y are the projection on the coordinates). By denoting $\nu_j = \pi^y_{\#}\zeta_j$ for $j = 1, 2$, we get

$$\begin{aligned} W_2^2\left(\rho, \mu_1 + \mu_2\right) &= \int_{\Omega^2} |x - y|^2 d\zeta \\ &= \int_{\Omega^2} |x - y|^2 d\zeta_1 + \int_{\Omega^2} |x - y|^2 d\zeta_2 \\ &\geq W_2^2\left(\nu_1, \mu_1\right) + W_2^2\left(\nu_2, \mu_2\right). \end{aligned} \tag{31}$$

By optimality of the pair (ρ_1, ρ_2), one has

$$W_2^2\left(\rho_1, \mu_1\right) + W_2^2\left(\rho_2, \mu_2\right) \leq W_2^2\left(\rho, \mu_1 + \mu_2\right). \tag{32}$$

By combining (28), (29), (32), we get $W_2^2\left(\rho, \mu_1 + \mu_2\right) = W_2^2\left(\rho_1 + \rho_2, \mu_1 + \mu_2\right)$. As the projection on K_1 is uniquely defined (proposition 2 in [15]), we get $\rho = \rho_1 + \rho_2$. □

Remark 2. From the previous lemma, we get that the sum of two densities that realize the distance to any pair of densities (μ_1, μ_2) can be computed by projecting $\mu_1 + \mu_2$ on K_1. However, there can be many pairs that realize this distance. Consider for instance in dimension 1 the case $\mu_1 = \mu_2 = \delta_0$. One has $P_{K_1}(\mu_1 + \mu_2) = \mathbf{1}_{[-1,1]}$. Let then be $f, g \in \mathbb{L}^\infty([-1, 1])$ two nonnegative functions such that $f + g = 1$ almost everywhere. Setting

$$\begin{aligned} \rho_1 &= f d\lambda \\ \rho_2 &= g d\lambda \end{aligned} \tag{33}$$

where λ is Lebesgue measure on $[-1, 1]$, (ρ_1, ρ_2) realizes the distance to (μ_1, μ_2) in the sense of (26).

Nevertheless, when every density is absolutely continuous with respect to Lebesgue measure we have uniqueness of the projection.

Proposition 1. *Let $\mu_1, \mu_2 \in W_2(\Omega)$ two absolutely continuous measures with respect to Lebesgue measure. There exists a unique couple (ρ_1, ρ_2) in K_2 that minimizes $W_2^2(\rho_1, \mu_1) + W_2^2(\rho_2, \mu_2)$.*

Proof. As every measure is absolutely continuous, all the transport plans are transport maps (in the sense of Monge). Let T be the unique transport map from $\mu_1+\mu_2$ to $\rho := P_{K_1}(\mu_1 + \mu_2)$ and ρ_1, ρ_2 that minimize the distance to K_2. Let T_1 and T_2 the optimal transport maps from μ_1 to ρ_1 and from μ_2 to ρ_2. The T_j are uniquely defined on the supports of the μ_j.

Let us show that $T_1 = T_2$ almost everywhere. Out of $\mathrm{supp}(\mu_1) \cap \mathrm{supp}(\mu_2)$, one can modify T_1 or T_2. Let $\gamma = (Id, T_1)_\# \mu_1 + (Id, T_2)_\# \mu_2$: the marginals of γ are $\mu_1 + \mu_2$ and $\rho_1 + \rho_2 = \rho$ from lemma 3.1. The cost of γ is

$$\begin{aligned} \int_\Omega |x-y|^2 d\gamma &= \int_\Omega |x - T_1(x)|^2 d\mu_1 + \int_\Omega |x - T_2(x)|^2 d\mu_2 \\ &= W_2^2(\rho_1, \mu_1) + W_2^2(\rho_2, \mu_2) \\ &= W_2^2(\rho, \mu_1 + \mu_2). \end{aligned} \tag{34}$$

γ is thus an optimal transport plan between its marginals. By uniqueness one has $\gamma = (Id, T)_\#(\rho_1 + \rho_2)$. Let $z \in \mathrm{supp}(\rho_1) \cap \mathrm{supp}(\rho_2)$. We have

- $(z, T_1(z)) \in \mathrm{supp}(\gamma)$
- $(z, T_2(z)) \in \mathrm{supp}(\gamma)$

so $T_1 = T_2 = T$ a.e., and $\rho_j = T_\# \mu_j$, for $j = 1, 2$. The ρ_j are thus uniquely determined. □

From the previous proof, we get the following corollary that gives the form of the K_2 projection from the K_1 projection of the sum of the densities:

Corollary 1. Let $\mu_1, \mu_2 \in W_2(\Omega)$ two absolutely continuous measures with respect to Lebesgue measure. Let T be an optimal transport map from $\mu_1 + \mu_2$ to its projection on K_1. Then the projection of (μ_1, μ_2) on K_2 is given by $(T_\# \mu_1, T_\# \mu_2)$.

Provided that the predicted densities in (25) are absolutely continuous, the correction step is thus uniquely defined. The remaining question is the following: what condition should be assumed to have a pair of predicted densities that are absolutely continuous with respect to Lebesgue measure?

3.2. **Well-posedness of the splitting scheme.** Let us start with a convenient definition.

