# A Wasserstein norm for signed measures, with application to non local transport equation with source term

Benedetto Piccoli, Francesco Rossi, Magali Tournus

# A Wasserstein norm for signed measures, with application to non-local transport equation with source term

Benedetto Piccoli [*] Francesco Rossi [†] Magali Tournus [‡]

**Abstract**

We introduce an optimal transportation interpretation of the Kantorovich norm on the space of signed Radon measures with finite mass, based on the generalized Wasserstein distance for measures with different masses.

With this new interpretation, we obtain new topological properties for this norm. We use these tools to prove existence and uniqueness for solutions to non-local transport equations with source terms, when the initial condition is a signed measure.

**Keywords.** Wasserstein distance, Transport equation, Signed measures, Kantorovich duality.

**AMS subject classifications.** 28A33, 35A01.

## 1 Introduction

The problem of optimal transportation, also called Monge-Kantorovich problem, has been intensively studied in the mathematical community. Related to this problem, Wasserstein distances in the space of probability measures have revealed to be powerful tools, in particular for dealing with dynamics of measures like the transport Partial Differential Equation (PDE in the following), see e.g. [1, 2]. For a complete introduction to Wasserstein distances, see [22, 23].

The main limit of this approach, at least for its application to dynamics of measures, is that the Wasserstein distances $W_p(\mu,\nu)$ ($p \geq 1$) are defined only if the two positive measures $\mu,\nu$ have the same mass. For this reason, the generalized Wasserstein distances $W_p^{a,b}(\mu,\nu)$ are introduced in [19, 20]: they combine the standard Wasserstein and total variation distances. In rough words, for $W_p^{a,b}(\mu,\nu)$ an infinitesimal mass $\delta\mu$ of $\mu$ can either be removed at cost $a|\delta\mu|$, or moved from $\mu$ to $\nu$ at cost $bW_p(\delta\mu,\delta\nu)$. An optimal transportation problem between densities with different masses has been studied in [7, 11], where only a given fraction $m$ of each density is transported. These works were motivated by a modeling issue: using the example of a resource that is extracted and that we want to distribute in factories, one aims to use only a certain given fraction of production and consumption capacity. In this approach and contrarily to the generalized Wasserstein distance [18], the mass that is leftover has no impact on the distance between the measures $\mu$ and $\nu$. In another context, for the purpose to interpret some reaction-diffusion equations not preserving masses as gradient flows, the authors of [12] define the distance $Wb_2$ between measures with different masses on a bounded domain. Further generalizations for positive measures with different masses, based

[*]Department of Mathematical Sciences, Rutgers University Camden, Camden, NJ, USA. Email: piccoli@camden.rutgers.edu.

[†]Dipartimento di Matematica "Tullio Levi-Civita" Università degli Studi di Padova, Via Trieste 63, 35121 Padova, Italy. Email: francesco.rossi@math.unipd.it

[‡]Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France. Email: mtournus@math.cnrs.fr.

on the Wasserstein distance and its Benamou-Brenier formulation, are introduced in [9, 15, 16]. See [8] for a unifying framewok for unbalanced optimal transport.

Such generalizations still have a drawback: both measures need to be positive. In the present paper we introduce a norm, parametrized by two positive numbers $(a, b)$, on the space of signed Radon measures with finite mass on $\mathbb{R}^d$. Such norm, based on an optimal transportation approach, induces a distance generalizing the Wasserstein distance to signed measures. We then prove that for $(a, b) = (1, 1)$ this norm corresponds to the extension of the so-called Kantorovich distance or Bounded-Lipschitz norm (BL norm) for finite signed Radon measures presented in [13] in the dual form

$$\|\mu\|_{BL} = \sup_{\|f\|_\infty \leq 1,\ \|f\|_{Lip} \leq 1} \int_{\mathbb{R}^d} f d\mu. \tag{1}$$

The first novelty brought by our paper is the statement of Theorem (1) and thus lies in the dual interpretation of the BL norm in the framework of optimal transportation. This new formulation enables us to prove new topological properties and characterizations of the BL norm.

The second main contribution of the paper is the statement of Theorem (2) and lies in the use of the $(a, b)$ norm to guarantee well-posedness of the following non local transport equation

$$\partial_t \mu_t(x) + \operatorname{div}(v[\mu_t](x)\mu_t(x)) = h[\mu_t](x), \qquad \mu_{|t=0}(x) = \mu_0(x), \tag{2}$$

for $x \in \mathbb{R}^d$ and $\mu_0 \in \mathcal{M}^s(\mathbb{R}^d)$, $h[\mu] \in \mathcal{M}^s(\mathbb{R}^d)$, where $\mathcal{M}^s(\mathbb{R}^d)$ is the space of signed Radon measures with finite mass on $\mathbb{R}^d$. Equation (2) has already been studied in the framework of positive measures, where it has been used for modeling several different phenomena, such as crowd motion and cell development in biology; see a review in [21]. From the modeling point of view, one of the interests of signed measures is that they can be used to model phenomena for which the measures under study are intrinsically signed. For instance, in a model coming from the hydrodynamic equations of Ginzburg-Landau vortices, the vortex density $\mu_t$ (which can be positive or negative depending on the local topological degree) in a domain occupied by a superconducting sample satisfies (2) with $h[\mu_t] = 0$, where $v[\mu_t]$ is the magnetic field induced in the sample (see [3, 17]).

Another motivation to study equation (2) in the framework of signed measures is the interpretation of $\mu_t$ as the spatial derivative of the entropy solution $\rho(x, t)$ to a scalar conservation law. A link between scalar conservation laws and non local transport equation has been initiated in [5, 14], but until now, studies are restricted to convex fluxes and monotonous initial conditions, so that the spatial derivative $\mu_t$ is a positive measure for all $t > 0$. To deal with generic scalar conservation laws, one needs a space of signed measures equipped with a metric of Wasserstein type, see e.g. [4].

Motivated by Ginzburg-Landau vortices, the authors of [3] suggested to extend the usual Wasserstein distance $W_1$ to the couples of signed measures $\mu = \mu^+ - \mu^-$ and $\nu = \nu^+ - \nu^-$ such that $|\mu^+| + |\nu^-| = |\mu^-| + |\nu^+|$ by the formula $\mathbb{W}_1(\mu, \nu) = W_1(\mu^+ + \nu^-, \mu^- + \nu^+)$. This procedure fails for $p \neq 1$, since triangular inequality is lost. A counter-example to the triangular inequality is provided in [3] for $d = 1$ and $p = 2$: taking $\mu = \delta_0$, $\nu = \delta_4$, $\eta = \delta_1 - \delta_2 + \delta_3$, we obtain $\mathbb{W}_2(\mu, \nu) = 4$ whereas $\mathbb{W}_2(\mu, \eta) + \mathbb{W}_2(\eta, \nu) = \sqrt{2} + \sqrt{2}$.

We use the same trick from [3] to turn the generalized Wasserstein distance $W_1^{a,b}$ into a norm for signed measures, by setting

$$\|\mu\|^{a,b} := W_1^{a,b}(\mu^+, \mu^-) = \inf_{\substack{\tilde{\eta}, \tilde{\nu} \in \mathcal{M}(\mathbb{R}^d) \\ |\tilde{\eta}| = |\tilde{\nu}|}} \left( a \left( |\mu^+ - \tilde{\eta}| + |\mu^- - \tilde{\nu}| \right) + b W_1(\tilde{\eta}, \tilde{\nu}) \right),$$

where $\mu^+, \mu^-$ are any positive finite measures such that $\mu = \mu^+ - \mu^-$, and where $|\mu|$ is the total variation of $\mu$ For the reason mentioned above, this construction only defines a norm for $p = 1$.

Notice that we need to restrict ourselves to Radon measures $\mu$ with finite mass.

The regularity assumptions on the vector field and on the source term are the following:

(H-1) There exists $K$ such that for all $\mu, \nu \in \mathcal{M}^s(\mathbb{R}^d)$ it holds

$$\|v[\mu] - v[\nu]\|_{\mathcal{C}^0(\mathbb{R}^d)} \leq K\|\mu - \nu\|^{a,b}. \tag{3}$$

(H-2) There exist $L, M$ such that for all $x, y \in \mathbb{R}^d$, for all $\mu \in \mathcal{M}^s(\mathbb{R}^d)$ it holds

$$|v[\mu](x) - v[\mu](y)| \leq L|x - y|, \qquad |v[\mu](x)| \leq M. \tag{4}$$

(H-3) There exist $Q, P, R$ such that for all $\mu, \nu \in \mathcal{M}^s(\mathbb{R}^d)$ it holds

$$\|h[\mu] - h[\nu]\|^{a,b} \leq Q\|\mu - \nu\|^{a,b}, \qquad |h[\mu]| \leq P, \qquad \mathrm{supp}(h[\mu]) \subset B_0(R). \tag{5}$$

The main results of the paper are the following.

**Theorem 1** (Kantorovich formulation for the BL norm)**.** *For $\mu \in \mathcal{M}^s(\mathbb{R}^d)$,*

$$\|\mu\|_{BL} = \|\mu\|^{1,1} = \inf_{\substack{\tilde{\eta},\tilde{\nu} \in \mathcal{M}(\mathbb{R}^d) \\ |\tilde{\eta}| = |\tilde{\nu}|}} \left(|\mu^+ - \tilde{\eta}| + |\mu^- - \tilde{\nu}| + W_1(\tilde{\eta}, \tilde{\nu})\right),$$

*where $\mu^+, \mu^-$ are any measures of $\mathcal{M}(\mathbb{R}^d)$ such that $\mu = \mu^+ - \mu^-$.*

**Theorem 2** (Existence, uniqueness and stability in $(\mathcal{M}^s(\mathbb{R}^d), \|.\|^{a,b})$)**.** *Let $v$ and $h$ satisfy (H-1)-(H-2)-(H-3) and $\mu_0 \in \mathcal{M}^s(\mathbb{R}^d)$ compactly supported be given. Then, there exists a unique distributional solution to* (2)*. In addition, for $\mu_0$ and $\nu_0$ in $\mathcal{M}^s(\mathbb{R}^d)$, denoting by $\mu_t$ and $\nu_t$ the corresponding solutions, we have the following property for $t \in [0, 1]$ of continuous dependence with respect to initial data:*

$$\|\mu_t - \nu_t\|^{a,b} \leq \|\mu_0 - \nu_0\|^{a,b} \exp(K_1 t), \qquad K_1 = 2L + 2bK(P + \min\{|\mu_0|, |\nu_0|\}) + Q,$$

*the following estimates on the mass and support:*

$$|\mu_t| \leq |\mu_0| + Pt, \qquad supp\{\mu_t\} \subset B(0, R' + tM) \;\; \text{for } R' \text{ such that } \; (supp\{\mu_0\} \cup B_0(R)) \subset B_0(R').$$

*Moreover, the solution is Lipschitz in time:*

$$\|\mu_{t+\tau} - \mu_t\|^{a,b} \leq K_2\tau, \quad K_2 = aP + bM(P + |\mu_0|), \quad \tau \geq 0, \qquad t + \tau \leq 1.$$

**Remark 3.** *We emphasize that the assumptions (H-2)-(H-3) are incompatible with a direct interpretation of the solution of* (2) *as the spatial derivative of a conservation law and need to be relaxed in a future work. Indeed, to draw a parallel between conservation laws and non-local equations, discontinuous vector fields need to be considered.*

The structure of the article is the following. In Section 2, we state and prove results of measure theory, which are needed for the rest of the paper. We also recall the definition of generalized Wasserstein distance. In Section 3, we to define the norm $\|\cdot\|^{a,b}$ on the space of signed Radon measures with finite mass. We state a Kantorovich-Rubinstain type duality for this norm, which proves Theorem 1. We end Section 3 by proving some topological properties for this norm. Section 4 is devoted the proof of Theorem 2.

## 2 Measure theory and the Generalized Wasserstein distance

In this section, we introduce the notations and state preliminary results. Throughout the paper, $\mathcal{B}(\mathbb{R}^d)$ is the space of Borel sets on $\mathbb{R}^d$, $\mathcal{M}(\mathbb{R}^d)$ is the space of Radon measures with finite mass (i.e. Borel regular, positive, and finite on every set).

### 2.1 Reminders on measure theory

In this section, $\mu$ and $\nu$ are in $\mathcal{M}(\mathbb{R}^d)$.

