Measure sweeping processes

S. Di Marino, B. Maury, F. Santambrogio

June 18, 2015

Abstract

We propose and analyze a natural extension of the Moreau sweeping process: given a family of moving convex sets $(C(t))_t$, we look for the evolution of a probability density ρ_t, constrained to be supported on $C(t)$. We describe in detail three cases: in the first particles do not interact with each other and stay at rest unless pushed by the moving boundary; in the second they interact via a maximal density constraint $\rho \leq 1$, so that they are not only pushed by the boundary, but also by the other particles; in the third case particles are submitted to Brownian diffusion, reflected along the moving boundary. We prove existence, uniqueness and approximation results by using techniques from optimal transport, and we provide numerical illustrations.

1 Introduction

J.J. Moreau introduced in 1977 (see [31]) the so-called *sweeping process*, that describes the motion of a point $t \mapsto q(t)$ in a Hilbert space subject to remain in a moving closed convex set $C(t)$. The point is assumed to move as little as possible: it stands still as far as it lies in the interior of the C, and it is *caught up* by its moving boundary otherwise. The problem takes the form of a differential inclusion:

$$q'(t) \in -\partial I_{C(t)}(q(t)), \tag{1}$$

where I_C is the indicatrix function of the set C (it vanishes for any $q \in C$, and takes the value $+\infty$ for $q \notin C$).

This model was initially motivated by applications in plasticity (see [30]), and it has since been used in various domains, in particular granular materials and rigid body mechanics ([41, 37]), and more recently crowd motion models ([22, 34]).

In this case, the following result, that we present here with strong regularity assumptions on the prescribed motion, is well-known. The regularity of the motion of the convex set $C(t)$ is expressed in terms of the Hausdorff distance d_H.

Theorem 1.1. *Suppose that $(C(t))_t$ is a family of compact convex sets in $\mathbb{R}^d$ with $d_H(C(t), C(s)) \leq L|t-s|$. Then, for every initial point $q_0 \in C(0)$, Equation (1) admits a unique absolutely continuous solution $t \mapsto q(t)$. This solution*

- *satisfies $|q'(t)| \leq L$;*

- *is obtained as a uniform limit as $\tau \to 0$ of the curves q^τ obtained in the following way: first define recursively q^τ_k via*

$$\begin{cases} q^\tau_0 = q_0, \\ q^\tau_{k+1} = P_{C((k+1)\tau)}[q^\tau_k], \end{cases} \tag{2}$$

where P_C denotes the projection on the compact convex set C, then define the curve q^τ on each interval $[k\tau, (k+1)\tau]$ to be the affine interpolant between q^τ_k and q^τ_{k+1}.

The iterated projection scheme in Theorem 1.1 is the main tool to study this sweeping process and is called *catching up* method. This seminal work has given rise to a huge literature, and the process has been generalized to many situations. In particular, since the catching-up process needs a well-defined projection only in a neighborhood of C, the convexity assumption can be relaxed, and the result applies to so-called *prox regular* sets, for which the projection is well-defined in the proximity of the set (see e.g. [42, 8, 12, 17]). The evolution problem with a forcing term has also been considered, and associated control problems have been studied ([16]). Abstract extensions to the non-Hilbertian situation have also been given ([6]).

In this paper we want to study similar problems in the class of probability measures on $C(t)$. We will look for evolutions $\rho(t) \in \mathcal{P}(C(t))$ which are "pushed" by the movement of the boundary of $C(t)$.

We will look at three different cases.

The first is the easiest generalization of the sweeping process by Moreau: the measure $\rho(t)$ represents a collection of particles, and each of them stays at rest except when pushed by the $\partial C(t)$. They are allowed to superpose and create concentration, and in general they create singular measures on the boundary. We will see that this motion can be considered as a superposition of several sweeping processes, one for each particle.

Then, we move to a subject closer to the recent researches of the second and third author: we add a density constraint. This means that each particle tries to stay at rest, but can be pushed either by the boundary or by the other particles which are in between. Indeed, we impose a maximal density constraint $\rho(t, x) \leq 1$ (here we identify the probability measure ρ with its density) and it is possible that particles which are not on $\partial C(t)$ must move in order to comply with this density constraint. No concentration of mass is allowed in this model, and the movement is ruled by a pressure, which is part of the unknown, and plays the role of a Lagrange multiplier for the constraint $\rho \leq 1$. We will see that this model fits well the sweeping framework by Moreau, that it can be obtained by iterated projections, and we will study the corresponding PDE in a moving domain.

Finally, we study a somehow different model, where particles instead of being passively pushed by the motion of the convex set $C(t)$ also have Brownian diffusion. Observe that there is no symmetry between possible Brownian effects in the motion of the particles and of the sets (see for instance [12]) and that here we are interested in the case where the movement of the convex set is smooth (i.e., Lipschitz) and the Brownian motion in only present at the level of the particles. Moreover, since we consider the global evolution of a large number of particles, these effects only translate into a diffusion term in the equation, which becomes a heat equation on a moving domain with suitable no-flux boundary conditions. From the point of view of the techniques, this model is no longer attacked by iterated projections. Yet, using the well-known approach by Jordan, Kinderlehrer and Otto in [19], we can provide a scheme which is not so different, where instead of minimizing at each time step the distance to the previous configuration we also add an entropy term.

All these equations can be interpreted as Moreau processes in the space of measures, with a set

$K(t)$ of admissible measures which moves in time. In the first case we have

$$K(t) = \{\rho \in \mathcal{P}(\mathbb{R}^d) \,:\, \mathrm{spt}(\rho) \subset C(t)\}, \tag{3}$$

while in the second we have

$$K^1(t) = \{\rho \in \mathcal{P}(\mathbb{R}^d) \,:\, \mathrm{spt}(\rho) \subset C(t), \rho \leq 1\} \tag{4}$$

(in this second case we only use absolutely continuous measures, and we identify measures with their densities). The third case is somehow different, because it corresponds to a forced Moreau process (which would read as $x'(t) \in -\partial I_{C(t)} + f_t$ in the Euclidean case). The forcing corresponds to the action of the diffusion.

2 Preliminaries and notations

2.1 Moving convex sets

Here we recall some intuitive facts about moving convex sets; for the interested reader we report the proof of the main formula we will need, namely equation (6). Moreover we prove also a technical fact (Corollary 2.6), that will be useful for some estimates in the sequel.

Lemma 2.1. *Let C be a bounded convex domain. For every set A let us denote by $(A)_\varepsilon$ the ε-neighbourhood of A, namely the set of points such that $d(x, A) < \varepsilon$. Then:*

- *the perimeter* $\mathrm{Per}(C)$*, equal to* $\mathcal{H}^{d-1}(\partial C)$*, is finite;*
- $\varepsilon^{-1}\mathcal{L}^d|_{(\partial C)_\varepsilon} \rightharpoonup \mu + \mathcal{H}^{d-1}|_{\partial C}$*, where* $\mu \leq \mathcal{H}^{d-1}|_{\partial C}$*.*

Proof. The fact that C is a set of finite perimeter is well known and we omit its proof. For the second part of the statement we split $(\partial C)_\varepsilon = (\partial C)^+_\varepsilon \cup (\partial C)^-_\varepsilon$ where $(\partial C)^+_\varepsilon = (\partial C)_\varepsilon \setminus C$ and $(\partial C)^-_\varepsilon = (\partial C)_\varepsilon \cap C$. Given a smooth test function ϕ, we start from the following computation

$$\int_{(\partial C)^+_\varepsilon} \phi(P_C(x))dx = \varepsilon \int_{\partial C} \phi(x) d\mathcal{H}^{d-1} + o(\varepsilon)$$

where P_C is the projection onto C, and the $o(\varepsilon)$ term is indeed a polynomial in ε (from Steiner's formula, it equals $\sum_{j=2}^d \varepsilon^j \int_{\partial C} \phi \, d\mu_j$, where μ_j are the curvature measures of ∂C, see [40]). In particular when $\phi = 1$ we obtain $|(\partial C)^+_\varepsilon| = \varepsilon \mathrm{Per}(C) + o(\varepsilon)$. Then we go on with

$$\begin{aligned} \int_{(\partial C)^+_\varepsilon} \phi(x)dx &= \textstyle\int_{(\partial C)^+_\varepsilon} \phi(P_C(x))dx + \int_{(\partial C)^+_\varepsilon} (\phi(x) - \phi(P_C(x)))dx \\ &= \varepsilon \textstyle\int_{\partial C} \phi(x) d\mathcal{H}^{d-1} + o(\varepsilon) + \varepsilon \mathrm{Lip}(\phi)|(\partial C)^+_\varepsilon|. \end{aligned} \tag{5}$$

For the $(\partial C)^-_\varepsilon$ part, we consider $\phi \geq 0$ and we note that, for convex polyhedra C_j, for every projection P_j on ∂C_j (the projection from the interior is no more unique), we have

$$\int_{(\partial C_j)^-_\varepsilon} \phi(P_j(x)) \, dx \leq \varepsilon \int_{\partial C_j} \phi \, d\mathcal{H}^{d-1};$$

Again, since $|P_j(x) - x| \leq \varepsilon$ for $x \in (\partial C_j)^-_\varepsilon$, we have

$$\int_{(\partial C_j)^-_\varepsilon} \phi(x) \, dx \leq \varepsilon \int_{\partial C_j} \phi \, d\mathcal{H}^{d-1} + \mathrm{Lip}(\phi)\mathrm{Per}(C_j)\varepsilon^2.$$

Furthermore P_j is a 1-Lipschitz function on ∂C if $C_j \subseteq C$ and so it is true that $(P_j)_\sharp \mathcal{H}^{d-1}|_{\partial C} \geq \mathcal{H}^{d-1}|_{\partial C_j}$. In particular $\int_{\partial C_j} \phi d\mathcal{H}^{d-1} \leq \int_{\partial C} (\phi \circ P_j) d\mathcal{H}^{d-1}$ and $\mathrm{Per}(C_j) \leq \mathrm{Per}(C)$. Hence we have

$$\begin{aligned}\int_{(\partial C_j)_\varepsilon^-} \phi(x)\,dx &\leq \varepsilon \int_{\partial C} \phi\, d\mathcal{H}^{d-1} + \mathrm{Lip}(\phi)\mathrm{Per}(C_j)\varepsilon^2 + \varepsilon \int_{\partial C} (\phi(P_j(x)) - \phi(x)) d\mathcal{H}^{d-1} \\ &\leq \varepsilon \int_{\partial C} \phi\, d\mathcal{H}^{d-1} + \varepsilon \mathrm{Per}(C)\mathrm{Lip}(\phi)(\varepsilon + d_H(C, C_j))\end{aligned}$$

Thus, it is sufficient to approximate C in the Hausdorff sense from the interior with polyhedra (for instance considering the convex hull of suitable finite sets of points $X_j = \{x_i^j\}_{i \leq I_j} \subset \partial C$). Letting $j \to \infty$ we obtain

$$\int_{(\partial C)_\varepsilon^-} \phi(x)\,dx \leq \varepsilon \int_{\partial C} \phi\, d\mathcal{H}^{d-1} + \varepsilon^2 \mathrm{Per}(C)\mathrm{Lip}(\phi)$$

which, together with (5), gives the claim. $\square$

Remark 2.2. *We note that if ∂C is a C^2 hyper-surface then it is of positive reach and so we can apply directly Steiner formula to $(\partial C)_\varepsilon$ as shown by Federer, obtaining in that case that $\mu = \mathcal{H}^{d-1}|_{\partial C}$. Yet, for the sake of our proofs the inequality is enough, and is easier to prove.*

Lemma 2.3. *Let $(C(t))_t$ be a curve of convex sets, contained in a larger convex set $\Omega \subset \mathbb{R}^d$, that is L-Lipschitz with respect to the Hausdorff distance. Then for every $\phi \in H^1(\mathbb{R}^d)$ the map $t \mapsto \int_{C_t} \phi$ is absolutely continuous and there exist a scalar field $V_t : \partial C(t) \to \mathbb{R}$ such that*

$$\frac{d}{dt} \int_{C(t)} \phi\, dx = \int_{\partial C(t)} V_t \phi\, d\mathcal{H}^{d-1} \qquad \textit{for a.e. } t. \tag{6}$$

Moreover $|V_t| \leq 2L$.