Definition 3.2. We define the partial order on the set of positive measures on Ω by

$$\mu_1 \leq_* \mu_2 \iff \mu_1 \ll \mu_2 \text{ and } \frac{d\mu_1}{d\mu_2} \leq 1 \ \mu_2 - \text{a.e.} \tag{35}$$

In other words, $\mu_1 \leq_* \mu_2$ if $\mu_2 - \mu_1$ defines a positive measure. The following lemma controls the image of Lebesgue measure by the prediction step.

Lemma 3.3. *Let Ω be a (large enough) open set of $\mathbb{R}^2$, U a C^1 velocity field on Ω and Ω_0 a compact set included in Ω. Assume that the partial derivatives of U are bounded by a constant $c > 0$. Then for every $\tau < (5c)^{-1}$, denoting λ the Lebesgue measure on $\mathbb{R}^2$,*

$$(Id + \tau U)_\# \lambda_{|\Omega_0} \leq_* (1 + 5c\tau) \lambda_{|\Omega_0{}^{\tau \|U\|_{\mathbb{L}^\infty(\Omega_0)}}}. \tag{36}$$

where ${\Omega_0}^{\tau\|U\|_{\mathbb{L}^\infty(\Omega_0)}}$ *is the set of points at distance lower than* $\tau\|U\|_{\mathbb{L}^\infty(\Omega_0)}$ *from* Ω_0.

Proof. Let $\tau > 0$ and

$$f : \left\{ \begin{array}{ccc} \Omega_0 & \longrightarrow & f(\Omega_0) \\ x & \longmapsto & x + \tau U(x). \end{array} \right. \tag{37}$$

f is C^1 and onto for $\tau < c^{-1}$: let $x, y \in \Omega$ such that $f(x) = f(y)$. One has

$$x - y = \tau \int_0^1 \mathrm{Jac}_U((1-t)x + ty) \cdot (y - x) dt \tag{38}$$

so $\|x - y\| \leq c\tau\|x - y\|$.

Let A a borelian set of Ω.

$$\begin{aligned} (Id + \tau U)_{\#}\lambda_{|\Omega_0}(A) &= \int_{\Omega_0} \mathbf{1}_A \left((Id + \tau U)(x)\right) d\lambda \\ &= \int_{\Omega_0} \mathbf{1}_A(f(x)) \frac{|\det(\mathrm{Jac}_f(x))|}{|\det(\mathrm{Jac}_f(x))|} d\lambda \end{aligned} \tag{39}$$

One has $|\det(\mathrm{Jac}_f)| \geq 1 - 2c\tau - 2c^2\tau^2$, and for $\tau < \dfrac{\sqrt{2}-1}{2c}$, one gets $\dfrac{1}{1 - 2c\tau - 2c^2\tau^2}$ $\leq 1 + 4c\tau + 4c^2\tau^2$. If $\tau \leq (4c)^{-1}$, one has $1 + 4c\tau + 4c^2\tau^2 \leq 1 + 5c\tau$ and

$$\begin{aligned} (Id + \tau U)_{\#}\lambda_{|\Omega_0}(A) &\leq (1 + 5c\tau) \int_{\Omega_0} \mathbf{1}_A(f(x)) |\det(\mathrm{Jac}_f(x))| d\lambda \\ &= (1 + 5c\tau) \int_{f(\Omega_0)} \mathbf{1}_A(x) d\lambda \\ &\leq (1 + 5c\tau)\lambda \left(A \cap \Omega_0^{\tau\|U\|_{\mathbb{L}^\infty(\Omega_0)}} \right). \end{aligned} \tag{40}$$

□

We are now able to state a well-posedness result for the splitting scheme at two types (25).

Theorem 3.1. *Let* $U_1, U_2 \in C^1(\mathbb{R}^2, \mathbb{R}^2)$ *two velocity fields. Assume that they are bounded by* $c_1 > 1$, *and their partial derivatives are bounded by a constant* c_2. *Let* ρ_1^0, ρ_2^0 *two absolutely continuous measures whose support is included in a ball* $B(x, r_0)$, $\tau > 0$ *a time step smaller than* $(5c_2)^{-1}$, *and a total time* T. *Then, setting* $\Omega = B(x, R)$ *for* R *large enough, the splitting scheme* (25) *is uniquely defined for* $n := \left\lceil \dfrac{T}{\tau} \right\rceil$ *iterations.*

Proof. From lemma 3.3, the first predicted pair of densities is supported in $B(x, r_0 + \tau c_1)$ and its sum has a density lower than $1 + 5c_2\tau$. From proposition 1, the first correction step is well-defined. We shall use the following lemma, whose proof is in appendix.

Lemma 3.4. *Let* Ω *be a compact set, and* μ_1, μ_2 *two absolutely continuous measures, such that* $\mu_1 \leq_* \mu_2$. *Then* $P_{K_1}(\mu_1) \leq_* P_{K_1}(\mu_2)$.