**Definition 4.** *We say that*

- $\mu << \nu$ *if* $\forall A \in \mathcal{B}(\mathbb{R}^d),\ (\nu(A) = 0) \Rightarrow (\mu(A) = 0)$
- $\mu \leq \nu$ *if* $\forall A \in \mathcal{B}(\mathbb{R}^d),\ \mu(A) \leq \nu(A)$
- $\mu \perp \nu$ *if there exists* $E \in \mathcal{B}(\mathbb{R}^d)$ *such that* $\mu(\mathbb{R}^d) = \mu(E)$ *and* $\nu(E) = 0$

The concept of largest common measure between measures is now recalled.

**Lemma 5.** *We consider $\mu$ and $\nu$ two measures in $\mathcal{M}(\mathbb{R}^d)$. Then, there exists a unique measure $\mu \wedge \nu$ which satisfies*

$$\mu \wedge \nu \leq \mu, \quad \mu \wedge \nu \leq \nu, \quad (\eta \leq \mu \text{ and } \eta \leq \nu) \Rightarrow \eta \leq \mu \wedge \nu. \tag{6}$$

*We refer to $\mu \wedge \nu$ as the largest common measure to $\mu$ and $\nu$. Moreover, denoting by $f$ the Radon Nikodym derivative of $\mu$ with respect to $\nu$, i.e. the unique measurable function $f$ such that $\mu = f\nu + \nu_\perp$, with $\nu_\perp \perp \nu$, we have*

$$\mu \wedge \nu = \min\{f, 1\}\nu. \tag{7}$$

*Proof.* The uniqueness is clear using (6). Existence is given by formula (7) as follows. First, it is obvious that $\min\{f, 1\}\nu \leq \nu$ and using $\mu = f\nu + \nu_\perp$, it is also clear that $\min\{f, 1\}\nu \leq \mu$. Let us now assume by contradiction the existence of a measure $\eta$ and of $A \in \mathcal{B}(\mathbb{R}^d)$ such that

$$\eta \leq \mu, \quad \eta \leq \nu, \quad \eta(A) > \int_A \min\{f, 1\} d\nu. \tag{8}$$

Since $\nu_\perp \perp \nu$, there exists $E \in \mathcal{B}(\mathbb{R}^d)$ such that $\nu(\mathbb{R}^d) = \nu(E)$ and $\nu_\perp(E) = 0$ (see Definition 4), then $\nu(A) = \nu(A \cap E)$ and $\nu_\perp(A) = \nu_\perp(A \cap E^c)$. Since $\eta \leq \nu$, we have $\eta(A \cap E^c) = 0$, thus $\eta(A \cap E) = \eta(A)$, and using (8)

$$\eta(A \cap E) > \int_{A \cap E} \min\{f, 1\} d\nu.$$

We define

$$B = A \cap E \cap \{f > 1\}.$$

Then

$$\begin{aligned}
\eta(B) + \eta((A \cap E) \setminus B) &= \eta(A \cap E) \\
&> \int_{A \cap E} \min\{f, 1\} d\nu(x) \\
&= \int_B \min\{f, 1\} d\nu(x) + \int_{(A \cap E) \setminus B} \min\{f, 1\} d\nu \\
&= \int_B 1 d\nu + \int_{(A \cap E) \setminus B} f d\nu \\
&= \nu(B) + \mu((A \cap E) \setminus B)
\end{aligned}$$

which contradicts the fact that both $\eta \leq \nu$ and $\eta \leq \mu$. This implies that $\eta$ satisfying (8) does not exist, and then Lemma 5 holds. □

### 2.2 Signed measures

We now introduce signed Radon measures, that are measures $\mu$ that can be written as $\mu = \mu_+ - \mu_-$ with $\mu_+, \mu_- \in \mathcal{M}(\mathbb{R}^d)$. We denote with $\mathcal{M}^s(\mathbb{R}^d)$ the space of such signed Radon measures.

For $\mu \in \mathcal{M}^s(\mathbb{R}^d)$, we define $|\mu| = |\mu_+^J| + |\mu_-^J|$ where $(\mu_+^J, \mu_-^J)$ is the unique Jordan decomposition of $\mu$, i.e. $\mu = \mu_+^J - \mu_-^J$ with $\mu_+^J \perp \mu_-^J$. Observe that $|\mu|$ is always finite, since $\mu_+^J, \mu_-^J \in \mathcal{M}(\mathbb{R}^d)$.

**Definition 6** (Push-forward)**.** *For $\mu \in \mathcal{M}^s(\mathbb{R}^d)$ and $T : \mathbb{R}^d \to \mathbb{R}^d$ a Borel map, the push-forward $T\#\mu$ is the measure on $\mathbb{R}^d$ defined by $T\#\mu(B) = \mu(T^{-1}(B))$ for any Borel set $B \subset \mathbb{R}^d$.*

We now remind the definition of tightness for a sequence in $\mathcal{M}^s(\mathbb{R}^d)$.

**Definition 7** (Tightness)**.** *A sequence $(\mu_n)_{n\in\mathbb{N}}$ of measures in $\mathcal{M}(\mathbb{R}^d)$ is tight if for each $\varepsilon > 0$, there is a compact set $K \subset \mathbb{R}^d$ such that for all $n \geq 0$, $\mu_n(\mathbb{R}^d \setminus K) < \varepsilon$. A sequence $(\mu_n)_{n\in\mathbb{N}}$ of signed measures of $\mathcal{M}^s(\mathbb{R}^d)$ is tight if the sequences $(\mu_n^+)_{n\in\mathbb{N}}$ and $(\mu_n^-)_{n\in\mathbb{N}}$ given by the Jordan decomposition are both tight.*

For a sequence of probability measures, different notions of weak convergences are equivalent. It is not the case for signed measures and we precise here what we call narrow and vague convergence. In the present paper, $\mathcal{C}^0(\mathbb{R}^d;\mathbb{R})$ is the set of continuous functions, $\mathcal{C}_b^0(\mathbb{R}^d;\mathbb{R})$ is the set of bounded continuous functions, and $\mathcal{C}_c^0(\mathbb{R}^d;\mathbb{R})$ is the set of continuous functions with compact support on $\mathbb{R}^d$.

**Definition 8** (Narrow and vague convergence for signed measures)**.**

*A sequence $(\mu_n)_{n\in\mathbb{N}}$ of measures in $\mathcal{M}^s(\mathbb{R}^d)$ is said to converge narrowly to $\mu$ if the following holds: for all $\varphi \in C_b^0(\mathbb{R}^d;\mathbb{R})$, $\int_{\mathbb{R}^d} \varphi(x) d\mu_n(x) \to \int_{\mathbb{R}^d} \varphi(x) d\mu(x)$.*

*A sequence $(\mu_n)_{n\in\mathbb{N}}$ of measures in $\mathcal{M}^s(\mathbb{R}^d)$ is said to converge vaguely to $\mu$ if the following holds: for all $\varphi \in C_c^0(\mathbb{R}^d;\mathbb{R})$, $\int_{\mathbb{R}^d} \varphi(x) d\mu_n(x) \to \int_{\mathbb{R}^d} \varphi(x) d\mu(x)$.*

**Lemma 9** (Weak compactness for positive measures)**.** *Let $\mu_n$ be a sequence of measures in $\mathcal{M}(\mathbb{R}^d)$ that are uniformly bounded in mass. We can then extract a subsequence $\mu_{\phi(n)}$ such that $\mu_{\phi(n)}$ converges vaguely to $\mu$ for some $\mu \in \mathcal{M}(\mathbb{R}^d)$.*

A proof can be found in [10, Theorem 1.41]. Notice that in [10], vague convergence is called weak convergence. In [13, 22] however, weak convergence refers to what we define here as narrow convergence. Notice that if a sequence of positive measures $\mu_n$ converges vaguely to $\mu$ and if $(\mu_n)_n$ is tight, then $\mu_n$ converges narrowly to $\mu$.

### 2.3 Properties of the generalized Wasserstein distance

In this section, we recall key properties of the standard and generalized Wasserstein distance. For more details on these topics, see [22, 20].

**Definition 10** (Transference plan)**.** *A transference plan between two positive measures $\mu$ and $\nu$ of same mass is a measure $\pi \in \mathcal{M}(\mathbb{R}^d, \mathbb{R}^d)$ which satisfies for all $A, B \in \mathcal{B}(\mathbb{R}^d)$*

$$\pi(A \times \mathbb{R}^d) = \mu(A), \quad \pi(\mathbb{R}^d \times B) = \nu(B).$$

Note that transference plans are *not* probability measures in general, as their mass is $|\mu| = |\nu|$, the common mass of both marginals. We denote by $\Pi(\mu, \nu)$ the set of transference plans between $\mu$ and $\nu$. The p-Wasserstein distance for positive Radon measures of same mass is defined for $p \geq 1$ as

$$W_p(\mu, \nu) = \left( \min_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\pi(x, y) \right)^{\frac{1}{p}}.$$

It was extended to positive measures having possibly different mass in [19, 20], where the authors introduce the distance $W_p^{a,b}$ on the space $\mathcal{M}(\mathbb{R}^d)$ of Radon measures with finite mass. The formal definition is the following.

**Definition 11** (Generalized Wasserstein distance [19]). *Let $\mu, \nu$ be two positive measures in $\mathcal{M}(\mathbb{R}^d)$. The generalized Wasserstein distance between $\mu$ and $\nu$ is given for $p \geq 1$, $a > 0$ and $b > 0$ by*

$$W_p^{a,b}(\mu, \nu) = \left( \inf_{\substack{\tilde{\mu}, \tilde{\nu} \in \mathcal{M}(\mathbb{R}^d) \\ |\tilde{\mu}| = |\tilde{\nu}|}} a^p (|\mu - \tilde{\mu}| + |\nu - \tilde{\nu}|)^p + b^p W_p^p(\tilde{\mu}, \tilde{\nu}) \right)^{1/p}. \tag{9}$$

**Proposition 12** (Scaling and dilation formulae for the generalized Wasserstein distance). *Consider $a > 0$ and $b > 0$ and let $\mu, \nu$ be two measures.*

1. *The following scaling formula holds for $p \geq 1$*

$$W_p^{\lambda a, \lambda b}(\mu, \nu) = \lambda W_p^{a,b}(\mu, \nu), \quad \lambda > 0. \tag{10}$$

2. *Define $D_\ell : x \to \ell x$ with $\ell > 0$ the dilation in $\mathbb{R}^n$, then it holds*

$$W_1^{a,b}(D_\ell \# \mu, D_\ell \# \nu) = W_1^{a, \ell b}(\mu, \nu). \tag{11}$$

*Proof.* The first statement is directly deduced from Definition 11. For the second statement, define

$$C^{a,b}(\bar{\mu}, \bar{\nu}, \pi; \mu, \nu) := a(|\mu - \bar{\mu}| + |\nu - \bar{\nu}|) + b \int |x - y| \, d\pi(x, y),$$

where $\pi$ is a transference plan in $\Pi(\bar{\mu}, \bar{\nu})$. It holds

$$\begin{aligned} C^{a,b}(D_\ell \# \bar{\mu}, D_\ell \# \bar{\nu}, &(D_\ell \times D_\ell) \# \pi; D_\ell \# \mu, D_\ell \# \nu) \\ &= a(|D_\ell \# \mu - D_\ell \# \bar{\mu}| + |D_\ell \# \nu - D_\ell \# \bar{\nu}|) + b \int |x - y| \, d(D_\ell \times D_\ell)\pi(x, y), \\ &= a(|\mu - \bar{\mu}| + |\nu - \bar{\nu}|) + b \int |\ell x - \ell y| \, d\pi(x, y) = C^{a, \ell b}(\bar{\mu}, \bar{\nu}, \pi; \mu, \nu). \end{aligned}$$

As a consequence, (11) holds. □

Notice that the first statement of Proposition 12 implies in particular that

$$W_p^{a,b} = \frac{b}{b'} W_p^{a',b'}, \quad \text{for } \frac{a}{b} = \frac{a'}{b'}. \tag{12}$$

The following lemma is useful to derive properties for the generalized Wasserstein distance.

**Lemma 13.** *The infimum in (9) is always attained. Moreover, there always exists a minimizer that satisfy the additional constraint $\tilde{\mu} \leq \mu$, $\tilde{\nu} \leq \nu$.*

The proof can be found in [19].