Proof. Thanks to the assumption we know that $C(t)\Delta C(t+h) \subset (\partial C(t))_h^- \cup (\partial C(t+h))_h^-$. In particular $|C(t)\Delta C(t+h)| \leq CLh$ for some $C = C(\Omega)$. Now we consider a continuous function ϕ. Let us consider the linear functional $F_\phi(t) = \int_{C(t)} \phi\, dx$. Since

$$|F_\phi(t) - F_\phi(s)| \leq \int_{C(t)\Delta C(s)} |\phi|\, dx \leq C\|\phi\|_\infty |t-s| L,$$

we obtain that F_ϕ is Lipschitz and so in particular the derivative exists for a.e. time t; we call it $L_t(\phi)$. Now the set of continuous functions in a bounded set is separable and so we can consider a countable dense set and we have that for almost every $t \in [0, T]$ there exists $L_t(\phi_n)$ for every n. Moreover it is obvious that L_t is linear and continuous and so we can extend it to the whole set of continuous functions. In particular L_t can be represented by a finite measure μ_t. In order to have an estimate we consider $t, h > 0$, and for every continuous function ϕ we have

$$\limsup_{h \to 0} \frac{|F_\phi(t+h) - F_\phi(t)|}{h} \leq \limsup_{h \to 0} \frac{1}{h} \int_{(\partial C(t))_{Lh}} |\phi| dx \leq 2L \int_{\partial C} |\phi|\, d\mathcal{H}^{d-1}$$

(for the last inequality we use the statement of Lemma 2.1)

In this way we obtain $|\mu_t| \leq 2L\mathcal{H}^{d-1}$, and so the thesis. The extension to H^1 function can be done in this way: we notice that the right hand side in (6) is a bounded linear functional in H^1 thanks to the trace inequality. Then the integral version of (6) is true by density. We observe that the Lipschitz bound holds in the integral sense and so $t \mapsto \int_{C_t} \phi$ is really Lipschitz, which gives the claim. $\square$

Remark 2.4. *For the sake of our proofs we only need an L^∞ bound on V_t, and we gave $2L$ which is easier to prove, but one actually expects the sharp bound to be L.*

Lemma 2.5. *Let $C(t)$ be a family of convex sets such that $\inf_t |C(t)| = c > 0$ and $C(t) \subset \Omega$ where Ω is a fixed bounded domain. Then there exists a radius r such that for every t there exists $x_t \in C(t)$ such that $B(x_t, r) \subset C(t)$.*

Proof. It is sufficient to prove that the inradius of $C(t)$ is bounded from below. We observe that the diameter of the the sets $C(t)$ is bounded by some $R > 0$. Now we know by John's Lemma (see for instance [26]) that for every convex set C there exists an ellipsoid E such that up to translation $E \subset C \subset dE$ and in particular, denoting by $\lambda_1 \leq \ldots \leq \lambda_n$ the lengths of the principal axis of E we have that $\omega_n n^n \lambda_1 \cdots \lambda_d \geq |C|$, $r \geq \lambda_1$ and $D \geq 2\lambda_n$ where r and D are respectively the inradius and the diameter of C. Using these inequalities together we find that

$$\omega_d d^d r (D/2)^{d-1} \geq |C|.$$

But, from the lower bound $|C(t)| \geq c$ and $D \leq R$, we find a bound from below on r. □

Corollary 2.6. *Let $(C(t))_t$ be a curve of convex sets that is L-Lipschitz with respect to the Hausdorff distance such that $\inf_t |C(t)| = c > 1$. Then, for every $t \in [0, T]$ and every $h > 0$, there exists a C^1 map $T : C(t) \to C(t + h)$ such that $||T - id||_{L^\infty} \leq MLh$ and $\det(DT) \geq 1 - MLh$, for some constant $M = M(c, d)$.*

Proof. We are in the same hypothesis of Lemma 2.5, and so we there exists $r > 0$ and $x_t \in C(t)$ such that $B(x_t, r) \subset C(t)$. Now we claim that $(1 - \frac{Lh}{r})C(t) + \frac{Lh}{r}x_t \subseteq C(t+h)$. If this was not the case, then there would be a point $x = (1 - \frac{Lh}{r})y + \frac{Lh}{r}x_t$, with $y \in C(t)$, such that $x \notin C(t+h)$. For every $z \in B(0, r)$ we would also have $x + \frac{Lh}{r}z = (1 - \frac{Lh}{r})y + \frac{Lh}{r}(x_t + z) \in C(t)$. Yet, from $x \notin C(t+h)$ we deduce that there is an hyperplane separating x and $C(t + h)$, which means that if we choose z in a suitable direction orthogonal to such an hyperplane, we get $d(x + \frac{Lh}{r}z, C(t + h)) > \frac{Lh}{r}|z|$. Choosing $|z| = r$ we get a contradiction to the Lipschitz behavior of $C(t)$.

Now that we know $(1 - \frac{Lh}{r})C(t) + \frac{Lh}{r}x_t \subseteq C(t + h)$ we can take $T(x) = (1 - \frac{Lh}{r})x + \frac{Lh}{r}x_t$, which satisfies all the required properties. □

Sometimes in the paper we will need some functional inequalities with constants independent of time. This proposition collects them all:

Proposition 2.7. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain and $c > 1$ be a real constant. Then there exists a constant β such that for every convex set $C \subseteq \Omega$ such that $|C| \geq c$ we have*

$$\int_{\partial C} u^2 \, d\mathcal{H}^{d-1} \leq \beta \int_C (u^2 + |\nabla u|^2)\, dx \qquad \forall u \in H^1(C)$$

$$\int_C u^2 \, dx \leq \beta \int_C |\nabla u|^2 \, dx \qquad \forall u \in H^1(C) \ s.t. \ |\{u = 0\}| \geq c - 1.$$

Proof. The first is the classical trace inequality while for the second one we refer to [10]. □

2.2 Optimal transport and curves of measures

We recall here the main notion that we will use from the Monge-Kantorovich optimal transport theory. We refer to [39] (Chapters 1, 5 and 7) and to [4, 44] for more details and complete proofs.

If two probabilities $\mu, \nu \in \mathcal{P}(\Omega)$ are given on a domain $\Omega \subset \mathbb{R}^d$ (that we take compact for simplicity), we consider

$$\min \left\{ \int_{\Omega\times\Omega} \frac{1}{2}|x-y|^2 \, d\gamma \; : \; \gamma \in \Pi(\mu,\nu) \right\},$$

where $\Pi(\mu,\nu)$ is the set of the so-called *transport plans*, i.e.

$$\Pi(\mu,\nu) = \{\gamma \in \mathcal{P}(\Omega\times\Omega) : \, (p^x)_\sharp \gamma = \mu, \, (p^y)_\sharp\gamma = \nu, \},$$

p^x and p^y being the two projections of $\Omega\times\Omega$ onto Ω. It is an extension of the Monge problem, which is

$$\inf \left\{ \int \frac{1}{2}|x-T(x)|^2 d\mu \; : \; T : \Omega \to \Omega, \; T_\sharp \mu = \nu \right\}$$

(in the sense that to any transport map T we can associate a transport plan γ_T by taking $\gamma_T = (id \times T)_\sharp \mu$, that the cost of T in the Monge problem is the same as that of γ_T in Kantorovich's one, and that, under some additional assumption on μ, the minimum over the transport plans is realized by a plan of the form γ_T).

For the above problem one can prove that the minimal value also equals the maximal value of a dual problem

$$\max \left\{ \int \varphi \, d\mu + \int \psi \, d\nu \; : \; \varphi(x) + \psi(y) \leq \frac{1}{2}|x-y|^2 \right\}, \tag{7}$$

and that the optimal function φ (called *Kantorovich potential*) may be used to construct an optimizer γ. Indeed, the optimal φ is locally lipschitz and semiconcave (in particular $x \mapsto \frac{1}{2}|x|^2 - \varphi(x)$ is convex) on $\mathrm{spt}(\mu)$ and differentiable $\mu-$a.e. if $\mu \ll \mathcal{L}^d$; one can define a map $\mathrm{T} : \Omega \to \Omega$ through $\mathrm{T}(x) = x - \nabla\varphi(x)$ and this map satisfies $\mathrm{T}_\sharp \mu = \nu$ and $\gamma_{\mathrm{T}} := (id, \mathrm{T})_\sharp \mu$ (i.e. the image measure of μ through the map $x \mapsto (x, \mathrm{T}(x))$) belongs to $\Pi(\mu,\nu)$ and is optimal in the above problem. Moreover, the map T is the gradient of the convex function u given by $u(x) = \frac{1}{2}|x|^2 - \varphi(x)$ and is called the optimal transport map from μ to ν. The fact that the optimal transport map T exists, is unique, and is the gradient of a convex function is known as Brenier Theorem (see [9]).

However, independently of the fact that the minimum is realized by a transport map or not, we can use the minimal value of the above problem to define the distance $W_2(\mu,\nu)$ between two measures μ and ν

$$W_2(\mu,\nu) := \sqrt{\min \left\{ \int |x-y|^2 \, d\gamma \; : \; \gamma \in \Pi(\mu,\nu) \right\}}.$$

When Ω is compact, this quantity may be proven to be a distance over $\mathcal{P}(\Omega)$ and to metrize the weak-* convergence of probability measures. The space $\mathcal{P}(\Omega)$ endowed with the distance W_2 is called Wasserstein space of order 2 and denoted in this paper by $\mathbb{W}_2(\Omega)$.

Another fact that we need to know concerning the Kantorovich potential ϕ and the Wasserstein distance is the following: the function ϕ also plays the role of the derivative of $\frac{1}{2}W_2^2(\cdot,\nu)$. Indeed, we have,

$$\frac{d}{d\varepsilon}\frac{1}{2}W_2^2(\mu + \varepsilon\chi, \nu)_{|\varepsilon=0} = \int \varphi d\chi$$

whenever $\mu + \varepsilon\chi \in \mathcal{P}(\Omega)$ and the Kantorovich potential φ in the transport from μ to ν is unique up to additive constants. This will be useful whenever we need to write optimality conditions for minimization problems involving W_2^2. It can be proven by using Monge-Kantorovich duality, and details are, for instance, in Chapter 7 of [39].

With the help of the Wasserstein distance, we can also study the continuity and absolute continuity of curves of measures. In particular, the following characterization has been proven in [4]: a curve

$[0,T] \ni t \mapsto \rho_t \in \mathbb{W}_2(\Omega)$ is absolutely continuous if and only if there exists a family of vector fields $(v_t)_t$ with $v_t \in L^2(\rho_t)$ and $\int_0^T ||v_t||_{L^2(\rho_t)} dt < +\infty$ which solves the continuity equation

$$\partial_t \rho_t + \nabla \cdot (\rho_t v_t) = 0$$

in the sense of distribution on $\mathbb{R}^d$. This means that $t \mapsto \rho_t$ is weakly continuous at $t = 0$ and $t = T$ and that, for every C^1, bounded and compactly supported function on $\mathbb{R}^d \times [0,T]$, we have

$$\int_0^T \int \partial_t \phi(t,x)\, d\rho_t + \int_0^T \int \nabla \phi(t,x) \cdot v_t\, d\rho_t = 0,$$

where the integrals in space are actually performed only on $\operatorname{spt} \rho_t \subset \Omega$, but the test functions are not required to vanish on $\partial\Omega$. Equivalently, we can use test functions only depending on x, and write for a.e. t

$$\frac{d}{dt} \int \phi(x)\, d\rho_t = \int \nabla \phi(x) \cdot v_t\, d\rho_t.$$

Using general notions from analysis in metric spaces (see for instance [5]), we define the metric derivative

$$|\rho'|(t) := \lim_{h \to 0} \frac{W_2(\rho_t, \rho_{t+h})}{|h|}$$

whenever the limit exists (and it exists for a.e. t if the curve is absolutely continuous). Then, the characterization above can be strengthened in the following way: for every family $(v_t)_t$ solving the continuity equation, we have $||v_t||_{L^2(\rho_t)} \geq |\rho'|(t)$ for a.e. t, and moreover there exists a (unique) "minimal" one, such that we have the equality $||v_t||_{L^2(\rho_t)} = |\rho'|(t)$. This vector field is called *velocity field* of the curve $(\rho_t)_t$.

A useful tool (see [4] or [39]) in the study of evolution PDEs in the form of continuity equations is the following. Let $(\rho_t^{(i)}, v_t^{(i)})$ for $i = 1,2$ be two solutions of the continuity equation $\partial_t \rho_t^{(i)} + \nabla \cdot (v_t^{(i)} \rho_t^{(i)}) = 0$ on a compact domain Ω and suppose that $\rho_t^{(i)} \ll \mathcal{L}^d$ for every t and that $\varrho^{(i)}$ are absolutely continuous curves in $\mathbb{W}_2(\Omega)$. Then we have,

$$\frac{d}{dt}\left(\frac{1}{2} W_2^2(\rho_t^{(1)}, \rho_t^{(2)})\right) = \int \nabla \varphi_t \cdot v_t^{(1)} \rho_t^{(1)}\, dx + \int \nabla \psi_t \cdot v_t^{(2)} \rho_t^{(2)}\, dx \tag{8}$$

for a.e. t, where (φ_t, ψ_t) is any pair of Kantorovich potentials in the transport between $\rho_t^{(1)}$ and $\rho_t^{(2)}$ for the cost $\frac{1}{2}|x-y|^2$.

Among the curves in the metric space $\mathbb{W}_2(\Omega)$, the geodesics in this space play an important role in the theory of optimal transport. If $\mu, \nu \in \mathcal{P}(\Omega)$ and $\mu \ll \mathcal{L}^d$, we define $\rho_t := ((1-t)id + tT)_\sharp \mu$, where T is the optimal trnasport from μ to ν. This curve ρ_t happens to be a constant speed geodesic for the distance W_2 connecting μ to ν (in case neither μ nor ν are absolutely continuous, it is possible to produce a geodesic by taking $\rho_t := (\pi_t)_\sharp \gamma$ where $\pi_t(x,y) = (1-t)x + ty$ and γ is optimal in the Kantorovich problem, which gives the same result if $\gamma = \gamma_T$). The velocity field corresponding to this curves (in the absolutely continuous case) is given by $v_t = (T - id) \circ ((1-t)id + tT)^{-1}$.

2.3 General construction for evolution problems in $\mathbb{W}_2(\Omega)$

We recall here the construction used to provide, via a time-discretization, a solution to many evolution equations in the Wasserstein space $\mathbb{W}_2(\Omega)$. We refer to [23] and to Chapter 8 in [39] for more technical details about this procedure, which has mainly been used for equations with a variational structure (gradient flows), but can be presented in general as we do here below. Equivalently, the reader can refer to [4], which presents a more abstract framework.

In many equations, we need to produce a curve $(\rho_t)_t$ which is a solution of a PDE (in the distributional sense on $\mathbb{R}^d$) of the form $\partial_t \rho_t + \nabla \cdot (\rho_t v_t) = 0$, where we require $v_t \in A(t, \rho_t)$, the set $A(t, \rho)$ being the set of vector fields which satisfy a certain compatibility condition with the density ρ. For instance, for the linear heat equation we take $A(t, \rho) = \{-\nabla \rho / \rho\}$ and for the linear continuity equation with given advection field u we take $A(t, \rho) = \{u_t\}$. In the crowd motion model studied in [23] we have $A(t, \rho) = \{u_t - \nabla p : p \geq 0, p(1 - \rho) = 0\}$. Note that, to provide a meaning to the above continuity equation, we need at least to require that $E = \rho v$ is a finite measure over $\Omega \times [0, T]$ (the variable E is called momentum), acting on functions ϕ via $\langle E, \phi \rangle := \int dt \int \phi(t, x) \cdot v_t \, d\rho_t$. This condition is equivalent to $\int_0^T ||v_t||_{L^1(\rho_t)} dt < +\infty$.