From this lemma and corollary 1, one gets that the corrected velocities are supported on the ball $B(x, r_1)$, with $r_1 = \sqrt{1 + 5c_2\tau}(r_0 + c_1\tau)$. One proceed by induction to show that the densities are well-defined at each step and supported at step k by a ball of radius

$$r_k = (1 + 5c_2\tau)^{\frac{k}{2}} \left(r_0 + \frac{c_1\tau}{\sqrt{1 + 5c_2\tau} - 1}\right) - \frac{c_1\tau}{\sqrt{1 + c_2\tau} - 1}. \tag{41}$$

Setting $R = r_n$ concludes the proof. □

Let us end this theoretical section with a continuity proposition for the projection on K_2.

Proposition 2. *Assume that Ω is compact. Then the restriction of the projection on K_2 to the set of absolutely continuous measures with respect to Lebesgue measure is continuous.*

Proof. Let $(\mu_1^n, \mu_2^n)_n$ that converges towards (μ_1, μ_2). Denote

$$\begin{aligned}(\rho_1^n, \rho_2^n) &= P_{K_2}(\mu_1^n, \mu_2^n)\\ (\rho_1, \rho_2) &= P_{K_2}(\mu_1, \mu_2).\end{aligned} \tag{42}$$

Let T^n an optimal transport map from $\mu_1^n + \mu_2^n$ to $\rho_1^n + \rho_2^n$ and T from $\mu_1 + \mu_2$ to $\rho_1 + \rho_2$. As P_{K_1} is continuous (proposition 2 in [15]), $(\rho_1^n + \rho_2^n)_n$ converges towards $\rho_1 + \rho_2$. Theorem 1.50 in [22] ensures (as Ω is compact) that $(Id, T^n)_\#(\mu_1^n + \mu_2^n)$ converges towards $(Id, T)_\#(\mu_1 + \mu_2)$. Let ν_1 be an accumulation point of $(T^n_\#\mu_1^n)_n$; we still denote $(T^n_\#\mu_1^n)_n$ the subsequence converging to ν_1. One has

$$\pi^y_\# \left((Id, T^n)_\#(\mu_1^n + \mu_2^n)\right) = T^n_\#\mu_1^n + T^n_\#\mu_2^n. \tag{43}$$

On the one hand, $\pi^y{}_\# \left((Id, T^n)_\#(\mu_1^n + \mu_2^n)\right)$ converges to $\rho_1 + \rho_2$. On the other hand, $(T^n_\#\mu_1^n)_n$ converges towards ν_1. Therefore,

$$\lim_{n\to\infty} T^n_\#\mu_2^n = \rho_1 + \rho_2 - \nu_1. \tag{44}$$

One thus has, by using twice lemma 3.1:

$$\begin{aligned}W_2^2(\mu_1, \nu_1) + W_2^2(\mu_2, \rho_1 + \rho_2 - \nu_1) &= \lim_{n\to\infty} W_2^2(\mu_1^n, T^n_\#\mu_1^n) + W_2^2(\mu_2^n, T^n_\#\mu_2^n)\\ &= \lim_{n\to\infty} W_2^2(\mu_1^n + \mu_2^n, T^n_\#(\mu_1^n + \mu_2^n))\\ &= W_2^2(\mu_1 + \mu_2, \rho_1 + \rho_2)\\ &= W_2^2(\mu_1, \rho_1) + W_2^2(\mu_2, \rho_2).\end{aligned} \tag{45}$$

Since $(\nu_1, \rho_1 + \rho_2 - \nu_1)$ is admissible, by uniqueness of the projection on K_2, $\nu_1 = \rho_1$. Thus $(\rho_1^n)_n$ has one unique accumulation point in $W^2(\Omega)$ which is compact: it converges towards ρ_1 (and similarly $(\rho_2^n)_n$ converges towards ρ_2). □

4. Numerical implementation and results.

In this section we detail the numerical scheme used to approximate the splitting scheme (25). It is largely inspired of the methods developed for the mono-type model in [15], [21].

4.1. Prediction step.

We present here two main methods to implement the prediction step. We consider only one density since the predictions are made independently for the two types.

Finite volume step. In [21], a conservative finite volume scheme in introduced, with an upwind operator to discretize the flux. Given a mesh of size h, we start by defining the velocities at the edges of the mesh, as on figure 2. We then introduce the upwind operator that computes the flux flowing through a given edge:

$$A^{up}(U, \rho_-, \rho_+) = \begin{cases} U\rho_- \text{ if } U \geq 0\\ U\rho_+ \text{ if } U < 0.\end{cases} \tag{46}$$

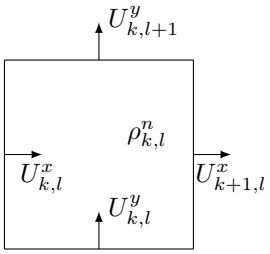

FIGURE 2. The cell of the mesh in position (k,l). The densities are defined inside the cell, whereas the velocities are defined at the edges.

The prediction step reads

$$\begin{aligned}\tilde{\rho}^{n+1}_{k,l} = \rho^n_{k,l} &- \frac{dt}{h}\left(A^{up}\left(U^x_{k+1,l},\rho^n_{k,l},\rho^n_{k+1,l}\right) - A^{up}\left(U^x_{k,l},\rho^n_{k-1,l},\rho^n_{k,l}\right)\right)\\ &- \frac{dt}{h}\left(A^{up}\left(U^y_{k,l+1},\rho^n_{k,l},\rho^n_{k,l+1}\right) - A^{up}\left(U^y_{k,l},\rho^n_{k,l-1},\rho^n_{k,l}\right)\right).\end{aligned} \tag{47}$$

This scheme is conservative, stable under the CFL condition $\frac{dt}{h} \leq \frac{1}{4\|U\|_\infty}$. However, it forces choosing small time steps as the mesh shrinks, and it is diffusive as we shall see on the following example in dimension 1.