For $f \in \mathcal{C}_c^0(\mathbb{R}^d;\mathbb{R})$, we define

$$\|f\|_\infty = \sup_{x\in\mathbb{R}^d} |f(x)|, \qquad \|f\|_{Lip} = \sup_{x\neq y} \frac{|f(x)-f(y)|}{|x-y|}.$$

We also denote by $\mathcal{C}_c^{0,Lip}(\mathbb{R}^d;\mathbb{R})$ the subset of functions $f \in \mathcal{C}_c^0(\mathbb{R}^d;\mathbb{R})$ for which it holds $\|f\|_{Lip} < +\infty$.

The following result is stated in [20, Theorem 13].

**Lemma 14** (Kantorovitch-Rubinstein duality for $W_1^{1,1}$)**.** *For $\mu$, $\nu$ in $\mathcal{M}(\mathbb{R}^d)$, it holds*

$$W_1^{1,1}(\mu,\nu) = \sup\left\{ \int_{\mathbb{R}^d} \varphi \, d(\mu-\nu);\ \varphi \in \mathcal{C}_c^{0,Lip}, \|\varphi\|_\infty \le 1, \|\varphi\|_{Lip} \le 1 \right\}.$$

**Lemma 15** (Properties of the generalized Wasserstein distance)**.** *Let $\mu,\nu,\eta,\mu_1,\mu_2,\nu_1,\nu_2$ be some positive measures in $\mathcal{M}(\mathbb{R}^d)$. The following properties hold for $p \ge 1$, $a > 0$ and $b > 0$*

1. $W_p^{a,b}(\mu_1+\mu_2,\nu_1+\nu_2) \le W_p^{a,b}(\mu_1,\nu_1) + W_p^{a,b}(\mu_2,\nu_2)$,
2. $W_1^{a,b}(\mu+\eta,\nu+\eta) = W_1^{a,b}(\mu,\nu)$.

*Proof.* The first property is taken from [19, Proposition 11]. For $a = b = 1$, the second statement is a direct consequence of the Kantorovitch-Rubinstein duality in Lemma 14 for $W_1^{1,1}$. For general $a > 0$, $b > 0$, we use the results of Proposition 12. Indeed, also applying Kantorovich-Rubinstein for $W_1^{1,1}$ and setting $a' = 1, b' = \ell = \frac{b}{a}$, it holds

$$\begin{aligned} &W_1^{a,b}(\mu+\eta,\nu+\eta) = aW_1^{1,\frac{b}{a}}(\mu+\eta,\nu+\eta) = aW_1^{1,1}(D_\ell\#\mu + D_\ell\#\eta, D_\ell\#\nu + D_\ell\#\eta) = \\ &= aW_1^{1,1}(D_\ell\#\mu, D_\ell\#\nu) = aW_1^{1,\frac{b}{a}}(\mu,\nu) = W_1^{a,b}(\mu,\nu). \end{aligned}$$

□

**Definition 16** (Image of a measure under a plan)**.** *Let $\mu$ and $\nu$ two measures in $\mathcal{M}(\mathbb{R}^d)$ of same mass and $\pi \in \Pi(\mu,\nu)$. For $\eta \le \mu$, we denote by $f$ the Radon-Nikodym derivative of $\eta$ with respect to $\mu$ and by $\pi_f$ the transference plan defined by $\pi_f(x,y) = f(x)\pi(x,y)$. Then, we define the image of $\eta$ under $\pi$ as the second marginal $\eta'$ of $\pi_f$.*

Observe that the second marginal satisfies $\eta' \le \nu$. Indeed, since $\eta \le \mu$, it holds $f \le 1$. Thus, for all Borel set $B$ of $\mathbb{R}^d$ we have

$$\eta'(B) = \pi_f(\mathbb{R}^d \times B) \le \pi(\mathbb{R}^d \times B) = \nu(B).$$

# 3 Generalized Wasserstein norm for signed measures

In this section, we define the generalized Wasserstein norm for signed measures and prove Theorem (1).

**Definition 17** (Generalized Wasserstein distance extended to signed measures)**.** *For $\mu,\nu$ two signed measures in $\mathcal{M}^s(\mathbb{R}^d)$, we define*

$$\mathbb{W}_1^{a,b}(\mu,\nu) = W_1^{a,b}(\mu_+ + \nu_-, \mu_- + \nu_+),$$

*where $\mu_+,\mu_-,\nu_+$ and $\nu_-$ are any measures in $\mathcal{M}(\mathbb{R}^d)$ such that $\mu = \mu_+ - \mu_-$ and $\nu = \nu_+ - \nu_-$.*

**Proposition 18.** *The operator* $\mathbb{W}_1^{a,b}$ *is a distance on the space* $\mathcal{M}^s(\mathbb{R}^d)$ *of signed measures with finite mass on* $\mathbb{R}^d$.

*Proof.* First, we prove that the definition does not depend on the decomposition. Indeed, if we consider two distinct decompositions, $\mu = \mu_+ - \mu_- = \mu_+^J - \mu_-^J$, and $\nu = \nu_+ - \nu_- = \nu_+^J - {\nu_-}^J$, with the second one being the Jordan decomposition, then we have $(\mu_+ + \nu_-) - (\mu_+^J + \nu_-^J) = (\mu_- + \nu_+) - (\mu_-^J + \nu_+^J)$, and this is a positive measure since $\mu_+ \geq \mu_+^J$ and $\nu_+ \geq \nu_+^J$. The second property of Lemma 15 then gives

$$
\begin{aligned}
&W_1^{a,b}(\mu_+^J + \nu_-^J, \mu_-^J + \nu_+^J) = \\
&W_1^{a,b}(\mu_+^J + \nu_-^J + (\mu_+ + \nu_-) - (\mu_+^J + \nu_-^J), \mu_-^J + \nu_+^J + (\mu_- + \nu_+) - (\mu_-^J + \nu_+^J)) = \\
&W_1^{a,b}(\mu_+ + \nu_-, \mu_- + \nu_+).
\end{aligned}
$$

We now prove that $\mathbb{W}_1^{a,b}(\mu, \nu) = 0$ implies $\mu = \nu$. By choosing the Jordan decomposition for both $\mu$ and $\nu$ and observing that $W_1^{a,b}$ is a distance, we obtain $\mu_+ + \nu_- = \mu_- + \nu_+$, thus $\mu = \nu$.

We now prove the triangle inequality. We have $\mathbb{W}_1^{a,b}(\mu, \eta) = W_1^{a,b}(\mu_+ + \eta_-, \mu_- + \eta_+)$. Using Lemma 15, we have

$$
\begin{aligned}
\mathbb{W}_1^{a,b}(\mu, \eta) &= W_1^{a,b}(\mu_+ + \eta_- + \nu_+ + \nu_-, \mu_- + \eta_+ + \nu_+ + \nu_-) \\
&\leq W_1^{a,b}(\mu_+ + \nu_-, \mu_- + \nu_+) + W_1^{a,b}(\eta_- + \nu_+, \eta_+ + \nu_-) \\
&= \mathbb{W}_1^{a,b}(\mu, \nu) + \mathbb{W}_1^{a,b}(\nu, \eta).
\end{aligned}
$$

□

We also state the following lemma about adding and removing masses.

**Lemma 19.** *Let* $\mu, \nu, \eta, \mu_1, \mu_2, \nu_1, \nu_2$ *be in* $\mathcal{M}^s(\mathbb{R}^d)$*, then the following properties hold*

- $\mathbb{W}_1^{a,b}(\mu + \eta, \nu + \eta) = \mathbb{W}_1^{a,b}(\mu, \nu)$,
- $\mathbb{W}_1^{a,b}(\mu_1 + \mu_2, \nu_1 + \nu_2) \leq \mathbb{W}_1^{a,b}(\mu_1, \nu_1) + \mathbb{W}_1^{a,b}(\mu_2, \nu_2)$.

*Proof.* The proof is direct. For the first item, it holds

$$
\begin{aligned}
\mathbb{W}_1^{a,b}(\mu + \eta, \nu + \eta) &= W_1^{a,b}([\mu_+ + \eta_+] + [\nu_- + \eta_-], [\nu_+ + \eta_+] + [\mu_- + \eta_-]) \\
&= W_1^{a,b}(\mu_+ + \nu_- + [\eta_+ + \eta_-], \nu_+ + \mu_- + [\eta_+ + \eta_-])
\end{aligned}
$$

which by Lemma 15 is equal to $W_1^{a,b}(\mu_+ + \nu_-, \mu_- + \nu_+) = \mathbb{W}_1^{a,b}(\mu, \nu)$.

For the second item, it holds

$$
\begin{aligned}
\mathbb{W}_1^{a,b}(\mu_1 + \mu_2, \nu_1 + \nu_2) &= W_1^{a,b}(\mu_{1,+} + \mu_{2,+} + \nu_{1,-} + \nu_{2,-}, \nu_{1,+} + \nu_{2,+} + \mu_{1,-} + \mu_{2,-}) \\
&\leq W_1^{a,b}(\mu_{1,+} + \nu_{1,-}, \nu_{1,+} + \mu_{1,-}) + W_1^{a,b}(\mu_{2,+} + \nu_{2,-}, \nu_{2,+} + \mu_{2,-}) \\
&= \mathbb{W}_1^{a,b}(\mu_1, \nu_1) + \mathbb{W}_1^{a,b}(\mu_2, \nu_2),
\end{aligned}
$$

where the inequality comes from Lemma 15. □

**Definition 20.** *For* $\mu \in \mathcal{M}^s(\mathbb{R}^d)$ *and* $a > 0,\ b > 0$*, we define*

$$
\|\mu\|^{a,b} = \mathbb{W}_1^{a,b}(\mu, 0) = W_1^{a,b}(\mu_+, \mu_-),
$$

*where* $\mu_+$ *and* $\mu_-$ *are any measures of* $\mathcal{M}(\mathbb{R}^d)$ *such that* $\mu = \mu_+ - \mu_-$.

It is clear that the definition of $\|\mu\|^{a,b}$ does not depend on the choice of $\mu_+, \mu_-$ as a consequence of the corresponding property for $W_1^{a,b}$.

**Proposition 21.** *The space of signed measures* $(\mathcal{M}^s(\mathbb{R}^d), \|.\|^{a,b})$ *is a normed vector space.*

*Proof.* First, we notice that $\|\mu\|^{a,b} = 0$ implies that $W_1^{a,b}(\mu_+, \mu_-) = 0$, which is $\mu_+ = \mu_-$ so that $\mu = \mu_+ - \mu_- = 0$. For triangular inequality, using the second property of Lemma 19, we have that for $\mu, \eta \in \mathcal{M}^s(\mathbb{R}^d)$,

$$\|\mu + \eta\|^{a,b} = \mathbb{W}_1^{a,b}(\mu + \eta, 0) \leq \mathbb{W}_1^{a,b}(\mu, 0) + \mathbb{W}_1^{a,b}(\eta, 0) = \|\mu\|^{a,b} + \|\eta\|^{a,b}.$$

Homogeneity is obtained by writing for $\lambda > 0$, $\|\lambda\mu\|^{a,b} = \mathbb{W}_1^{a,b}(\lambda\mu, 0) = W_1^{a,b}(\lambda\mu_+, \lambda\mu_-)$ where $\mu = \mu_+ - \mu_-$. Using Lemma 14 combined with Definition 11 and the notation of Proposition 12 we have

$$\begin{aligned}
W_1^{a,b}(\lambda\mu_+, \lambda\mu_-) &= aW_1^{1,\frac{b}{a}}(\lambda\mu_+, \lambda\mu_-) \\
&= aW_1^{1,1}\left(D_{\frac{b}{a}}\#\lambda\mu_+, D_{\frac{b}{a}}\#\lambda\mu_-\right) \\
&= a \sup\left\{\int_{\mathbb{R}^d} \varphi \, d\left(D_{\frac{b}{a}}\#\lambda\mu_+, D_{\frac{b}{a}}\#\lambda\mu_-\right);\ \varphi \in \mathcal{C}_c^{0,Lip}, \|\varphi\|_\infty \leq 1, \|\varphi\|_{Lip} \leq 1\right\} \\
&= \lambda\ a \sup\left\{\int_{\mathbb{R}^d} \varphi \, d\left(D_{\frac{b}{a}}\#\mu_+, D_{\frac{b}{a}}\#\mu_-\right);\ \varphi \in \mathcal{C}_c^{0,Lip}, \|\varphi\|_\infty \leq 1, \|\varphi\|_{Lip} \leq 1\right\} \\
&= \lambda W_1^{a,b}(\mu_+, \mu_-).
\end{aligned}$$

□

We provide here an example that illustrates the competition between cancellation and transportation. This example is used later in the paper.

