DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **41**, Number **11**, November **2021**
doi:10.3934/dcds.2021066
pp. **5009–5036**

MODELING OF CROWDS IN REGIONS WITH MOVING OBSTACLES

Nadezhda Maltugueva and Nikolay Pogodaev*

Matrosov Institute for System Dynamics and Control Theory
134 Lermontov St., Irkutsk, 664033, Russia

(Communicated by Rinaldo M. Colombo)

Abstract. We present a model of crowd motion in regions with moving obstacles, which is based on the notion of measure sweeping process. The obstacle is modeled by a set-valued map, whose values are complements to r-prox-regular sets. The crowd motion obeys a nonlinear transport equation outside the obstacle and a normal cone condition (similar to that of the classical sweeping processes theory) on the boundary. We prove the well-posedness of the model, give an application to environment optimization problems, and provide some results of numerical computations.

1. **Introduction.** Moving crowds are usually modeled, at the macroscopic level, by evolution PDEs with nonlocal terms [5, 9, 7, 8, 17, 16, 19]. States of these equations are measures (or densities) which describe the distribution of individuals (also called agents) on some configuration space, typically, the space of agents' positions or position-velocity pairs. Nonlocal terms appear due to the fact that the behavior of each agent depends on the positions of other agents. Such equations can often be expressed either as Wasserstein gradient flows [2] or nonlocal transport equations [18]. Each framework has its own advantages: the first one allows to deal with various diffusion terms, the second one admits vector fields that do not possess the gradient structure. Stationary obstacles in both cases are handled by imposing either the Neumann $\rho v \cdot n = 0$ or the Dirichlet $\rho = 0$ boundary condition. The latter condition is more demanding: to achieve it one has to adjust nonlocal terms [8] or introduce specific distances in the space of measures [13].

Measures evolving inside moving domains were considered, probably for the first time, by Di Marino, Maury and Santambrogio in [11]. They described, in particular, how a measure ρ_t supported on a time dependent convex set $\boldsymbol{C}(t)$ evolves when it is pushed by the boundary of $\boldsymbol{C}(t)$. To deal with the problem they introduced a notion of measure sweeping process and extend the classical Moreau catching-up scheme to the space of measures.

Our goal is to extend the approach of [11] in two respects: we admit here non-convex (more precisely, r-prox-regular) driving sets $\boldsymbol{C}(t)$ and measures ρ_t which are

2020 *Mathematics Subject Classification.* Primary: 35F31; Secondary: 35R37, 49J53, 49Q20.
Key words and phrases. Crowd dynamics, sweeping process, Wasserstein space, well-posedness, catching-up scheme, moving boundary problem.
This work was partially supported by the Russian Foundation for Basic Research under contracts no. 18-31-20030 (the first author) and no. 18-01-00026 (the second author).
* Corresponding author.

not only pushed by $\partial\boldsymbol{C}(t)$ but also drift along a given nonlocal vector field $\mathscr{V}(\rho_t)$. One can naturally think of such evolution models as *perturbed* measure sweeping processes with perturbation given by $\mathscr{V}$. We shall see that these modifications turn the concept of measure sweeping process into a usable tool of crowd dynamics. Moreover, we shall prove that any solution of the measure sweeping process satisfies the underlying PDE together with the Neumann type condition $\rho(n_t + v\cdot n_x) = 0$ on the boundary of the time dependent domain. Let us stress that below we deal only with Lipschitz non-local vector field, so basically we stay within the framework of [18].

We define a moving obstacle by a set-valued map $\boldsymbol{O}\colon [0,T] \rightrightarrows \mathbb{R}^d$ taking values among open subsets of $\mathbb{R}^d$. Instead of dealing with $\boldsymbol{O}(t)$ we prefer to look at its complement $\boldsymbol{C}(t) = \mathbb{R}^d \setminus \boldsymbol{O}(t)$, which we call the viability region. The values of $\boldsymbol{C}\colon [0,T] \rightrightarrows \mathbb{R}^d$ are assumed to be closed bounded r-prox-regular sets, for a given $r > 0$. The perturbed measure sweeping process is the system of the form

$$\begin{cases} \partial_t\rho_t + \nabla\cdot(v_t\rho_t) = 0, \\ v_t(x) - \mathscr{V}(\rho_t)(x) \in -N_{\boldsymbol{C}(t)}(x), \\ \operatorname{spt}(\rho_t) \subset \boldsymbol{C}(t). \end{cases} \tag{1}$$

where $\mathscr{V}$ maps measures to vector fields and $N_A(x)$ denotes the proximal normal cone to $A \subset \mathbb{R}^d$ at x. In what follows, $\mathcal{P}_2(\mathbb{R}^d)$ denotes the space of probability measures with finite second moments equipped with the Wasserstein distance W_2.

Definition 1.1. An absolutely continuous curve $\rho\colon [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ is said to be a solution of the measure sweeping process (1) if

- there exists a Borel vector field $(t,x) \mapsto v_t(x)$ such that (ρ, v) satisfies
$$\partial_t\rho_t + \nabla\cdot(v_t\rho_t) = 0$$
in the sense of distributions;
- the normal cone condition
$$v_t(x) - \mathscr{V}(\rho_t)(x) \in -N_{\boldsymbol{C}(t)}(x)$$
holds for a.e. $t \in [0,T]$ and ρ_t-a.e. $x \in \mathbb{R}^d$;
- $\operatorname{spt}(\rho_t) \subset \boldsymbol{C}(t)$ for all $t \in [0,T]$.

Remark that the normal cone condition implies that $v_t(x) = \mathscr{V}(\rho_t)(x)$ if x lies in the interior of $\boldsymbol{C}(t)$, which means that inside $\boldsymbol{C}(t)$ the crowd moves according to the nonlocal transport equation

$$\partial_t\rho_t + \nabla\cdot(\mathscr{V}(\rho_t)\rho_t) = 0.$$

The inclusion $\operatorname{spt}\rho_t \subset \boldsymbol{C}(t)$ guaranties that the crowd never leaves the viability region.

Throughout the paper, we impose the following assumptions:

($\mathbf{A_1}$) There exist $L > 0$ such that $\mathscr{V}\colon \mathcal{P}_2(\mathbb{R}^d) \to C(\mathbb{R}^d;\mathbb{R}^d)$ satisfies
$$\|\mathscr{V}(\rho_1) - \mathscr{V}(\rho_2)\|_\infty \le LW_2(\rho_1,\rho_2) \quad \forall \rho_1,\rho_2 \in \mathcal{P}_2(\mathbb{R}^d)$$
$$|\mathscr{V}(\rho)(x) - \mathscr{V}(\rho)(y)| \le L|x-y|, \quad |\mathscr{V}(\rho)(x)| \le L \quad \forall x,y \in \mathbb{R}^d,\ \forall \rho \in \mathcal{P}_2(\mathbb{R}^d),$$

($\mathbf{A_2}$) There exist $M > 0$ and $r > 0$ such that the set-valued map $\boldsymbol{C}\colon [0,T] \rightrightarrows \mathbb{R}^d$ is M-Lipschitz in the Hausdorff distance d_H:
$$d_H(\boldsymbol{C}(t), \boldsymbol{C}(s)) \le M|t-s| \quad \forall t,s \in [0,T],$$
and its values are compact r-prox-regular sets.

Now we are ready to state the main result of the paper.

Theorem 1.2. *Let $\mathscr{V}$ satisfy* $(\mathbf{A_1})$ *and* $\boldsymbol{C}$ *satisfy* $(\mathbf{A_2})$. *Then the following assertions hold:*

(1) *For any $\vartheta \in \mathcal{P}_2(\mathbb{R}^d)$ such that $\operatorname{spt}\vartheta \subset \boldsymbol{C}(0)$, there exists a unique solution $\rho\colon [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ of the measure sweeping process* (1) *with $\rho_0 = \vartheta$.*
(2) *The corresponding vector field v satisfies*

$$|v_t(x)| \le 2L + M, \quad \text{for a.e. } t \text{ and } \rho_t\text{-a.e. } x.$$

In particular, $t \mapsto \rho_t$ is $(2L+M)$-Lipschitz.
(3) *If* graph $\boldsymbol{C}$ *is r'-prox-regular for some $r' > 0$ then*

$$\xi + v_t(x)\cdot\eta = 0 \quad \forall(\xi,\eta) \in N_{\operatorname{graph}\boldsymbol{C}}(t,x)$$

for a.e. t and ρ_t-a.e. x.
(4) *If $\tilde{\boldsymbol{C}}$ is another set-valued map satisfying* $(\mathbf{A_2})$ *and $\tilde\rho$ is a solution of the corresponding measure sweeping process then the estimate*

$$\boldsymbol{r}(t) \le \Big(\boldsymbol{r}(0) + (6L+2M)\int_0^t \Delta(s)\,ds\Big) e^{\left(4L + \frac{3L+M}{2r}\right)t} \tag{2}$$

holds for all $t \in [0,T]$, where $\boldsymbol{r}(t) = \frac{1}{2}W_2^2(\rho_t,\tilde\rho_t)$ and $\Delta(t) = d_H(\boldsymbol{C}(t), \tilde{\boldsymbol{C}}(t))$.

To prove the existence part we use the following catching-up scheme, which yields, for every natural N, a sequence of probability measures $\rho^\tau_{k\tau}$, $k = 0,\dots,2N$.

Algorithm 1 The catching-up scheme (see Figure 1)

1: Split $[0,T]$ into $2N$ segments of length $\tau := T/(2N)$, then set $\rho_0 := \vartheta$, $k := 0$.
2: **while** $2k\tau < T$ **do**
3: Solve the linear continuity equation

$$\partial_t\mu_t + \nabla\cdot(2\mathscr{V}(\rho_{2k\tau})\mu_t) = 0, \quad \mu_{2k\tau} = \rho^\tau_{2k\tau}, \tag{3}$$

on the segment $[2k\tau,(2k+1)\tau]$ and set $\rho^\tau_{(2k+1)\tau} := \mu_{(2k+1)\tau}$.
4: Project $\rho^\tau_{(2k+1)\tau}$ onto $\boldsymbol{C}((2k+2)\tau)$ and set $\rho^\tau_{(2k+2)\tau}$ to be equal to this projection.
5: $k := k+1$.
6: **end while**

With $\rho^\tau_{k\tau}$ at hand, we can construct two curves on $\mathcal{P}_2(\mathbb{R}^d)$:

- a *continuous* one ρ^τ, by connecting $\rho^\tau_{2k\tau}$, $\rho^\tau_{(2k+1)\tau}$ with a (unique) trajectory of (3) and $\rho^\tau_{(2k+1)\tau}$, $\rho^\tau_{(2k+2)\tau}$ with a (unique) Wasserstein geodesic;
- a *piecewise constant* one $\bar\rho^\tau$, which equals to $\rho^\tau_{2k\tau}$ on $[2k\tau,(2k+1)\tau]$ and $\rho^\tau_{(2k+1)\tau}$ on $[(2k+1)\tau,(2k+2)\tau]$.

It can be shown that ρ^τ converges to some ρ as $\tau \to 0$. The latter curve, being absolutely continuous, satisfies $\partial_t\rho_t + \nabla\cdot(v_t\rho_t) = 0$, for some velocity v, and $\operatorname{spt}\rho_t \subset \boldsymbol{C}(t)$, for all t. The piecewise constant curve $\bar\rho^\tau$, which has the same limit as ρ^τ, is used to prove the normal cone condition.

Assertion (4) (and, thus, the uniqueness part) follows from the standard representation for the time derivative of the squared Wasserstein distance along a pair of absolutely continuous curves (see [2] or Appendix C).

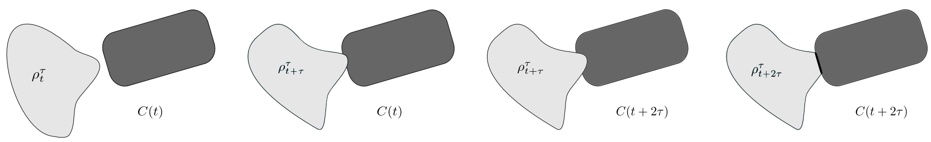


FIGURE 1. One step of the catching-up algorithm. Here $t = 2k\tau$, the dark rounded rectangle represents an obstacle.

Structure of the paper. In Section 2 we introduce the notation, recall basic properties of prox-regular sets and standard results concerning the geometry of $\mathcal{P}_2(\mathbb{R}^d)$. Section 3, the most technical one, contains a proof of the existence part of Theorem 1.2. The well-posedness part is proven in Section 4. Then, in Section 5, we give an application of Theorem 1.2 to environment optimization problems. Finally, Section 6 contains some numerical computations for the measure sweeping process (1).

2. **Preliminaries.** We begin with notation, then recall some useful facts about prox-regular sets and the Wasserstein space.

2.1. **Notation.** Below, X and Y are metric spaces, U is an open subset of $\mathbb{R}^d$.

$\mathcal{P}(X)$	the space of probability measures on X
$\mathcal{P}_2(\mathbb{R}^d)$	the space of probability measures μ on $\mathbb{R}^d$ with $\int \lvert x\rvert^2\, d\mu < \infty$
$\mathcal{M}(X;\mathbb{R}^d)$	the space of finite Radon vector measures on X
$C(X;Y)$	the space of continuous maps $f\colon X \to Y$
$C^k(U)$	the space of k times continuously differentiable maps $f\colon U \to \mathbb{R}$
$C^k_c(U)$	the space of all compactly supported maps from $C^k(U)$
$\mathcal{K}(\mathbb{R}^d)$	the space of nonempty compact subsets of $\mathbb{R}^d$
$\mathcal{K}_r(\mathbb{R}^d)$	the space of compact r-prox-regular subsets of $\mathbb{R}^d$
$\rightharpoonup$	the weak (narrow) convergence on $\mathcal{M}(X;\mathbb{R}^d)$
W_2	the quadratic Wasserstein distance on $\mathcal{P}_2(\mathbb{R}^d)$
d_H	the Hausdorff distance on $\mathcal{K}(\mathbb{R}^d)$
$\lVert\cdot\rVert_\infty$	the supremum norm on $C(X;Y)$
$\mathrm{Lip}(f)$	the minimal Lipschitz constant of $f \in C(X;Y)$
$\boldsymbol{B}$	the closed unit ball in $\mathbb{R}^d$ centered at 0
$a + r\boldsymbol{B}$	the closed ball in $\mathbb{R}^d$ with center $a \in \mathbb{R}^d$ and radius $r \geq 0$
$d_A(x)$	the distance between a compact set $A \subset \mathbb{R}^d$ and a point $x \in \mathbb{R}^d$
$P_A(x)$	the projection map, i.e., $P_A(x) = \{a \in A :\ \lvert x - a\rvert = d_A(x)\}$
A°	the interior of $A \subset \mathbb{R}^d$
∂A	the boundary of $A \subset \mathbb{R}^d$
A^c	the complement of $A \subset \mathbb{R}^d$
$N_A(x)$	the proximal normal cone to $A \subset \mathbb{R}^d$ at x

2.2. **Prox-regular sets.** We collect here, for the future references, some basic properties of *prox-regular sets* (also called sets with positive reach).