Example 2. Consider for $\Omega = [0,1]$ a single density $\rho = \epsilon \mathbf{1}_{\left[\frac{1}{4},\frac{3}{4}\right]}$ subject to a velocity $U = \frac{5}{2} - 2x$ (so $U\left(\frac{1}{4}\right) = 2$ and $U\left(\left(\frac{3}{4}\right) = 1\right)$. One can pick a maximal time step $dt = \frac{h}{2}$ to preserve positivity. At step n, one can see that the n cells at the right of $x = \frac{3}{4}$ have some mass. The front of the density is thus traveling at speed $\frac{\delta_x}{\delta_t} = \frac{nh}{ndt} = 2$. The scheme thus presents a quicker diffusion than the model, that prescribes a front speed equal to 1.

Lagrangian transport. In order to have a scheme that can use large time step, we present a Lagrangian version of the prediction step, introduced in [15]. The key idea is to transport the whole cell $C_{k,l} = [kh,(k+1)h] \times [lh,(l+1)h]$ by computing the image of its center by $T(x) := x + dtU(x)$. One then dispatches the transported cell

$$B_{\|\cdot\|_\infty}\left(T(x), \frac{h}{2}\right) \tag{48}$$

on the cells of the mesh, as on figure 3.

Remark 3. This scheme is conservative, stable and preserves the positivity for any time step. The front speed of the density studied in example 2 is now for this scheme $1 + \frac{dh}{dt}$, so it can be reduced by considering a large time step or a small mesh size.

4.2. **Correction step.** Let us first recall how the projection on K_1 is obtained in the mono-type case in [15]. We shall then use corollary 1 to derive an algorithm to compute the projection on K_2.

Starting from a density μ discretized on a given squared mesh, the projection on K_1 is approximated by the following stochastic algorithm.

- Start from a random oversaturated cell and transport the excess of mass $\mu - 1$ with a random walk,

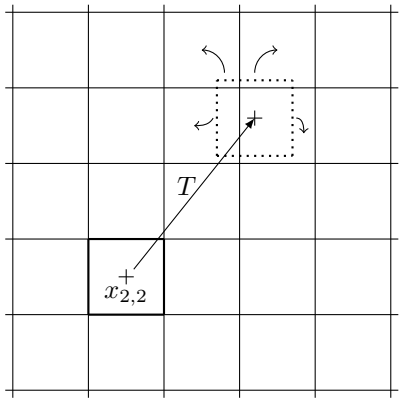

FIGURE 3. The distribution of the image of the cell $(2,2)$ by T. We first compute the image of the center, then draw a box of size h. The lower left part is lifted on the cell $(3,4)$, the upper left part on $(3,5)$, the upper right part on $(4,5)$ and the lower right part on $(4,4)$.

- When an undersaturated cell is reached, deliver as much mass as possible,
- When the total mass has been distributed, restart the process starting from another oversaturated cell.

We then average this random projection on a great number of realizations in order to limit the error introduced in the computation. We refer to [15] or the heuristics used to derive this scheme. Let us illustrate on a toy example that in practical situations, despite the randomness of this projection step, the obtained density does not fluctuate much as we compute it several times with independent replicates.

Example 3. On a 2D squared grid of $N = 80$ subdivisions along each direction, we computed several times the projection of an oversatured disc carrying a density $\rho = 1.01$ on the area

$$\left(x - \frac{1}{2}\right)^2 + \left(y - \frac{1}{2}\right)^2 \leq 0.2^2. \tag{49}$$

For each computation, we run $\mathrm{N}_{\mathrm{test}} = 50$ parallel computations and take the average. We estimate so 20 different projections and then compute the 2-Wasserstein distances between the pairs of projected densities. The histogram of the computed distances is represented on figure 3. One sees that the distances lie in the range $[0.03, 0.05]$, whereas the total mass is $\rho_{\mathrm{tot}} = 800.93$ and the total moved mass is $\delta_\rho = 7.93$. The error is thus small regarding the amount of mass moved.

Whereas the convergence of the previous scheme is far beyond the scope of this paper and could be in itself an interesting research focus, let us recall a partial result of convergence in dimension 1 [21].

Theorem 4.1 (Theorem 5.2.1 in [21])**.** *Let μ be a positive density on $[-L, L]$, $\epsilon > 1$ and $\rho_\epsilon = \mathbf{1}_{[-L,L]} + \epsilon\mu$. Denote ν_ϵ the mean of the random projection of ρ_ϵ as the mesh size tends to 0 (we assume it converges). Then*

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} W_2(\nu_\epsilon, P_{K_1}(\rho_\epsilon)) = 0. \tag{50}$$

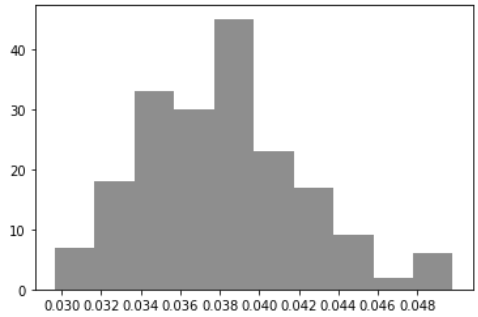

FIGURE 4. Distribution of the 2-Wasserstein distances between pairs of estimated projections of the density in example 3.