**Example 22.** *Take* $\mu = \delta_x - \delta_y$. *Then*

$$\|\mu\|^{a,b} = W_1^{a,b}(\delta_x, \delta_y) = \inf_{\substack{\tilde{\mu}, \tilde{\nu} \in \mathcal{M}(\mathbb{R}^d) \\ |\tilde{\mu}| = |\tilde{\nu}|}} \left\{a(|\delta_x - \tilde{\mu}| + |\delta_y - \tilde{\nu}|) + bW_1(\tilde{\mu}, \tilde{\nu})\right\}.$$

*Using Lemma 13, the minimum is attained and it can be written as* $\tilde{\mu} = \epsilon\delta_x$ *and* $\tilde{\nu} = \epsilon\delta_y$ *for some* $0 \leq \epsilon \leq 1$. *Then*

$$\|\mu\|^{a,b} = \min_{0 \leq \epsilon \leq 1} \left\{2a(1 - \epsilon) + b\epsilon|x - y|\right\}.$$

*The minimizers depend on the distance between* $x$ *and* $y$. *For* $b|x - y| < 2a$, *the minimum is attained for* $\epsilon = 1$ *and* $\|\mu\|^{a,b} = b|x - y|$. *In that case, we say that all the mass is transported. On the contrary, for* $b|x - y| > 2a$, *the minimum is attained for* $\epsilon = 0$ *and* $\|\mu\|^{a,b} = 2a$, *and we say that all the mass is cancelled (or removed). For* $b|x - y| = 2a$, *we can both transport and cancel.*

## 3.1 Topological properties

In this section, we study the topological properties of the norm introduced above. In particular, we prove that it admits a duality formula, that indeed coincides with (1). We first prove that the topology of $\|.\|^{a,b}$ does not depend on $a, b > 0$.

**Proposition 23.** *For* $a > 0$, $b > 0$, *the norm* $\|.\|^{a,b}$ *is equivalent to* $\|.\|^{1,1}$.

*Proof.* For $\mu \in \mathcal{M}^s(\mathbb{R}^d)$ denote by $(m_+^{a,b}, m_-^{a,b})$ the positive measures such that

$$\|\mu\|^{a,b} = a|\mu_+ - m_+^{a,b}| + a|\mu_- - m_-^{a,b}| + bW_1(m_+^{a,b}, m_-^{a,b}),$$

and similarly define $(m_+^{1,1}, m_-^{1,1})$. Their existence is guaranteed by Lemma 13. By definition of the minimizers, we have

$$\begin{aligned}\|\mu\|^{a,b} &= a|\mu_+ - m_+^{a,b}| + a|\mu_- - m_-^{a,b}| + bW_1(m_+^{a,b}, m_-^{a,b}) \\ &\leq a|\mu_+ - m_+^{1,1}| + a|\mu_- - m_-^{1,1}| + bW_1(m_+^{1,1}, m_-^{1,1}) \leq \max\{a,b\}\|\mu\|^{1,1},\end{aligned}$$

In the same way, we obtain

$$\min\{a,b\}\|\mu\|^{1,1} \leq \|\mu\|^{a,b} \leq \max\{a,b\}\|\mu\|^{1,1}.$$

□

We now give an equivalent Kantorovich-Rubinstein duality formula for the new distance. For $f \in \mathcal{C}_b^0(\mathbb{R}^d;\mathbb{R})$, similarly to $\mathcal{C}_c^0(\mathbb{R}^d;\mathbb{R})$, we define the following

$$\|f\|_\infty = \sup_{x\in\mathbb{R}^d} |f(x)|, \qquad \|f\|_{Lip} = \sup_{x\neq y} \frac{|f(x)-f(y)|}{|x-y|},$$

and we introduce

$$\mathcal{C}_b^{0,Lip} = \{f \in \mathcal{C}_b^0(\mathbb{R}^d;\mathbb{R}) \mid \|f\|_{Lip} < \infty\}.$$

In the next proposition, we express the Kantorovich-Rubinstein duality for the norm $\mathbb{W}_1^{1,1}$. This shows that $\mathbb{W}_1^{1,1}$ coincides with the Bounded Lipschitz norm introduced in [13], also called Fortet Mourier distance in [23].

**Proposition 24** (Kantorovitch-Rubinstein duality for $\mathbb{W}_1^{1,1}$)**.** *The signed generalized Wasserstein norm $\|.\|^{1,1}$ coincides with the bounded Lipschitz norm: for $\mu$ in $\mathcal{M}^s(\mathbb{R}^d)$, it holds*

$$\|\mu\|^{1,1} = \sup\left\{\int_{\mathbb{R}^d}\varphi\, d\mu;\ \varphi \in \mathcal{C}_b^{0,Lip}, \|\varphi\|_\infty \leq 1, \|\varphi\|_{Lip} \leq 1\right\}.$$

We emphasize that Proposition 24 does not coincide with Lemma 14, since it involves non-compactly supported test functions.

*Proof.* By using Lemma 14 we have

$$\|\mu\|^{1,1} = W_1^{1,1}(\mu_+,\mu_-) = \sup\left\{\int_{\mathbb{R}^d}\varphi\, d\mu;\ \varphi \in \mathcal{C}_c^{0,Lip}, \|\varphi\|_\infty \leq 1, \|\varphi\|_{Lip} \leq 1\right\}.$$

We denote by

$$S = \sup\left\{\int_{\mathbb{R}^d}\varphi\, d\mu;\ \varphi \in \mathcal{C}_b^{0,Lip}, \|\varphi\|_\infty \leq 1, \|\varphi\|_{Lip} \leq 1\right\}.$$

First observe that $S < +\infty$. Indeed, it holds $\int_{\mathbb{R}^d}\varphi\, d\mu \leq \|\varphi\|_\infty|\mu| < +\infty$. Denote with $\varphi_n$ a sequence of functions of $\mathcal{C}_b^{0,Lip}$ such that $\int_{\mathbb{R}^d}\varphi_n\, d\mu \to S$ as $n \to \infty$. Consider a sequence of functions $\rho_n$ in

$\mathcal{C}_c^{0,Lip}$ such that $\rho_n(x) = 1$ for $x \in B_0(n)$, $\rho_n(x) = 0$ for $x \notin B_0(n+1)$ and $\|\rho_n\|_\infty \le 1$. For the sequence $\psi_n = \varphi_n \rho_n$ of functions of $\mathcal{C}_c^{0,Lip}$, it holds

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \psi_n \, d\mu - S \right| &\le \left| \int_{\mathbb{R}^d} (\psi_n - \varphi_n) \, d\mu \right| + \left| \int_{\mathbb{R}^d} \varphi_n \, d\mu - S \right| \\ &\le 2 \left| \int_{\mathbb{R}^d \setminus B_0(n)} d\mu \right| + \left| \int_{\mathbb{R}^d} \varphi_n \, d\mu - S \right| \end{aligned}$$

since $\|\varphi_n\|_\infty \le 1$. The first term goes to zero with $n$, since $\mu$ being of finite mass is tight, and the second term goes to zero with $n$ by definition of $S$ and $\varphi_n$. Then

$$S = \sup \left\{ \int_{\mathbb{R}^d} \varphi \, d\mu; \; \varphi \in \mathcal{C}_c^{0,Lip}, \|\varphi\|_\infty \le 1, \|\varphi\|_{Lip} \le 1 \right\},$$

and Proposition 24 is proved. $\square$

This proves Theorem (1). In the rest of the section, we state topological properties for the norm $\|.\|^{a,b}$.

**Remark 25.** *We observe that a sequence $\mu_n$ of $\mathcal{M}^s(\mathbb{R})$ which satisfies $\|\mu_n\|^{a,b} \underset{n\to\infty}{\to} 0$ is not necessarily tight, and its mass is not necessarily bounded. For instance, we have that*

$$\nu_n = \delta_n - \delta_{n+\frac{1}{n}}$$

*is not tight, whereas it satisfies for $n$ sufficiently large*

$$\|\nu_n\|^{a,b} = \frac{b}{n} \underset{n\to\infty}{\to} 0.$$

*See Example 22 for the details of the calculation. Now take the sequence*

$$\mu_n = n \, \delta_{\frac{1}{n^2}} - n \, \delta_{-\frac{1}{n^2}}.$$

*As explained in Example 22, depending on the sign of $2a - \frac{2b}{n^2}$, we either cancel the mass or transport it. For $n$ large enough, $2a \ge \frac{2b}{n^2}$, so we transport the mass. Thus for $n$ sufficiently large it holds*

$$\|\mu_n\|^{a,b} = \frac{2bn}{n^2} \underset{n\to\infty}{\to} 0$$

*whereas $|\mu_n| = 2n$ is not bounded.*

**Remark 26.** *Norm $\|.\|^{1,1}$ does not metrize narrow convergence, contrarily to what is stated in [13]. Indeed, take $\mu_n = \delta_{\sqrt{2\pi n + \frac{\pi}{2}}} - \delta_{\sqrt{2\pi n + \frac{3\pi}{2}}}$. We have*

$$\|\mu_n\|^{1,1} \le \left| \sqrt{2\pi n + \frac{\pi}{2}} - \sqrt{2\pi n + \frac{3\pi}{2}} \right| \underset{n\to\infty}{\to} 0,$$

*even though for $\varphi(x) = \sin(x^2)$ in $\mathcal{C}_b^0(\mathbb{R})$, we have*

$$\int_{\mathbb{R}} \varphi d\mu_n = 2, \quad n \in \mathbb{N}.$$

**Remark 27.** *We have as a direct consequence of Proposition 24 that*

$$\|\mu_n - \mu\|^{a,b} \underset{n\to\infty}{\to} 0 \qquad \Rightarrow \quad \forall \varphi \in \mathcal{C}_b^{0,Lip}(\mathbb{R}^d),\ \int_{\mathbb{R}^d} \varphi d\mu_n \underset{n\to\infty}{\to} \int_{\mathbb{R}^d} \varphi d\mu. \tag{13}$$

*However, the reciprocal statement of* (13) *is false: define*

$$\mu_n := n\cos(nx)\chi_{[0,\pi]}.$$

*For*

$$\varphi_n := \frac{1}{n}\cos(nx),$$

*it is clear that*

$$\int_{\mathbb{R}} \varphi_n\, d\mu_n = \int_0^\pi \cos^2(nx)\, dx = \frac{\pi}{2} \not\to 0.$$

*In particular,*

$$\sup_{\varphi\in\mathcal{C}_b^{0,Lip}(\mathbb{R})} \int_{\mathbb{R}} \varphi\, d(\mu_n - 0) \geq \frac{\pi}{2},$$

*hence by Proposition 24,* $\|\mu_n - 0\| \geq \frac{\pi}{2}$ *does not converge to zero. We now prove that, for each* $\varphi$ *in* $\mathcal{C}_b^{0,Lip}(\mathbb{R})$*, it holds* $\int_{\mathbb{R}} \varphi\, d\mu_n \to 0$*. Given* $\varphi \in \mathcal{C}_b^{0,Lip}(\mathbb{R})$*, define*

$$f(x) := \begin{cases} \varphi(-x), & \text{when } \ x \in [-\pi, 0], \\ \varphi(x), & \text{when } \ x \in [0, \pi], \end{cases}$$

*and extend* $f$ *as a* $2\pi$*-periodic function on* $\mathbb{R}$*. We have*

$$\int_{\mathbb{R}} \varphi\, d\mu_n = \int_{\mathbb{R}} f\, d\mu_n.$$

*Since* $f$ *is a* $2\pi$*-periodic function, it also holds* $\int f\, d\mu_n = na_n$*, where* $a_n$ *is the* $n$*-th cosine coefficient in the Fourier series expansion of* $f$*. We then prove* $na_n \to 0$ *for any* $2\pi$*-periodic Lipschitz function* $f$*, following the ideas of [24, p. 46, last line]. Since* $f$ *is Lipschitz, then its distributional derivative is in* $L^\infty[-\pi,\pi]$ *and thus in* $L^1[-\pi,\pi]$*. Then*

$$a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\, dx = -\frac{1}{2n\pi}\int_{-\pi}^{\pi} f'(x)\sin(nx)\, dx = -\frac{b'_n}{n},$$

*where* $b'_n$ *is the* $n$*-th sine coefficient of* $f'$*. As a consequence of the Riemann-Lebesgue lemma,* $b'_n \to 0$*, and this implies* $na_n \to 0$*.*

We recall from [20] that the space $(\mathcal{M}(\mathbb{R}^d), W_p^{a,b})$ is a complete metric space. The proof is based on the fact that a Cauchy sequence of positive measures is both uniformly bounded in mass and tight. This is not true anymore for a Cauchy sequence of signed measures.