Definition 2.1 ([12, 20]). A closed set $S \subset \mathbb{R}^d$ is called r**-prox-regular**, for $r \in (0, +\infty]$, if the projection map P_S is single-valued and continuous within the open spherical neighborhood $S + r\boldsymbol{B}^\circ = \{x :\ d_S(x) < r\}$.

A closely related notion of *proximal normal* is defined as follows.

Definition 2.2 ([6, Chapter 11])**.** Let S be a closed set and $x \in S$. A vector $v \in \mathbb{R}^d$ is called a **proximal normal** to S at x if there exists $\sigma = \sigma(x,v) \geq 0$ such that

$$\langle v, y-x\rangle \leq \sigma\,|y-x|^2 \quad \forall y \in S.$$

The set $N_S(x)$ consisting of such v defines the **proximal normal cone** to S at x.

The prox-regular sets can be characterized in several equivalent ways (see [24]) gathered in the proposition below.

Proposition 1. *The following assertions are equivalent:*

(*a*) *S is r-prox-regualar;*

(*b*) *for any $x \in S$, each nonzero proximal normal $v \in N_S(x)$ is realized by an r-ball, i.e.,*

$$\langle v, y-x\rangle \leq \frac{|v|}{2r}\,|x-y|^2 \quad \forall y \in S;$$

(*c*) *d_S^2 is continuously differentiable over $S + r\boldsymbol{B}^\circ$.*

Moreover, one has

$$\frac{1}{2}\nabla d_S^2(x) = x - P_S(x) \quad \forall x \in S + r\boldsymbol{B}^\circ,$$

and, for any positive $r' < r$,

$$|P_S(x) - P_S(y)| \leq \frac{r}{r-r'}\,|x-y| \quad \forall x,y \in S + r'\boldsymbol{B}^\circ.$$

2.3. **Space of measures.** Here we briefly recall basic facts about the Wasserstein space $\mathcal{P}_2(\mathbb{R}^d)$. The corresponding proofs can be found, e.g., in [2, 25].

Space of measures as a metric space. Recall that $\mathcal{P}_2(\mathbb{R}^d)$ is a *complete separable metric space* when equipped with the quadratic *Wasserstein distance* [25, Chapter 6]:

$$W_2(\mu,\nu) = \left(\inf_{\pi\in\Gamma(\mu,\nu)} \int |x-y|^2\,d\pi(x,y)\right)^{1/2}. \tag{4}$$

The infimum above is taken over the set $\Gamma(\mu,\nu)$ of all transport plans between the measures μ and ν. Recall that a transport plan $\Pi \in \Gamma(\mu,\nu)$ is a probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ whose projections on the first and the second factor are μ and ν, respectively. In other words,

$$\pi^1_\sharp \Pi = \mu, \quad \pi^2_\sharp \Pi = \nu,$$

where π^1, π^2 are the projection maps on the factors. Here $\sharp$ is the pushforward functor, which works as follows: for any Borel map $f\colon X \to Y$ between metric spaces, it generates a map $f_\sharp\colon \mathcal{P}(X) \to \mathcal{P}(Y)$ by the rule

$$(f_\sharp\mu)(A) \doteq \mu\left(f^{-1}(A)\right), \quad \text{for all Borel sets } A \subset Y.$$

Remark that the minimum in (4) can always be achieved. Any transport plan that provides the minimum is called *optimal*. A plan π is optimal if and only if its support $\operatorname{spt}\pi$ is contained in a *cyclically monotone set* $A \subset \mathbb{R}^d \times \mathbb{R}^d$, which means that any finite collection of points $(x_i, y_i) \in A$, $i = 1, \ldots, k$, satisfies

$$\langle y_1, x_2 - x_1\rangle + \langle y_2, x_3 - x_2\rangle + \cdots + \langle y_k, x_1 - x_k\rangle \leq 0.$$

The convergence in the Wasserstein distance is slightly stronger then the weak convergence of measures. More precisely, $W_2(\rho_k, \rho) \to 0$ if and only if

$$\int \varphi\,d\rho_k \to \int \varphi\,d\rho,$$

for any continuous φ satisfying $\varphi(x) \le C\left(1+|x|^2\right)$ for some $C>0$.

Space of measures as a length space. Recall that the *length* of a continuous curve $\rho\colon [a,b] \to \mathcal{P}_2(\mathbb{R}^d)$ is given by

$$L(\rho) = \sup \sum_{i=1}^{N} W_2(\rho_{t_{i-1}}, \rho_{t_i}),$$

where the supremum is taken among all finite partitions $a = t_0 \le t_1 \le \cdots \le t_N = b$ of the interval $[a,b]$.

The space $\mathcal{P}_2(\mathbb{R}^d)$ is a *strictly intrinistic length space*, meaning that any two measures $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ can be connected with a continuous curve whose length is $W_2(\mu,\nu)$ (see [4]). Such a curve is called a *minimal geodesic* from μ to ν. Any minimal geodesic joining μ and ν can be uniquely parametrized by $t \in [0,1]$ so that

$$W_2(\rho_t, \rho_s) = |t-s| W_2(\mu,\nu).$$

In what follows, by saying that ρ is a *geodesic* joining μ and ν we mean that ρ is a minimal geodesic from μ to ν parametrized in this way.

Any geodesic ρ joining μ and ν takes the form $\rho_t = \left((1-t)\pi^1 + t\pi^2\right)_\sharp \Pi$, where Π is an optimal plan between μ and ν. In particular, if the optimal plan is unique, the geodesic is unique as well. If an optimal plan $\Pi \in \Gamma(\mu,\nu)$ is realized by a transport map $F\colon \mathbb{R}^d \to \mathbb{R}^d$, i.e.,

$$F_\sharp \mu = \nu, \quad W_2^2(\mu,\nu) = \int |x - F(x)|^2 \, d\mu(x),$$

then $\Pi = (\mathbf{id}, F)_\sharp \mu$ and thus the geodesic takes the form $\rho_t = \left((1-t)\mathbf{id} + tF\right)_\sharp \mu$.

Curves in the space of measures. A map $\rho\colon [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ is called Borel measurable if $t \mapsto \rho_t(B)$ is Borel measurable for any Borel set $B \subset \mathbb{R}^d$. Any Borel measurable ρ produces a measure $\boldsymbol{\rho}$ on $[0,T] \times \mathbb{R}^d$ by the rule

$$\int_{[0,T]\times\mathbb{R}^d} \varphi(t,x)\, d\boldsymbol{\rho}(t,x) \doteq \int_0^T \int_{\mathbb{R}^d} \varphi(t,x)\, d\rho_t(x)\, dt, \quad \forall \varphi \in C_c^\infty\left((0,T)\times\mathbb{R}^d\right).$$

Below, we never distinguish ρ from $\boldsymbol{\rho}$.

Lemma 2.3. *If $\rho^k \to \rho$ in $C\left([0,T];\mathcal{P}_2(\mathbb{R}^d)\right)$ then $\rho^k \rightharpoonup \rho$ in $\mathcal{M}\left([0,T]\times\mathbb{R}^d;\mathbb{R}\right)$*

Proof. For any $\varphi \in C_b\left([0,T]\times\mathbb{R}^d\right)$, one has

$$\int \varphi(\, d\rho^k - d\rho) = \int_0^T \int \varphi(t,x)\left(\, d\rho_t^k(x) - \, d\rho_t(x)\right)\, dt.$$

Since φ is bounded and $\rho_t^k \to \rho_t$ in $\mathcal{P}_2(\mathbb{R}^d)$ for each $t \in [0,T]$, we conclude that $f^k(t) \doteq \int \varphi(t,x)\left(\, d\rho_t^k(x) - \, d\rho_t(x)\right)$ converges to 0 for all $t \in [0,T]$. On the other hand, $|f^k| \le 2\|\varphi\|_\infty$, hence the assertion follows from Lebesgue's dominated convergence theorem. □

A curve $\rho\colon [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ is called *absolutely continuous* if there exists a function $g \in L^1([0,T])$ such that

$$W_2(\rho_t, \rho_s) \le \int_s^t g(\tau)\, d\tau, \quad \forall t \ge s, \quad s,t \in [0,T].$$

If $\rho\colon [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ is absolutely continuous, the limit

$$|\rho'|(t) \doteq \lim_{\varepsilon\to 0} \frac{W_2(\rho_{t+\varepsilon}, \rho_t)}{|\varepsilon|}$$

exists for a.e. $t \in [0,T]$ and is called the *speed* (or the *metric derivative*) of ρ at t. Absolutely continuous curves has finite length which can be expressed as

$$L(\rho) = \int_0^T |\rho'|(t)\, dt.$$

The following theorem from [2, Section 8.3] shows that absolutely continuous curves on $\mathcal{P}_2(\mathbb{R}^d)$ are completely characterized by the continuity equations. In the statement, $L^2_\mu(\mathbb{R}^d;\mathbb{R}^d)$ denotes the space of μ-measurable vector fields $v\colon \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\|v\|_\mu \doteq \left(\int_{\mathbb{R}^d} |v|^2\, d\mu\right)^{1/2} < \infty.$$

As usual, two maps v and v' are considered equivalent if they coincide for μ-a.e. x.

Theorem 2.4 ([2, Chapter 8.3])**.** *Let $\rho\colon [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve. Then there exists a Borel vector field $(t,x) \mapsto v_t(x)$ such that*

$$v_t \in L^2_{\rho_t}(\mathbb{R}^d;\mathbb{R}^d), \quad \|v_t\|_{\rho_t} \le |\rho'|(t) \quad \text{for a.e. } t \in [0,T],$$

and the continuity equation

$$\partial_t \rho_t + \nabla \cdot (v_t \rho_t) = 0 \tag{5}$$

holds in the sense of distributions. Conversely, if a Borel map $\rho\colon [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ satisfies equation (5) for some Borel vector field v with

$$\int_0^T \|v_t\|^2_{\rho_t}\, dt < \infty, \tag{6}$$

then $t \mapsto \rho_t$ admits an absolutely continuous representative $t \mapsto \varrho_t$ with

$$|\varrho'|(t) \le \|v_t\|_{\varrho_t} \quad \text{for a.e. } t \in [0,T].$$

Any Borel vector field v satisfying (5) is called a *velocity* of the absolutely continuous curve ρ. If a velocity v is sufficiently regular, e.g.,

$$v_t \in C(\mathbb{R}^d;\mathbb{R}^d) \quad \forall t \in [0,T] \quad \text{and} \quad \int_0^T (\|v_t\|_\infty + \mathrm{Lip}(v_t))\, dt < \infty, \tag{7}$$

it generates a map $\Psi\colon [0,T] \times [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ by the rule

$$\Psi_{s,t}(x) \doteq y(t), \quad s,t \in [0,T], \quad x \in \mathbb{R}^d,$$

where $y\colon [0,T] \to \mathbb{R}^d$ is a unique solution of the Cauchy problem

$$\dot y(t) = v_t\left(y(t)\right), \quad y(s) = x.$$

This map Ψ, called the *flow* of v, is well-defined and satisfies the identities

$$\Psi_{s,s} = \mathbf{id}, \quad \Psi_{s,t} = \Psi_{\tau,t} \circ \Psi_{s,\tau} \quad \forall s,t,\tau \in [0,T].$$

Moreover, for any $s,t \in [0,T]$, the map $\Psi_{s,t}\colon \mathbb{R}^d \to \mathbb{R}^d$ is a homeomorphism with

$$\mathrm{Lip}(\Psi_{s,t}) \le \exp\left(\int_s^t \mathrm{Lip}(v_\tau)\, d\tau\right).$$

Absolutely continuous curves generated by regular vector fields admit a nice representation formula given in following theorem (see [1, Proposition 4] and [2, Proposition 8.1.7]).

Theorem 2.5. *Let $\rho\colon [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve that starts at ϑ and whose velocity v satisfies (7). Then*

(*a*) $\rho_t = \Psi_{0,t\sharp}\vartheta$ *for all* $t \in [0,T]$,
(*b*) ρ *is a unique solution of* (5) *that satisfies* $\rho_0 = \vartheta$.

The following result shows that the boundedness of a regular vector field implies the Lipschitz continuity of the corresponding curve.

Lemma 2.6. *Let* Ψ *be the flow of a bounded vector field* w *satisfying* (7), *and* $\vartheta \in \mathcal{P}_2(\mathbb{R}^d)$. *Then* $\rho\colon [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ *defined by* $\rho_t = \Psi_{0,t\sharp}\vartheta$ *is* $\|w\|_\infty$*-Lipschitz.*

Proof. Take two time moments $s, t \in [a,b]$ such that $s < t$. Since $\Psi_{s,t}$ is a transport map between ρ_s and ρ_t, we have

$$W_2^2(\rho_t, \rho_s) \le \int |x - \Psi_{s,t}(x)|^2 \, d\rho_s(x).$$

On the other hand,

$$\Psi_{s,t}(x) - x = \int_s^t w_\tau\left(\Psi_{s,\tau}(x)\right) \, d\tau.$$

Hence the boundedness of w implies the Lipschitz continuity. ☐

General absolutely continuous curves admit another useful representation formula. To describe it, we first define, for every $t \in [0,T]$, the evaluation map $e_t\colon \mathbb{R}^d \times C\left([0,T];\mathbb{R}^d\right) \to \mathbb{R}^d$ by $e_t(x,\gamma) = \gamma(t)$.

Theorem 2.7 ([2, Chapter 8.2])**.** *Let* $\rho\colon [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ *be an absolutely continuous curve and* v *be its velocity field such that* (6) *holds. Then* $\rho_t = (e_t)_\sharp\eta$ *for a suitable Borel probability measure* η *on* $\mathbb{R}^d \times C\left([0,T];\mathbb{R}^d\right)$. *This measure is concentrated on the set* Γ *of pairs* (x,γ) *such that* γ *is an absolutely continuous solutions of the equation* $\dot{x}(t) = v_t\left(x(t)\right)$, *for a.e.* $t \in [0,T]$, *with* $\gamma(0) = x$.