From this very partial convergence result and the stability of the projection illustrated in example 3, we expect this random method to prescribe a projection close to the actual K_1 projection and that does not depend much on the random realization of the process.

The correction step is thus computed with the following scheme :

Correction scheme for two types

- Project the sum of the predicted densities $\mu_1 + \mu_2$ on K_1 with the stochastic algorithm:
$$\rho = P_{K_1}(\mu_1 + \mu_2),$$
.
- Compute with Sinkhorn's algorithm the transport plan γ^ϵ from $\mu_1+\mu_2$ to ρ,
- Compute $f = \dfrac{d\mu_1}{d(\mu_1+\mu_2)}$ and $g = \dfrac{d\mu_2}{d(\mu_1+\mu_2)}$,
- Setting $\gamma_1^\epsilon(x,y) = f(x)\gamma^\epsilon(x,y)$ and $\gamma_2^\epsilon(x,y) = g(x)\gamma^\epsilon(x,y)$, the projection on K_2 is given by
$$(\pi^y_\# \gamma_1^\epsilon, \pi^y_\# \gamma_2^\epsilon) \tag{51}$$
where π^y is the projection on the y coordinate.

Remark 4. Let us compute (51) when $\gamma_T = (Id, T)_\# \mu$ is derived from a transport map and $\mu = \mu_1 + \mu_2$. One can compute that in this case
$$T_\# \mu_1 = \pi^y_\# \left(\frac{d\mu_1}{d\mu}(x) d\gamma_T(x,y) \right), \tag{52}$$
so the exact projection $(T_\#\mu_1)$ is indeed given by the previous algorithm when one has access to the actual transport map. In the case of regularized transport computed with Sinkhorn's algorithm, the plan is generically not a transport map. However, it is known [7] that the regularized transport plan γ^ϵ converges towards

the exact transport map γ_T as the regularization parameter ϵ is small. One thus has that γ_j^ϵ converges towards $\frac{d\mu_j}{d\mu}(x)d\gamma_T(x,y)$ and therefore that the approximated projection does so.

4.3. **More general velocities.** The splitting scheme allows the velocities to be more general than two time-constant velocity fields that are generic when using a JKO scheme. In the implementation we shall consider, we allow a mixture of different terms:

- a constant velocity field U_i,
- attraction towards an external potential: $-\nabla\phi_i$,
- attraction towards a chemoattractant emitted by the particles: ∇c where the concentration c satisfies a Segel equation with instantaneous diffusion, i.e.

$$-\kappa\Delta c = \rho_1 + \rho_2, \tag{53}$$

- attraction/repulsion between the particles: $\nabla(V * (\rho_1 + \rho_2))$ where $V(x,y) = \psi(\|x-y\|)$ is a potential of interaction.

4.4. **Numerical results.** A full implementation of the previous scheme is available[1] with a ready-to-use notebook demo that reproduces every simulation of this section. One can in particular find every parameter and initial condition used to generate these simulations.

On figure 5 is depicted the case of two saturated discs of radius $r = 0.15$ subject to constant fields that make the densities cross. One sees that the crowd finds a compromise to avoid oversaturation by spreading on a wider area while crossing. The desired velocities chosen are here constants

$$\begin{aligned} u_1 &= e_x \\ u_2 &= -e_x. \end{aligned} \tag{54}$$

On figure 6, chemoattraction is introduced. After spreading in order to cross, the two densities aggregate back into saturated areas. The desired velocities are taken here of the form

$$\begin{aligned} u_1 &= e_x + \nabla c_1 \\ u_2 &= -e_x + \nabla c_2 \\ -\kappa\Delta c_1 &= \rho_1 \\ -\kappa\Delta c_2 &= \rho_2. \end{aligned} \tag{55}$$

A third example of more complex behavior is illustrated on figure 7. In the absence of external velocities field or potential, the motion is only driven here by chemoattractraction and short-range interactions with the same type. The short range interaction potential has been chosen of the form

$$V(x,y) = \mathbf{1}_{\|x-y\|<R}\left(1 - \frac{\|x-y\|^2}{R^2}\right)^3. \tag{56}$$

We added a repulsion for the chemoattractant of the other type. One sees that the crowd reaches a compromise to separate into distinct phases.