**Remark 28.** *Observe that* $(\mathcal{M}^s(\mathbb{R}^d), \|.\|^{a,b})$ *is not a Banach space. Indeed, take the sequence*

$$\mu_n = \sum_{i=1}^{n} \left(\delta_{i+\frac{1}{2^i}} - \delta_{i-\frac{1}{2^i}}\right).$$

*It is a Cauchy sequence in $(\mathcal{M}^s(\mathbb{R}^d), \|.\|^{a,b})$: indeed, by choosing to transport all the mass from $\mu_n^+ + \mu_{n+k}^-$ onto $\mu_{n+k}^+ + \mu_n^-$ with the cost $b$, it holds*

$$\mathbb{W}_1^{a,b}(\mu_n, \mu_{n+k}) \leq 2b \sum_{i=n+1}^{n+k} \frac{1}{2^i} \leq 2b \sum_{i=n+1}^{+\infty} \frac{1}{2^i} \underset{n\to\infty}{\to} 0.$$

*However, the sequence $(\mu_n)_n$ does not converge in $(\mathcal{M}^s(\mathbb{R}^d), \|.\|^{a,b})$. As seen in Remark 27, the convergence for the norm $\|.\|^{a,b}$ implies the convergence in the sense of distributions. In the sense of distributions we have*

$$\mu_n \rightharpoonup \mu^* := \sum_{i=1}^{+\infty} \left( \delta_{i+\frac{1}{2^i}} - \delta_{i-\frac{1}{2^i}} \right) \notin \mathcal{M}^s(\mathbb{R}).$$

*Indeed, for all $\varphi \in \mathcal{C}_c^\infty(\mathbb{R})$, since $\varphi$ is compactly supported, it holds*

$$\int_{\mathbb{R}} \varphi(x)\ (d\mu_n(x) - d\mu^*(x)) = \sum_{i=n+1}^{+\infty} \left( \varphi\left(i + \frac{1}{2^i}\right) - \varphi\left(i - \frac{1}{2^i}\right)\right) \underset{n\to\infty}{\to} 0.$$

*The measure $\mu^*$ does not belong to $\mathcal{M}^s(\mathbb{R})$, as it has infinite mass.*

Nevertheless, we have the following convergence result.

**Theorem 29.** *Let $\mu_n$ be a Cauchy sequence in $(\mathcal{M}^s(\mathbb{R}^d), \|.\|^{a,b})$. If $\mu_n$ is tight (in the sense of Definition 7) and has uniformly bounded mass, then it converges in $(\mathcal{M}^s(\mathbb{R}^d), \|.\|^{a,b})$.*

*Proof.* Take a tight Cauchy sequence $(\mu_n)_n \in \mathcal{M}^s(\mathbb{R}^d)$ such that the sequences given by the Jordan decomposition $|\mu_n^+|$ and $|\mu_n^-|$ are uniformly bounded. Then, by Lemma 9, there exists $\mu^+$ and $\mu^-$ in $\mathcal{M}(\mathbb{R}^d)$ and $\varphi_1$ non decreasing such that, $\mu^+_{\varphi_1(n)} \underset{n\to\infty}{\rightharpoonup} \mu^+$ vaguely. Then, $|\mu_n^-|$ being uniformly bounded, there exists $\varphi_2$ non decreasing such that for $\varphi = \varphi_1 \circ \varphi_2$ it holds

$$\mu^-_{\varphi(n)} \underset{n\to\infty}{\rightharpoonup} \mu^- \quad \text{vaguely.}$$

Since $\mu_n^+$ and $\mu_n^-$ are assumed to be tight, the sequences $\mu^-_{\varphi(n)}$ and $\mu^+_{\varphi(n)}$ also converge to $\mu^-$ and $\mu^+$ narrowly, and it holds $W_1^{a,b}(\mu^+_{\varphi(n)}, \mu^+) \underset{n\to\infty}{\to} 0$ and $W_1^{a,b}(\mu^-_{\varphi(n)}, \mu^-) \underset{n\to\infty}{\to} 0$, due to [19, Theorem 13]. We then have

$$\begin{aligned}
\|\mu_n - (\mu^+ - \mu^-)\|^{a,b} \leq & \|\mu_n - \mu_{\varphi(n)}\|^{a,b} + \|\mu_{\varphi(n)} - (\mu^+ - \mu^-)\|^{a,b} \\
& \leq \|\mu_n - \mu_{\varphi(n)}\|^{a,b} + W_1^{a,b}(\mu^+_{\varphi(n)} + \mu^-, \mu^-_{\varphi(n)} + \mu^+) \\
& \leq \|\mu_n - \mu_{\varphi(n)}\|^{a,b} + W_1^{a,b}(\mu^+_{\varphi(n)}, \mu^+) + W_1^{a,b}(\mu^-_{\varphi(n)}, \mu^-) \underset{n\to\infty}{\to} 0.
\end{aligned}$$

Here, we used the fact that $(\mu_n)_n$ is a Cauchy sequence. □

We end this section with a characterization of the convergence for the norm. If a sequence $\mu_n$ of signed measures converges toward $\mu \in \mathcal{M}^s(\mathbb{R}^d)$, then for any decomposition of $\mu_n$ into two positive measures $\mu_n = \mu_n^+ - \mu_n^-$ (not necessarily the Jordan decomposition), we have that each $\mu_n^+, \mu_n^-$ is the sum of two positive measures: $m_n^+, z_n^+$ and $m_n^-, z_n^-$, respectively. The measures $m_n^+$ and $m_n^-$ are the parts that converge respectively to $\mu^+$ and $\mu^-$. Both $m_n^+$ and $m_n^-$ are uniformly bounded and tight. The measures $z_n^+$ and $z_n^-$ are the residual terms that may be unbounded and not tight. They compensate each other in the sense that $W_1^{a,b}(z_n^+, z_n^-)$ vanishes for large $n$.

# 4 Application to the transport equation with source term

This section is devoted to the use of the norm introduced in Definition 20 to guarantee existence, uniqueness, and stability with respect to initial condition for the transport equation (2). We denote the set of test functions (i.e. $C^\infty$ with compact support) on a given space $X$ by $\mathcal{D}(X)$.

**Definition 30** (Measure-valued weak solution)**.** *A measure-valued weak solution to* (2) *is a continuous map with respect to the weak-* topology of measures (i.e $\mu \in \mathcal{C}^0([0,1];\mathcal{M}^s(\mathbb{R}^d))$ such that for all $\varphi \in \mathcal{D}([0,1) \times \mathbb{R}^d)$ it holds*

$$\int_0^1 dt \int_{\mathbb{R}^d} (d\mu_t(\partial_t\varphi(t,x) + v[\mu_t]\cdot\nabla\varphi(t,x)) + dh[\mu_t]\,\varphi(t,x)) = -\int_{\mathbb{R}^d} d\mu_0\varphi(0,\cdot). \tag{14}$$

*Equivalently, $\mu$ satisfies $\mu_{t=0} = \mu_0$ and for all $\varphi \in \mathcal{D}(\mathbb{R}^d)$ it holds*

$$\frac{d}{dt}\langle\mu_t,\varphi\rangle = \langle\mu_t, v[\mu_t]\cdot\nabla\varphi\rangle + \langle h[\mu_t],\varphi\rangle, \tag{15}$$

*where*

$$\langle\mu_t,\varphi\rangle := \int_{\mathbb{R}^d}\varphi(x)d\mu_t(x).$$

The equivalence of definitions is classical, see e.g. [2, Chap 8]. We will also use the following classical fact: for $\mu_t$ solving the transport equation with source (2), any interval $[t_1,t_2] \subset [0,1)$ and all $\varphi \in \mathcal{D}([t_1,t_2]\times\mathbb{R}^d)$, it holds

$$\int_{t_1}^{t_2} dt \int_{\mathbb{R}^d} (d\mu_t(\partial_t\varphi(t,x) + v[\mu_t]\cdot\nabla\varphi(t,x)) + dh[\mu_t],\varphi(t,x)) = \int_{\mathbb{R}^d} d\mu_{t_2}\varphi(t_2,\cdot) - \int_{\mathbb{R}^d} d\mu_{t_1}\varphi(t_1,\cdot). \tag{16}$$

## 4.1 Estimates of the norm under flow action

In this section, we extend the action of flows on probability measures to signed measures, and state some estimates about the evaluation of $\|\mu\|^{a,b}$ under a flow action on $\mu$. Notice that for $\mu \in \mathcal{M}^s(\mathbb{R}^d)$ and $T$ a map, we have $T\#\mu = T\#\mu^+ - T\#\mu^-$, where $\mu = \mu^+ - \mu^-$ is any decomposition of $\mu$. Observe that in general, given $\mu \in \mathcal{M}^s(\mathbb{R}^d)$ and $T : \mathbb{R}^d \mapsto \mathbb{R}^d$ a Borel map, it only holds

$$|T\#\mu| \le |\mu|, \tag{17}$$

even by choosing the Jordan decomposition for $(\mu^+,\mu^-)$, since it may hold that $T\#\mu^+$ and $T\#\mu^-$ are not orthogonal. However, if $T$ is injective (as it will be in the rest of the paper), it holds $T\#\mu^+ \perp T\#\mu^-$, hence $|T\#\mu| = |\mu|$.

**Lemma 31.** *For $v(t,x)$ measurable in time, uniformly Lipschitz in space, and uniformly bounded, we denote by $\Phi_t^v$ the flow it generates, i.e. the unique solution to*

$$\frac{d}{dt}\Phi_t^v = v(t,\Phi_t^v), \qquad \Phi_0^v = I_d.$$

*Given $\mu_0 \in \mathcal{M}^s(\mathbb{R}^d)$, then, $\mu_t = \Phi_t^v\#\mu_0$ is the unique solution of the linear transport equation*

$$\begin{cases} \dfrac{\partial}{\partial t}\mu_t + \nabla\cdot(v(t,x)\mu_t) = 0, \\ \mu_{|t=0} = \mu_0 \end{cases}$$

*in $\mathcal{C}([0,T],\mathcal{M}^s(\mathbb{R}^d))$.*

*Proof.* The proof is a direct consequence of [22, Theorem 5.34] combined with [6, Theorem 2.1.1]. ☐

**Lemma 32.** *Let $v$ and $w$ be two vector fields, both satisfying for all $t \in [0,1]$ and $x, y \in \mathbb{R}^d$ the following properties:*

$$|v(t,x) - v(t,y)| \leq L|x-y|, \qquad |v(t,x)| \leq M.$$

*Let $\mu$ and $\nu$ be two measures of $\mathcal{M}^s(\mathbb{R}^d)$. Then*

- $\|\Phi_t^v \# \mu\|^{a,b} \leq e^{Lt} \|\mu\|^{a,b}$
- $\|\mu - \Phi_t^v \# \mu\|^{a,b} \leq b\, t\, M|\mu|,$
- $\|\Phi_t^v \# \mu - \Phi_t^w \# \mu\|^{a,b} \leq b|\mu| \frac{(e^{Lt}-1)}{L} \|v - w\|_{L^\infty(0,1;\, \mathcal{C}^0)}$
- $\|\Phi_t^v \# \mu - \Phi_t^w \# \nu\|^{a,b} \leq e^{Lt}\|\mu - \nu\|^{a,b} + b \min\{|\mu|, |\nu|\} \frac{(e^{Lt}-1)}{L} \|v - w\|_{L^\infty(0,1;\, \mathcal{C}^0)}$

*Proof.* The first three inequalities follow from [20, Proposition 10]. For the first inequality, we write

$$\begin{aligned}
\|\Phi_t^v \# \mu\|^{a,b} &= W_1^{a,b}(\Phi_t^v \# \mu^+, \Phi_t^v \# \mu^-) \\
&\leq e^{Lt} W_1^{a,b}(\mu^+, \mu^-) \quad \text{by [20, Prop. 10]} \\
&= e^{Lt} \|\mu\|^{a,b}.
\end{aligned}$$

For the second inequality,

$$\begin{aligned}
\|\mu - \Phi_t^v \# \mu\|^{a,b} &= W_1^{a,b}(\mu^+ + \Phi_t^v \# \mu^-, \mu^- + \Phi_t^v \# \mu^+) \\
&\leq W_1^{a,b}(\mu^+, \Phi_t^v \# \mu^+) + W_1^{a,b}(\mu^-, \Phi_t^v \# \mu^-) \quad \text{(Lemma 15)} \\
&\leq b\, t\, \|v\|_{\mathcal{C}^0}(|\mu^+| + |\mu^-|) \quad \text{by [20, Prop. 10]} \\
&= b\, t\, \|v\|_{L^\infty(0,1;\, \mathcal{C}^0(\mathbb{R}))} |\mu| \quad \text{since } \mu = \mu^+ - \mu^- \text{ is the Jordan decomposition.}
\end{aligned}$$

The third inequality is given by

$$\begin{aligned}
\|\Phi_t^v \# \mu - \Phi_t^w \# \mu\|^{a,b} &= W_1{}^{a,b}(\Phi_t^v \# \mu^+ + \Phi_t^w \# \mu^-, \Phi_t^w \# \mu^+ + \Phi_t^v \# \mu^-) \\
&\leq W_1^{a,b}(\Phi_t^v \# \mu^+, \Phi_t^w \# \mu^+) + W_1^{a,b}(\Phi_t^w \# \mu^-, \Phi_t^v \# \mu^-) \\
&\leq b \left( W_1(\Phi_t^v \# \mu^+, \Phi_t^w \# \mu^+) + W_1(\Phi_t^w \# \mu^-, \Phi_t^v \# \mu^-) \right) \\
&\leq b\, (|\mu^+| + |\mu^-|) \frac{(e^{Lt} - 1)}{L} \|v - w\||_{L^\infty(0,1;\, \mathcal{C}^0(\mathbb{R}))}
\end{aligned}$$

by using [20, Prop. 10] with $\mu = \nu$.