2.4. **Projecting measures on sets.** Let $C \subset \mathbb{R}^d$ be a bounded r-prox-regular set. Consider all measures ρ supported in C:

$$\mathscr{C} = \left\{\rho \in \mathcal{P}_2(\mathbb{R}^d) : \ \operatorname{spt}\rho \subset C\right\}.$$

Lemma 2.8. *Let* $\vartheta \in \mathcal{P}_2(\mathbb{R}^d)$ *satisfy* $\operatorname{spt}\vartheta \subset C + r\boldsymbol{B}^\circ$. *Then*
(i) there exists $\vartheta_C \in \mathscr{C}$ *such that* $W_2(\vartheta, \vartheta_C) = \inf_{\rho\in\mathscr{C}} W_2(\vartheta, \rho)$;
(ii) ϑ_C *is unique and given by* $\vartheta_C = (P_C)_\sharp\vartheta$;
(iii) ϑ *and* ϑ_C *are connected with the unique geodesic* $\rho_t = \left((1-t)\mathbf{id} + tP_C\right)_\sharp\vartheta$.

Proof. First note that P_C is single-valued and continuous on $C + r\boldsymbol{B}^\circ$ by the definition of r-prox-regularity. Consider the transport plan Π given by

$$\Pi(A) = \vartheta\left(\{x \in C + r\boldsymbol{B}^\circ \ : \ (x, P_C(x)) \in A\}\right).$$

Its support belongs to graph P_C. Hence the cyclical monotonicity of graph P_C would imply the optimality of Π. Cyclical monotonicity can be expressed as follows:

$$\sum_i |x_i - P_C(x_i)|^2 \le \sum_i |x_i - P_C(x_{i+1})|^2 \quad \text{for any} \quad x_1, \dots, x_k \in C + r\boldsymbol{B}^\circ.$$

This property clearly holds because

$$|x_i - P_C(x_i)| = d_C(x_i) \le |x_i - P_C(x_{i+1})|.$$

Let $\vartheta_C' \in \mathscr{C}$ satisfy $W_2(\vartheta_C, \vartheta) = W_2(\vartheta_C', \vartheta)$ and Π' be the corresponding optimal transport plan. Since $\pi^1_\sharp\Pi' = \vartheta$, the previous identity can be rewritten as follows:

$$\int \left(|x-y|^2 - |x - P_C(x)|^2\right) \, d\pi'(x,y) = 0.$$

The integrand is nonnegative on $\operatorname{spt}\Pi'$ because $|x-y| \geq |x - P_C(x)|$ for all $x \in C + r\boldsymbol{B}^\circ$ and $y \in C$. Therefore, for Π'-a.e. (x,y), we have

$$|x-y|^2 = |x - P_C(x)|^2.$$

Since in the open r-neighborhood of C the projection P_C is unique, we conclude that $y = P_C(x)$, for Π'-a.e. (x,y). In other words, Π' is supported on $\operatorname{graph} P_C$. Now from $\pi^1_\sharp \Pi = \pi^1_\sharp \Pi' = \vartheta$ it follows that $\Pi = \Pi'$.

There are two consequences of this fact: 1) $\vartheta_C = \vartheta'_C$ (so we established uniqueness) and 2) the optimal plan Π between ϑ_C and ϑ is unique. Recall that any geodesic between ϑ and ϑ_C takes the form $\rho_t = \left((1-t)\pi^1 + t\pi^2\right)_\sharp \Pi$, where Π is an optimal plan between ϑ and ϑ_C. Thus, the geodesic is unique due to the uniqueness of the optimal plan. □

Definition 2.9. The measure ϑ_C defined in the previous lemma is called a **projection** of ϑ on C.

3. **Existence.**

In this section we shall prove the existence part of Theorem 1.2.

3.1. **Continuous approximation.**

We consider two *processes* on $\mathbb{R}^d$. The first one $\Phi^\vartheta_{s,t}$ is the flow of the vector field $2\mathscr{V}(\vartheta)$. The second one $\Psi^\tau_{s,t}$ is defined only for $x \in \boldsymbol{C}(s+\tau) + r\boldsymbol{B}^\circ$ and $t \in [s, s+\tau]$ by

$$\Psi^\tau_{s,t}(x) = \left[\left(1 - \frac{t-s}{\tau}\right)\mathbf{id} + \frac{t-s}{\tau} P_{\boldsymbol{C}(s+\tau)}\right](x).$$

Both processes generate maps in the space of measures:

$$\widetilde{\Phi}_{s,t}(\vartheta) = \left(\Phi^\vartheta_{s,t}\right)_\sharp \vartheta, \quad \widetilde{\Psi}^\tau_{s,t}(\vartheta) = \left(\Psi^\tau_{s,t}\right)_\sharp \vartheta.$$

We merge these maps to construct a curve $t \mapsto \rho^\tau_t$ in the following way:

$$\rho^\tau_0 = \vartheta, \quad \rho^\tau_t = \begin{cases} \widetilde{\Phi}_{2k\tau,t}\left(\rho^\tau_{2k\tau}\right), & t \in [2k\tau, (2k+1)\tau], \\ \widetilde{\Psi}^\tau_{(2k+1)\tau,t}\left(\rho^\tau_{(2k+1)\tau}\right) & t \in [(2k+1)\tau, (2k+2)\tau], \end{cases} \quad k \in \mathbb{N}. \tag{8}$$

Let us find a velocity of this curve. To that end, take

$$w^\tau_s(x) \doteq \frac{P_{\boldsymbol{C}(s+\tau)}(x) - x}{\tau}, \quad x \in \boldsymbol{C}(s+\tau) + r\boldsymbol{B}^\circ,$$

and note that

$$\Psi^\tau_{s,t}(x) = x + (t-s)w^\tau_s(x).$$

Hence the time dependent vector field $(t,x) \mapsto w^\tau_{s,t}(x)$ generating the map $(t,x) \mapsto \Psi^\tau_{s,t}(x)$ satisfies

$$\frac{d}{dt}\Psi^\tau_{s,t}(x) = w^\tau_{s,t}\left(\Psi^\tau_{s,t}(x)\right) = w^\tau_s(x).$$

Thus we conclude that ρ^τ satisfies the continuity equation with the vector field given by

$$v^\tau_t = \begin{cases} 2\mathscr{V}\left(\rho^\tau_{2k\tau}\right), & t \in [2k\tau, (2k+1)\tau), \\ w^\tau_{(2k+1)\tau} \circ \left[\Psi^\tau_{(2k+1)\tau,t}\right]^{-1} & t \in [(2k+1)\tau, (2k+2)\tau], \end{cases} \quad k \in \mathbb{N}. \tag{9}$$

3.2. **Properties of the continuous approximation.**

Lemma 3.1. *For all sufficiently small τ, the curve $\rho^\tau\colon [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ is well-defined and Lipschitz with constant $2(L+M)$. Moreover, $|v^\tau_t(x)| \le 2(L+M)$ for all t and ρ^τ_t-a.e. x.*

Proof. **1.** Assume that ρ^τ is well-defined up to a time moment $s = 2k\tau$ (we can always choose $k=0$). Then $\operatorname{spt}\rho^\tau_s \subset \boldsymbol{C}(s)$ because the image of $\Psi^\tau_{s-\tau,s}$ belongs to $\boldsymbol{C}(s)$. In order to construct ρ^τ on the interval $[s, s+2\tau]$, we must show that $\operatorname{spt}\rho^\tau_{s+\tau}$ lies in the domain of $\Psi^\tau_{s+\tau,t}$ for all sufficiently small τ. Indeed, by Lipschitz continuity of $\boldsymbol{C}$,

$$\boldsymbol{C}(s) \subset \boldsymbol{C}\left(s+2\tau\right) + 2\tau M\boldsymbol{B}.$$

This inclusion together with $\operatorname{spt}\rho^\tau_s \subset \boldsymbol{C}(s)$ implies

$$\operatorname{spt}\rho^\tau_{s+\tau} \subset \Phi_{s,s+\tau}\left(\boldsymbol{C}(s)\right) \subset \boldsymbol{C}(s) + 2\tau L\boldsymbol{B} \subset \boldsymbol{C}\left(s+2\tau\right) + 2\tau(L+M)\boldsymbol{B}. \tag{10}$$

Thus we conclude that $\operatorname{spt}\rho^\tau_{s+\tau} \subset \boldsymbol{C}\left(s+2\tau\right) + r\boldsymbol{B}^\circ$, whenever τ is sufficiently small. This proves that ρ^τ is well-defined on $[0,T]$.

2. Let us estimate v^τ. For each $t \in [s, s+\tau)$ (here again $s = 2k\tau$), we have

$$|v^\tau_t(x)| = |2\mathscr{V}(\rho^\tau_s)(x)| \le 2L \quad \forall x \in \mathbb{R}^d.$$

Now suppose that $t \in [s+\tau, s+2\tau]$ and let

$$\alpha^* \doteq \inf\left\{\alpha :\ \operatorname{spt}\rho^\tau_{s+\tau} \subset \boldsymbol{C}(s+2\tau) + \alpha\boldsymbol{B}\right\}.$$

Then, by construction,

$$|v^\tau_t(x)| \le \frac{\alpha^*}{\tau} \quad \forall x \in \boldsymbol{C}(s+2\tau) + \alpha^*\boldsymbol{B}.$$

It follows from (10) that $\alpha^* \le 2\tau(L+M)$. Thus, for each $t \in [s+\tau, s+2\tau]$, we have

$$|v^\tau_t(x)| \le 2(L+M) \quad \forall x \in \boldsymbol{C}(s+2\tau) + \alpha^*\boldsymbol{B}.$$

Since $\operatorname{spt}\rho^\tau_t \subset \boldsymbol{C}(s+2\tau) + \alpha^*\boldsymbol{B}$, for any $t \in [s+\tau, s+2\tau]$, we get the desired estimate on v^τ. The lipschitzeanity of ρ^τ now follows from Lemma 2.6. ☐

3.3. **Piecewise constant approximation.** Let us introduce the map

$$R^\tau(t) = k\tau, \quad t \in [k\tau, (k+1)\tau),$$

which can be roughly thought as a "projection" of t on the mesh $\{k\tau\}_{k=0}^\infty$.

Taking the curves $t \mapsto \rho^\tau_t$ and $t \mapsto v^\tau_t$, we construct two piecewise constant curves $t \mapsto \bar\rho^\tau_t$ and $t \mapsto \bar v^\tau_t$ in the following way:

$$\bar\rho^\tau_t = \rho^\tau_{R^\tau(t)}, \quad \bar v^\tau_t = v^\tau_{R^\tau(t)}, \quad t \in [0,T].$$

The next lemmas establish the relationship between the continuous and the piecewise constant approximations.

Lemma 3.2. *Consider two families of vector measures $E^\tau = v^\tau\rho^\tau$ and $\bar E^\tau_t = E^\tau_{R^\tau(t)}$, where ρ^τ and v^τ are defined by (8) and (9). Assume that $E^\tau \rightharpoonup E$, $\bar E^\tau \rightharpoonup \bar E$, $\rho^\tau \rightharpoonup \rho$, $\bar\rho^\tau \rightharpoonup \bar\rho$. Then*

(1) *$E = \bar E$ and $\rho = \bar\rho$;*
(2) *$E = v\rho$ for some v.*

Proof. **1.** Since $C_c(\mathbb{R}^{d+1})$ is dense in $C_0(\mathbb{R}^{d+1})$, two Radon measures E and $\bar{E}$ coincide if $\int \varphi \cdot dE = \int \varphi \cdot d\bar{E}$ for any $\varphi \in C_c^\infty(\mathbb{R}^{d+1})$. Fix $\varphi \in C_c^\infty(\mathbb{R}^{d+1})$ and note that

$$\int \varphi \cdot dE^\tau - \int \varphi \cdot d\bar{E}^\tau = \int_0^T \Big(\int \varphi \cdot v_t^\tau \, d\rho_t^\tau - \int \varphi \cdot \bar{v}_t^\tau \, d\bar{\rho}_t^\tau \Big) \, dt.$$

If $t \in [(2k+1)\tau, (2k+2)\tau)$, the definitions of ρ^τ and v^τ allow us to write

$$\begin{aligned} \int \varphi \cdot v_t^\tau \, d\rho_t^\tau &= \int \varphi \cdot w_{(2k+1)\tau} \circ \Big[\Psi^\tau_{(2k+1)\tau,t}\Big]^{-1} \, d\Big[\Psi^\tau_{(2k+1)\tau,t}\Big]_\sharp \rho^\tau_{(2k+1)\tau} \\ &= \int \varphi \circ \Psi^\tau_{(2k+1)\tau,t} \cdot w_{(2k+1)\tau} \, d\rho^\tau_{(2k+1)\tau}, \\ \int \varphi \cdot \bar{v}_t^\tau \, d\bar{\rho}_t^\tau &= \int \varphi \cdot w_{(2k+1)\tau} \, d\rho^\tau_{(2k+1)\tau}. \end{aligned}$$

Hence, by Lemma 3.1,

$$\begin{aligned} \Big| \int \varphi \cdot v_t^\tau \, d\rho_t^\tau - \int \varphi \cdot \bar{v}_t^\tau \, d\bar{\rho}_t^\tau \Big| &\le \int \Big| \varphi \circ \Psi^\tau_{(2k+1)\tau,t} - \varphi \Big| \cdot \big| w_{(2k+1)\tau} \big| \, d\rho^\tau_{(2k+1)\tau} \\ &\le 2\mathrm{Lip}(\varphi)(L+M) \int \Big| \Psi^\tau_{(2k+1)\tau,t}(x) - x \Big| \, d\rho^\tau_{(2k+1)\tau}(x), \end{aligned}$$

Similarly, for $t \in [2k\tau, (2k+1)\tau)$, we have

$$\begin{aligned} \int \varphi \cdot v_t^\tau \, d\rho_t^\tau &= 2 \int \varphi \cdot \mathscr{V}(\rho_{2k\tau}) \, d\big[\Phi^\tau_{2k\tau,t}\big]_\sharp \rho^\tau_{2k\tau} \\ &= 2 \int \varphi \circ \Phi^\tau_{2k\tau,t} \cdot \mathscr{V}(\rho_{2k\tau}) \circ \Phi^\tau_{2k\tau,t} \, d\rho^\tau_{2k\tau}, \\ \int \varphi \cdot \bar{v}_t^\tau \, d\bar{\rho}_t^\tau &= 2 \int \varphi \cdot \mathscr{V}(\rho_{2k\tau}) \, d\rho^\tau_{2k\tau}. \end{aligned}$$

Again by Lemma 3.1,

$$\begin{aligned} \Big| \int \varphi \cdot v_t^\tau \, d\rho_t^\tau - \int \varphi \cdot \bar{v}_t^\tau \, d\bar{\rho}_t^\tau \Big| &\le 2 \Big| \int \varphi \circ \Phi^\tau_{2k\tau,t} \cdot \big(\mathscr{V}(\rho_{2k\tau}) \circ \Phi^\tau_{2k\tau,t} - \mathscr{V}(\rho_{2k\tau}) \big) \, d\rho^\tau_{2k\tau} \Big| \\ &\quad + 2 \Big| \int \big(\varphi \circ \Phi^\tau_{2k\tau,t} - \varphi \big) \cdot \mathscr{V}(\rho_{2k\tau}) \, d\rho^\tau_{2k\tau} \Big| \\ &\le 2\|\varphi\|_\infty \int |\mathscr{V}(\rho^\tau_{2k\tau}) \circ \Phi_{2k\tau,t} - \mathscr{V}(\rho^\tau_{2k\tau})| \, d\rho^\tau_{2k\tau} \\ &\quad + 2L \int |\varphi \circ \Phi_{2k\tau,t} - \varphi| \, d\rho^\tau_{2k\tau} \\ &\le 2L \left(\|\varphi\|_\infty + \mathrm{Lip}(\varphi) \right) \int |\Phi_{2k\tau,t}(x) - x| \, d\rho^\tau_{2k\tau}(x). \end{aligned}$$

Recalling that

$$\begin{aligned} |\Phi_{2k\tau,t}(x) - x| &\le 2\tau L, \quad t \in [2k\tau, (2k+1)\tau), \\ \big|\Psi_{(2k+1)\tau,t}(x) - x\big| &\le 2\tau(L+M), \quad t \in [(2k+1)\tau, (2k+2)\tau), \end{aligned}$$

we conclude that

$$\Big| \int \varphi \cdot v_t^\tau \, d\rho_t^\tau - \int \varphi \cdot \bar{v}_t^\tau \, d\bar{\rho}_t^\tau \Big| \le C\tau \quad \forall t \in [0,T],$$

for some $C > 0$ that does not depend on τ. As a consequence,

$$\left| \int \varphi \cdot dE^\tau - \int \varphi \cdot d\bar{E}^\tau \right| \leq TC\tau,$$

which implies $E = \bar{E}$. By the same arguments, one can show that $\rho = \bar{\rho}$.