The method that we present consists in a time-discretization. For each $\tau > 0$, we build a sequence $(\rho_k^\tau)_k$. We also define a sequence of velocities $v_k^\tau = (id - \mathrm{T})/\tau$, taking as T the optimal transport from ρ_k^τ to ρ_{k-1}^τ. We need to choose the sequence $(\rho_k^\tau)_k$ so that we have, at least in some approximate sense, $v_k^\tau \in A(k\tau, \rho_k^\tau)$.

Then, we build at least two interesting curves in the space of measures:

- first we can define some piecewise constant curves, i.e. $\overline{\rho}_t^\tau := \rho_{k+1}^\tau$ for $t \in]k\tau, (k+1)\tau]$; associated to this curve we also define the velocities $\overline{v}_t^\tau = v_{k+1}^\tau$ for $t \in]k\tau, (k+1)\tau]$ and the momentum variable $\overline{E}^\tau = \overline{\rho}^\tau \overline{v}^\tau$;
- then, we can also consider the densities $\widehat{\rho}_t^\tau$ that interpolate the discrete values $(\rho_k^\tau)_k$ along geodesics:
$$\widehat{\rho}_t^\tau = \big((k\tau - t)v_k^\tau + id\big)_\sharp \rho_k^\tau, \quad \text{for } t \in](k-1)\tau, k\tau[; \tag{9}$$
the velocities $\widehat{v}_t^\tau$ are defined so that $(\widehat{\rho}^\tau, \widehat{v}^\tau)$ satisfy the continuity equation, taking
$$\widehat{v}_t^\tau = v_t^\tau \circ \big((k\tau - t)v_k^\tau + id\big)^{-1};$$
as before, we define: $\widehat{E}^\tau = \widehat{\rho}^\tau \widehat{v}^\tau$.

After these definitions we look for a priori bounds on the curves and the velocities that we defined. In many cases it is possible to obtain
$$\sum_k \tau \left(\frac{W_2(\rho_k^\tau, \rho_{k-1}^\tau)}{\tau} \right)^2 \leq C, \tag{10}$$
which is the discrete version of an H^1 estimate. As for $\widehat{\rho}_t^\tau$, it is an absolutely continuous curve in the Wasserstein space and its velocity on the time interval $[(k-1)\tau, k\tau]$ is given by the ratio $W_2(\rho_{k-1}^\tau, \rho_k^\tau)/\tau$. Hence, the L^2 norm of its velocity on $[0, T]$ is given by
$$\int_0^T |(\widehat{\rho}^\tau)'|^2(t) dt = \sum_k \frac{W_2^2(\rho_k^\tau, \rho_{k-1}^\tau)}{\tau}, \tag{11}$$
and, thanks to (10), it admits a uniform bound independent of τ. In our case, thanks to results on the continuity equation and the Wasserstein metric, this metric derivative is also equal to $||\widehat{v}_t^\tau||_{L^2(\widehat{\rho}_t^\tau)}$. This gives compactness of the curves $\widehat{\rho}^\tau$, as well as an Hölder estimate on their variations (since $H^1 \subset C^{0,1/2}$). The characterization of the velocities $\overline{v}^\tau$ and $\widehat{v}^\tau$ allows to deduce bounds on these vector fields from the bounds on $W_2(\rho_{k-1}^\tau, \rho_k^\tau)/\tau$.

Considering all these facts, one obtains the following situation.

- The norm $\int ||\overline{v}_t^\tau||^2_{L^2(\overline{\rho}_t^\tau)} dt$ is τ-uniformly bounded.
- In particular, the bound is valid in L^1 as well, which implies that $\overline{E}^\tau$ is bounded in the space of measures over $[0,T] \times \Omega$.
- The very same estimates are true for $\widehat{v}^\tau$ and $\widehat{E}^\tau$.
- The curves $\widehat{\rho}^\tau$ are bounded in $H^1([0,T], \mathbb{W}_2(\Omega))$ and hence compact in $C^0([0,T], \mathbb{W}_2(\Omega))$.
- Up to a subsequence, one has $\widehat{\rho}^\tau \to \rho$, as $\tau \to 0$, uniformly according to the W_2 distance.
- From the estimate $W_2(\overline{\rho}_t^\tau, \widehat{\rho}_t^\tau) \leq C\tau^{1/2}$ one gets that ρ^τ converges to the same limit ρ in the same sense.
- If we denote by E a weak limit of $\widehat{E}^\tau$, since $(\widehat{\rho}^\tau, \widehat{E}^\tau)$ solves the continuity equation, by linearity, passing to the weak limit, also (ρ, E) solves the same equation.
- It is possible to prove (see [23] Section 3.2, Step 1, or Chapter 8 in [39]) that the weak limits of $\widehat{E}^\tau$ and $\overline{E}^\tau$ are the same.
- From the bounds in L^2 one gets that also the measure E is absolutely continuous w.r.t. ρ and has an L^2 density, so that we have for a.e. time t a measure E_t of the form $\rho_t v_t$.
- It is only left to prove that one has $v_t \in A(t, \rho_t)$ for a.e. t. This is done by passing to the limit in a suitable sense as $\tau \to 0$. It is crucial in this step to consider the limit of $(\overline{\rho}^\tau, \overline{E}^\tau)$ instead of $(\widehat{\rho}^\tau, \widehat{E}^\tau)$ and exploit the properties of ρ_k^τ and v_k^τ.

Summarizing, it is possible to produce a solution of the PDE

$$\partial_t \rho_t + \nabla \cdot (\rho_t v_t) = 0, \qquad \text{with } v_t \in A(t, \rho_t)$$

whenever we have a discrete sequence (ρ_k^τ) with velocities v_k^τ obtained as above such that

- we have $\sum_k \tau \left(\frac{W_2(\rho_k^\tau, \rho_{k-1}^\tau)}{\tau} \right)^2 \leq C$
- we have, at least in some approximate sense, $v_k^\tau \in A(k\tau, \rho_k^\tau)$
- we can prove at the limit $v_t \in A(t, \rho_t)$ for a.e. t.

In every concrete PDE example, these are the conditions that we need to check.

We notice that the above scheme, with the choice $v_k^\tau = (id - \mathrm{T})/\tau$, is only able to produce vector fields v_t which have a gradient structure, since $\mathrm{T}(x) = x - \nabla\phi(x)$ is always a gradient. This is enough for the scopes of this paper, but different choices are possible. Indeed, if one takes a different transport map T with $\mathrm{T}_\sharp(\rho_k^\tau) = \rho_{k-1}^\tau$, the same scheme can be performed but we need to check

$$\sum_k \tau ||v_k^\tau||^2_{L^2(\rho_k^\tau)} \leq C$$

instead of the estimate with the Wasserstein distances.

3 Sweeping of a measure

We are interested in this section in the movement of a collection of particles "pushed" by the moving domain $C(t)$. We require that all particles are contained, at every instant of time, in the closed domain $C(t)$ (they can possibly be on the boundary). Then, the velocity $v(t,x)$ of a particle located at $x \in C(t)$ at time t is required to satisfy $v(t,x) \in -\partial I_{C(t)}(x)$, which implies in particular $v(t,x) = 0$ whenever x is in the interior of $C(t)$.

Mathematically, we look for a family of probability densities $(\rho_t)_{t\in[0,T]}$, satisfying the continuity equation

$$\begin{cases} \partial_t \rho_t + \nabla \cdot (\rho_t v_t) = 0, \\ v_t(x) \in -\partial I_{C(t)}(x), \\ \mathrm{spt}(\rho_t) \subset C(t). \end{cases} \tag{12}$$

We require a very mild regularity on the curve $t \mapsto \rho_t$, i.e. $\int_0^T ||v_t||_{L^1(\rho_t)} dt < +\infty$, which is a minimal requirement to give a meaning, as a measure, to the term inside the divergence.

We observe that this equation can be put in the framework of Section 2.3, considering as admissible vector fields

$$A(t,\rho) = \{v \in L^1(C(t);\mathbb{R}^d) \,:\, v(x) \in -\partial I_{C(t)}(x)\}.$$

When we want to see this evolution as a Moreau process in the Wasserstein space, it corresponds to the choice

$$K(t) = \{\rho \in \mathcal{P}(\Omega) \,:\, \mathrm{spt}\,\rho \subset C(t)\}, \tag{13}$$

where Ω denotes a large domain containing all the convex sets $C(t)$ for $t \in [0,T]$.

We will prove the following result.

Theorem 3.1. *Suppose that $(C(t))_t$ is a family of compact convex sets with $d_H(C(t),C(s)) \leq L|t-s|$. Then, for every initial datum $\rho_0 \in \mathcal{P}(C(0))$ the solution to the Equation* (12)

- *exists;*
- *is unique (in the sense that ρ_t is uniquely defined, as a measure, for every t);*
- *satisfies $|v_t(x)| \leq L$ for ρ_t-almost every x;*
- *is obtained as $\rho_t := (Q_t)_\sharp \rho_0$, where $Q_t(x)$ denotes the value at time t of the solution $q(t)$ of the pointwise Moreau process* (1) *with $q(0) = x$;*
- *is obtained as a limit as $\tau \to 0$ of the curves $\widehat{\rho}^\tau$ obtained in the following way: first define recursively ρ_k^τ via*

  $$\begin{cases} \rho_0^\tau = \rho_0, \\ \rho_{k+1}^\tau = P_{K((k+1)\tau)}[\rho_k^\tau], \end{cases}$$

  where $P_{K(t)}$ denotes the projection, according to the Wasserstein distance W_2, onto the set $K(t)$, then define the curve $\widehat{\rho}^\tau$ on each interval $[k\tau,(k+1)\tau]$ to be the geodesic interpolation between ρ_k^τ and ρ_{k+1}^τ.

Proof. First of all, we define the recursive sequence as in the last point of the claim, and we observe that, whenever we take $K = \{\rho \in \mathcal{P}(\Omega) \,:\, \mathrm{spt}\,\rho \subset C\}$, we have

$$P_K[\rho] = (P_C)_\sharp \rho,$$

which means that projecting a probability ρ onto the set of probabilities supported on C is equivalent to projecting onto C each particle of ρ, and taking the image measure. Moreover, P_C is obviously a transport map (it is indeed the optimal one) from ρ to $P_K[\rho]$, and hence we have

$$W_2(\rho, P_K[\rho]) \le ||P_C - id||_{L^2(\rho)} \le d_H(C, \tilde{C}),$$

whenever $\operatorname{spt}\rho \subset \tilde{C}$. If we apply this estimate to the case $\rho = \rho_k^\tau$, $\tilde{C} = C(k\tau)$, $C = C((k+1)\tau)$ we get $W_2(\rho_{k+1}^\tau, \rho_k^\tau) \le L\tau$.

We define a vector field v_k^τ as

$$v_k^\tau(x) = \frac{P_{C(k\tau)}(x) - x}{\tau}$$

and we observe that we have $|v_k^\tau| \le L$.

We can use $(v_k^\tau)_k$ and $(\rho_k^\tau)_k$ to produce a piecewise constant pair $(\overline{\rho}_t^\tau, \overline{v}_t^\tau)$ defined as

$$(\overline{\rho}_t^\tau, \overline{v}_t^\tau) = (\rho_k^\tau, v_k^\tau) \ \text{ if } t \in]k\tau, (k+1)\tau].$$

We also define another pair, by setting

$$\mathrm{T}_t := \frac{(k+1)\tau - t}{\tau} id + \frac{t - k\tau}{\tau} P_{C((k+1)\tau)} \quad \text{for } t \in]k\tau, (k+1)\tau]$$

and taking

$$\widehat{\rho}_t^\tau = (\mathrm{T}_t)_\sharp \rho_k^\tau;\ \widehat{v}_t^\tau := v_k^\tau \circ (\mathrm{T}_t)^{-1}.$$

We also set $\overline{E}_t^\tau := \overline{\rho}_t^\tau v^\tau$ and $\widehat{E}_t^\tau := \widehat{\rho}_t^\tau \widehat{v}_t^\tau$. The curve $\widehat{\rho}_t^\tau$ is composed of geodesics in the space $\mathcal{P}(\Omega)$ (where Ω is a large comapct set containing all the domains $C(t)$) endowed with the W_2 distance, and it is uniformly (w.r.t. τ) Lipschitz in this space. It solves the continuity equation together with $\widehat{v}^\tau$, i.e. $\partial_t \widehat{\rho}_t^\tau + \nabla \cdot \widehat{E}_t^\tau = 0$. The measures $\overline{E}_t^\tau$ and $\widehat{E}_t^\tau$ are uniformly bounded in the space of vector measures over $[0,T] \times \Omega$ (we have $|\overline{E}_t^\tau| \le L\rho_t^\tau$ and $|\widehat{E}_t^\tau| \le L\widehat{\rho}_t^\tau$). Hence, it is possible to extract a subsequence $(\tau_j)_j$ such that $\overline{\rho}_t^\tau, \widehat{\rho}_t^\tau, \overline{E}_t^\tau$ and $\widehat{E}_t^\tau$ have a limit as $\tau \to 0$.

As standard in this kind of proofs (see [23], Chapter 8 in [39] and Section 2.3 of this paper), we can prove that the limit of $\overline{\rho}_t^\tau$ and $\widehat{\rho}_t^\tau$ are the same (we call it ρ_t), as those of $\overline{E}_t^\tau$ and $\widehat{E}_t^\tau$ are the same (we call it E_t). The continuity equation obviously passes to the limit and we have $\partial_t \rho_t + \nabla \cdot E_t = 0$. From $|\overline{E}_t^\tau| \le L\rho_t^\tau$ we infer $|E_t| \le L\rho_t$ and hence $E_t = \rho_t v_t$ with $|v_t| \le L$.