The last inequality is deduced from the first and the third ones using triangular inequality. ☐

### 4.2 A scheme for computing solutions of the transport equation

In this section, we define an approximation scheme for solutions to (2). This will be useful to prove existence of solutions. We then prove Theorem 2.

Fix $\mu_0 \in \mathcal{M}^s(\mathbb{R}^d)$ such that $\text{supp}(\mu_0) \subset \mathcal{K}$, with $\mathcal{K}$ compact. Let $v \in \mathcal{C}^{0,Lip}(\mathcal{M}^s(\mathbb{R}^d), \mathcal{C}^{0,Lip}(\mathbb{R}^d))$ and $h \in \mathcal{C}^{0,Lip}(\mathcal{M}^s(\mathbb{R}^d), \mathcal{M}^s(\mathbb{R}^d))$ satisfy (H-1)-(H-2)-(H-3). We now define a sequence $(\mu_t^k)_k$ of approximated solutions for (2) through the following Euler-explicit-type iteration scheme. For simplicity of notations, we define a solution on the time interval $[0,1]$ only.

Scheme

**Initialization.** Fix $k \in \mathbb{N}$. Define $\Delta t = \dfrac{1}{2^k}$. Set $\mu_0^k = \mu_0$.
**Induction.** Given $\mu_{i\Delta t}^k$ for $i \in \{0, 1, \ldots, 2^k - 1\}$, define $v_{i\Delta t}^k := v[\mu_{i\Delta t}^k]$ and

$$\mu_t^k = \Phi_{t-i\Delta t}^{v_{i\Delta t}^k} \# \mu_{i\Delta t}^k + (t - i\Delta t) h[\mu_{i\Delta t}^k], \qquad t \in [i\Delta t, (i+1)\Delta t]. \tag{18}$$

**Remark 33.** *The flow $\Phi_{t-i\Delta t}$ encodes the transport part $\partial_t \mu + \operatorname{div}(v\mu) = 0$ while $(t - i\Delta t)h$ encodes the reaction $\partial_t \mu = h$.*

We now prove equi-Lipschitz continuity of the sequence $(\mu_t^k)_k$. We also define the following sup-norm on curves in $\mathcal{C}^0([0,1], \mathcal{M}^s(\mathbb{R}^d))$ by writing

$$\|\mu\| := \sup_{t \in [0,1]} \|\mu_t\|^{a,b}.$$

**Proposition 34.** *The sequence $(\mu_t^k)_k \in \mathcal{C}^0([0,1], \mathcal{M}^s(\mathbb{R}^d), \|.\|)$ is equi-Lipschitz with respect to time, i.e. there exists $L' = aP + bM(P + |\mu_0|)$ independent on $k$ such that for all $t, s \in [0,1]$ it holds*

$$\|\mu_t^k - \mu_s^k\|^{a,b} \leq L'|t - s|. \tag{19}$$

*Moreover, the sequence is uniformly bounded in mass and compactly supported, i.e.*

$$|\mu_t^k| \leq Pt + |\mu_0|, \qquad \mathit{supp}\{\mu_t\} \subset B(0, R' + M) \tag{20}$$

*for $R'$ such that $(\mathit{supp}\{\mu_0\} \cup B_0(R)) \subset B_0(R')$.*

**Remark 35.** *Estimates* (20) *are expected at the discrete level from the PDE* (2) *with the assumptions (H-1), (H-2), (H-3). Indeed, the transport part preserves mass, while the reaction term gives a mass growth that is at most linear. Likewise, the support estimate is expected from the PDE since $h$ has support in $B_0(R)$ (no mass created out of this ball) and transport cannot expand the support with a speed faster than $|v| \leq M$.*

*Proof.* We first prove (20). The sequence built by the scheme satisfies

$$|\mu_t^k| \leq Pt + |\mu_0|, \quad t \in [0,1], \tag{21}$$

where $P$ is such that $|h[\mu]| \leq P$ by (H-3). Indeed, it holds directly from (18) and from (H-3) that

$$|\mu_{(i+1)\Delta t}^k| \leq |\Phi_{\Delta t}^{v_{i\Delta t}^k} \# \mu_{i\Delta t}^k| + \Delta t |h[\mu_{i\Delta t}^k]| \leq |\mu_{i\Delta t}^k| + \Delta t P,$$

then by induction on $i$ (for $k$ fixed), we have

$$|\mu_{i\Delta t}^k| \leq P i \Delta t + |\mu_0|. \tag{22}$$

Thus for $t \in [i\Delta t, (i+1)\Delta t]$, using again (18) and (H-3)

$$|\mu_t^k| \leq |\Phi_{t-i\Delta t}^{v_{i\Delta t}^k} \# \mu_{i\Delta t}^k| + (t - i\Delta t)|h[\mu_{i\Delta t}^k]| \leq |\mu_{i\Delta t}^k| + (t - i\Delta t)P \leq |\mu_0| + Pt,$$

using (22) for the last inequality. This proves the first estimate of (20), as $t \leq 1$. We now prove the second statement of (20). First observe that $\operatorname{supp}\{\mu\} = \operatorname{supp}\{\mu^+\} \cup \operatorname{supp}\{\mu^-\}$, where $(\mu^+, \mu^-)$ is the Jordan decomposition of $\mu$. Choose $\mathcal{K}$ such that $\text{supp}\{\mu_0\} \subset \mathcal{K}$ and use (18) and (H-2)-(H-3) to write

$$\text{supp}\{\mu_t^k\} \subseteq \mathcal{K}_{t,M,R},$$

with

$$\mathcal{K}_{t,M,R} := \{x \in \mathbb{R}^d,\ x = x_{\mathcal{K},R} + x',\ x_{\mathcal{K},R} \in \mathcal{K} \cup B_0(R),\ \|x'\| \leq tM\}.$$

Take now $R'$ such that $\mathcal{K} \cup B_0(R) \subset B_0(R')$. Then, it holds $\mathcal{K}_{t,M,R} \subset B(0, R' + tM)$. Again by recalling $t \leq 1$, we have the second statement of (20).

We now prove that $(\mu_t^k)_k$ is Lipschitz with respect to time. We have two cases:

- Let $t, s \in [i\Delta t, (i+1)\Delta t]$ for some $i \in \{0, 1, \ldots, 2^k - 1\}$. By applying (18), the triangular inequality and Lemma 32, we have

$$\begin{aligned}
\|\mu_t^k - \mu_s^k\|^{a,b} &\leq \|\Phi_{t-i\Delta t}^{v_{i\Delta t}^k} \# \mu_{i\Delta t}^k - \Phi_{s-i\Delta t}^{v_{i\Delta t}^k} \# \mu_{i\Delta t}^k\|^{a,b} + \|(t - i\Delta t)h[\mu_{i\Delta t}^k] - (s - i\Delta t)h[\mu_{i\Delta t}^k]\|^{a,b} \\
&= \|\Phi_{t-s}^{v_{i\Delta t}^k} \# \nu_1 - \nu_1\|^{a,b} + \|(t-s)h[\mu_{i\Delta t}^k] + \nu_2 - \nu_2\|^{a,b} \\
&\leq |t-s| bM |\nu_1| + a|t-s|\, |h[\mu_{i\Delta t}^k]| \leq |t-s|(bM|\mu_{i\Delta t}^k| + aP) \\
&\leq |t-s|(bM(Ps + |\mu_0|) + aP),
\end{aligned}$$

  where $\nu_1 = \Phi_{s-i\Delta t}^{v_{i\Delta t}^k} \# \mu_{i\Delta t}^k$ and $\nu_2 = (s - i\Delta t)h[\mu_{i\Delta t}^k]$. Recall that $s \in [0,1]$ and observe that this implies Lipschitz continuity on $[i\Delta t, (i+1)\Delta t]$.

- Choose now any $t, s \in [0,1]$ and assume $t < s$ with no loss of generality. Then choose $i, j \in \{0, 1, \ldots, 2^k - 1\}$ the unique indexes so that

$$i\Delta t \leq t < (i+1)\Delta t < \ldots < (j-1)\Delta t < s \leq j\Delta t.$$

  By applying triangular inequality and the estimate of the previous case on each term, it holds

$$\begin{aligned}
\|\mu_t^k - \mu_s^k\|^{a,b} &\leq \|\mu_t^k - \mu_{(i+1)\Delta t}^k\|^{a,b} + \|\mu_{(i+1)\Delta t}^k - \mu_{(i+2)\Delta t}^k\|^{a,b} + \ldots + \|\mu_{(j-1)\Delta t}^k - \mu_s^k\|^{a,b} \\
&\leq L'((i+1)\Delta t - t + (i+2)\Delta t - (i+1)\Delta t + \ldots + s - (j-1)\Delta t) = L'(s-t).
\end{aligned}$$

This proves uniform Lipschitz continuity. □

We now prove that $\mu_t^k$ is an approximated solution of (2).

**Proposition 36.** *There exists $L''$ such that, for each $k$ and $\varphi \in \mathcal{D}([0,1) \times \mathbb{R}^d)$ satisfying*

$$\|\varphi(t,\cdot)\|_\infty \leq 1, \qquad \|\varphi(t,\cdot)\|_{Lip} \leq 1, \qquad \textit{for all } \ t \in [0,1], \tag{23}$$

*it holds*

$$\left| \int_0^1 dt \int_{\mathbb{R}^d} \left( d\mu_t^k (\partial_t \varphi(t,\cdot) + v[\mu_t^k] \cdot \nabla \varphi(t,\cdot)) + dh[\mu_t^k]\, \varphi(t,\cdot) \right) + \int_{\mathbb{R}^d} d\mu_0 \varphi(0,\cdot) \right| \leq \frac{L''}{2^k}. \tag{24}$$

*Proof.* By using the formulation (16), for each interval $[i\Delta t, (i+1)\Delta t]$ and $\varphi \in \mathcal{D}([i\Delta t, (i+1)\Delta t] \times \mathbb{R}^d)$ it holds

$$\begin{aligned}
&\int_{i\Delta t}^{(i+1)\Delta t} dt \int_{\mathbb{R}^d} d\Phi_{t-i\Delta t}^{v[\mu_{i\Delta t}^k]} \# \mu_{i\Delta t}^k (\partial_t \varphi(t,\cdot) + v[\mu_{i\Delta t}^k] \cdot \nabla \varphi(t,\cdot)) = \\
&\int_{\mathbb{R}^d} d\Phi_{\Delta t}^{v[\mu_{i\Delta t}^k]} \# \mu_{i\Delta t}^k \varphi((i+1)\Delta t, \cdot) - \int_{\mathbb{R}^d} d\mu_{i\Delta t}^k \varphi(i\Delta t, \cdot)
\end{aligned}$$

and

$$\int_{i\Delta t}^{(i+1)\Delta t} dt \int_{\mathbb{R}^d} \Big( d((t-i\Delta t)h[\mu^k_{i\Delta t}])(\partial_t \varphi(t,\cdot) + dh[\mu^k_{i\Delta t}]\,\varphi(t,\cdot) \Big) =$$
$$\int_{\mathbb{R}^d} d(\Delta t\, h[\mu^k_{i\Delta t}])\varphi((i+1)\Delta t,\cdot) - \int_{\mathbb{R}^d} d0\,\varphi(i\Delta t,\cdot).$$