2. We derive the last assertion from the properties of the Benamou-Brenier functional $\mathcal{B}_2$ (see Appendix A). Since $E^\tau = v^\tau \rho^\tau$, Proposition 3(i) and Lemma 3.1 yield

$$\mathcal{B}_2(\rho^\tau, E^\tau) = \frac{1}{2} \int |v^\tau|^2 \, d\rho^\tau \leq C,$$

for some $C > 0$ which does not depend on τ. Hence the lower semicontinuity of $\mathcal{B}_2$ implies

$$\mathcal{B}_2(\rho, E) \leq \liminf_{\tau \downarrow 0} \mathcal{B}_2(\rho^\tau, E^\tau) < +\infty,$$

which means, by Proposition 3(iv), that there exists v such that $E = v\rho$. □

Lemma 3.3. *Let $u_t^\tau = \mathscr{V}(\rho_t^\tau)$ and $\bar{u}_t^\tau = \mathscr{V}(\bar{\rho}_t^\tau)$. If $\|\rho^\tau - \rho\|_\infty \to 0$ then $\|\bar{\rho}^\tau - \rho\|_\infty \to 0$ and the sequences of vector measures $u^\tau \rho^\tau$ and $\bar{u}^\tau \bar{\rho}^\tau$ converge to $u\rho$ with $u_t = \mathscr{V}(\rho_t)$.*

Proof. According to Lemma 3.1, one has

$$\begin{aligned} \|\bar{\rho}^\tau - \rho\|_\infty &\leq \|\bar{\rho}^\tau - \rho^\tau\|_\infty + \|\rho^\tau - \rho\|_\infty \\ &= \sup_{t \in [0,T]} W_2(\rho^\tau_{R^\tau(t)}, \rho_t^\tau) + \|\rho^\tau - \rho\|_\infty \\ &\leq 2(L + M)\tau + \|\rho^\tau - \rho\|_\infty. \end{aligned}$$

This proves the first assertion.

Let us show that the limit of $u^\tau \rho^\tau$ is $u\rho$. Indeed, for any $\varphi \in C_b(\mathbb{R}^{d+1})$, one has

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} \varphi(t,x) \cdot \mathscr{V}(\rho_t^\tau)(x) \, d\rho_t^\tau(x) \, dt - \int_0^T \int_{\mathbb{R}^d} \varphi(t,x) \cdot \mathscr{V}(\rho_t)(x) \, d\rho_t(x) \, dt \\ = \int_0^T \int_{\mathbb{R}^d} \varphi(t,x) \cdot \left[\mathscr{V}(\rho_t^\tau)(x) - \mathscr{V}(\rho_t)(x) \right] \, d\rho_t^\tau(x) \, dt \\ + \int_0^T \int_{\mathbb{R}^d} \varphi(t,x) \cdot \mathscr{V}(\rho_t)(x) \left[d\rho_t^\tau(x) - d\rho_t(x) \right] \, dt. \end{aligned}$$

The absolute value of the first integral from the right-hand side is estimated by

$$L \|\varphi\|_\infty \sup_{t \in [0,T]} W_2(\rho_t^\tau, \rho_t),$$

and thus tends to 0. The second integral converges to 0, because $(t,x) \mapsto u_t(x)$ is continuous and bounded.

It remains to check that $\bar{u}^\tau \bar{\rho}^\tau \rightharpoonup u\rho$. For any $\varphi \in C_b(\mathbb{R}^{d+1})$, one has

$$\begin{aligned} \int \varphi \cdot \bar{u}^\tau \, d\bar{\rho}^\tau - \int \varphi \cdot u \, d\rho = \int \varphi \cdot u \left(d\bar{\rho}^\tau - d\rho \right) + \int \varphi \cdot \left(\bar{u}^\tau - u^\tau \right) \, d\bar{\rho}^\tau \\ + \int \varphi \cdot (u^\tau - u) \, d\bar{\rho}^\tau. \end{aligned}$$

The first integral from the right-hand side, which can be rewritten as

$$\int_0^T \int_{\mathbb{R}^d} \varphi(t,x) \cdot u_t(x) \left(d\bar{\rho}_t^\tau - d\rho_t^\tau \right) \, dt,$$

converges to 0 since $\bar{\rho}_t^\tau \rightharpoonup \rho_t$ for a.e. $t \in [0,T]$. The absolute values of the last two integrals can be estimated by

$$L\|\varphi\|_\infty \sup_{t\in[0,T]} W_2(\bar{\rho}_t^\tau, \rho_t) \quad \text{and} \quad L\|\varphi\|_\infty \sup_{t\in[0,T]} W_2(\rho_t^\tau, \rho_t),$$

respectively. This proves the last assertion. □

3.4. **Normal cone inclusion.** The curves ρ^τ and the vector fields v^τ are constructed so that

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

The sequence ρ^τ, being uniformly Lipschitz by Lemma 3.1, converges (up to a subsequence) to a Lipschitz map ρ in $C\left([0,T]; \mathcal{P}_2(\mathbb{R}^d)\right)$. Lemmas 2.3 and 3.2 yield that $\rho^\tau \rightharpoonup \rho$ and $v^\tau\rho^\tau \rightharpoonup v\rho$ for some Borel map v. In particular, it follows that

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

The inclusion $\operatorname{spt} \rho_t \subset \boldsymbol{C}(t)$ is a direct consequence of the following lemma.

Lemma 3.4. *Let $\mu_k \in \mathcal{P}_2(\mathbb{R}^d)$ and $\operatorname{spt} \mu_k \subset A + r_k \boldsymbol{B}$, where A is compact. If $\mu_k \rightharpoonup \mu$ and $r_k \to 0$ then $\operatorname{spt} \mu \subset A$.*

Proof. The set $U_n = (A + r_n \boldsymbol{B})^c$ is open. Hence, for each n, we have

$$0 = \liminf_{k\to\infty} \mu_k(U_n) \geq \mu(U_n).$$

This means that $\mu\left(\bigcup_n U_n\right) = 0$. It remains to note that

$$\bigcup_n U_n = \left(\bigcap_n (A + r_n \boldsymbol{B})\right)^c = A^c,$$

completing the proof. □

To prove that ρ is a solution of (1), it remains to establish the following Proposition 2, whose proof heavily relies on the properties of the piecewise constant approximation.

Proposition 2. *For a.e. $t \in [0,T]$, there exists a set $A_t \subset \mathbb{R}^d$ such that*

$$v_t(x) - \mathscr{V}(\rho_t)(x) \in -N_{\boldsymbol{C}(t)}(x) \quad \forall x \in A_t \tag{11}$$

and $\rho_t(A_t) = 1$.

Consider a set-valued map $t \mapsto \bar{\boldsymbol{C}}^\tau(t)$ defined by

$$\bar{\boldsymbol{C}}^\tau(t) = \boldsymbol{C}\left((2k+2)\tau\right), \quad t \in (2k\tau, (2k+2)\tau].$$

Given a measurable selection $y(t) \in \boldsymbol{C}(t)$, we define in the same way

$$\bar{y}^\tau(t) = y\left((2k+2)\tau\right) = y_{(2k+2)\tau}, \quad t \in (2k\tau, (2k+2)\tau].$$

Let us introduce the integral

$$J^\tau = \int_0^T \int_{\mathbb{R}^d} a(t)b(x)\Big(\left\langle \bar{v}_t^\tau(x) - \mathscr{V}\left(\bar{\rho}_t^\tau\right)(x), P_{\bar{\boldsymbol{C}}^\tau(t)}(x) - \bar{y}^\tau(t) \right\rangle - \frac{1}{2r}\left|\bar{v}_t^\tau(x) - \mathscr{V}\left(\bar{\rho}_t^\tau\right)(x)\right| \cdot \left|P_{\bar{\boldsymbol{C}}^\tau(t)}(x) - \bar{y}^\tau(t)\right|^2\Big)\, d\bar{\rho}_t^\tau(x)\, dt, \tag{12}$$

where a and b are nonnegative bounded Lipschitz functions. Our aim is to pass to the limit in the integral as $\tau \to 0$. Without loss of generality, we can consider only those τ that satisfy $(2N+2)\tau = T$ for some $N \in \mathbb{N}$.

Recall that

$$\bar{v}_t^\tau - \mathscr{V}(\bar{\rho}_t^\tau) = \begin{cases} \mathscr{V}(\rho_{2k\tau}^\tau), & t \in [2k\tau, (2k+1)\tau], \\ w_{(2k+1)\tau}^\tau - \mathscr{V}\left(\rho_{(2k+1)\tau}^\tau\right) & t \in [(2k+1)\tau, (2k+2)\tau]. \end{cases}$$

Hence J^τ can be rewritten as

$$\sum_{k=0}^N \int_{2k\tau}^{(2k+1)\tau} a(t)\,dt \int b(x)\Big(\big\langle \mathscr{V}\left(\rho_{2k\tau}^\tau\right)(x), P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\big\rangle$$
$$- \frac{1}{2r}\big|\mathscr{V}\left(\rho_{2k\tau}^\tau\right)(x)\big| \cdot \big|P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\big|^2\Big)\, d\rho_{2k\tau}^\tau(x)$$
$$+ \sum_{k=0}^N \int_{2k\tau}^{(2k+1)\tau} a(t+\tau)\,dt \int b(x)\Big(\big\langle w_{(2k+1)\tau}^\tau(x) - \mathscr{V}\left(\rho_{(2k+1)\tau}^\tau\right)(x),$$
$$P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\big\rangle$$
$$- \frac{1}{2r}\big|w_{(2k+1)\tau}^\tau(x) - \mathscr{V}\left(\rho_{(2k+1)\tau}^\tau\right)(x)\big| \cdot \big|P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\big|^2\Big)\, d\rho_{(2k+1)\tau}^\tau(x).$$

Since

$$\big|w_{(2k+1)\tau}^\tau(x) - \mathscr{V}\left(\rho_{(2k+1)\tau}^\tau\right)(x)\big| \geq \big|w_{(2k+1)\tau}^\tau(x)\big| - \big|\mathscr{V}\left(\rho_{(2k+1)\tau}^\tau\right)(x)\big|,$$

we conclude that $J^\tau \leq J_1^\tau + J_2^\tau + J_3^\tau$, where

$$J_1^\tau = \sum_{k=0}^N \bigg(\int_{2k\tau}^{(2k+1)\tau} a(t)\,dt \int b(x)\big\langle \mathscr{V}\left(\rho_{2k\tau}^\tau\right)(x), P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\big\rangle\, d\rho_{2k\tau}^\tau(x)$$
$$- \int_{2k\tau}^{(2k+1)\tau} a(t+\tau)\,dt \int b(x)\big\langle \mathscr{V}\left(\rho_{(2k+1)\tau}^\tau\right)(x),$$
$$P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\big\rangle\, d\rho_{(2k+1)\tau}^\tau(x)\bigg),$$

$$J_2^\tau = -\frac{1}{2r}\sum_{k=0}^N \bigg(\int_{2k\tau}^{(2k+1)\tau} a(t)\,dt \int b(x)\big|\mathscr{V}\left(\rho_{2k\tau}^\tau\right)(x)\big|$$
$$\cdot \big|P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\big|^2\, d\rho_{2k\tau}^\tau(x)$$
$$- \int_{2k\tau}^{(2k+1)\tau} a(t+\tau)\,dt \int b(x)\big|\mathscr{V}\left(\rho_{(2k+1)\tau}^\tau\right)(x)\big|$$
$$\cdot \big|P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\big|^2\, d\rho_{(2k+1)\tau}^\tau(x)\bigg),$$

$$J_3^\tau = \sum_{k=0}^N \int_{2k\tau}^{(2k+1)\tau} a(t+\tau)\,dt \int b(x)\Big(\big\langle w_{(2k+1)\tau}^\tau(x), P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\big\rangle$$
$$- \frac{1}{2r}\big|w_{(2k+1)\tau}^\tau(x)\big| \cdot \big|P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\big|^2\Big)\, d\rho_{(2k+1)\tau}^\tau(x).$$

We are going to show that $J_1^\tau = O(\tau)$ and $J_2^\tau = O(\tau)$.

Lemma 3.5. *Given $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^d)$ and a Lipschitz function $b \in C(\mathbb{R}^d)$, suppose that*

(a) *there exists $K \in \mathcal{K}_r(\mathbb{R}^d)$ such that* $\operatorname{spt}\mu_1 \cup \operatorname{spt}\mu_2 \subset K + \frac{r}{2}\boldsymbol{B}$*;*

(b) *there exist a Borel measurable map* $\psi\colon \mathbb{R}^d \to \mathbb{R}^d$ *and* $C > 0$ *such that* $\mu_2 = \psi_\sharp \mu_1$ *and* $|x - \psi(x)| \le C\tau$ *for all* x.

Then there exists $C_1 > 0$ *such that*

$$\Big| \int b(x) \langle \mathscr{V}(\mu_1)(x), P_K(x) - y \rangle \, d\mu_1(x) \\ - \int b(x) \langle \mathscr{V}(\mu_2)(x), P_K(x) - y \rangle \, d\mu_2(x) \Big| \le C_1 \tau,$$

$$\Big| \int b(x) \left| \mathscr{V}(\mu_1)(x) \right| \cdot \left| P_K(x) - y \right|^2 \, d\mu_1(x) \\ - \int b(x) \left| \mathscr{V}(\mu_2)(x) \right| \cdot \left| P_K(x) - y \right|^2 \, d\mu_2(x) \Big| \le C_1 \tau,$$

for all $y \in K + \frac{r}{2}\boldsymbol{B}$.