If we prove $v_t(x) \in -\partial I_{C(t)}(x)$ for a.e. t and ρ_t-a.e x we have found a solution of (12). In order to prove this, we start from the characterization of $v_t^\tau(x)$ in terms of the projection. For every $x \in C(k\tau)$ and every $y \in C((k+1)\tau)$ we have

$$(y - P_{C((k+1)\tau)}[x]) \cdot v_k^\tau(x) \ge 0. \tag{14}$$

In particular, for every positive smooth function $a : [0,T] \times \Omega \to \mathbb{R}_+$ and every point $y \in \mathbb{R}^d$ we have

$$\int_0^T \int_\Omega a(t,x)(P_{C(R_\tau(t))}[y] - P_{C(R_\tau(t))}[x]) \cdot d\overline{E}^\tau(t,x) \ge 0,$$

where R_τ is a rounding operator: $R_\tau(t) := (k+1)\tau$ for every $t \in]k\tau, (k+1)\tau]$. Passing to the limit as $\tau \to 0$, using $\overline{E}_t^\tau \to E := v_t\rho_t$ and the Lipschitz behaviour of $C(t)$, so that $P_{C(R_\tau(t))}$ uniformly converges to $P_{C(t)}$, we get

$$\int_0^T \int_\Omega a(t,x)(P_{C(t)}[y] - x) \cdot v(t,x) d\rho_t(x) \ge 0.$$

From the arbitrariness of the function a we get that, for every y, the quantity $(P_{C(t)}[y]-x)\cdot v(t,x)$ is positive for a.e. t and ρ_t-a.e. x. Using a dense and countable set of possible points y, we deduce $(z-x)\cdot v(t,x)\geq 0$ for every $z\in C(t)$, which is exactly the desired property.

What we showed so far proves the first, third and fifth points of the claim. We need now to discuss the uniqueness of the solution and its connection with the pointwise Moreau process. In order to do that, we recall the so-called superposition principle (see for example Theorem 12 in [3]), which gives that every solution ρ of a continuity equation $\partial_t\rho_t+\nabla\cdot(v_t\rho_t)=0$ with $\int_0^T ||v_t||_{L^1(\rho_t)}dt<+\infty$ can be written as $\rho_t=(e_t)_\sharp\eta$ where η is a probability measure on $W^{1,1}([0,T];\Omega)$, $e_t:W^{1,1}([0,T])\to\Omega$ is the evaluation map $e_t(\omega)=\omega(t)$ and η can be taken concentrated on curves which solve the ODE $\omega'(t)=v_t(\omega(t))$ in the a.e. sense.

Here we use the fact that we know the behavior of the ODE $\omega'(t)=v_t(\omega(t))$, when v_t is such that $v_t(x)\in-\partial I_{C(t)}$. For every initial point there is only one solution, which is the solution to the Moreau sweeping process.

This provides at the same time uniqueness and $\rho_t=(Q_t)_\sharp\rho_0$. □

Note that in this model concentration of the measure, with creation of singular parts, can really occurr. This is due to the projection which tends to concentrate mass on the boundary. Yet, the singular part of ρ_t is not limited to $\partial C(t)$. Indeed, it is possible that some mass is created on $\partial C(t)$ but, after the movement of $C(t)$, the very same points of the support of this singular part are no more on the boundary.

4 Maximal density constraint

In this section we want to consider the sweeping of a probability measure in the moving compact convex set $C(t)$ with the additional constraint that ρ_t is absolutely continuous with respect to the Lebesgue measure and its density is less or equal to 1. So the set of probability measures we are interested in is now

$$K^1(t)=\{\rho\in\mathcal{P}(\Omega)\ ,\ \ \mathrm{spt}\,\rho\subset C(t)\ :\ \rho\leq 1\},\tag{15}$$

where Ω is a large bounded domain containing all the $C(t)'$ s. In the spirit of the previous works of the second and third author, given a measure $\rho\in\mathcal{P}(C)$ such that $\rho\leq 1$, we can describe heuristically a set of "admissible" velocities $\mathrm{adm}(\rho,C)$, saying that we require $\mathrm{div}(v)\geq 0$ on the set $\{\rho=1\}$ in order to preserve the density constraint. Moreover, since we are moving the set which contains the support of ρ_t one should also take care of the fact that, on the boundary, the inward normal velocity of the particles must be at least that of the boundary: this sums up as $v\cdot n\leq v_{\partial C}\cdot n$ where $v_{\partial C}$ is the boundary velocity (we will denote $V=v_{\partial C}\cdot n$). It is clear that without any regularity assumption these conditions don't make sense and so we have to weaken our hypothesis by duality: we denote by Π_ρ the so-called set of admissible pressures:

$$\Pi_\rho(C)=\{p\in H^1(C)\ :\ p\geq 0,\ p(1-\rho)=0\ \text{ a.e. on } C\}.$$

Then we consider the following formal computation, for $p\in\Pi_\rho(C)$:

$$\int_C v\cdot\nabla p\,dx=-\int_C \mathrm{div}(v)p\,dx+\int_{\partial C}p\,v\cdot n\,d\mathcal{H}^{d-1}\leq\int_{\partial C}p\,V\,d\mathcal{H}^{d-1};$$

this leads us to the following definition:

$$\mathrm{adm}(\rho,C)=\left\{v\in L^2(C,\rho)\ :\ \int_C v\cdot\nabla p\,dx\leq\int_{\partial C}p\,V\,d\mathcal{H}^{d-1},\quad\forall p\in\Pi_\rho(C)\right\}.$$

Now the evolution equation we want to solve becomes a continuity equation where the velocity has to be admissible, and have minimal L^2 norm:

$$\begin{cases} \partial_t \rho_t + \nabla \cdot (\rho_t v_t) = 0, \\ v_t = P_{\mathrm{adm}(\rho_t, C(t))}(0), \\ \mathrm{spt}(\rho_t) \subset C(t), \rho_t \leq 1. \end{cases} \tag{16}$$

In particular we see that this problem fits in the general construction done in Section 2.3 with the choice $A(t,\rho) = \{P_{\mathrm{adm}(\rho, C(t))}(0)\}$. In order to understand better the equation, we want to present a simple lemma that will be useful in the sequel, where we characterize the element of minimal norm in $\mathrm{adm}(\rho, C)$:

Lemma 4.1. *Let C be a bounded domain with Lipschitz boundary and ρ be a probability density on C such that $\rho \leq 1$. Then we have*

$$P_{\mathrm{adm}(\rho,C)}(0) = \operatorname{argmin}\left\{ \int_C |v|^2 \, dx \ : \ v \in \mathrm{adm}(\rho, C) \right\} = -\nabla p,$$

where $p \in \Pi_\rho(C)$ is the unique pressure that satisfies the following conditions

$$\int_C -\nabla p \cdot \nabla q \, dx \geq \int_{\partial C} q \, V \, d\mathcal{H}^{d-1} \qquad \forall q \in \Pi_\rho(C); \tag{17}$$

$$\int_C -|\nabla p|^2 \, dx = \int_{\partial C} p \, V \, d\mathcal{H}^{d-1}. \tag{18}$$

Proof. Let us consider

$$p^* = \operatorname{argmin}\left\{ \frac{1}{2} \int_C |\nabla q|^2 \, dx + \int_{\partial C} q \, V \, d\mathcal{H}^{d-1} \ : \ q \in \Pi_\rho(C) \right\};$$

the minimum exists thanks to the fact that Proposition 2.7 guarantees coercivity of this quadratic functional. As $\Pi_\rho(C)$ is a cone, the optimality conditions imply that p^* satisfies (17) and (18). Now, we notice that for every $v \in \mathrm{adm}(\rho, C)$ we have $\int_C v \cdot \nabla p^* \leq \int_{\partial C} p^* \, V \, d\mathcal{H}^{d-1} = -\int_C |\nabla p^*|^2$ and so we can conclude that

$$\begin{aligned} \int_C |v|^2 &= \int_C |v + \nabla p^*|^2 \, dx - \int_C |\nabla p^*|^2 \, dx - 2\int_C v \cdot \nabla p^* \, dx \\ &\geq \int_C |\nabla p^*|^2 \, dx + \int_C |v + \nabla p^*|^2 \, dx \\ &\geq \int_C |\nabla p^*|^2 \, dx. \end{aligned}$$

Eventually we notice that all these inequalities are equalities when $\nabla p^* = -v$, and we have equality only in this case (thanks to the last inequality). In order to conclude it is sufficient to observe that (17) guarantees that $-\nabla p^* \in \mathrm{adm}(\rho, C)$ and so we are finished. □

Thanks to this lemma we can say that if we solve (16) then we can find also a solution to

$$\begin{cases} \partial_t \rho_t = \Delta p_t, \\ p_t \in \Pi_{\rho_t}(C(t)) \\ \mathrm{spt}(\rho_t) \subset C(t), \rho_t \leq 1. \end{cases} \tag{19}$$

Also this problem fits in the general construction with $\overline{A}(t,\rho) = \{-\nabla p \ : \ p \in \Pi_\rho(C(t))\}$. Notice that $A(t,\rho) \subseteq \overline{A}(t,\rho)$, but indeed at the end of this section we will show that every solution of (19) is in fact a solution of (16). The main result is however existence and uniqueness for (16).

Theorem 4.2. *Suppose that $(C(t))_t$ is a family of compact convex sets with $d_H(C(t),C(s)) \leq L|t-s|$, and $\inf_t |C(t)| = c > 1$. Then, for every initial datum $\rho_0 \in \mathcal{P}(C(0))$ with $\rho_0 \leq 1$ the solution to the Equation* (16)

- *exists;*
- *is unique*
- *is obtained as a limit as $\tau \to 0$ of the curves $\widehat{\rho}^\tau$ obtained in the following way: first define recursively ρ_k^τ via*

$$\begin{cases} \rho_0^\tau = \rho_0, \\ \rho_{k+1}^\tau = P_{K^1((k+1)\tau)}[\rho_k^\tau], \end{cases}$$

  then define the curve $\widehat{\rho}^\tau$ on each interval $[k\tau,(k+1)\tau]$ to be the geodesic interpolation between ρ_k^τ and ρ_{k+1}^τ.

Lemma 4.3 (Wasserstein estimate)**.** *Let $C(t)$ be a curve of convex sets that is L-Lipschitz with respect to the Hausdorff distance, with $\inf_t |C(t)| = c > 1$. Let $K^1(t)$ be the corresponding set of measures defined in* (4) *and μ be a measure in $K^1(t)$. Then*

$$\inf \{W_2(\mu,\rho) \ : \ \rho \in K^1(t+h)\} \leq MLh$$

for some constant $M = M(c,d)$.

Proof. First we use Corollary 2.6 which provides a transport map $T : C(t) \to C(t+h)$ such that the density of $T_\sharp\mu$ is bounded by $(1-CLh)^{-1}$ and $W_2(\mu, T_\sharp\mu) \leq ||T-id||_{L^2(\mu)} \leq ||T-id||_{L^\infty} \leq MLh$. Then we use Theorem B.1 in [24] which associates to $T_\sharp\mu$ a new measure $\rho \in \mathcal{P}(C(t+h))$ with $\rho \leq 1$ and $W_2(\rho, T_\sharp\mu) \leq Mh$. □

Remark 4.4. *This last lemma can be seen as a "lifting property" from $\mathbb{R}^d$ to $\mathbb{W}_2(\mathbb{R}^d)$: in fact we are saying that there exists a constant C such that if $d_H(C(s),C(t)) \leq r$ then $d_H(K^1(s),K^1(t)) \leq Mr$.*

4.1 Main construction and existence

The main proposition here is the following, which provides us with a discrete version of the characterization of the velocity as the negative gradient of an admissible pressure.

Proposition 4.5. *Let Ω be a bounded domain of $\mathbb{R}^d$, $C \subset \Omega$ a compact set, with $|C| \geq 1$. For every absolutely continuous $\rho \in \mathcal{P}(\Omega)$ there exists a unique minimizer $\rho_C = P_{K^1}[\rho]$ to the problem*

$$\min_{\eta \in K^1} W_2(\rho,\eta),$$

where $K^1 = \{\eta \in \mathcal{P}(\Omega) \ , \ \operatorname{spt}\rho \subset C \ , \ \eta \leq 1\}$. Moreover the unique optimal transport map T *between ρ_C and ρ satisfies $T(x) = x + \nabla p_C(x)$ where $p_C \in \Pi_{\rho_C}(C)$.*

Proof. The existence is obvious by compactness, while the uniqueness follows from the fact that $\eta \mapsto W_2^2(\rho, \eta)$ is strictly convex whenever ρ is absolutely continuous (see [13] for a proof of this fact, and of the uniqueness of the projection under more general assumptions). Now let us consider the maximizer ρ_C, and an arbitrary competitor $\eta \in K^1$. By considering perturbations $\eta_\varepsilon = \rho_C + \varepsilon(\eta - \rho_C)$, we can write optimality conditions using the first variation of the functional $\eta \mapsto \frac{1}{2}W_2^2(\rho, \eta)$, which involves the Kantorovich potentials (see Section 2.3). In this way we can obtain that there exists a Kantorovich potential φ from ρ_C to ρ such that, for every $\eta \in K^1$, we have

$$\int \varphi \, d\rho_C \leq \int \varphi \, d\eta \qquad \forall \eta \in K^1 \tag{20}$$

(the reader may look at Lemma 3.1 and 3.3 in [23] for a rigorous treatment of this optimality condition).