By adding on both sides and recalling the definition of $\mu^k$ in (18), it holds

$$\int_{i\Delta t}^{(i+1)\Delta t} dt \int_{\mathbb{R}^d} \Big( d\mu^k_t \partial_t\varphi(t,\cdot) + d\Phi^{v[\mu^k_{i\Delta t}]}_{t-i\Delta t} \# \mu^k_{i\Delta t} v[\mu^k_{i\Delta t}]\cdot\nabla\varphi(t,\cdot)) + dh[\mu^k_{i\Delta t}]\,\varphi(t,\cdot)\Big) =$$
$$\int_{\mathbb{R}^d} d\mu^k_{(i+1)\Delta t}\varphi((i+1)\Delta t,\cdot) - \int_{\mathbb{R}^d} d\mu^k_{i\Delta t}\varphi(i\Delta t,\cdot).$$

Recall that $\mu_t = \Phi^{v[\mu^k_{i\Delta t}]}_{t-i\Delta t}\#\mu^k_{i\Delta t} + (t-i\Delta t)h[\mu^k_{i\Delta t}]$ and sum all terms $i=0,\ldots,2^k-1$. By recalling that $\varphi(1,\cdot)=0$, we have

$$\sum_{i=0}^{2^k-1}\int_{i\Delta t}^{(i+1)\Delta t} dt \int_{\mathbb{R}^d} \Big( d\mu^k_t(\partial_t\varphi(t,\cdot) + v[\mu^k_{i\Delta t}]\cdot\nabla\varphi(t,\cdot)) + dh[\mu^k_{i\Delta t}]\,\varphi(t,\cdot)\Big) =$$
$$-\sum_{i=0}^{2^k-1}\int_{i\Delta t}^{(i+1)\Delta t} dt \int_{\mathbb{R}^d} (t-i\Delta t)h[\mu^k_{i\Delta t}]\,v[\mu^k_{i\Delta t}]\cdot\nabla\varphi(t,\cdot)) - \int_{\mathbb{R}^d} d\mu_0\varphi(0,\cdot). \tag{25}$$

Recall that (19) implies that for each $t\in[i\Delta t,(i+1)\Delta t]$ it holds $\|\mu^k_t - \mu^k_{i\Delta t}\|^{a,b} \leq L'\Delta t$, hence by (H-1)-(H-3) it holds

$$\|v[\mu^k_{i\Delta t}] - v[\mu^k_t]\|_{\mathcal{C}^0(\mathbb{R}^d)} \leq KL'\Delta t, \qquad \|h[\mu^k_{i\Delta t}] - h[\mu^k_t]\|^{a,b} \leq QL'\Delta t.$$

By using (25), we have

$$\left|\int_0^1 dt \int_{\mathbb{R}^d} \Big( d\mu^k_t(\partial_t\varphi(t,\cdot) + v[\mu^k_t]\cdot\nabla\varphi(t,\cdot)) + dh[\mu^k_t]\,\varphi(t,\cdot)\Big) + \int_{\mathbb{R}^d} d\mu_0\varphi(0,\cdot)\right| \leq$$
$$\sum_{i=0}^{2^k-1}\left|\int_{i\Delta t}^{(i+1)\Delta t} dt \int_{\mathbb{R}^d} (t-i\Delta t)h[\mu^k_{i\Delta t}]\,v[\mu^k_{i\Delta t}]\cdot\nabla\varphi(t,\cdot))\right| +$$
$$\sum_{i=0}^{2^k-1}\left|\int_{i\Delta t}^{(i+1)\Delta t} dt \int_{\mathbb{R}^d} \Big( d\mu^k_t(v[\mu^k_t] - v[\mu^k_{i\Delta t}])\cdot\nabla\varphi(t,\cdot)) + d(h[\mu^k_t] - h[\mu^k_{i\Delta t}])\,\varphi(t,\cdot)\Big)\right| \leq$$
$$\int_0^1 dt\Delta t PM\|\nabla\varphi(t,\cdot)\|_{\mathcal{C}^0(\mathbb{R}^d)} + \int_0^1 dt \int_{\mathbb{R}^d} d|\mu^k_t| L'K\Delta t\|\nabla\varphi(t,\cdot)\|_{\mathcal{C}^0(\mathbb{R}^d)} + L'Q\Delta t\|\varphi\|_{\mathcal{C}^0(\mathbb{R}^d)} \leq$$
$$\Delta t PM + \Delta t L'((P+|\mu_0|)K + Q).$$

Here we used that $h[\mu^k_{i\Delta t}]$ and $\mu^k_t$ have bounded mass, see (H-3)-(20), as well as bounded $\mathcal{C}^0$ norm of $v$, due to (H-2). Observe that $\|\nabla\varphi(t,\cdot)\|_{\mathcal{C}^0(\mathbb{R}^d)} = \|\varphi(t,\cdot)\|_{Lip}$ and recall $\Delta t = \frac{1}{2^k}$. By choosing $L'' := PM + L'((P+|\mu_0|)K+Q)$ not depending on $k$, we have the result. □

## 4.3 Proof of Theorem 2

In this section, we prove Theorem 2, stating existence and uniqueness of the solution to the Cauchy problem associated to (2). The proof is based on the proof of the same result for positive measures written in [20]. We first focus on existence.

**Step 1. Existence.** Recall Proposition 34: the sequence given by the scheme $(\mu_t^k)_k$ is uniformly Lipschitz continuous, uniformly bounded in mass and tight. Since $\mu_0^k = \mu_0$ for all $k$, this implies that the sequence is also uniformly bounded and equi-continuous. By Ascoli-Arzelà Theorem, this implies that the sequence is relatively compact in $\mathcal{C}^0\left([0,1], \mathcal{M}^s(\mathbb{R}^d)\right)$, hence there exist converging sub-sequences. By passing to one of such subsequences, for which we use the index $j$, we define

$$\mu_t := \lim_{j\to\infty} \mu_t^j, \tag{26}$$

We now prove that $\mu_t$ satisfies (14) for all $\varphi \in \mathcal{D}([0,1) \times \mathbb{R}^d)$. Observe that each $\varphi$ satisfies $\sup_{t\in[0,1]} \|\varphi(t,\cdot)\|_{\mathcal{C}^0(\mathbb{R}^d)}, \|\varphi(t,\cdot)\|_{Lip} < +\infty$. Moreover, by homogeneity of (14) and of $\|\varphi(t,\cdot)\|_{\mathcal{C}^0(\mathbb{R}^d)}$, $\|\varphi(t,\cdot)\|_{Lip}$, $\|\partial_t\varphi(t,\cdot)\|_{\mathcal{C}^0(\mathbb{R}^d)}$, $\|\partial_t\varphi(t,\cdot)\|_{Lip}$ with respect to $\varphi \to \lambda\varphi$, it is sufficient to prove that $\mu_t$ satisfies (14) for all $\varphi \in \mathcal{D}([0,1) \times \mathbb{R}^d)$ with the additional constraint

$$\|\varphi(t,\cdot)\|_{\mathcal{C}^0(\mathbb{R}^d)},\ \|\varphi(t,\cdot)\|_{Lip},\ \|\partial_t\varphi(t,\cdot)\|_{\mathcal{C}^0(\mathbb{R}^d)},\ \|\partial_t\varphi(t,\cdot)\|_{Lip} \le 1 \ \text{ for all } t \in [0,1). \tag{27}$$

Observe that for each $\varphi \in \mathcal{D}([0,1) \times \mathbb{R}^d)$ satisfying (27) it holds

$$C := \left|\int_0^1 dt \int_{\mathbb{R}^d} (d\mu_t(\partial_t\varphi(t,x) + v[\mu_t]\cdot\nabla\varphi(t,x)) + dh[\mu_t]\,\varphi(t,x)) + \int_{\mathbb{R}^d} d\mu_0\varphi(0,\cdot)\right| \le \frac{L''}{2^j} +$$
$$\left|\int_0^1 dt \int_{\mathbb{R}^d} d(\mu_t - \mu_t^j)\partial_t\varphi(t,\cdot) + d\mu_t v[\mu_t]\cdot\nabla\varphi(t,\cdot) - d\mu_t^j v[\mu_t^j]\cdot\nabla\varphi(t,\cdot)) + d(h[\mu_t] - h[\mu_t^j])\,\varphi(t,\cdot)\right|$$

Since such estimate holds for any $j$, it is sufficient to prove that the right-hand side tends to zero for $j \to +\infty$. We have

$$\begin{aligned}
C \quad \le \quad & \frac{L''}{2^j} + \|\mu - \mu^j\| \sup_{t\in[0,1]} \max\{\|\partial_t\varphi(t,\cdot)\|_\infty, \|\partial_t\varphi(t,\cdot)\|_{Lip}\} + \left|\int_0^1 dt \int_{\mathbb{R}^d} d(\mu_t - \mu_t^j) v[\mu_t]\cdot\nabla\varphi(t,\cdot)\right| + \\
& \left|\int_0^1 dt \int_{\mathbb{R}^d} d\mu_t^j(v[\mu_t] - v[\mu_t^j])\cdot\nabla\varphi(t,\cdot))\right| + \|h[\mu] - h[\mu^j]\| \le \frac{L''}{2^j} + \|\mu - \mu^j\| + \\
& \|\mu - \mu^j\| M \sup_{t\in[0,1]} \|\varphi(t,\cdot)\|_{Lip} + \int_0^1 dt \int_{\mathbb{R}^d} d|\mu_t^j| L \|\mu_t - \mu_t^j\|^{a,b} \|\varphi(t,\cdot)\|_{Lip} + Q\|\mu - \mu^j\| \\
\le \quad & \frac{L''}{2^j} + \|\mu - \mu^j\|\left(1 + M + (P + |\mu_0|)L + Q\right).
\end{aligned}$$

By letting $j \to +\infty$ and recalling that (26) is equivalent to $\|\mu - \mu^j\| \to 0$, we have the result.

**Remark 37.** *From this construction, we do not prove uniqueness of the limit for the sequence $\mu^k$. Yet, we will prove uniqueness of the solution to* (2) *in Step 4, that will in turn ensure uniqueness of the limit.*

**Step 2. Any weak solution to** (2) **is Lipschitz in time.** In this step, we prove that any weak solution in the sense of Definition 30 to the transport equation (2) is Lipschitz with respect to time, since it satisfies

$$\|\mu_{t+\tau} - \mu_t\|^{a,b} \le L'\tau, \qquad t \ge 0,\ \tau \ge 0, \tag{28}$$

with $L'$ defined in Proposition (34). To do so, we consider a solution $\mu_t$ to (2). We define the vector field $w(t,x) := v[\mu_t](x)$ and the signed measure $b_t = h[\mu_t]$. The vector field $w$ is uniformly Lipschitz and uniformly bounded with respect to $x$, since $v$ is so. The field $w$ is also measurable in time, since by definition, $\mu_t$ is continuous in time. Then, $\mu_t$ is the unique solution of

$$\partial_t\mu_t(x) + \mathrm{div}\,.(w(t,x)\mu_t(x)) = b_t(x), \qquad \mu_{|t=0}(x) = \mu_0(x). \tag{29}$$

Uniqueness of the solution of the linear equation (29) is a direct consequence of Lemma 31. Moreover, the scheme presented in Section 4.2 can be rewritten for the vector field $w$ and the source $b$, in which dependence with respect to time is added and dependence with respect to the measure is dropped. Thus, the unique solution $\mu$ to (29) can be obtained as the limit of this scheme $\mu^k$.

For each $k \geq 0$ it holds

$$\|\mu_{t+\tau} - \mu_t\|^{a,b} \leq \|\mu_t - \mu_t^k\|^{a,b} + \|\mu_t^k - \mu_{t+\tau}^k\|^{a,b} + \|\mu_{t+\tau}^k - \mu_{t+\tau}\|^{a,b} \leq 2\|\mu - \mu^k\| + L'\tau,$$

where we used (19) for Lipschitz continuity of $\mu^k$. By letting $k \to +\infty$ we have $\|\mu - \mu^k\| \to 0$, thus (28) holds.