Proof. **1.** We start with the first inequality. Note that

$$\int b(x) \langle \mathscr{V}(\mu_2)(x), P_K(x) - y \rangle \, d\mu_2(x) \\ = \int b \circ \psi(x) \langle \mathscr{V}(\mu_2) \circ \psi(x), P_K \circ \psi(x) - y \rangle \, d\mu_1(x).$$

Now taking $x_1 = x$ and $x_2 = \psi(x)$, we may write

$$\begin{aligned} &b(x_1) \langle \mathscr{V}(\mu_1)(x_1), P_K(x_1) - y \rangle - b(x_2) \langle \mathscr{V}(\mu_2)(x_2), P_K(x_2) - y \rangle \\ &\quad = (b(x_1) - b(x_2)) \cdot \langle \mathscr{V}(\mu_1)(x_1), P_K(x_1) - y \rangle \\ &\quad + b(x_2) \left(\langle \mathscr{V}(\mu_1)(x_1), P_K(x_1) - y \rangle - \langle \mathscr{V}(\mu_2)(x_2), P_K(x_2) - y \rangle \right). \end{aligned}$$

The first term from the right-hand side can be estimated by

$$\mathrm{Lip}(b) L |x_1 - x_2| \cdot |P_K(x_1) - y|.$$

To deal with the second term, consider the difference

$$\begin{aligned} \langle \mathscr{V}(\mu_1)(x_1), P_K(x_1) - y \rangle &- \langle \mathscr{V}(\mu_2)(x_2), P_K(x_2) - y \rangle \\ &= \langle \mathscr{V}(\mu_1)(x_1) - \mathscr{V}(\mu_2)(x_1), P_K(x_1) - y \rangle \\ &+ \langle \mathscr{V}(\mu_2)(x_1) - \mathscr{V}(\mu_2)(x_2), P_K(x_1) - y \rangle \\ &+ \langle \mathscr{V}(\mu_2)(x_2), P_K(x_1) - P_K(x_2) \rangle . \end{aligned}$$

Our assumptions imply that the first term from the right-hand side can be estimated by $L W_2(\mu_1, \mu_2) |P_K(x_1) - y|$, the second term by $L|x_1 - x_2| \cdot |P_K(x_1) - y|$, and the third term by $L \left| P_K(x_1) - P_K(x_2) \right|$.

Since b is bounded on $K + \frac{r}{2}\boldsymbol{B}$ and, for all $x_1 \in \operatorname{spt} \mu_1$, we have

$$\begin{aligned} &|x_1 - x_2| = |x - \psi(x)| \le C\tau, \\ &W_2(\mu_1, \mu_2) \le \Big(\int |x - \psi(x)|^2 \, d\mu_1(x) \Big)^{1/2} \le C\tau, \\ &|P_K(x_1) - P_K(x_2)| \le 2|x_1 - x_2| \le 2C\tau, \quad \text{(by Proposition 1)} \qquad (13) \\ &|P_K(x_1) - y| \le \operatorname{diam} K + r. \end{aligned}$$

After combining all the estimates, we get the desired inequality.

2. The second inequality can be proven in the similar way. We begin with the identity

$$\int b(x)\,|\mathscr{V}(\mu_2)(x)|\cdot|P_K(x)-y|^2\,d\mu_2(x) = \int b\circ\psi(x)\,|\mathscr{V}(\mu_2)\circ\psi(x)|\cdot|P_K\circ\psi(x)-y|^2\,d\mu_1(x).$$

Again, by taking $x_1 = x$ and $x_2 = \psi(x)$, we get

$$\begin{aligned} &b(x_1)\,|\mathscr{V}(\mu_1)(x_1)|\cdot|P_K(x_1)-y|^2 - b(x_2)\,|\mathscr{V}(\mu_2)(x_2)|\cdot|P_K(x_2)-y|^2 \\ &\quad= (b(x_1)-b(x_2))\cdot|\mathscr{V}(\mu_1)(x_1)|\cdot|P_K(x_1)-y|^2 \\ &\quad+ b(x_2)\left(|\mathscr{V}(\mu_1)(x_1)|\cdot|P_K(x_1)-y|^2 - |\mathscr{V}(\mu_2)(x_2)|\cdot|P_K(x_2)-y|^2\right). \end{aligned}$$

The first term from the right-hand side is estimated by $\mathrm{Lip}(b)CL\,(\mathrm{diam}\,K+r)^2\,\tau$. As for the second term, we have

$$\begin{aligned} &|\mathscr{V}(\mu_1)(x_1)|\cdot|P_K(x_1)-y|^2 - |\mathscr{V}(\mu_2)(x_2)|\cdot|P_K(x_2)-y|^2 \\ &\quad= \big(\,|\mathscr{V}(\mu_1)(x_1)| - |\mathscr{V}(\mu_2)(x_2)|\,\big)\cdot|P_K(x_1)-y|^2 \\ &\quad+ |\mathscr{V}(\mu_2)(x_2)|\cdot\big(\,|P_K(x_1)-y|^2 - |P_K(x_2)-y|^2\,\big) \end{aligned} \tag{14}$$

We can easily estimate the first term in (14) because

$$\begin{aligned} \big|\,|\mathscr{V}(\mu_1)(x_1)| - |\mathscr{V}(\mu_2)(x_2)|\,\big| &\le |\mathscr{V}(\mu_1)(x_1) - \mathscr{V}(\mu_2)(x_2)| \\ &\le L|x_1-x_2| + LW_2(\mu_1,\mu_2) \le 2LC\tau. \end{aligned}$$

Thanks to (13) and the identity

$$|P_K(x_1)-y|^2 - |P_K(x_2)-y|^2 = |P_K(x_1)-P_K(x_2)|\cdot|P_K(x_1)+P_K(x_2)-2y|\,,$$

we can estimate the second term in (14) by $4LC\tau\,(\mathrm{diam}\,K+r)$. Combining all the estimates above, we obtain the desired inequality. □

Lemma 3.6. *Let $a\colon[0,T]\to\mathbb{R}$ be Lipschitz, $s\in[0,T-\tau]$, and $\alpha,\beta\in\mathbb{R}$. Then*

$$\left|\alpha\int_s^{s+\tau} a(t)\,dt - \beta\int_s^{s+\tau} a(t+\tau)\,dt\right| \le |\alpha|\mathrm{Lip}(a)\tau^2 + |\alpha-\beta|\cdot\|a\|_\infty\tau.$$

Proof. After rearranging the left-hand side can be written as follows:

$$\left|\alpha\int_s^{s+\tau}[a(t)-a(t+\tau)]\,dt + (\alpha-\beta)\int_s^{s+\tau} a(t+\tau)\,dt\right|.$$

Now the required estimate easily follows from the Lipschitz continuity of a. □

Lemma 3.7. *One has $J_1^\tau = O(\tau)$, $J_2^\tau = O(\tau)$.*

Proof. **1.** We begin with J_1^τ. Let us take

$$\begin{aligned} \alpha &= \int b(x)\left\langle \mathscr{V}\left(\rho_{2k\tau}^\tau\right)(x), P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\right\rangle d\rho_{2k\tau}^\tau(x), \\ \beta &= \int b(x)\left\langle \mathscr{V}\left(\rho_{(2k+1)\tau}^\tau\right)(x), P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\right\rangle d\rho_{(2k+1)\tau}^\tau(x). \end{aligned}$$

We know that $\mathrm{spt}\,\rho_{(2k+2)\tau} \subset \boldsymbol{C}\left((2k+2)\tau\right)$. Hence if τ is small enough then

$$\mathrm{spt}\,\rho_{2k\tau}\cup\mathrm{spt}\,\rho_{(2k+1)\tau} \subset \boldsymbol{C}\left((2k+2)\tau\right) + \frac{r}{2}\boldsymbol{B}.$$

Lemma 3.5 implies that $|\alpha - \beta| \le C_1\tau$ for some $C_1 > 0$. Now from Lemma 3.6 it follows that

$$\alpha \int_{2k\tau}^{(2k+1)\tau} a(t)\,dt - \beta \int_{2k\tau}^{(2k+1)\tau} a(t+\tau)\,dt = O(\tau^2). \tag{15}$$

This gives $J_1^\tau = (N+1)O(\tau^2) = \frac{T}{2\tau}O(\tau^2) = O(\tau)$.

2. To deal with J_2^τ we take

$$\alpha = \int b(x)\left|\mathscr{V}\left(\rho_{2k\tau}^\tau\right)(x)\right| \cdot \left|P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\right|^2 d\rho_{2k\tau}^\tau(x),$$

$$\beta = \int b(x)\left|\mathscr{V}\left(\rho_{(2k+1)\tau}^\tau\right)(x)\right| \cdot \left|P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\right|^2 d\rho_{(2k+1)\tau}^\tau(x).$$

Then Lemma 3.6 gives (15) and, as a consequence, $J_2^\tau = O(\tau)$. □

Since $-w_{(2k+1)\tau}^\tau(x)$ is a proximal normal to $\boldsymbol{C}\left((2k+2)\tau\right)$ at $P_{\boldsymbol{C}(2k+2)\tau}(x)$, we conclude that

$$\left\langle w_{(2k+1)\tau}^\tau(x), P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\right\rangle - \frac{1}{2r}\left|w_{(2k+1)\tau}^\tau(x)\right| \cdot \left|P_{\boldsymbol{C}((2k+2)\tau)}(x) - y_{(2k+2)\tau}\right|^2 \le 0,$$

for all $x \in \boldsymbol{C}\left((2k+2)\tau\right) + r\boldsymbol{B}^\circ$. This means that $J_3^\tau \le 0$. So we have

$$J^\tau + O(\tau) \le 0. \tag{16}$$

Lemma 3.8. *Let $y(\cdot)$ be a Lipschitz continuous selection of $\boldsymbol{C}(\cdot)$ and $a \in C(\mathbb{R})$, $b \in C(\mathbb{R}^d)$ be nonnegative bounded Lipschitz functions. Then*

$$\int_0^T \int_{\mathbb{R}^d} a(t)b(x)\Big(\langle v_t(x) - \mathscr{V}(\rho_t)(x), x - y(t)\rangle - \sigma(t,x)\cdot\left|x - y(t)\right|^2\Big)\, d\rho_t(x)\,dt \le 0,$$

for some nonnegative Borel map $\sigma\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$.

Proof. **1.** We shall prove the lemma by passing to the limit in (16) as $\tau \to 0$. But first, let us show that $\bar{E}^\tau \doteq \bar{\sigma}^\tau\bar{\rho}^\tau$ with

$$\bar{\sigma}^\tau(t,x) = \frac{1}{2r}\left|\bar{v}_t^\tau(x) - \mathscr{V}\left(\bar{\rho}_t^\tau\right)(x)\right|$$

tends to $\sigma\rho$ for some Borel map σ. Since all $\bar{E}^\tau$ are supported on the compact set

$$\mathcal{C}_r = \left\{(t,x):\ x \in \boldsymbol{C}(t) + \frac{r}{2}\boldsymbol{B},\ t \in [0,T]\right\}$$

and, by Lemma 3.1, their total variations are uniformly bounded:

$$\|\bar{E}^\tau\| \doteq \int \bar{\sigma}^\tau\, d\bar{\rho}^\tau \le \frac{1}{2r}\int_0^T (3L + 2M)\,dt,$$

we conclude that $\bar{E}^\tau$ weakly converges (up to a subsequence) to some nonnegative measure E. As in Lemma 3.2, the corresponding Benamou-Brenier functional is uniformly bounded:

$$\mathcal{B}_2(\bar{\rho}^\tau, \bar{E}^\tau) = \frac{1}{2}\int |\bar{\sigma}^\tau|^2\, d\bar{\rho}^\tau \le \int_0^T (3L+2M)^2\, dt.$$

Hence the lower semicontinuity of $\mathcal{B}_2$ (Proposition 3) implies that $\mathcal{B}_2(\rho, E) < +\infty$, and therefore $E = \sigma\rho$, for a Borel map σ.

2. Let us show that we get the desired limit if we replace $P_{\bar{\boldsymbol{C}}^\tau(t)}(x) - \bar{y}^\tau(t)$ with $f(t,x) = P_{\boldsymbol{C}(t)}(x) - y(t)$. Indeed, if τ is small then $\operatorname{spt}\bar{\rho}^\tau \subset \mathcal{C}_r$. The function f

is continuous inside $\mathcal{C}_r$ thanks to Lemma B.1. Recalling Lemmas 3.2 and 3.3, we obtain

$$\int_0^T \int_{\mathbb{R}^d} a(t)b(x) \left(\langle v_t(x) - \mathscr{V}(\rho_t)(x), f(t,x) \rangle - \sigma(t,x) f^2(t,x) \right) d\rho_t(x)\, dt$$

in the limit. Since $\operatorname{spt} \rho_t \subset \boldsymbol{C}(t)$, we have $P_{\boldsymbol{C}(t)}(x) = x$ for all $x \in \operatorname{spt} \rho_t$. This gives the desired inequality.

3. The function f, being defined on a compact set, is uniformly continuous. In particular, for any $\varepsilon > 0$ there exists δ such that for all $\tau < \delta$

$$\left| f(t,x) - f(R^{2\tau}(t), x) \right| \le \varepsilon \quad \forall (t,x) \in \mathcal{C}_r.$$

Hence letting $\bar{u}_t^\tau = \bar{v}_t^\tau - \mathscr{V}(\bar{\rho}_t^\tau)$ we get

$$\left| \iint a(t)b(x) \left[f(R^{2\tau}(t), x) - f(t,x) \right] \cdot d(\bar{u}^\tau \bar{\rho}^\tau)(t,x) \right| \to 0.$$

Similarly using uniform continuity of $|f|^2$ we can show that

$$\left| \iint a(t)b(x) \left[|f(R^{2\tau}(t), x)|^2 - |f(t,x)|^2 \right] \cdot d(\bar{\sigma}^\tau \bar{\rho}^\tau)(t,x) \right| \to 0,$$

which completes the proof. □

To proceed, we need one more technical lemma.

Lemma 3.9. *Let μ be a Borel measure on $\mathbb{R}^d$ with compact support and $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ be a bounded Borel measurable function. If for any smooth function $a\colon \mathbb{R}^d \to [0,1]$ we have*

$$\int a(x)\varphi(x)\, d\mu(x) \le 0$$

then $\varphi(x) \le 0$ for μ-a.e. x.