From this, it is easy to see that there exists a threshold l such that

$$\begin{cases} \rho_C(x) = 1 & \text{if } \varphi(x) < l \\ 0 \leq \rho_C \leq 1 & \text{if } \varphi(x) = l \\ \rho_C = 0 & \text{if } \varphi(x) > l. \end{cases}$$

Now we can consider $p_C = (l - \varphi)_+$ and observe $\nabla\varphi = -\nabla p_C$ ρ_C-a.e. Then, the optimal transport between ρ_C and ρ is $\mathrm{T}(x) = x - \nabla\varphi(x) = x + \nabla p_C(x)$. Moreover, we observe that by definition we have $p_C \geq 0$ and $p_C(1 - \rho_C) = 0$. □

In order to construct a solution to our PDE, we make use of the general construction described in 2.3. Hence, we want to construct a discrete (in time) sequence, and we fix a time step $\tau > 0$. We will build a sequence of measures ρ_k^τ defined recursively as

$$\begin{cases} \rho_0^\tau = \rho_0 \\ \rho_{k+1}^\tau = P_{K^1((k+1)\tau)}[\rho_k^\tau] \end{cases}$$

Then we will consider two curves in the measure-momentum space as done in Section 2.3: the piecewise constant $(\overline{\rho}_t^\tau, \overline{E}_t^\tau)$ and the geodesic one $(\widehat{\rho}_t, \widehat{E}_t^\tau)$. Thanks to Proposition 4.5 we have the peculiar structure for $v_k^\tau = -\nabla p_k^\tau$, where $p_k^\tau(1 - \rho_k^\tau) = 0$ in $C(k\tau)$. The main estimate is the fact that thanks to Lemma 4.3 we have $W_2(\rho_k^\tau, \rho_{k-1}^\tau) \leq C\tau$ and in particular we have the discrete H^1 estimate

$$\sum_{k=1}^{T/\tau} \tau \left(\frac{W_2(\rho_k^\tau, \rho_{k-1}^\tau)}{\tau} \right)^2 \leq TC.$$

So, thanks to the general strategy that we evoked, we have convergence of $\widehat{\rho}^\tau$ and $\overline{\rho}^\tau$ towards the same continuous curve ρ_t (which in this case is also Lipschiz). We want to prove $\mathrm{supp}(\rho_t) \subset C(t)$ and this can be done using the information that $\mathrm{supp}(\rho_k^\tau) \subset C(k\tau)$. Indeed, we have that $\mathrm{supp}(\overline{\rho}_t^\tau) \subset (C(t))_{L\tau}$ and so in the limit we get the thesis. Also the condition $\overline{\rho}_t^\tau \leq 1$ easily passes to the limit.

Lemma 4.6. *We have $E_t = v_t\rho_t$ and moreover there exists $p \in L^\infty([0,T]; H^1(\mathbb{R}^n))$ such that $v_t = -\nabla p_t$ on $C(t)$ and $p_t(1 - \rho_t) = 0$ on $C(t)$, that is $v_t \in \overline{A}(t, \rho_t)$.*

Proof. We will proceed in a similar fashion as done in [23] and explained in Section 2.3. By the lower semicontinuity of $\iint |E|^2/\rho$ we know that $E_t = v_t\rho_t$ for some velocity field. First of all let us extend the pressure functions p_k^τ to a larger domain Ω containing all the sets $C(t)$. Thanks to

Proposition 2.7 and the fact that $|\{p^\tau_k = 0\} \cap C(k\tau)| \ge c - 1 > 0$ we can construct these extensions in such a way that

$$\|p^\tau_k\|_{H^1(\Omega)} \le C\|p^\tau_k\|_{H^1(C(k\tau))} \le C\|\nabla p^\tau_k\|_{L^2(C(k\tau))} = C\frac{W_2(\rho^\tau_{k-1}, \rho^\tau_k)}{\tau} \le CL. \tag{21}$$

Let us consider the piecewise constant function $\overline{p}^\tau_t = p^\tau_k$ for $t \in](k-1)\tau, k\tau]$. By equation (21) we have that $\overline{p}^\tau_t$ is bounded in $L^\infty([0,T], H^1(\Omega))$, and so there exists a subsequence weakly converging to some p.

It is easy to check that $\overline{E}^\tau_t = -\mathbb{1}_{C(k\tau)}\nabla\overline{p}^\tau_t \rightharpoonup -\mathbb{1}_{C(t)}\nabla p_t$ weakly as measures (because of the weak convergence of ∇p^τ and of the strong convergence of $\mathbb{1}_{C((k+1)\tau))}$) and in particular we have $E_t = -\mathbb{1}_{C(t)}\nabla p_t$.

It remains to prove that $p_t(1-\rho_t) = 0$ for a.e. $x \in C(t)$. This is done as in [24] paying attention to the fact that we have moving domains: we know that $\overline{p}^\tau_t(1-\overline{\rho}^\tau_t) = 0$ in $C(R_\tau(t))$ (here $R_\tau(t) = \tau\lfloor t/\tau\rfloor + \tau$). So we can write

$$\frac{1}{|a-b|}\int_a^b \int_{(C(t))_{L\tau}} \overline{p}^\tau_t(1-\overline{\rho}^\tau_t)\,dx\,dt = \frac{1}{|a-b|}\int_a^b \int_{(C(t))_{L\tau}\setminus C(R_\tau(t))} \overline{p}^\tau_t\,dx\,dt.$$

Now we have $|(C(t))_{L\tau} \setminus C(R_\tau(t))| \le C\tau$ and so in particular, thanks to the uniform estimate on $\|\overline{p}^\tau_t\|_{L^2(\Omega)}$, we have

$$\frac{1}{|a-b|}\int_a^b \int_{(C(t))_{L\tau}} \overline{p}^\tau_t(1-\overline{\rho}^\tau_t)\,dx\,dt \le C\sqrt{\tau}.$$

Now we can use the estimate $|\int p\,d\rho - \int p\,d\eta| \le \|\nabla p\|_2 W_2(\rho,\eta)$ that holds whenever $\rho, \eta \le 1$ and $p \in H^1$ (see [23] or Section 5.5.2 in [39], or even [21]), to obtain that

$$\int_{(C(t))_{L\tau}} \overline{p}^\tau_t\overline{\rho}^\tau_t\,dx - \int_{(C(a))_{L\tau}} \overline{p}^\tau_t\overline{\rho}^\tau_a\,dx \le C\|\nabla p_t\|_2(|t-a| + \tau).$$

Summing up we obtain:

$$\frac{1}{|a-b|}\int_a^b \int_{(C(a))_{L\tau}} \overline{p}^\tau_t(1-\overline{\rho}^\tau_a)\,dx\,dt \le C(\sqrt{\tau} + 2|a-b| + \tau).$$

Now we can let $\tau \to 0$ and use the fact that $\overline{\rho}^\tau_a \to \rho_a$ weakly in $L^\infty(\mathbb{R}^d)$, $\overline{p}^\tau$ converges weakly to p in $L^2([0,T]; H^1(\mathbb{R}^d))$, and hence $\int_a^b \overline{p}^\tau \to \int_a^b p$ strongly in $L^2(\mathbb{R}^d)$. Then, letting also $b \to a$ for weak Lebesgue point of p_t, we obtain $\int_{C(a)} p_a(1-\rho_a)\,dx = 0$ and since $p_a(1-\rho_a) \ge 0$, we get $p_t \in \Pi_{\rho_t}(C(t))$. $\square$

Now we proved that there exists a solution (ρ_t, p_t) of (19). In order to check that this provides also a solution to (16) we have to prove that for a.e. t we have $-\nabla p_t = P_{\mathrm{adm}(\rho_t, C(t))}(0)$. As we already know that $p_t \in \Pi_{\rho_t}(C(t))$, thanks to Lemma 4.1 it is sufficient to prove that (17) and (18) are satisfied. We will do this in the form of a general implication: (19) implies (16)

Proposition 4.7. *Let (ρ_t, p_t) be a solution of* (19) *with $\iint |\nabla p_t|^2\,dxdt < \infty$. Then $(\rho_t, -\nabla p_t)$ is a solution of* (16).

Proof. Fix a time t and then fix any admissible pressure $q \in \Pi_{\rho_t}(C(t))$; we can extend q to the whole Ω and, from $\rho_s \le \mathbb{1}_{C(s)}$, we have

$$\int_\Omega q\,d\rho_s \le \int_{C(s)} q\,dx \qquad \forall s \in (t-\delta, t+\delta).$$

Since we have equality when $s = t$ we obtain that the derivatives at time t of the left and right hand sides above are equal. This gives

$$-\int_{C(t)} \nabla p_t \cdot \nabla q \, d\rho_t = \int_{\partial C(t)} q \cdot a_t \, d\mathcal{H}^{d-1} \qquad \forall q \in \Pi_{\rho_t}(C(t)), \tag{22}$$

where in the left-hand side we used the continuity equation $\partial_t \rho_t = \nabla \cdot (\nabla p_t \rho_t)$ and in the right hand side we used Lemma 2.3. Once we have this equality we are done since this implies both (17) and (18) and so $-\nabla p_t = P_{\mathrm{adm}(\rho_t, C_t)}(0)$ as we wanted to show.

Now there is a subtle point here since we have that for every $q \in H^1(\mathbb{R}^d)$ these derivatives exist for almost every time, but this negligible set can depend on q; in particular it can happen that for every time t, for whatever $q \in \Pi_{\rho_t}(C(t))$ the derivative at time t doesn't exists. However we can show a set of times where the derivative for the left hand side exists for every $q \in H^1(\Omega)$: this can be taken as the set of times where we have a L^2-weak Lebesgue point of the momentum, i.e. for t such that we have $\frac{1}{s-t}\int_s^t \nabla p_r \rho_r dr \rightharpoonup \nabla p_t \rho_t$ weakly in L^2 (it is sufficient to have convergence for a dense set and then use the density and the bounds on the norms). For the right hand side a similar reasoning can be applied taking in consideration points t where $\frac{1}{s-t}\int_s^t a_r \mathcal{H}^{n-1}|_{\partial C(r)} dr \rightharpoonup a_t \mathcal{H}^{n-1}|_{\partial C(t)}$ weakly in duality with $H^1(\Omega)$ (we recall that the trace inequality is uniform). □

Now we can summarize the proof of the main theorem of the section:

Proof. (of Theorem 4.2) As for existence, it is sufficient to observe that the construction given by the catching up algorithm provided a solution for (19) thanks to Lemma 4.6. Moreover Proposition 4.7 implies that this solution is also solution of (16). As for the uniqueness the very same reasoning present in [14] can be applied: let us consider two solutions $(\rho_t^{(1)}, -\nabla p_t^{(1)})$, $(\rho_t^{(2)}, -\nabla p_t^{(2)})$ of (16) with a minimal integrability property that is $\int \|\nabla p_t^{(i)}\|_{L^2} \, dt < \infty$. Then it is easy to see that $\rho_t^{(i)}$ are absolutely continuous curves in the Wasserstein space and so we can differentiate their distance using (8):

$$\frac{d}{dt}\left(\frac{1}{2}W_2^2(\rho_t^{(1)}, \rho_t^{(2)})\right) = -\int_{C(t)} \nabla \varphi_t \cdot \nabla p_t^{(1)} \, dx - \int_{C(t)} \nabla \psi_t \cdot \nabla p_t^{(2)} \, dx. \tag{23}$$

Now it is sufficient to use a technical lemma (for example Lemma 2.1 in [14] or Lemma 4.3.13 in [29], notice that we do not care about the fact that the sets $C(t)$ are moving) that says that both the terms in the right hand side are negative. We obtain that the equation leads to a contraction in the Wasserstein space (as it happens also in the Hilbertian sweeping process of Moreau). □

5 Diffusion

In this section we consider the motion of a collection of particles "pushed" by the moving domain $C(t)$, but also subject to Brownian diffusion. Compare to Section 3, where we considered the law of a process satisfying $dX_t = v_t(X_t)dt$ with $X_0 \sim \rho_0$ and $v_t(x) \in -\partial I_{C(t)}(x)$: here we consider, roughly speaking, the law of a process given by

$$dX_t + \partial I_{C(t)}(X_t) \ni dB_t$$

where B_t is a d-dimensional Brownian motion. In other words, we require $dX_t = v_t(X_t)dt + dB_t$ where again $v_t \in -\partial I_{C(t)}(x)$. We refer to [7] for a precise description of this reflected Brownian motion in a moving domain and we do not enter into extra details here. We observe that, due to the diffusion, the density ρ_t will never develop singular parts and hence we will expect $v_t = 0$ for a.e. particles. The motion will hence be driven by the Brownian diffusion only, and the attachment to the moving convex set only depends on the boundary conditions.

Without entering into details about the translation of this individual motion into a PDE, which is however standard in stochastic analysis, we directly write the PDE that we want to solve.

Precisely, we look for a family of probability densities $(\rho_t)_{t\in[0,T]}$, satisfying the continuity equation

$$\begin{cases} \partial_t \rho_t - \Delta \rho_t = 0 & \text{in } C(t), \\ -\nabla \rho_t \cdot n = \rho_t V_t & \text{on } \partial C(t), \\ \operatorname{spt}(\rho_t) \subset C(t), \end{cases} \tag{24}$$

where V_t is the normal velocity of the boundary. As usual, the above equation is to be intended in a distributional sense on $\mathbb{R}^d$, and the Neumann condition which accompanies it is just a formal expression of the boundary conditions which express the fact that the equation is satisfied by a density which is concentrated on a moving domain $C(t)$. In case of smooth solutions, Lemma 2.3 allows to see that this fact exactly gives the boundary condition $-\nabla \rho_t \cdot n = \rho_t V_t$ (where V_t is the normal velocity of the boundary).

As we underlined in the introduction, this equation corresponds to a forced version of a Moreau process in $\mathcal{P}(\Omega)$ with a moving set $K(t) := \{\rho \in \mathcal{P}(\Omega) \,:\, \operatorname{spt} \rho \subset C(t)\}$.