[1] https://gitlab.math.u-psud.fr/bourdin/macroscopic-cell-motion

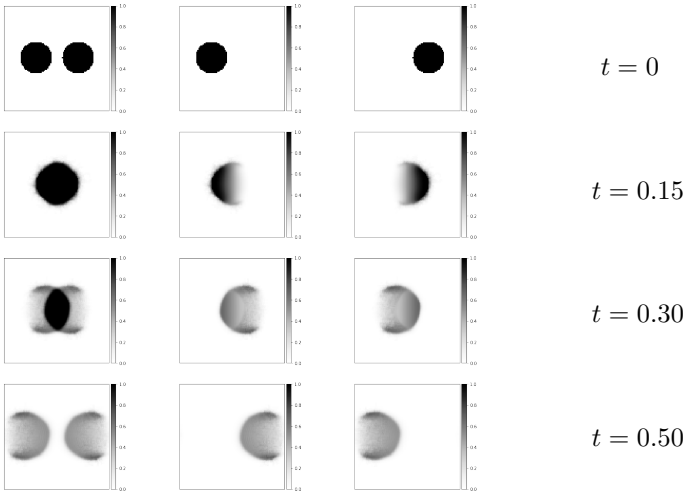

FIGURE 5. The motion of two crossing discs. The first column represents the sum of the two densities, and the other two the separated densities. The total time of the simulation is $T = 0.5$ for a timestep $dt = 0.01$ and a mesh size $N = 100$. The random projection step is averaged on 50 experiments.

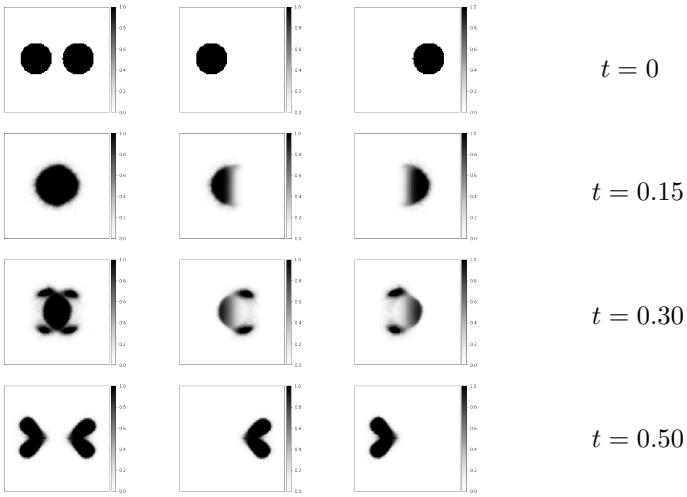

FIGURE 6. The motion of two crossing discs in the presence of chemoattraction. We chose the same parameters that in the previous simulation, with $\kappa = 10$.

The desired velocities write here

$$\begin{aligned} u_1 &= e_x + \nabla c_1 - \alpha \nabla c_2 + \eta \nabla (V * \rho_1) \\ u_2 &= -e_x + \nabla c_2 - \alpha \nabla c_2 + \eta \nabla (V * \rho_2) \\ -\kappa \Delta c_1 &= \rho_1 \\ -\kappa \Delta c_2 &= \rho_2. \end{aligned} \tag{57}$$

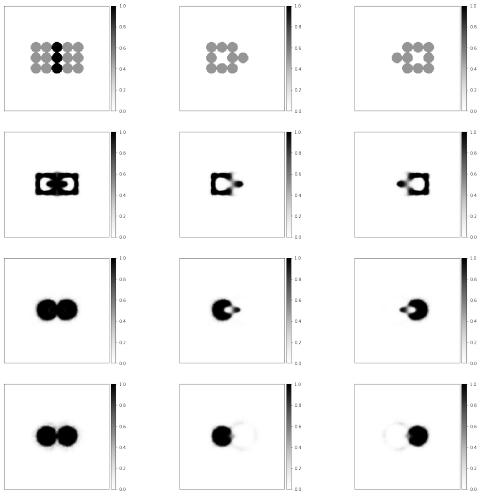

FIGURE 7. Aggregation of a composite crowd driven by chemoattraction and short-range interactions. For $t = 0.50$, we see small numerical diffusion due to the stochastic projection on K_2. The parameters here are $N = 150$, $\kappa = 8$, $\alpha = 0.2$, $\eta = 0.1$, $R = 0.04$. The random projection step is averaged on 100 runs and the mesh size is $N = 150$. The time parameters are still $T = 0.5$ and $dt = 0.01$.

Appendix.

Proof of lemma 3.4. The proof of the following lemma was suggested by Filippo Santambrogio.

Lemma 3.4. *Let Ω be a compact set, and μ_1, μ_2 two absolutely continuous measures, such that $\mu_1 \leq_* \mu_2$. Then $P_{K_1}(\mu_1) \leq_* P_{K_1}(\mu_2)$.*

Proof. Let $\mu_1 \leq_* \mu_2$ two absolutely continuous measures. Let $f_n : \mathbb{R} \longrightarrow \mathbb{R}_+$ a sequence of strictly convex C^1 functions, such that $f_n'(0) = f_n(0) = 0$, and $f_n'(\infty) = n$. We choose f_n such that f_n converges towards $f = \infty \mathbf{1}_{]1,\infty[}$. It is possible to choose $\|f_n - g_n\|_\infty \leq \frac{1}{n}$, with $g_n(x) = n(x-1)\mathbf{1}_{x>1}$, as represented on figure 8.