Proof. Let $A = \{x \colon \varphi(x) > 0\}$. Since A is measurable and μ is regular then, for any $\varepsilon > 0$, there exist a compact set $F_\varepsilon \subset A$ and an open set $G_\varepsilon \supset A$ such that $\mu(A \setminus F_\varepsilon) < \varepsilon$ and $\mu(G_\varepsilon \setminus A) < \varepsilon$. By Urysohn's lemma there exists a smooth function $a_\varepsilon \colon \mathbb{R}^d \to [0,1]$ which is 1 on F_ε and 0 outside of G_ε. Consider the obvious identity

$$\int a_\varepsilon \varphi\, d\mu = \int_{F_\varepsilon} a_\varepsilon \varphi\, d\mu + \int_{A \setminus F_\varepsilon} a_\varepsilon \varphi\, d\mu + \int_{G_\varepsilon \setminus A} a_\varepsilon \varphi\, d\mu.$$

Since $\varphi > 0$ on $A \setminus F_\varepsilon$ and $\varphi \le 0$ on $G_\varepsilon \setminus A$, we obtain

$$\begin{aligned} &\int_{F_\varepsilon} a_\varepsilon \varphi\, d\mu = \int_{F_\varepsilon} \varphi\, d\mu \ge \int_A \varphi\, d\mu - c\mu(A \setminus F_\varepsilon), \\ &\int_{A \setminus F_\varepsilon} a_\varepsilon \varphi\, d\mu \ge 0, \\ &\int_{G_\varepsilon \setminus A} a_\varepsilon \varphi\, d\mu \ge \int_{G_\varepsilon \setminus A} \varphi\, d\mu \ge -c\mu(G_\varepsilon \setminus A), \end{aligned}$$

where c is chosen so that $|\varphi(x)| \le c$ for all $x \in \mathbb{R}^d$. These inequalities imply

$$\int a_\varepsilon \varphi\, d\mu \ge \int_A \varphi\, d\mu - 2c\varepsilon. \tag{17}$$

Suppose that $\mu(A) > 0$. Then $\int_A \varphi\, d\mu > 0$. Indeed, A contains a density point y of φ (see, e.g., [3, Theorem 5.8.8]) and from $\varphi(y) = \lim_{r \downarrow 0} \frac{1}{\mu(y + r\boldsymbol{B})} \int_{y + r\boldsymbol{B}} \varphi\, d\mu > 0$

it follows that $\int_A \varphi\, d\mu \geq \int_{y+r\boldsymbol{B}} \varphi\, d\mu > 0$ for some r. Thus, choosing ε small enough makes the right-hand side of (17) strictly positive and leads to a contradiction. $\square$

Proof of Proposition 2. Take a countable dense subset subset $\{t_n\}_n$ of $[0,T]$. Then, for each t_n, choose a countable dense subset $\{x_n^k\}_k$ of $\boldsymbol{C}(t_n)$. The set of pairs $\{(t_n, x_n^k)\}_{n,k}$ is also countable. For each (t_n, x_n^k) we consider the map $y_{n,k}\colon [0,T] \to \mathbb{R}^d$ defined by

$$-\dot{y}_{n,k}(t) \in N_{\boldsymbol{C}(t)}, \quad y_{n,k}(t_n) = x_n^k.$$

This map is uniquely defined and Lipschitz continuous. We state that $\{y_{n,k}(t)\}_{n,k}$ is dense in $\boldsymbol{C}(t)$ for each $t \in [0,T]$. Indeed, since $\boldsymbol{C}$ is lower semicontinuous, for any t and any open ball $x + \varepsilon \boldsymbol{B}^\circ$ such that $\boldsymbol{C}(t) \cap \{x + \varepsilon \boldsymbol{B}^\circ\} \neq \varnothing$ there exists t_n such that $\boldsymbol{C}(t_n) \cap \{x + \varepsilon \boldsymbol{B}^\circ\} \neq \varnothing$. The latter set has nonempty interior and we can select from it some x_n^k. Since t_n can be arbitrary close to t then $y_{n,k}(t) \in x + \varepsilon \boldsymbol{B}^\circ$, as desired.

Now, for each $y_{n,k}$, apply Lemma 3.9 to the inequality established in Lemma 3.8. Then we get

$$\int_{\mathbb{R}^d} b(x)\Big(\langle v_t(x) - \mathscr{V}(\rho_t)(x), x - y_{n,k}(t)\rangle - \sigma(t,x) \cdot |x - y_{n,k}(t)|^2 \Big)\, d\rho_t(x) \leq 0$$

for all $t \in [0,T] \setminus I_{n,k}$, where each $I_{n,k}$ is a set of Lebesgue measure zero. The union I of these sets also has measure zero. Since $y_{n,k}(t)$ are dense in $\boldsymbol{C}(t)$, we have

$$\int_{\mathbb{R}^d} b(x) \max_{y \in \boldsymbol{C}(t)} \Big(\langle v_t(x) - \mathscr{V}(\rho_t)(x), x - y\rangle - \sigma(t,x) \cdot |x - y|^2 \Big)\, d\rho_t(x) \leq 0,$$

for all $t \in [0,T] \setminus I$. Using again Lemma 3.9, we obtain that for ρ_t-a.e. x

$$\langle v_t(x) - \mathscr{V}(\rho_t)(x), x - y\rangle \leq \sigma(t,x) \cdot |x - y|^2 \quad \forall y \in \boldsymbol{C}(t).$$

This completes the proof. $\square$

4. **Continuous dependence.** Before passing to the continuous dependence, let us prove assertions (2) and (3) of Theorem 1.2.

Lemma 4.1. *For each solution ρ of (1) assertions (1) and (2) of Theorem 1.2 hold.*

Proof. **1.** Since the velocity v of ρ can be tweaked on a ρ-negligible set without changing the solution of the continuity equation, we may assume that (11) holds for all t and x.

2. Let us show that $\rho = E_\sharp(\lambda \times \boldsymbol{\eta})$, where $\boldsymbol{\eta}$ is defined as in Theorem 2.7, λ is the one dimensional Lebesgue measure, and $E\colon (t,x,\gamma) \mapsto (t,\gamma(t))$. Indeed, take $A \subset [0,T] \times \mathbb{R}^d$ and denote by A_t its slice $\{\xi : (t,\xi) \in A\}$. Then

$$\rho(A) = \int_0^T \rho_t(A_t)\, dt = \int_0^T \boldsymbol{\eta}\left(e_t^{-1}(A_t)\right)\, dt = \int_0^T \boldsymbol{\eta}(\tilde{A}_t)\, dt = (\lambda \times \boldsymbol{\eta})(\tilde{A}),$$

where $\tilde{A} = \{(t,x,\gamma) : (t,\gamma(t)) \in A\}$. It remains to note that $\tilde{A} = E^{-1}(A)$.

3. Let Γ be defined as in Theorem 2.7 and $\tilde{A} \subset [0,T] \times \Gamma$ be the set of all triples (t,x,γ) such that $\dot{\gamma}(t)$ exists and equals to $v_t(\gamma(t))$. We are going to show that $\tilde{A}$ is a set of full measure $\lambda \times \boldsymbol{\eta}$. By Fubini's theorem,

$$(\lambda \times \boldsymbol{\eta})(\tilde{A}) = \int_\Gamma \lambda(\tilde{A}_{(x,\gamma)})\, d\boldsymbol{\eta}(x,\gamma), \quad \text{where} \quad \tilde{A}_{(x,\gamma)} = \{t : (t,x,\gamma) \in \tilde{A}\},$$

Now we obtain $(\lambda \times \boldsymbol{\eta})(\tilde{A}) = T$ because $\lambda(\tilde{A}_{(x,\gamma)}) = T$, for all x and γ.

4. Since $\rho = E_\sharp(\lambda \times \boldsymbol{\eta})$, we conclude that $E(\tilde{A})$ is a set of full measure ρ. In other words, for ρ-a.e. (t,x) there exists a solution y of the sweeping process

$$\dot{y}(t) \in \mathscr{V}(\rho_t)(y(t)) - N_{\boldsymbol{C}(t)}(y(t)), \quad \text{for a.e. } t \in [0,T],$$

such that $y(t) = x$, $\dot{y}(t)$ exists and equals to $v_t(x)$.

5. Now we deduce from [23, Theorem 2.4] that $|v_t(x)| \le 2L + M$ for ρ-a.e. (t,x) and from Proposition 4 that

$$\xi + \eta \cdot v_t(x) = 0 \quad \forall (\xi,\eta) \in N_{\operatorname{graph} \boldsymbol{C}}(t,x)$$

for ρ-a.e. (t,x), when $\operatorname{graph} \boldsymbol{C}$ is r'-prox-regular. □

Let $\rho^1, \rho^2 \colon [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ be solutions of the sweeping processes (1) corresponding to the set-valued maps $\boldsymbol{C}^1, \boldsymbol{C}^2 \colon [0,T] \to \mathcal{K}_r(\mathbb{R}^d)$, respectively. By v_t^1, v_t^2 we denote their velocity fields.

In order to prove the continuous dependence, we are going to differentiate the function $\boldsymbol{r}(t) \doteq \frac{1}{2} W_2^2(\rho_t^1, \rho_t^2)$. Since both curves ρ^1 and ρ^2 are absolutely continuous, we can use the formula

$$\frac{d}{dt} W_2^2(\rho_t^1, \rho_t^2) = 2 \iint \left\langle v_t^1(x) - v_t^2(y), x - y \right\rangle \, d\Pi_{\rho_t^1, \rho_t^2}(x,y),$$

whose proof repeats that of Theorem 8.4.7 [2] (we put it in Appendix C, for completeness). The measure $\Pi_{\rho_t^1,\rho_t^2}$ in the right-hand side denotes an *optimal* plan between ρ_t^1 and ρ_t^2.

Let $i = 1, 2$. By definition, $\mathscr{V}(\rho_t^i)(x) - v_t^i(x) \in N_{\boldsymbol{C}^i(t)}(x)$, for a.e. $t \in [0,T]$ and ρ_t^i-a.e. $x \in \mathbb{R}^d$. Since the values of $\boldsymbol{C}^i$ are r-prox-regular, Proposition 1(b) implies that

$$\left\langle v_t^i(x) - \mathscr{V}(\rho_t^i)(x), x - y \right\rangle \le \frac{1}{2r} \left| v_t^i(x) - \mathscr{V}(\rho_t^i)(x) \right| \, |x - y|^2, \tag{18}$$

for a.e. $t \in [0,T]$, ρ_t^i-a.e. $x \in \mathbb{R}^d$, and all $y \in \boldsymbol{C}^i(t)$.

According to Lemma C.1, we have

$$\begin{aligned} \frac{1}{2}\frac{d}{dt} W_2^2(\rho_t^1, \rho_t^2) &= \iint \left\langle v_t^1(x) - v_t^2(y), x - y \right\rangle \, d\Pi_{\rho_t^1,\rho_t^2}(x,y) \\ &= \iint \left\langle v_t^1(x) - \mathscr{V}(\rho_t^1)(x), x - y \right\rangle \, d\Pi_{\rho_t^1,\rho_t^2}(x,y) \\ &+ \iint \left\langle v_t^2(y) - \mathscr{V}(\rho_t^2)(y), y - x \right\rangle \, d\Pi_{\rho_t^1,\rho_t^2}(x,y) \\ &+ \iint \left\langle \mathscr{V}(\rho_t^1)(x) - \mathscr{V}(\rho_t^2)(y), x - y \right\rangle \, d\Pi_{\rho_t^1,\rho_t^2}(x,y) \\ &= I_1 + I_2 + I_3. \end{aligned} \tag{19}$$

for a.e. $t \in [0,T]$.

We split the first integral I_1 as follows:

$$\begin{aligned} I_1 = &\iint \left\langle v_t^1(x) - \mathscr{V}(\rho_t^1)(x), x - P_{\boldsymbol{C}^1(t)}(y) \right\rangle d\Pi_{\rho_t^1,\rho_t^2}(x,y) \\ &+ \iint \left\langle v_t^1(x) - \mathscr{V}(\rho_t^1)(x), P_{\boldsymbol{C}^1(t)}(y) - y \right\rangle d\Pi_{\rho_t^1,\rho_t^2}(x,y). \end{aligned} \tag{20}$$

Note that $t \mapsto P_{\boldsymbol{C}^1(t)}(y)$ is, in general, a set-valued map. Here, slightly abusing the notation, we denoted by $P_{\boldsymbol{C}^1(t)}(y)$ its measurable selection, which always exists[1].

Taking into account the inclusions

$$\operatorname{spt} \Pi_{\rho_t^1,\rho_t^2} \subset \operatorname{spt} \rho_t^1 \times \operatorname{spt} \rho_t^2 \subset \boldsymbol{C}^1(t) \times \boldsymbol{C}^2(t),$$

we deduce from (18) and assertion (2) of Theorem 1.2 that

$$\iint \left\langle v_t^1(x) - \mathscr{V}(\rho_t^1)(x), x - P_{\boldsymbol{C}^1(t)}(y) \right\rangle d\Pi_{\rho_t^1,\rho_t^2}(x,y) \leq \frac{3L+M}{2r} W_2^2(\rho_1,\rho_2),$$
$$\iint \left\langle v_t^1(x) - \mathscr{V}(\rho_t^1)(x), P_{\boldsymbol{C}^1(t)}(y) - y \right\rangle d\Pi_{\rho_t^1,\rho_t^2}(x,y) \leq (3L+M)\Delta(t),$$

where $\Delta(t) \doteq d_H\left(\boldsymbol{C}^1(t), \boldsymbol{C}^2(t)\right)$. This gives

$$I_1 \leq (3L+M)\Delta(t) + \frac{3L+M}{4r}\boldsymbol{r}(t).$$

The same inequality holds for I_2.

We rewrite the last integral I_3 as the sum

$$\int \left\langle \mathscr{V}(\rho_t^1)(x) - \mathscr{V}(\rho_t^2)(x), x - y \right\rangle d\Pi_{\rho_t^1,\rho_t^2}(x,y)$$
$$+ \int \left\langle \mathscr{V}(\rho_t^2)(x) - \mathscr{V}(\rho_t^2)(y), x - y \right\rangle d\Pi_{\rho_t^1,\rho_t^2}(x,y).$$

The first integral above is bounded by

$$\left(\int \left|\mathscr{V}(\rho_t^1)(x) - \mathscr{V}(\rho_t^2)(x)\right|^2 d\rho_t^1(x)\right)^{1/2} \cdot \left(\int |x-y|^2 d\Pi_{\rho_t^1,\rho_t^2}(x,y)\right)^{1/2} \leq LW_2^2(\rho_t^1,\rho_t^2),$$

thanks to L-Lipschitz continuity of $\mathscr{V}\colon \mathcal{P}_2(\mathbb{R}^d) \to C(\mathbb{R}^d;\mathbb{R}^d)$. The second one is bounded by $LW_2^2(\rho_t^1,\rho_t^2)$ due to L-Lipschitz continuity of $\mathscr{V}(\rho_t^2) \in C(\mathbb{R}^d;\mathbb{R}^d)$. Thus, we have $I_3 \leq 4L\boldsymbol{r}(t)$.