We will prove the following result, where we denote by $\mathcal{E}$ the logarithmic entropy defined via

$$\mathcal{E}(\rho) := \begin{cases} \int \rho(x) \ln(\rho(x)) dx & \text{if } \rho \ll \mathcal{L}^d, \\ +\infty & \text{otherwise.} \end{cases}$$

Theorem 5.1. *Suppose that $(C(t))_t$ is a family of compact convex sets with $d_H(C(t), C(s)) \leq L|t-s|$. Then, for every initial datum $\rho_0 \in \mathcal{P}(C(0))$ with $\mathcal{E}(\rho_0) < +\infty$, the solution to the Equation* (24)

- *exists;*
- *is unique;*
- *is obtained as a limit as $\tau \to 0$ of the curves $\widehat{\rho}^\tau$ obtained in the following way: first define recursively ρ_k^τ via*

$$\begin{cases} \rho_0^\tau = \rho_0, \\ \rho_{k+1}^\tau = \operatorname{argmin} \left\{ \mathcal{E}(\rho) + \frac{W_2^2(\rho, \rho_k^\tau)}{\tau} \,:\, \rho \in K((k+1)\tau) \right\}, \end{cases}$$

  then define the curve $\widehat{\rho}^\tau$ on each interval $[k\tau, (k+1)\tau]$ to be the geodesic interpolation between ρ_k^τ and ρ_{k+1}^τ.

Proof. The existence and the approximation will be done as detailed in Section 2.3. The optimality conditions in the problem defining ρ_{k+1}^τ read as $\rho_{k+1}^\tau > 0$ a.e. (see Chapter 8 in [39] for details, for instance) and

$$\ln(\rho_{k+1}^\tau) + \frac{\varphi}{\tau} = const.$$

Passing to the gradients, this gives

$$v_k^\tau = \frac{\nabla \varphi}{\tau} = -\frac{\nabla \rho_{k+1}^\tau}{\rho_{k+1}^\tau} \qquad \text{on } C((k+1)\tau).$$

This may be written, in terms of the momentum variable, as $E_k^\tau = -\nabla\rho_k^\tau \mathbb{1}_{C(k\tau)}$. It is a linear condition, and it is clear that it passes to the limit as $\tau \to 0$, since the first term converges weakly while the second converges strongly. Hence, any limit of the discrete scheme will be a solution of

$$\partial_t\rho_t - \nabla\cdot(\nabla\rho_t \mathbb{1}_{C(t)}) = 0,$$

which exactly means (in a weak sense) $\partial_t\rho_t = \Delta\rho_t$ in $C(t)$, with the boundary conditions coming from the fact that, by construction, $\mathrm{spt}(\rho_t) \subset C(t)$ and from the derivative computed in Lemma 2.3.

Hence, we only need to estimate $W_2^2(\rho_{k+1}^\tau, \rho_k^\tau)$ to be able to show compactness and pass to the limit.

For this, we use Corollary 2.6 to produce a measure $\mu = T_\sharp\rho_k^\tau$ such that $\mathrm{spt}\,\mu \subset C((k+1)\tau)$ and $\mathcal{E}(\mu) \le \mathcal{E}(\rho_k^\tau) + Ch$ and $W_2(\mu, \rho_k^\tau) \le C\tau$. Indeed, every time that we have $\mu = T_\sharp\rho$ for a C^1 injective map T, we have (identifying absolutely continuous measures with their densities)

$$\mu(T(x)) = \frac{\rho(x)}{\det(DT(x))}$$

hence

$$\begin{aligned}
\mathcal{E}(\mu) &= \int \ln(\mu(y))d\mu(y) = \int \ln(\mu(T(x))d\rho(x) \\
&= \int \rho(x)[\ln(\rho(x)) - \ln(\det(DT(x)))]dx \le \mathcal{E}(\rho) + ||(\ln(\det(DT)))_-||_{L^\infty}.
\end{aligned}$$

In the present case, from $\det(DT) \ge 1 - CLh$, we have $\mathcal{E}(\mu) \le \mathcal{E}(\rho_k^\tau) + Ch$.

Then, the optimality of ρ_{k+1}^τ implies

$$\mathcal{E}(\rho_{k+1}^\tau) + \frac{W_2^2(\rho_{k+1}^\tau, \rho_k^\tau)}{\tau} \le \mathcal{E}(\mu) + \frac{W_2^2(\mu, \rho_k^\tau)}{\tau} \le \mathcal{E}(\rho_k^\tau) + C\tau.$$

This gives

$$\frac{W_2^2(\rho_{k+1}^\tau, \rho_k^\tau)}{\tau} \le \mathcal{E}(\rho_k^\tau) - \mathcal{E}(\rho_{k+1}^\tau) + C\tau,$$

which gives a converging series and provides the desired compactness, and hence existence.

We are only left with the uniqueness of the solution. Suppose that $\rho^{(1)}$ and $\rho^{(2)}$ are two solutions starting from the same initial datum ρ_0. We differentiate the quantity $\frac{1}{2}W_2^2(\rho_t^{(1)}, \rho_t^{(2)})$ and obtain

$$\frac{d}{dt}\left(\frac{1}{2}W_2^2(\rho_t^{(1)}, \rho_t^{(2)})\right) = -\int_{C(t)} \left(\nabla\rho_t^{(1)}\cdot\nabla\varphi_t + \nabla\rho_t^{(2)}\cdot\nabla\psi_t\right)dx. \tag{25}$$

The peculiar point, which is well-known and provides contractivity in $\mathbb{W}_2(\Omega)$ for the Heat equation, is the fact that the right-hand side is always positive. This can be obtained in several ways, either by using geodesic convexity of the entropy, or using the more general inequality proven in Section 3 of [13]. It is interesting to notice that this inequality holds on arbitrary convex domains, and the fact that our convex domains $C(t)$ are moving does not affect the proof. □

We can observe that the assumption $\mathcal{E}(\rho_0) < +\infty$ in the above Theorem can be dropped, at least as far as existence and uniqueness are concerned. This is a standard procedure in gradient evolutions, when there is contractivity (see [4]).

Theorem 5.2. *Suppose that $(C(t))_t$ is a family of compact convex sets with $d_H(C(t), C(s)) \leq L|t-s|$. Then, for every initial datum $\rho_0 \in \mathcal{P}(C(0))$ there exists unique a solution to the Equation* (24).

Proof. In the proof of Theorem 5.1 we saw that the W_2 distance between two arbitrary solutions of (24) is decreasing in time. This provides uniqueness independently of the initial datum. Moreover, we can take a sequence $(\rho_0^n)_n$ of initial data supported on $C(0)$, such that $W_2(\rho^n, \rho_0) \to 0$, and define the corresponding solutions ρ_t^n. As we have $W_2(\rho_t^n, \rho_t^m) \leq W_2(\rho_0^n, \rho_0^m)$, for each time t the sequence $(\rho_t^n)_n$ is a Cauchy sequence in $\mathbb{W}_2(\Omega)$. Since the Wasserstein space on the compact domain Ω is compact itself, and hence complete, these sequences converge. We have hence a limit curve ρ_t with initial datum ρ_0. This curve solves (24) because the PDE is linear and easily passes to the limit. This provides the required existence result. □

Unfortunately, it is not easy in Theorem 5.2 to add the convergence of the JKO scheme, as we did in Theorem 5.1. This would require a double limit procedure, some sort of discrete contractivity, or the choice of a suitable approximation ρ_0^n. Yet, the fact that the domain is moving makes it difficult to prove any estimate of this kind, even if we believe that they are true. In particular, they are quite easy once one proves a uniform bound on $\mathcal{E}(\rho_k^\tau)$ on positive times, i.e. for $k\tau > t_0 > 0$.

6 Numerical illustrations

We propose here to illustrate the previous considerations by numerical approximations of the measure sweeping process.

6.1 Numerical scheme for the sweeping process with diffusion

In the case with diffusion, we consider a smooth motion $t \longmapsto C(t) \subset \mathbb{R}^2$, and an initial measure ρ_0 supported in $C(0)$. We want to solve

$$\left\{ \begin{array}{rcll} \partial_t \rho_t - \kappa \Delta \rho_t & = & 0 & \text{in } C(t) \\ \kappa \dfrac{\partial \rho_t}{\partial n} + \rho V_t & = & 0 & \text{on } \partial C(t) \end{array} \right. \tag{26}$$

The time discretization strategy is based on a splitting between catching-up and diffusion. More precisely, for a given time step $\tau > 0$, sucessive approximations ρ_1, $\rho_2, \dots \rho_k$ (we omit the dependence on τ, as in this numerical section we will take it as fixed; also, for simplicity we will write in this section C_k for $C(k\tau)$ and K_k for $K(k\tau)$) are built according to the following scheme.

1. The current density is projected (in the Wasserstein sense) on the set K_{k+1} of measures supported in C_{k+1}:
$$\widetilde{\rho}_{k+1} = P_{K_{k+1}} \rho_k.$$
2. A step of the discretized heat equation with homogeneous Neuman B.C.'s, and with initial condition the projected measure, is then solved:
$$\left\{ \begin{array}{rcll} \dfrac{\rho_{k+1} - \widetilde{\rho}_{k+1}}{\tau} - \kappa \Delta \rho_{k+1} & = & 0 & \text{in } C_{k+1}, \\ \kappa \dfrac{\partial \rho_{k+1}}{\partial n} & = & 0 & \text{on } \partial C_{k+1}. \end{array} \right. \tag{27}$$

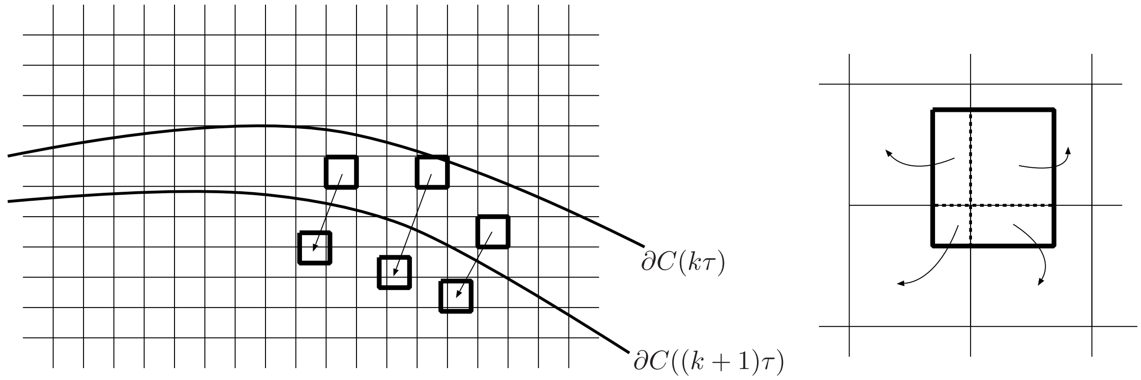


Figure 1: Catching-up at the discrete level

As for space discretization, the overall domain (a square in our case) is covered by a fixed cartesian mesh, and a finite volume approach is considered: each density field is represented by a piecewise constant function according to the underlying grid. Step 1 (catching-up) of the previous scheme is approximated as follows. Let x_i be the center of a cell containing some mass at time $k\tau$, while lying outside of C_{k+1}. We denote by z_i the projection of x_i on C_{k+1}. The mass in cell i is then translated by

$$z_i - x_i + h\frac{z_i - x_i}{|z_i - x_i|},$$

where h is the mesh step size. The obtained square of mass is then distributed over the four cells that it intersects (see Fig. 1). Thanks to the extra term above (the one that contains the mesh step size h), in only affects cells that are contained in C_{k+1}.

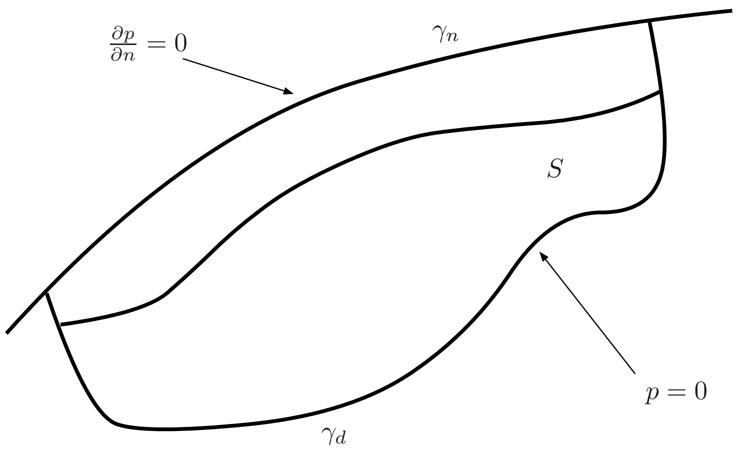
Step 2 is solved by a standard finite volume discretization of the heat equation based on the underlying cartesian grid. To account for the impervious condition on the boundary of the domain, we simply consider that the diffusion vanishes between cells that share an edge that lies outside of C_{k+1}.