We consider the following optimization problems

$$\rho_j^n \in \underset{\rho \in W_2(\Omega)}{\operatorname{argmin}} W_2^2(\rho, \mu_j) + \int_\Omega f_n(\rho(x))dx. \tag{58}$$

with the convention ∞ if ρ has no density. The optimality conditions write

$$\varphi_j + f_n'(\rho_j^n) = c_j \tag{59}$$

where φ_j is a Kantorovich potential from ρ_j^n to μ_j and c_j is a constant. We aim at showing $\rho_1^n \leq_* \rho_2^n$. As f_n' is strictly increasing, it is sufficient to show that $f_n'(\rho_n^1) \leq f_n'(\rho_n^2)$ a.e. Denote

$$m = \inf_{y \in \Omega} f_n'(\rho_2^n(y)) - f_n'(\rho_1^n(y)). \tag{60}$$

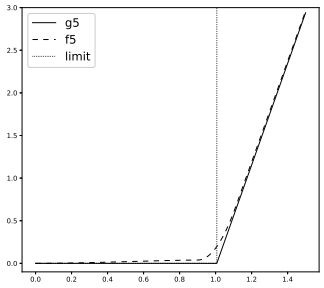

FIGURE 8. In solid line, the function g_5. In dashed, its approximation by a strictly convex smooth function f_5. In dotted line is displayed the common limit to f_n and g_n.

One can write

$$m = \inf_{y\in\Omega} c_2 - c_1 + \varphi_1 - \varphi_2. \tag{61}$$

m is realized on some $y_0 \in \Omega$ such that:

$$\begin{cases} \nabla\varphi_1(y_0) & = & \nabla\varphi_2(y_0) \\ \mathrm{Hess}(\varphi_1 - \varphi_2)_{|y_0} & \geq & 0 \end{cases} \tag{62}$$

Moreover, for $y \in \Omega$ one has

$$\rho_j^n(y) = |\det(I - D^2\varphi_j)_{|y}| \times \mu_j\left(y - \nabla\varphi_j(y)\right), \tag{63}$$

for the transport map from ρ_j^n to μ_j is of the form $T(x) = x - \nabla\phi_j(x)$. Assume $\operatorname{infess} \rho_2^n - \rho_1^n < 0$. Then $\operatorname{infess} f_n'(\rho_2^n) - f_n'(\rho_1^n) < 0$ so $m < 0$. Thus,

$$|\det(I - D^2\varphi_2)_{|y_0} \times |\mu_2\left(y_0 - \nabla\varphi_2(y_0)\right) < |\det(I - D^2\varphi_1)_{|y_0}| \times \mu_1\left(y_0 - \nabla\varphi_1(y_0)\right). \tag{64}$$

On the one hand, one has

$$\begin{aligned} \mu_2(y_0 - \nabla\varphi_2(y_0)) &= \mu_2(y_0 - \nabla\varphi_1(y_0)) \\ &\geq \mu_1(y_0 - \nabla\varphi_1(y_0)). \end{aligned} \tag{65}$$

On the other hand, denote $A = Id - D^2\varphi_2(y_0)$, $B = Id - D^2\varphi_1(y_0)$, $a = \det(A)$, $b = \det(B)$. A and B are Hessian matrix of convex functions so belong to $S_2^+(\mathbb{R})$.

Let

$$\gamma \begin{cases} [0,1] & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \det(B + t(A - B)). \end{cases} \tag{66}$$

The derivative of γ can be expressed with the comatrix of B: $\gamma'(t) = 1\mathrm{Tr}({}^T\mathrm{Com}(B)(A - B))$. It is nonnegative from (62). We get $a \geq b$ which is absurd with (64) and (65). One thus has $\rho_1^n \leq \rho_2^n$ a.e. Let us then show the convergence of the ρ_j^n. By optimality of ρ_j^n, one has

$$\begin{aligned} W_2^2(\rho_j^n, \mu_j) &\leq W_2^2(\rho_j^n, \mu_j) + \int_\Omega f_n(\rho_j^n) \\ &\leq W_2^2(\rho_j, \mu_j) + \int_\Omega f_n(\rho_j) \end{aligned} \tag{67}$$

where $\rho_j \in \underset{W_2(\Omega)}{\operatorname{argmin}} W_2^2(\rho_j, \mu_j) + \int_\Omega f(\rho_j)$.

By comparing f_n and g_n (increasing towards f), we get:

$$W_2^2(\rho_j^n, \mu_j) \leq W_2^2(\rho_j, \mu_j) + \int_\Omega f(\rho_j) + \frac{|\Omega|}{n}. \tag{68}$$

Let $\tilde{\rho}_j$ an accumulation point of ρ_j^n. $\tilde{\rho}_j$ satisfies:

$$W_2^2(\tilde{\rho}_j, \mu_j) \leq W_2^2(\rho_j, \mu_j). \tag{69}$$

Let us show that $\tilde{\rho}_j \in K_1$. One will then have by optimality $\tilde{\rho}_j = \rho_j$ and thus

$$\rho_1 = \lim_{n\to\infty} \rho_1^n \leq_* \lim_{n\to\infty} \rho_2^n = \rho_2. \tag{70}$$

Let $\epsilon > 0$ and $A = \{\tilde{\rho}_j > 1 + \epsilon\}$. Assume $\lambda(A) > 0$. One has

$$\limsup_{n\to\infty} \rho_j^n(A) \geq \tilde{\rho}_j(A) \geq (1+\epsilon)\lambda(A). \tag{71}$$

One can assume (by extracting)

$$\int_\Omega \rho_j^n(x)dx \geq \left(1 + \frac{\epsilon}{2}\right)\lambda(A). \tag{72}$$