Plugging the above estimates into (19) gives

$$\dot{\boldsymbol{r}}(t) \leq (6L+2M)\Delta(t) + \left(4L + \frac{3L+M}{2r}\right)\boldsymbol{r}(t).$$

By Grönwall's lemma, we obtain (2), thus completing the proof of assertion (4). Finally, note that uniqueness in assertion (1) is a direct consequence of the above estimate.

5. **Application to environment optimization.** An important task of crowd dynamics is to understand how environment affects the crowd motion. Consider a specific question: can an obstacle, such as a column, placed at the right spot help the crowd to evacuate a room? We know that under some circumstances it happens in the real life [14]. Numerical experiments (see Section 6) show that this phenomenon, called Braess's paradox [15], can be reproduced in our model. But can we find the best shape and position of the obstacle?

Let us formulate this problem within our framework. Suppose that r is a fixed positive constant, Ω a compact r-prox-regular set that represents the region where the crowd can move, ϑ a compactly supported measure on Ω which defines agents' distribution. We assume that Ω consists of two parts: the safe S and the dangerous

[1] Since $P_{\boldsymbol{C}^1(t)}(y) = \{y + d_{\boldsymbol{C}^1(t)}(y) \cdot \boldsymbol{B}\} \cap \boldsymbol{C}^1(t)$, $P_{\boldsymbol{C}^1}$ is measurable as an intersection of two measurable set-valued maps; hence it has a measurable selection.

D regions. The crowd leaves the dangerous region moving along a given nonlocal vector field $v_t = \mathscr{V}(\rho_t)$. Our aim is to place an obstacle $O \subset \Omega$ so that the number of agents staying in D by a time moment T were minimal. As was discussed before, each obstacle defines the corresponding viability region $C = \Omega \setminus O$. We assume that admissible viability regions C belong to the set

$$\mathcal{C} = \left\{ C \in \mathcal{K}_r(\mathbb{R}^d) \,:\, C \subset \Omega,\ \vartheta(C) = 1 \right\}.$$

The following theorem says that among all admissible viability regions one can always choose an optimal one.

Theorem 5.1. *Let $\rho^C \colon [0,T] \to \mathcal{P}_2(\mathbb{R}^d)$ denote the trajectory of (1) which corresponds to $\mathbf{C}(t) \equiv C$, for $C \in \mathcal{C}$. If D is open then the minimization problem*

$$\min \left\{ \rho_T^C(D) \,:\, C \in \mathcal{C} \right\}$$

admits a solution.

Proof. We know that $C \mapsto \rho_T^C$ is continuous as a map $\mathcal{K}_r(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$. Since D is open, we conclude, by the Portmanteau theorem, that $C \mapsto \rho_T^C(D)$ is lower semicontinuous as $\mathcal{K}_r(\mathbb{R}^d) \mapsto \mathbb{R}$. To complete the proof, it suffices to show that $\mathcal{C}$ is compact.

We can always extract from any sequence $C_n \in \mathcal{C}$ a subsequence converging to some compact set $C \subset \Omega$ (see, e.g., [21, p. 120]). By Theorem 4.13 in [12], $C \in \mathcal{K}_r(\mathbb{R}^d)$. Hausdorff convergence implies that for any $k \in \mathbb{N}$ one may find $n(k)$ such that $C_{n(k)} \subset C + \frac{1}{k}\mathbf{B}$. This means that $\vartheta\left(C + \frac{1}{k}\mathbf{B}\right) = 1$, for each $k \in \mathbb{N}$. Therefore, $1 = \lim\limits_{k\to\infty} \vartheta\left(C + \frac{1}{k}\boldsymbol{B}\right) = \vartheta\Big(\bigcap_{k=1}^{\infty}(C + \frac{1}{k}\boldsymbol{B})\Big) = \vartheta(C)$. ☐

6. **Numerical computations.** While continuous dependence on the moving set leads to existence results in environment optimization problems, continuous dependence on the initial measure provides an algorithm for computing trajectories of (1). Indeed, let ρ be a trajectory issuing from $\vartheta \in \mathcal{P}_2(\mathbb{R}^d)$. We can always approximate ϑ by a discrete measure $\vartheta_N \doteq \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ (because such measures are dense in $\mathcal{P}_2(\mathbb{R}^d)$ [25]). The corresponding trajectory ρ_N, being absolutely continuous, consists of discrete measures as well (note that several δ-functions could be glued into one along the way, but they can never be split again). Now, we can easily compute ρ_N by applying the catching-up scheme. Theorem 1.2 shows that $\rho_N \to \rho$ in $C([0,T];\mathcal{P}_2(\mathbb{R}^d))$ as $N \to \infty$.

Below we provide computations for two simple models of crowd dynamics taken from [17] and [9].

Example 1 (Attraction/repulsion model)**.** The first model [17] corresponds to

$$\mathscr{V}(\mu)(x) = w(x) + \int K(x-y)\,d\mu(y),$$

where $w \colon \mathbb{R}^2 \to \mathbb{R}^2$ is a drift and K the attraction/repulsion kernel of the form

$$K(x) = -\frac{A_a x}{2a^2}\exp\Big(-\frac{|x|^2}{2a^2}\Big) + \frac{A_r x}{2r^2}\exp\Big(-\frac{|x|^2}{2r^2}\Big),$$

Here a and r determine the attraction and repulsion ranges, A_a and A_r the attraction and repulsion intensities. It is common to take $r < a$, so agents repulse each other at short distances and attract at large ones. One can easily verify that $\mathscr{V}$ satisfies our assumptions if w is bounded and Lipschitz.

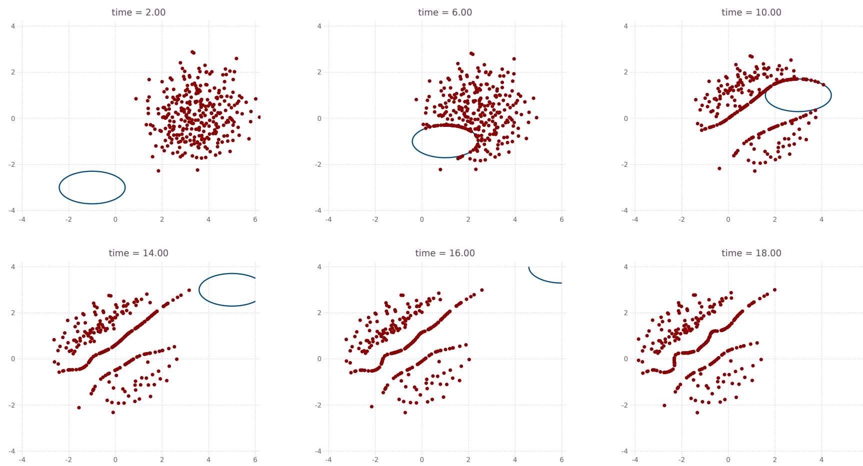


FIGURE 2. Solutions to the attraction/repulsion model (Example 1) at time moments $t = 2, 6, 10, 14, 16, 18$. The initial measure ϑ is the Gaussian probability distribution on $\mathbb{R}^2$ with mean $(4, 0)$ and variance **id**. The obstacle is the blue ellipse moving from the bottom left to the top right corner. Parameters of the model: $A_a = 4$, $A_r = 7$, $a = 1/\sqrt{2}$, $r = 0.5$, $w \equiv -0.3$, parameters of discretization: $\tau = 0.01$, $N = 300$.

For the computations presented in Figure 2, we choose $A_a = 4$, $A_r = 7$, $a = 1/\sqrt{2}$, $r = 0.5$, $w \equiv -0.3$, $\tau = 0.01$. The moving set is given by $\boldsymbol{C}(t) = \{x \in \mathbb{R}^2 : f(t, x) \leq 0\}$ with

$$f(t, x) = -(x_1 - 0.5t + 2)^2 - 4(x_2 - 0.5t + 4)^2 + 2,$$

that is, our obstacle is an ellipse crossing the crowd. We approximate the initial measure ϑ (the Gaussian measure with mean $(4, 0)$ and variance **id**) by a discrete measure $\frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}$ with x_i randomly distributed according to ϑ, $N = 300$.

Example 2 (Congestion model)**.** The second model [9] corresponds to the choice

$$\mathscr{V}(\mu)(x) = w(x) \cdot \psi \left(\int \eta(|x - y|) \, d\mu(y) \right),$$

where $w \colon \mathbb{R}^2 \to \mathbb{R}^2$ is a given vector field, $\eta \colon \mathbb{R} \to \mathbb{R}_+$ is a smooth bell-shaped function, $\psi \colon \mathbb{R}_+ \to [0, 1]$ is Lipschitz and non-increasing. The idea behind this model is that the velocity of an agent located at x decreases as the number of agents around x (estimated by $\int \eta(|x - y|) \, d\mu(y)$) grows.

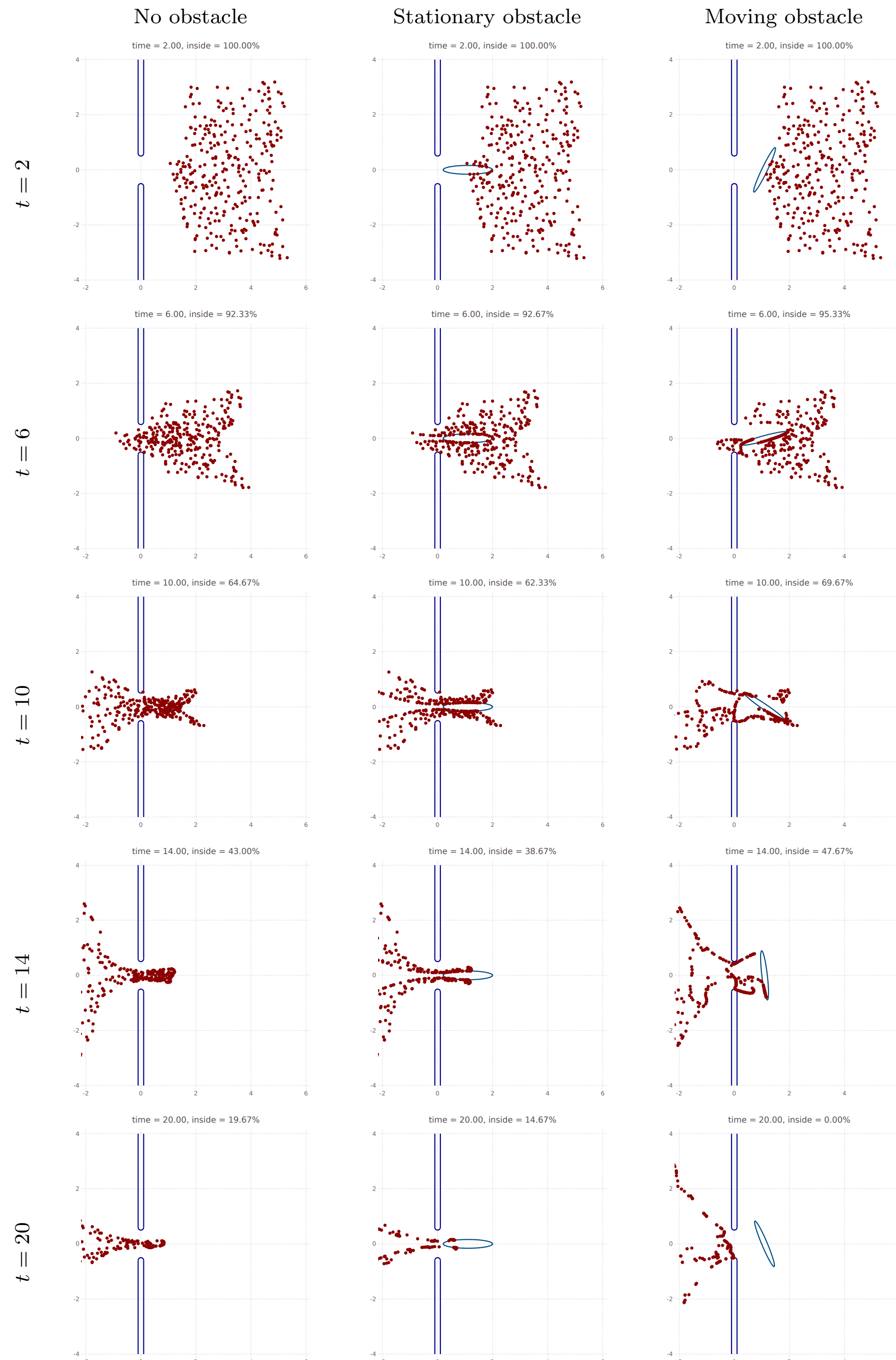
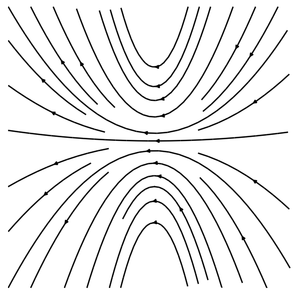


FIGURE 3. Solutions to the congestion model (Example 2) at time moments $t = 2, 6, 10, 14, 20$. First column: no obstacle ($c = (100, 100)$, $a = (0.9, 0.16)$, $\omega = 0$), second column: a stationary obstacle ($c = (1.1, 0)$, $a = (0.9, 0.16)$, $\omega = 0$), third column: a moving obstacle ($c = (1.1, 0)$, $a = (0.9, 0.1)$, $\omega = 1$). The initial measure ϑ is absolutely continuous with density $\frac{1}{32}\mathbf{1}_{[2,6]\times[-4,4]}$. Parameters of the model: $b = 0.6$, $\delta = 0.1$, parameters of discretization: $\tau = 0.01$, $N = 300$.