6.2 Numerical scheme for the sweeping process with congestion

We propose again a splitting strategy. The first step consists in catching up the mass that is outside of the domain, like in the diffusive case (see previous section). The obtained density is very likely to violate the constraint, and the second step is meant to push it down to the constrained space (i.e. with $\rho \leq 1$). This second step consists in projecting a density to the set K^1 of densities supported in a given closed set C, and that are bounded by 1. Since the projection is meant in the Wasserstein sense, it does not fit into standard numerical approaches. In [25], for a similar problem (macroscopic crowd motion model) we proposed a stochastic algorithm to approximate this projection. In the present context, the approach is based on the following stochastic interpretation of the Poisson problem on a domain S (see Fig. 2, top)

$$-\Delta p = \nu,$$

with homonegenous Neuman boundary conditions on a part of the boundary γ_n, and homogeneous Dirichlet boundary conditions on γ_d, where ν is a probability measure (see Fig. 2). Now pick a random point X_0 in S according to the law defined by ν, run a a Browian motion X_t starting from X_0, with reflection on γ_n. Define $t^\star$ as the smallest time (that is almost surely finite) at which



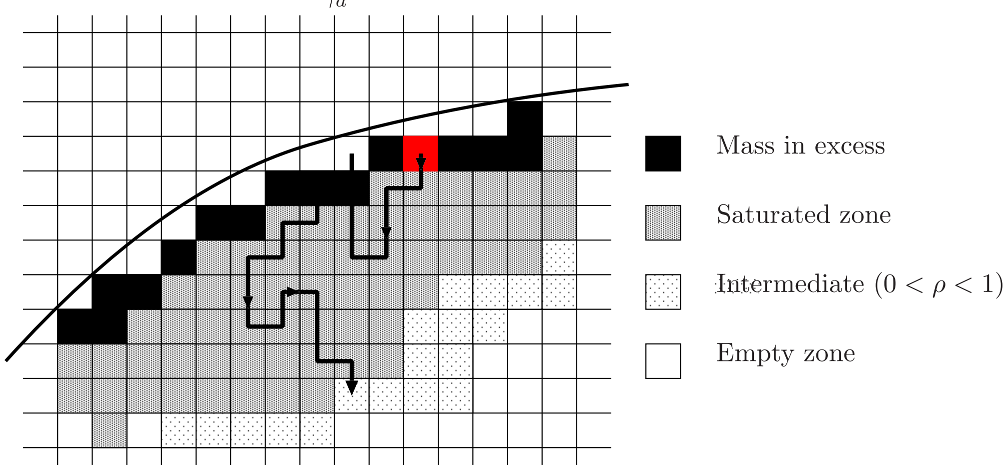


Figure 2: Stochastic algorithm

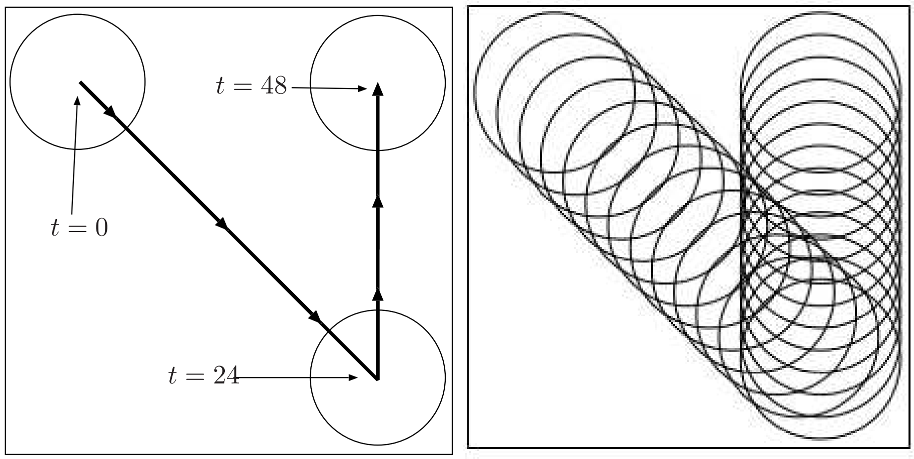


Figure 3: Moving disc

X_t hits the Dirichlet part of the boundary. Consider then $X_{t^\star}$ as a random point with values in γ_d: its law is the harmonic measure associated to ν, and its associated measure on γ_d is $-\partial p/\partial n$ (see e.g. [15]). Our numerical strategy is a straight transposition of this property at the discrete level. After the discrete catching-up that we previously described, we obtain a discrete measure $\widetilde{\rho}$ supported in the current set C_{k+1}. The cells contained in C_{k+1} are of 4 different types: empty ($\widetilde{\rho} = 0$), intermediate ($\widetilde{\rho} \in (0,1)$), saturated ($\widetilde{\rho} = 1$), and over-saturated ($\widetilde{\rho} > 1$). If the latter cell sub-population is empty, the density verifies the contraints, and no projection is needed. If not, for each of those overweighted cells, we run the following procedure: denoting by m the local mass in excess (that is $\widetilde{\rho} - 1$ multiplied by the volume of the cell), we start a random walk from the given cell, subject to jump on cells that are contained in C_{k+1} only. As far as saturated or oversaturated cells are visited, the walk goes on. When a intermediate (or empty) cell is met, then

1. If the space available is larger that (or equal to) m, the excess mass is put here, and the random walk stops;

2. If the space available is smaller that m, the density is set to 1, the excess mass is reduced by the corresponding amount, and the random walk continues, with a reduced amount of mass to get rid of.

Actually, this stochastic approach is replaced here by a deterministic version of the random walk, based on the *Router Rotor Model* (see e.g. [20]). It can be seen as a cellular automata algorithm: an arrow, pointing to one of the four cardinal directions, is associated to each cell. The arrows are initially distributed randomly. In the procedure decribed above, when a walk (that is no longer random) is in a cell, it goes to the cell to which the current arrow is pointing. The arrow is then rotated by 90°.

The test case we propose is presented in Fig. 4: a rigid disc moves on a square at constant speed, starting from the top-left corner, heading to the bottom-right corner, then it goes up along the righd hand-side. We simulate the corresponding sweeping process in three situations: diffusive case with $\kappa = 0.1$ (Fig. 4), diffusive case with $\kappa = 1.e-4$ (Fig. 5), and congested case (Fig. 6).

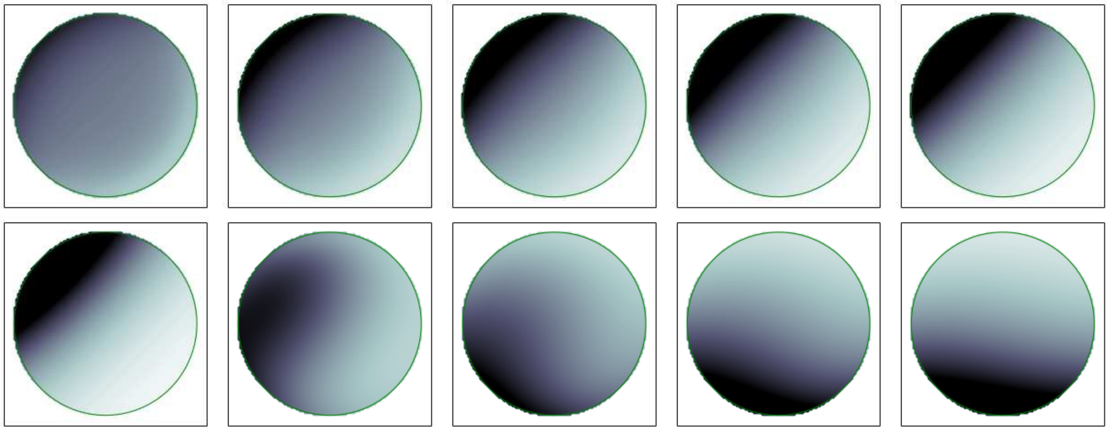


Figure 4: Catching up with diffusion, $\kappa = 0.1$, at times 1, 2, 3, 4, 5, 24, 26, 28, 30, 32.

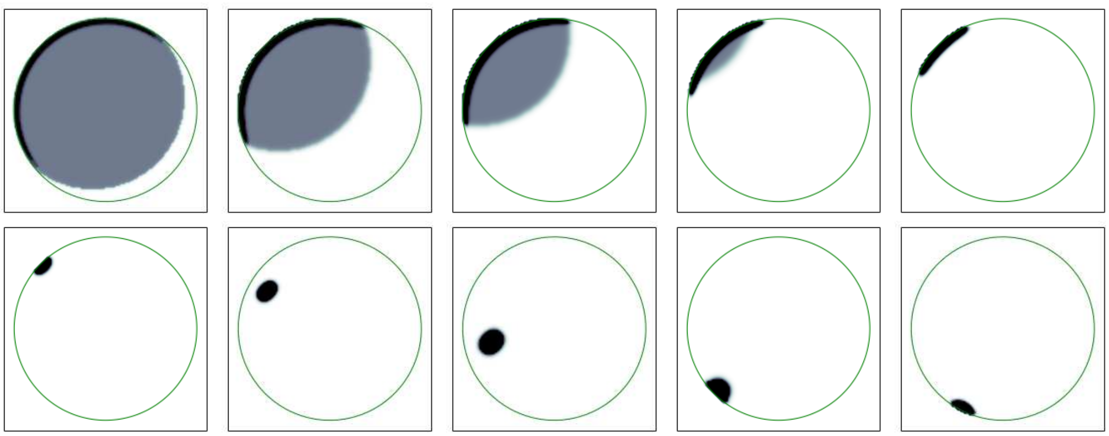


Figure 5: Catching up with diffusion, $\kappa = 1.e-4$, at times 1, 4, 6, 9, 12, 24, 26, 30, 34, 38.

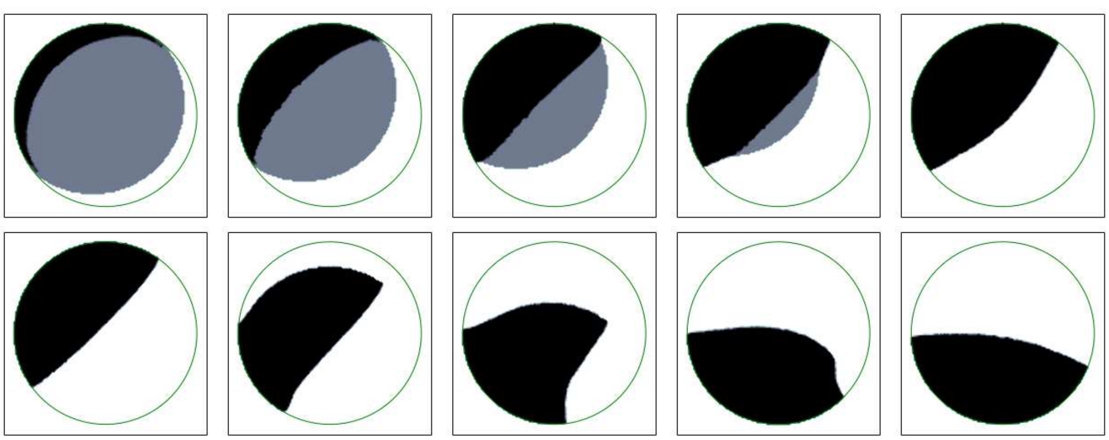


Figure 6: Catching up with congestion, at times 1, 2, 3, 4, 5, 24, 26, 29, 32, 35.

7 Extensions, perspectives, questions

7.1 Extensions

As in the Hilbertian case, some of the assumptions we made in order to prove our results could be relaxed. In particular, the convexity assumption could be replaced by a uniform prox-regularity assumption on the moving sets $C(t)$. If some of the arguments we used could be easily adapted, we point out that we repeatedly used convexity of the domain when establishing the main inequalities that gave contractivity, and hence uniqueness. For instance, the fact that the right-hand sides of (23) and (25) are negative used convexity of $C(t)$. Concerning the estimates providing existence, we used convexity in the proof of Corollary 2.6, but this was just a technical issue: any reasonable evolution of sets smooth enough allows to get the existence of a map T with those properties.

As for regularity in time, it is easy to generalize to an absolutely continuous curve of sets; it is more difficult if we try to do it for a continuous BV curve of sets.

Finally, we studied separately congestion and diffusion, but they can be treated together (see [28] for an evolution process merging diffusion and congestion, in the case of a fixed domain). Other external forcing terms could be accounted for, like the desired velocity field U of a crowd represented by the density ρ_t, subject to evolve in a moving environment (e.g. moving vehicles crossing a crowded area).

7.2 Asymptotic limits

Some links can be established between the different problems that we presented, in terms of asymptotic limits, but they are only informal observations for the moment.

Singular limit for the diffusive case. In the numerical section, we displayed an approximate solution of Equation (24) in a simple moving domain together with the solution when we drastically reduce the diffusion coefficient κ: this obviously provides a good approximation of the solution of the problem without diffusion, i.e. a solution to Equation (12). The relative weights of diffusion and sweeping can be quantified by a dimensionless number, commonly used in physics to characterize advection-diffusion processes, namely the *Péclet* number. In the present context, a natural way to define this number is

$$\mathrm{Pe} = \frac{U\ell}{\kappa},$$

where U is the order of magnitude of the velocity of the boundary (e.g. the maximal of the metric derivative for the Hausdorff distance over some time interval), ℓ is a typical lenght (e.g. the maximum of the diameters of the convex sets), and κ the diffusion coefficient. When Pe goes to 0, the diffusion dominates, and one can expect that the mass will tend to distribute uniformly over the moving set, at each instant. When Pe goes to infinity, a singular limit appears: the problem is expected to tend to the one without diffusion (i.e. pure sweeping, as addressed in Section 3). Yet, a rigorous proof of these convergence results (especially in the singular limit $\mathrm{Pe} \to \infty$) goes beyond the scope of this paper.

Non-linear diffusion. Another interesting limit result, that we present here as a conjecture, can be formulated as follows. We could possibly replace the linear diffusion equation

$$\partial_t \rho_t - \Delta \rho_t = 0$$

by a non-linear diffusion equation, i.e. we consider the so-called porous medium equation (see [33])

$$\partial_t \rho_t - \Delta(\rho_t^m) = 0.$$

Following [1], the solution of the congested problem (i.e. Equation (16)) can be expected to tend to the solution ot the congested problem (i.e. Equation (16)) when $m \to \infty$ (in [1] there is also a quantitative estimate of this convergence). This is due to the fact that the free energy, replacing the entropy $\mathcal{E}$, associated to the diffusion term $\Delta(\rho_t^m)$ is given by $\frac{1}{m-1}\rho^m$. When $m \to \infty$ this energy tends to 0 if $\rho \leq 1$ and to $+\infty$ otherwise, which is exactly the constraint which is used in Section 4.