Let $A_n = A \cap \left\{\rho_j^n > 1\right\}$, one has

$$\left(1 + \frac{\epsilon}{2}\right)\lambda(A) \leq \int_{A_n} \rho_j^n(x)dx + \int_{A\backslash A_n} 1dx. \tag{73}$$

For any $y > 1$, $y \leq \dfrac{f_n(y)}{n} + 1$, so

$$\left(1 + \frac{\epsilon}{2}\right)\lambda(A) \leq \int_{A_n} \frac{f_n(\rho_j^n(x))}{n}dx + \lambda(A_n) + \lambda\left(A\backslash A_n\right). \tag{74}$$

We get

$$\begin{aligned} \frac{n\epsilon}{2}\lambda(A) &\leq \int_{A_n} f_n(\rho_j^n(x))dx \\ &\leq \int_\Omega f_n(\rho_j^n(x))dx + W_2^2(\rho_j^n, \mu_j). \end{aligned} \tag{75}$$

Let χ a probability measure whose density is lower than 1 almost everywhere. By optimality, one has

$$\begin{aligned} \frac{n\epsilon}{2}\lambda(A) &\leq \int_\Omega f_n(\chi(x))dx + W_2^2(\chi, \mu_j) \\ &\leq \frac{\lambda(\Omega)}{n} + W_2^2(\chi, \mu_j) \end{aligned} \tag{76}$$

which is absurd when $n \to \infty$. Thus $\lambda(A) = 0$ and $\tilde{\rho}_j$ is admissible. □

REFERENCES

[1] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metricspaces and in the Space of Proba-bility Measures*, Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[2] T. M. Blackwell and P. Bentley, Don't push me! Collision-avoiding swarms, *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600)*, IEEE, **2** (2002).

[3] V. J. Blue and J. L. Adler, Cellular automata microsimulation for modeling bi-directional pedestrian walkways, *Transportation Research Part B: Methodological*, **35** (2001), 293–312.

[4] C. E. Brennen, *Fundamentals of Multiphase flow*, 2005.

[5] C. Burstedde, et al., Simulation of pedestrian dynamics using a two-dimensional cellular automaton, *Physica A: Statistical Mechanics and its Applications*, **295** (2001), 507–525.

[6] C. Cancés, T. O. Gallouét and L. Monsaingeon, Incompressible immiscible multiphase flows in porous media: A variational approach, *Anal. PDE*, **10** (2017), 1845–1876.

[7] G. Carlier, et al., Convergence of entropic schemes for optimal transport and gradient flows, *SIAM J. Math. Anal.*, **49** (2017), 1385–1418.

[8] J. A. Carrillo, M. P. Gualdani and G. Toscani, Finite speed of propagation in porous media by mass transportation methods, *C. R. Math.*, **338** (2004), 815–818.

[9] C. T. Crowe, *Multiphase Flow Handbook*, CRC press, 2005.

[10] M. Cuturi, Sinkhorn distances: Lightspeed computation of optimal transport, *Advances in Neural Information Processing Systems*, **26** (2013), 2292–2300.

[11] D. Helbing and P. Molnar, Social force model for pedestrian dynamics, *Physical review E*, **51** (1995), 4282.

[12] R. Jordan, D. Kinderlehrer and F. Otto, The variational formulation of the Fokker–Planck equation, *SIAM J. Math. Anal.*, **29** (1998), 1–17.

[13] I. Kim and A. R. Mészáros, On nonlinear cross-diffusion systems: An optimal transport approach, *Calc. Var. Partial Differential Equations*, **57** (2018), 1–40.

[14] B. Maury, A. Roudneff-Chupin and F. Santambrogio, A macroscopic crowd motion model of gradient flow type, *Math. Models Methods Appl. Sci.*, **20** (2010), 1787–1821.

[15] B. Maury, A. Roudneff-Chupin, F. Santambrogio and J. Venel, Handling congestion in crowd motion modeling, *Netw. Heterog. Media*, **6** (2011), 485–519.

[16] A. Mogilner and L. Edelstein-Keshet, A non-local model for a swarm, *J. Math. Biol.*, **38** (1999), 534–570.

[17] D. Morale, V. Capasso and K. Oelschläger, An interacting particle system modelling aggregation behavior: From individuals to populations, *J. Math. Biol.*, **50** (2005), 49–66.

[18] J. J. Moreau, Evolution problem associated with a moving convex set in a Hilbert space, *J. Differential Equations*, **26** (1977), 347–374.

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

[20] G. Peyré, Entropic approximation of Wasserstein gradient flows, *SIAM J. Imaging Sci.*, **8** (2015), 2323–2351.

[21] A. Roudneff-Chupin, Modélisation macroscopique de mouvements de foule, *Phdthesis, PhD Thesis, Université Paris-Sud XI*, 2011.

[22] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, Birkhäuser/Springer, Cham, 2015.

[23] R. H. Silsbee, Focusing in collision problems in solids, *J. Appl. Physics*, **28** (1957), 1246–1250.

[24] V. Shvetsov and D. Helbing, Macroscopic dynamics of multilane traffic, *Physical Review E*, **29** (1999), 6328.

Received January 2022; revised May 2022; early access June 2022.

E-mail address: felicien.bourdin@ens.fr