To define the non-local vector field we choose the following functions:



$$w(x) = -\frac{1}{2|x|}(1 + x_1^2, 2x_1x_2),$$
$$\psi(r) = 1 - \frac{2}{\pi}\arctan \kappa x^2,$$
$$\eta(r) = \begin{cases} \frac{1}{\beta} e^{\frac{1}{(r/\varepsilon)^2-1}}, & r < \varepsilon, \\ 0, & \text{otherwise}, \end{cases}$$
$$\varepsilon = 0.3,\ \kappa = 1000,\ \beta = 0.466.$$

Field lines of w are the parabolas depicted above. The moving set is given by

$$\boldsymbol{C}(t) = \{x \in \mathbb{R}^2 \,:\, f(t,x) \le 0\} \setminus (I + \delta \boldsymbol{B}^\circ), \text{ where } I = \big\{x = 0,\ |y| \ge b\big\},\ b, \delta > 0,$$
$$f(t,x) = -\Big(\frac{(x_1 - c_1)\cos\omega t - (x_2 - c_2)\sin\omega t}{a_1}\Big)^2 - \Big(\frac{(x_1 - c_1)\sin\omega t + (x_2 - c_2)\cos\omega t}{a_2}\Big)^2 + 1.$$

Here $I + \delta\boldsymbol{B}^\circ$ models a wall with an exit and f an elliptic obstacle with semi-axis a_1, a_2 rotating around its center $c = (c_1, c_2)$. In our case, $b = 0.6$, $\delta = 0.1$, $\tau = 0.01$. The initial measure is absolutely continuous with density $\frac{1}{32}\mathbf{1}_{[2,6]\times[-4,4]}$. We approximate it by a discrete measure $\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}$, where x_i are uniformly distributed on the rectangle $[2,6]\times[-4,4]$, $N = 300$. Solutions of (1) for various c, a, ω are presented in Figure 3. Note that, by the time moment $T = 20$, the dangerous region $D = \{x > 0\}$ contains 19.67% of the total mass if there are no obstacles, 14.67% for the stationary obstacle, 0.00% for the moving obstacle. Hence Braess's paradox may indeed occur in (1).

Appendix A. The Benamou-Brenier functional. For any couple (ρ, E), where $\rho \in \mathcal{M}(X;\mathbb{R})$ is a measure and $E \in \mathcal{M}(X;\mathbb{R}^d)$ is a vector measure, we correspond the number

$$\mathcal{B}_2(\rho, E) = \sup\left\{\int_X a(x)\,d\rho(x) + \int_X b(x)\cdot dE(x) \,:\, (a,b) \in C_b(X;K_2)\right\},$$

where

$$K_2 = \left\{(a,b) \in \mathbb{R}\times\mathbb{R}^d \colon a + \frac{1}{2}|b|^2 \le 0\right\}.$$

Proposition 3 (Proposition 5.18 [22])**.** *The map $\mathcal{B}_2$ is convex and lower semicontinuous on $\mathcal{M}(X;\mathbb{R})\times\mathcal{M}(X;\mathbb{R}^d)$. Moreover,*

(i) $\mathcal{B}_2 \ge 0$,
(ii) $C_b(X;K_2)$ can be replaced with $L^\infty(X;K_2)$ in the definition of $\mathcal{B}_2$,
(iii) if ρ and E are absolutely continuous with respect to a positive measure λ then

$$\mathcal{B}_2(\rho, E) = \int f_2(\rho(x), E(x))\,d\lambda(x),$$

where

$$f_2(t,x) = \sup_{(a,b)\in K_2} (at + b\cdot x) = \begin{cases} \frac{1}{2t}|x|^2 & \text{if } t = 0, \\ 0 & \text{if } t = 0,\ x = 0, \\ +\infty & \text{otherwise}; \end{cases}$$

(iv) $\mathcal{B}_2(\rho, E) < +\infty$ *only if* $\rho \geq 0$ *and* $E \ll \rho$,

(v) for $\rho \geq 0$ *and* $E \ll \rho$, *we have* $E = v\rho$ *and* $\mathcal{B}_2(\rho, E) = \frac{1}{2}\int |v|^2\, d\rho$.

Appendix B. **Continuity of the projection map.**

Lemma B.1. *Let* $A_n \xrightarrow{d_H} A$ *and* $x_n \to x$. *If the projections* $P_A(x)$ *and* $P_{A_n}(x_n)$ *are unique then* $P_{A_n}(x_n) \to P_A(x)$.

Proof. Consider the following functions

$$F_n(y) = \chi_{A_n}(y) + |x_n - y|^2, \quad F(y) = \chi_A(y) + |x - y|^2,$$

where χ_A denotes the indicator function of A. From [10, Proposition 4.15] it follows that $F = \Gamma\text{-}\lim F_n$. By our assumptions, each $P_{A_n}(x_n)$ is a unique minimizer of F_n and $P_A(x)$ is a unique minimizer of F. Since all $P_{A_n}(x_n)$ belong to a compact set (because $\{A_n\}$ is convergent), one may extract a converging subsequence. By one of the key properties of Γ-convergence [10, Corollary 7.17], its limit is a minimizer of F, i.e., $P_A(x)$. This means that $P_{A_n}(x_n) \to P_A(x)$. □

Appendix C. **Derivative of the squared Wasserstein distance.**

Lemma C.1. *Let* μ_t, ν_t *be two absolutely continuous curves in* $\mathcal{P}_2(\mathbb{R}^d)$ *and* u_t, v_t *be their velocity vector fields. Then, for a.e.* t, *one has*

$$\frac{d}{dt} W_2^2(\mu_t, \nu_t) = 2 \iint \langle u_t(x) - v_t(y), x - y\rangle \, d\Pi_{\mu_t,\nu_t}(x,y),$$

where Π_{μ_t,ν_t} *is an optimal transport plan from* μ_t *to* ν_t.

Proof. We shall prove the formula for all t satisfying the following conditions: 1) $W_2^2(\mu_t,\nu_t)$ is differentiable, 2) $\lim_{h\to 0} \frac{W_2(\mu_{t+h}, S^h_{t\sharp}\mu_t)}{h} = 0$, 3) $\lim_{h\to 0} \frac{W_2(\nu_{t+h}, P^h_{t\sharp}\mu_t)}{h} = 0$. Here

$$S_t^h = \mathbf{id} + hu_t, \quad P_t^h = \mathbf{id} + hv_t.$$

Proposition 8.4.6 from [2] says that all such t compose a set of full measure.

First, we show that

$$\frac{d}{ds} W_2^2(\mu_s, \nu_s)|_{s=t} = \lim_{h\to 0} \frac{W_2^2(S^h_{t\sharp}\mu_t, P^h_{t\sharp}\nu_t) - W_2^2(\mu_t,\nu_t)}{h}.$$

Indeed,

$$\begin{aligned} W_2^2(\mu_{t+h}, \nu_{t+h}) - W_2^2(\mu_t,\nu_t) &= W_2^2(S^h_{t\sharp}\mu_t, P^h_{t\sharp}\nu_t) - W_2^2(\mu_t,\nu_t) \\ &\quad + W_2^2(\mu_{t+h}, \nu_{t+h}) - W_2^2(S^h_{t\sharp}\mu_t, P^h_{t\sharp}\nu_t), \end{aligned}$$

so if we show that

$$\lim_{h\to 0} \frac{W_2^2(\mu_{t+h}, \nu_{t+h}) - W_2^2(S^h_{t\sharp}\mu_t, P^h_{t\sharp}\nu_t)}{h} = 0,$$

we are done. Let us note that

$$\left|W_2^2(\mu_{t+h}, \nu_{t+h}) - W_2^2(S^h_{t\sharp}\mu_t, P^h_{t\sharp}\nu_t)\right| \leq C\left|W_2(\mu_{t+h}, \nu_{t+h}) - W_2(S^h_{t\sharp}\mu_t, P^h_{t\sharp}\nu_t)\right|,$$

for some $C > 0$. Now,

$$\begin{aligned}\left|W_2(\mu_{t+h},\nu_{t+h}) - W_2(S^h_{t\sharp}\mu_t, P^h_{t\sharp}\nu_t)\right| &\le \left|W_2(\mu_{t+h},\nu_{t+h}) - W_2(S^h_{t\sharp}\mu_t, \nu_{t+h})\right| \\ &\quad + \left|W_2(S^h_{t\sharp}\mu_t,\nu_{t+h}) - W_2(S^h_{t\sharp}\mu_t, P^h_{t\sharp}\nu_t)\right| \\ &\le W_2(\mu_{t+h}, S^h_{t\sharp}\mu_t) + W_2(\nu_{t+h}, P^h_{t\sharp}\nu_t).\end{aligned}$$

It remains to apply properties 2) and 3).

Choose any optimal plan Π between μ_t and ν_t. Note that the following plan $(S^h_t \circ \pi^1, P^h_t \circ \pi^2)_\sharp \Pi$ transports $S^h_{t\sharp}\mu_t$ to $P^h_{t\sharp}\nu_t$. Hence

$$\begin{aligned}W_2^2\left(S^h_{t\sharp}\mu_t, P^h_{t\sharp}\nu_t\right) &\le \int |x + hu_t(x) - y - hv_t(y)|^2 \, d\Pi(x,y) \\ &\le W_2^2(\mu_t,\nu_t) + h^2 \int |u_t(x) - v_t(y)|^2 \, d\Pi(x,y) \\ &+ 2h \int \langle u_t(x) - v_t(y), x - y\rangle \, d\Pi(x,y).\end{aligned}$$

Therefore, if $h > 0$, we get

$$\frac{d}{ds} W_2^2(\mu_s,\nu_s)|_{s=t} \le 2 \int \langle u_t(x) - v_t(y), x - y\rangle \, d\Pi(x,y).$$

If $h < 0$, we get the opposite inequality. □

Appendix D. **No-flux property.** Here we prove a simple property of the perturbed sweeping process

$$\dot y(t) \in v_t(y(t)) - N_{\boldsymbol{C}(t)}(y(t)) \quad \text{for a.e. } t \in [0,T]. \tag{21}$$

that we failed to find in the literature. Below $\langle (s,x),(t,y)\rangle$ denotes the scalar product in $\mathbb{R}^{d+1}$ and $x \cdot y$ the scalar product in $\mathbb{R}^d$.

Proposition 4. *Let $(t,x) \mapsto v_t(x)$ be measurable in t, L-Lipschitz in x and L-bounded, $\boldsymbol{C}\colon [0,T] \rightrightarrows \mathbb{R}^d$ satisfy $(\mathbf{A_2})$ and have r'-prox-regular graph, $r' > 0$. Let y be a solution of (21). If $t_0 \in (0,T)$ is so that $\dot y(t_0)$ exists then $(1, \dot y(t_0))$ is tangent to* graph $\boldsymbol{C}$ *at $(t_0, y(t_0))$ in the sense that*

$$\xi + \dot y(t_0)\cdot\eta = 0 \quad \forall (\xi,\eta) \in N_{\operatorname{graph}\boldsymbol{C}}(t_0, y(t_0)).$$

Proof. Pick some $(\xi,\eta) \in N_{\operatorname{graph}\boldsymbol{C}}(t_0, y(t_0))$. By Proposition 1(b), we have

$$\frac{|(\xi,\eta)|}{2r'}\, |(t,y(t)) - (t_0,y(t_0))|^2 - \langle(\xi,\eta),(t,y(t)) - (t_0,y(t_0))\rangle \ge 0, \tag{22}$$

for all t sufficiently close to t_0. Since for $t = t_0$ the function on the left-hand side of (22) becomes 0, we conclude that t_0 is its extremal point. By Fermat's rule, $\xi + \dot y(t_0)\cdot\eta = 0$. □

REFERENCES

[1] L. Ambrosio and G. Crippa, Continuity equations and ODE flows with non-smooth velocity, *Proc. R. Soc. Edinb., Sect. A, Math.*, **144** (2014), 1191–1244.

[2] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Basel: Birkhäuser, 2005.

[3] V. I. Bogachev, *Measure Theory. Vol. I, II*, Springer-Verlag, Berlin, 2007.

[4] D. Burago, Y. Burago and S. Ivanov, A Course in Metric Geometry, vol. 33, Providence, RI: American Mathematical Society (AMS), 2001.

[5] J. A. Carrillo, M. Fornasier, G. Toscani and F. Vecil, Particle, kinetic, and hydrodynamic models of swarming, in *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, Boston, MA: Birkhäuser, 2010, 297–336.
[6] F. Clarke, *Functional Analysis, Calculus of Variations and Optimal Control*, vol. 264 of Graduate Texts in Mathematics, Springer London, London, 2013.
[7] R. M. Colombo, M. Garavello and M. Lécureux-Mercier, Non-local crowd dynamics, *Comptes Rendus Mathematique*, **349** (2011), 769–772.
[8] R. M. Colombo and E. Rossi, Nonlocal conservation laws in bounded domains, *SIAM J. Math. Anal.*, **50** (2018), 4041–4065.
[9] R. Colombo, Control of the continuity equation with a non local flow, *ESAIM: Control, Optimisation and Calculus of Variations*, **17** (2011), 353–379, http://journals.cambridge.org/abstract{_}S1292811910000072.
[10] G. Dal Maso, *An Introduction to* Γ*-convergence*, vol. 8, Basel: Birkhäuser, 1993.
[11] S. Di Marino, B. Maury and F. Santambrogio, Measure sweeping processes, *J. Convex Anal.*, **23** (2016), 567–601.
[12] H. Federer, Curvature measures, *Trans. Am. Math. Soc.*, **93** (1959), 418–491.
[13] A. Figalli and N. Gigli, A new transportation distance between non-negative measures, with applications to gradients flows with Dirichlet boundary conditions, *Journal des Mathematiques Pures et Appliquees*, **94** (2010), 107–130.
[14] D. Helbing, I. J. Farkas and T. Vicsek, Crowd disasters and simulation of panic situations, *The Science of Disasters*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2002, 330–350,
[15] R. L. Hughes, The flow of human crowds, *Annual Review of Fluid Mechanics*, **35** (2003), 169–182.
[16] B. Maury and S. Faure, *Crowds in Equations. An Introduction to the Microscopic Modeling of Crowds*, Hackensack, NJ: World Scientific, 2019.
[17] A. Mogilner and L. Edelstein-Keshet, A non-local model for a swarm, *Journal of Mathematical Biology*, **38** (1999), 534–570.
[18] B. Piccoli and F. Rossi, Transport equation with nonlocal velocity in Wasserstein spaces: convergence of numerical schemes, *Acta Appl. Math.*, **124** (2013), 73–105.
[19] B. Piccoli and F. Rossi, Measure-theoretic models for crowd dynamics, in *Crowd Dynamics, Volume 1. Theory, Models, and Safety Problems*, Cham: Birkhäuser, 2018, 137–165.
[20] R. A. Poliquin, R. T. Rockafellar and L. Thibault, Local differentiability of distance functions, *Trans. Amer. Math. Soc.*, **352** (2000), 5231–5249.
[21] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, vol. 317 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 1998.
[22] F. Santambrogio, *Optimal Transport for Applied Mathematicians. Calculus of Variations, PDEs, and Modeling*, vol. 87, Cham: Birkhäuser/Springer, 2015.
[23] M. Sene and L. Thibault, Regularization of dynamical systems associated with prox-regular moving sets., *J. Nonlinear Convex Anal.*, **15** (2014), 647–663.
[24] L. Thibault, Regularization of nonconvex sweeping process in Hilbert space, *Set-Valued Analysis*, **16** (2008), 319–333.
[25] C. Villani, *Optimal Transport. Old and New*, vol. 338, Berlin: Springer, 2009.

Received for publication February 2021.

E-mail address: nmaltugueva@gmail.com
E-mail address: nickpogo@gmail.com