7.3 Sweeping process in the Wasserstein space

The problems that we addressed in this paper can be seen as particular instances of a wider class of problems, namely sweeping processes in the Wasserstein space. An archetypal problem of this kind can be formulated as follows.

Given a family $(K(t))_t$ of sets in $W_2(\mathbb{R}^d)$, an initial measure $\rho_0 \in K_0$, find $t \mapsto \rho_t \in K(t)$ together with a family of velocity fields $(v_t)_t$, such that ρ_t is transported by v_t, i.e.

$$\partial_t \rho_t + \nabla \cdot (\rho_t v_t) = 0$$

over $\mathbb{R}^d$, with $v_t \in -\partial^W I_{K(t)}(\rho_t)$ for a.e. t, where the subdifferential is defined in the Wasserstein sense (see [4]). In the present situation, the W-subdifferential of an indicatrix function can be written

$$\partial^W I_K(\rho) = \{v \in L^2(\rho)\ ,\ \int (\mathrm{T}(x) - x) \cdot v(x)\, d\rho(x) \leq o(\|\mathrm{T} - id\|_{L^2(\rho)})\ ,\ \ \mathrm{T} \in \mathcal{A}_\rho^K\},$$

where $\mathcal{A}_\rho^K$ is the set of admissible maps, i.e. of all those maps verifying $T_\sharp \rho \in K$. All cases treated in this paper fit formally in this framework with $K(t)$ defined, respectively, by (13) for the pure measure sweeping, and by (15) for the maximal density constraint. In the case with diffusion, $K(t)$ is again defined by (13), but the evolution inclusion is supplemented by a forcing term that accounts for diffusion. The latter contribution can be integrated into the subdifferential: if $\mathcal{E}(\rho)$ is the logarithmic entropy: $\mathcal{E}(\rho) = \int \rho \ln \rho$, the velocity is subject to

$$v_t \in -\partial^W \left(I_{K(t)} + \mathcal{E}\right)(\rho_t).$$

One could also consider sweeping problems that are no longer related to a moving set in the underlying euclidean case. For instance, setting

$$K(t) = \left\{\rho \in \mathcal{P}(\mathbb{R}^d)\ ,\ \ \rho \leq 1/t\right\},$$

is a natural way to formalize in the Wasserstein framework the Hele-Shaw problem (see e.g. [18] and [1] for connections with crowd motion). One could replace the uniform threshold $1/t$ with a more general $f(t)$ (which involves no extra difficulties), or even $f(t,x)$, which is more delicate.

As for the abstract problem described above, it is tempting to carry out the catching-up approach, and to define approximate measures by transposing the scheme (2) in the Wasserstein setting:

$$\begin{cases} \rho_0^\tau = \rho_0, \\ \rho_{k+1}^\tau = P_{K((k+1)\tau)}[\rho_k^\tau], \end{cases} \quad . \tag{28}$$

In a second step , define v_{k+1}^τ as

$$v_{k+1}^\tau = \frac{id - T_{k+1}}{\tau},$$

where T_{k+1} transports ρ^τ_{k+1} to ρ^τ_k in an optimal way. Then, use compactness to build limits ρ_t and v_t that verify the transport equation. And finally, show that v_t lies in $-\partial^W I_{K(t)}$ for a.e. t. To be carried out, such an approach requires two main ingredients. Firstly, the distance to K has to be attained, in order to build the next density at each time of the catching-up algorithm. This can be ensured be requiring standard closedness assumptions on K. Secondly, in order to adapt the proofs we proposed, a Wasserstein counterpart of (14), that is some kind of dual characterization of the outward normal cone, has to be established in the present context. It requires in particular the existence of an optimal *map* (and not plan) from ρ_{k+1} to ρ_k. This can be ensured by requiring K to contain only measures that are absolutely continuous. Beyond these considerations on technical assumptions that would have to be prescribed to obtain a general abstract result, we also disregarded this fully abstract approach in the present paper because in general there is no obvious way to apply it to particular situations, like the ones we studied. Indeed, as already pointed out in [25], the W-subdifferential is commonly very difficult to identify in a precise way, and the overall approach does not make much sense if one is unable in general to describe this subdifferential, i.e. to translate $v_t \in -\partial^W I_{K_t}(\rho_t)$ in terms of characterization of the actual velocity, and finally obtain an evolution PDE.

7.3.1 Control issues, long-time behavior

Consider the pure sweeping process, with no congestion nor diffusion. It may happen that the measure ρ_t becomes a Dirac mass $\delta_{x(t)}$ after a finite time. An interesting control problem that could be investigated is how to move the convex set $C(t)$ (under constraints on its deformation and on its speed, for instance by translation only) so as to transform ρ_0 into a Dirac mass as fast as possible. It is also interesting to understand if, and in which sense, ρ_t generically becomes a Dirac mass, and how long does it take to observe it.

Concerning the situation with congestion, it is common sense to expect that any initial measure will be swept into a measure that saturates the constraints (i.e. the density ρ_t is the characteristic function of a subset of the moving set $C(t)$) as soon as the set has entirely "swept itself", i.e. every point of the initial set $C(0)$ has been visited by the moving boundary. It is also to be expected that this feature will remain forever (patches stay patches). Yet, there is no obvious proof of these properties. Note that in the discrete evolution, for fixed $\tau > 0$, it is true that patches stay patches. Indeed, it is well known that the projection of a saturated density is again a saturated density (i.e taking only values in $\{0, 1\}$, see for instance [13]); the problem is the limit as $\tau \to 0$, since it can destroy this property, which is not closed under weak convergence. A possible solution would be to prove stronger bounds (e.g., BV bounds, again as in [13]), in order to obtain strong L^1 convergence, but these estimates are complicated by the fact that the domain $C(t)$ is moving. One can produce easy examples where the perimeter of a patch increases by projection (think of a ball which is pushed by the moving boundary, and cannot be a ball anymore). A natural conjecture would be that the perimeter of a patch never goes beyond the sum of the perimeter of the initial patch and of the perimeter of $C(t)$, but this has to be proven.

Still in the context of congested measure sweeping, a natural problem is the following: given a a "shaper" defined by a domain C_0, and a goal shape (i.e. a given convex domain $\tilde{C} \subset C_0$), is there a way to obtain $\tilde{C}$ as the support of a saturated density ρ_T, solution to Problem (16) with a given initial density, and with a family $(C(t))$ obtained from C_0 by a smooth familly of rigid motions? Note that the realizable shapes may be very different from the shaper C_0. In particular, non convex shapes can be obtained with a convex shaper C_0. If the task is achievable in finite time, is there an optimal way to obtain the desirable shape, e.g. minimizing the time while prescribing an upper bound to the metric derivative of the moving sets?

References

[1] D. Alexander, I. Kim, Y. Yao, Quasi-static evolution and congested crowd transport. *Nonlinearity*, 27, 1-36, 2014.

[2] L. Ambrosio, Minimizing movements, *Rend. Accad. Naz. Sci. XL Mem. Mat. Sci. Fis. Natur.*, 19, 191–246, 1995.

[3] L. Ambrosio and G. Crippa, Existence, uniqueness, stability and differentiability properties of the flow associated to weakly differentiable vector fields. In *Transport Equations and Multi-D Hyperbolic Conservation Laws*, Lecture Notes of the Unione Matematica Italiana, Vol. 5, Springer (2008).

[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] L. Ambrosio and P. Tilli, *Topics on analysis in metric spaces.* Oxford Lecture Series in Mathematics and its Applications (25). Oxford University Press, Oxford, 2004.

[6] F. Bernicot and J. Venel, Differential inclusions with proximal normal cones in Banach spaces, *J. Convex Anal.,* 17, 451–484, 2010.

[7] F. Bernicot, J. Venel, Stochastic perturbation of sweeping process and a convergence result for an associated numerical scheme *Journal of Differential Equations* 251 (4-5), 1195–1224, 2011.

[8] M. Bounkhel and L. Thibault, Nonconvex sweeping process and prox-regularity in Hilbert space, *J. Nonlinear Convex Anal.,* 6, 359–374, 2005.

[9] Y. Brenier, Décomposition polaire et réarrangement monotone des champs de vecteurs. (French) *C. R. Acad. Sci. Paris Sér. I Math.* 305 (19), 805–808, 1987.

[10] A. Boulkhemair and A. Chakib, On the uniform Poincaré inequality. *Comm. Part. Diff. Eq.,* 32, 1439–1447, 2007.

[11] G. Colombo and V. V. Goncharov, The sweeping processes without convexity, *Set-Valued Anal.,* 7, 357–374, 1999.

[12] G. Colombo and M. D. P. Monteiro Marques, Sweeping by a continuous prox-regular set, *J. Differential Equations*, 187, 46–62, 2003.

[13] G. De Philippis, A. R. Mészáros, F. Santambrogio and B. Velichkov, BV estimates in optimal transportation and applications, preprint available at `http://cvgmt.sns.it/paper/2559/`

[14] S. Di Marino, A. R. Mészáros, Uniqueness issues for evolutive equations with density constraints. In preparation

[15] J. L. Doob, *Classical Potential Theory and Its Probabilistic Counterpart,* Grundlehren der Mathematischen Wissenschaften, 262, Springer-Verlag, New York, 1984.

[16] J. F. Edmond and L. Thibault, Relaxation of an optimal control problem involving a perturbed sweeping process, *Math. Program,* 104, 347–373, 2005.

[17] J. F. Edmond and L. Thibault, BV solutions of nonconvex sweeping process differential inclusion with perturbation, *J. Differential Equations,* 226, 135–179, 2006.

[18] B. Gustafsson and M. Sakai, Properties of some balayage operators, with applications to quadrature domains and moving boundary problems, *Nonlinear Analysis,* 22, 1221–1245, 1994.

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

[20] L. Levine, Y. Peres, Spherical asymptotics for the rotor-router model in Zd, *Indiana Univ. Math. J.* 57 (1) 431–449, 2008.

[21] G. Loeper, Uniqueness of the solution to the Vlasov-Poisson system with bounded density. *J. Math. Pures Appl.* 86 (1), 68–79, 2006.

[22] B. Maury and J. Venel, A discrete contact model for crowd motion, *ESAIM Mathematical Modelling and Numerical Analysis*, 45, 145–168, 2011.

[23] B. Maury, A. Roudneff-Chupin and F. Santambrogio, A macroscopic crowd motion model of gradient flow type, *Mathematical Models and Methods in Applied Sciences,* 20, 1787–1821, 2010.

[24] B. Maury, A. Roudneff-Chupin and F. Santambrogio, Congestion-driven dendritic growth, *Discrete and Continuous Dynamical Systems,* 34 (4), 1575–1604, 2014.

[25] B. Maury, A. Roudneff-Chupin, F. Santambrogio, J. Venel, Handling Congestion in Crowd Motion Modeling, *Networks and Heterogeneous Media,* 6 (3), 485–519, 2011.

[26] J. Matousek, *Lectures on Discrete Geometry* Springer Gr. Texts Math., 2002.

[27] R. J. Mc Cann, A convexity principle for interacting gases. *Adv. Math.* 128 (1), 153–159, 1997.

[28] A. R. Mészáros, F. Santambrogio, A diffusive model for macroscopic crowd motion with density constraints, preprint available at `http://cvgmt.sns.it/paper/2644/`

[29] A. R. Mészáros, Mean Field Games with Density Constraints, MSc thesis, É. Polytechnique, Palaiseau.

[30] J. J. Moreau, On unilateral constraints, friction and plasticity, in *New Variational Techniques in Mathematical Physics* (G. Capriz and G. Stampacchia, Ed.), 173- -322, C.I.M.E. II Ciclo 1973, Edizioni Cremonese, Roma, 1974.

[31] J.-J. Moreau, Evolution problem associated with a moving convex set in a Hilbert space, *J. Differential Equations* 26, 346–374, 1977.

[32] J.-J. Moreau, Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires, *C. R. Acad. Sci. Paris,* 255, 238–240, 1962.

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

[34] P. Pécol, S. Dal Pont, S. Erlicher, P. Argoul, Smooth/non-smooth contact modeling of human crowds movement: numerical aspects and application to emergency evacuations, *Physica A: Statist. Mech. Appl.* 388, 3600–3608, 2009.

[35] R. A. Poliquin and R. T. Rockafellar, Prox-regular functions in variational analysis, *Trans. Amer. Math. Soc.*, 348, 1805–1838, 1996.

[36] R. A. Poliquin, R. T. Rockafellar and L. Thibault, Local differentiability of distance functions, *Trans. Amer. Math. Soc.,* 352, 5231–5249, 2000.

[37] *Discrete Numerical Modeling of Ganular Materials,* Arang Radjai, Frederic Dubois Eds., ISTE Ltd and John Wiley & Sons Inc, 2011.

[38] R. T. Rockafellar and R. Wets, *Variational Analysis,* Grundlehren der Mathematischen, Wissenschaften, 317, Springer-Verlag, Berlin, 1998.

[39] F. Santambrogio, *Optimal Transport for Applied Mathematicians.* Birkäuser, NY, 2015.

[40] R. Schneider Curvature measures of convex bodies, *Annali di Matematica Pura ed Applicata*, 116 (1), 101–134, 1978.

[41] D.E. Stewart, Convergence of a time-ttepping scheme for rigid-body dynamics and resolution of Painlevé's problem, *Arch. Rational Mech. Anal.* 145, 215–260, 1998.

[42] L. Thibault, Sweeping process with regular and nonregular sets *J. Differential Equations,* 193, 1–26, 2003.

[43] J. Venel, A numerical scheme for a class of sweeping processes, *Numerische Mathematik* 118 (2), 367–400, 2011.

[44] C. Villani, *Topics in Optimal Transportation,* Grad. Stud. Math., 58, AMS, Providence, RI, 2003.