**Step 3. Any weak solution to** (2) **satisfies the operator splitting estimate**: there exist $K', \tau' > 0$ such that for all $t \in [0,1)$ and $\tau \in (0, \tau')$ saisfying $t + \tau \leq 1$, it holds

$$\|\mu_{t+\tau} - (\Phi_\tau^{v[\mu_t]} \# \mu_t + \tau h[\mu_t])\|^{a,b} \leq K'\tau^2. \tag{30}$$

To this end , consider $\mu_t$ as the solution to the non-autonomous linear equation (29), as in Step 2: this allows to define the trajectory $\mu_t^k$ given by the scheme presented in Section 4.2 and to prove that $\mu_t$ is the limit of $\mu_t^k$. We now prove an estimate similar to (30) for $\mu_t^k$.

Fix $k \in \mathbb{N}^*$ and let $t = i\Delta t$, $\tau = n\Delta t < \log(2)/L$ for some $i \in \{0, \ldots, 2^k - 1\}$, $n \in \mathbb{N}$ and $t + \tau \leq 1$ This ensures $e^{L\tau} \leq 1 + 2L\tau$. Define

$$a_n := \|\mu_{(i+n)\Delta t}^k - (\Phi_{n\Delta t}^{v[\mu_{i\Delta t}^k]} \# \mu_{i\Delta t}^k + n\Delta t h[\mu_{i\Delta t}^k])\|^{a,b}.$$

Observe that it holds $a_1 = 0$, while for $n \geq 2$ it holds

$$\begin{aligned}
a_n &\leq \|\Phi_{\Delta t}^{v[\mu_{(i+n-1)\Delta t}^k]} \# \mu_{(i+n-1)\Delta t}^k - (\Phi_{\Delta t}^{v[\mu_{i\Delta t}^k]} \# (\Phi_{(n-1)\Delta t}^{v[\mu_{i\Delta t}^k]} \# \mu_{i\Delta t}^k + (n-1)\Delta t h[\mu_{i\Delta t}^k]))\|^{a,b} + \\
&\|\Delta t h[\mu_{(i+n-1)\Delta t}^k] - \Delta t h[\mu_{i\Delta t}^k]\|^{a,b} + \|\Phi_{\Delta t}^{v[\mu_{i\Delta t}^k]} \# (n-1)\Delta t h[\mu_{i\Delta t}^k] - (n-1)\Delta t h[\mu_{i\Delta t}^k]\|^{a,b} \\
&\leq e^{L\Delta t}\|\mu_{(i+n-1)\Delta t}^k - (\Phi_{(n-1)\Delta t}^{v[\mu_t^k]} \# \mu_{i\Delta t}^k + (n-1)\Delta t h[\mu_{i\Delta t}^k])\|^{a,b} + \\
&b|\mu_{i\Delta t}^k| \frac{e^{L\Delta t} - 1}{L} \|v[\mu_{(i+n-1)\Delta t}^k] - v[\mu_{i\Delta t}^k]\|_{\mathcal{C}^0} + \Delta t Q \|\mu_{(i+n-1)\Delta t}^k - \mu_{i\Delta t}^k\|^{a,b} \\
&+(n-1)\Delta t(\Delta t b M |h[\mu_{i\Delta t}^k]| \\
&\leq (1 + 2L\Delta t)a_{n-1} + (|\mu_0| + P)2\Delta t a_{n-1} + \Delta t Q a_{n-1} + (n-1)\Delta t^2 MP \\
&\leq (1 + K_1'\Delta t)a_{n-1} + K_2'\tau\Delta t.
\end{aligned}$$

Here, we used (H-1)-(H-2)-(H-3), Lemma (32) as well as Lipschitz continuity and boundedness of mass proved in Proposition 34. Observe that $K_1', K_2'$ do not depend on $n$ or $k$, with $K_1 > L$. Thus, choose $\tau' = \log(2)/K_1$ independent on $k$ and observe that for all $\tau \in (0, \tau')$ it holds

$$\|\mu_{(i+n)\Delta t}^k - (\Phi_{n\Delta t}^{v[\mu_{i\Delta t}^k]} \# \mu_{i\Delta t}^k + n\Delta t h[\mu_{i\Delta t}^k])\|^{a,b} \leq K_2\tau\Delta t \frac{(1 + K_1'\Delta t)^n - 1}{K_1'\Delta t} = \frac{K_2'}{K_1'}\tau(e^{K_1'\tau} - 1) \leq 2K_2'\tau^2. \tag{31}$$

We are now ready to prove (30). For $t \in [0,1)$ fixed, build the sequence $i_k \in \{0, \ldots 2^k - 1\}$ such that $|t - i_k 2^{-k}| < 2^{-k}$. Similarly, for $\tau \in (0, \tau')$ fixed, build the sequence $n_k \in \{0, \ldots 2^k\}$ such that $|t - (i_k + n_k)2^{-k}| < 2^{-k}$. Observe that $\lim_{k \to +\infty} \|\mu - \mu^k\| = 0$, together with (19), implies

$$\|\mu_{t+\tau} - \mu_{(i_k+n_k)2^{-k}}^k\|^{a,b} \leq \|\mu_{t+\tau} - \mu_{t+\tau}^k\|^{a,b} + \|\mu_{t+\tau}^k - \mu_{(i_k+n_k)2^{-k}}^k\|^{a,b} \leq \|\mu - \mu^k\| + L'2^{-k} \to 0.$$

and similarly $\|\mu_t - \mu^k_{i_k 2^{-k}}\|^{a,b} \to 0$. By using (H-1)-(H-2)-(H-3) and Lemma 32, this in turn ensures

$$\|\Phi^{v[\mu^k_{i_k 2^{-k}}]}_{n_k 2^{-k}} \# \mu^k_{i_k 2^{-k}} - \Phi^{v[\mu_t]}_{\tau} \# \mu_t\|^{a,b} \leq \|\Phi^{v[\mu^k_{i_k 2^{-k}}]}_{n_k 2^{-k}} \# \mu^k_{i_k 2^{-k}} - \Phi^{v[\mu^k_{i_k 2^{-k}}]}_{\tau} \# \mu^k_{i_k 2^{-k}}\|^{a,b} +$$
$$\|\Phi^{v[\mu^k_{i_k 2^{-k}}]}_{\tau} \# \mu^k_{i_k 2^{-k}} - \Phi^{v[\mu_t]}_{\tau} \# \mu^k_{i_k 2^{-k}}\|^{a,b} + \|\Phi^{v[\mu_t]}_{\tau} \# \mu^k_{i_k 2^{-k}} - \Phi^{v[\mu_t]}_{\tau} \# \mu_t\|^{a,b}$$
$$\leq |\mu^k_{i_k 2^{-k}}| M |n_k 2^{-k} - \tau| + |\mu^k_{i_k 2^{-k}}| \frac{e^{L\tau} - 1}{L} K \|\mu_t - \mu^k_{i_k 2^{-k}}\|^{a,b} + e^{L\tau} \|\mu_t - \mu^k_{i_k 2^{-k}}\|^{a,b} \to 0$$

as well as

$$\|n_k 2^{-k} h[\mu^k_{i_k 2^{-k}}] - \tau h[\mu_t]\|^{a,b} \leq \|n_k 2^{-k} h[\mu^k_{i_k 2^{-k}}] - \tau h[\mu^k_{i_k 2^{-k}}]\|^{a,b} + \tau \|h[\mu^k_{i_k 2^{-k}}] - \tau h[\mu_t]\|^{a,b} \leq$$
$$P|n_k 2^{-k} - \tau| + \tau Q \|\mu^k_{i_k 2^{-k}} - \mu_t\|^{a,b} \to 0.$$

Since $n_k 2^{-k} \to \tau \in (0, \tau')$, for $k$ sufficiently large one can apply (31), thus

$$\lim_{k \to +\infty} \|\mu_{t+\tau} - (\Phi^{v[\mu_t]}_{\tau} \# \mu_t + \tau h[\mu_t])\|^{a,b} \leq \lim_{k \to +\infty} \Big( \|\mu_{t+\tau} - \mu^k_{(i_k + n_k) 2^{-k}}\|^{a,b} +$$
$$\|\mu^k_{(i_k+n_k)2^{-k}} - (\Phi^{v[\mu^k_{i_k 2^{-k}}]}_{n_k 2^{-k}} \# \mu^k_{i_k 2^{-k}} + n_k 2^{-k} h[\mu^k_{i_k 2^{-k}}])\|^{a,b} + \|\Phi^{v[\mu^k_{i_k 2^{-k}}]}_{n_k 2^{-k}} \# \mu^k_{i_k 2^{-k}} - \Phi^{v[\mu_t]}_{\tau} \# \mu_t\|^{a,b}$$
$$+ \|n_k 2^{-k} h[\mu^k_{i_k 2^{-k}}] - \tau h[\mu_t]\|^{a,b} \Big) \leq 0 + 2K_2' \tau^2 + 0 + 0.$$

that is (30) for $K' = 2K_2'$.

**Step 4. Uniqueness of the solution to** (2) **and continuous dependence.** Assume that $\mu_t$ and $\nu_t$ are two solutions to (2) with initial condition $\mu_0, \nu_0$, respectively. Define $\varepsilon(t) := \|\mu_t - \nu_t\|^{a,b}$, that is a Lipschitz function by Step 2. We denote

$$R_\mu(t, \tau) = \mu_{t+\tau} - (\Phi^{v[\mu_t]}_{\tau} \# \mu_t + \tau h[\mu_t]), \quad R_\nu(t, \tau) = \nu_{t+\tau} - (\Phi^{v[\nu_t]}_{\tau} \# \nu_t + \tau h[\nu_t]).$$

Fix $\tau < \tau' < \log(2)/L$, that ensures $e^{L\tau} \leq 1 + 2L\tau$. By Step 3, it holds

$$\begin{aligned}
\varepsilon(t + \tau) &= \|\mu_{t+\tau} - \nu_{t+\tau}\|^{a,b} = \|\Phi^{v[\mu_t]}_{\tau} \# \mu_t + \tau h[\mu_t] + R_\mu(t, \tau) - \Phi^{v[\nu_t]}_{\tau} \# \nu_t - \tau h[\nu_t] - R_\nu(t, \tau)\|^{a,b} \\
&\leq \|\Phi^{v[\mu_t]}_{\tau} \# \mu_t - \Phi^{v[\nu_t]}_{\tau} \# \nu_t\|^{a,b} + \tau \|h[\mu_t] - h[\nu_t]\|^{a,b} + \|R_\mu(t, \tau)\|^{a,b} + \|R_\nu(t, \tau)\|^{a,b} \\
&\leq e^{L\tau} \|\mu_t - \nu_t\|^{a,b} + (P + \min\{|\mu_0|, |\nu_0|\}) \frac{e^{L\tau} - 1}{L} \|v[\mu_t] - v[\nu_t]\|_{\mathcal{C}^0} + \tau Q \|\mu_t - \nu_t\|^{a,b} + 2K' \tau^2 \\
&\leq \big(e^{L\tau} + b(P + \min\{|\mu_0|, |\nu_0|\}) 2\tau K + \tau Q\big) \varepsilon(t) + 2K' \tau^2 \leq (1 + \tau K_3) \varepsilon(t) + 2K' \tau^2
\end{aligned}$$

for $K_2 = 2L + 2Kb(P + \min\{|\mu_0|, |\nu_0|\}) + Q$. By letting $\tau \to 0$, we deduce $\varepsilon'(t) \leq K_3 \varepsilon(t)$ almost everywhere. Then, $\varepsilon(t) \leq \varepsilon(0) \exp(K_2 t)$, that implies continuous dependence with respect to the initial data.

Moreover, if $\mu_0 = \nu_0$, then $\varepsilon(0) = 0$, thus $\varepsilon(t) = 0$ for all $t$. Since $\|.\|^{a,b}$ is a norm, this implies $\mu_t = \nu_t$ for all $t$, that corresponds to uniqueness of the solution.

## References


[1] Luigi Ambrosio and Wilfrid Gangbo. Hamiltonian odes in the wasserstein space of probability measures. *Communications on Pure and Applied Mathematics*, 61(1):18–53, 2008.

[2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures.* Springer Science & Business Media, 2008.

[3] Luigi Ambrosio, Edoardo Mainini, and Sylvia Serfaty. Gradient flow of the Chapman-Rubinstein-Schatzman model for signed vortices. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 28(2):217–246, 2011.

[4] François Bolley, Yann Brenier, and Grégoire Loeper. Contractive metrics for scalar conservation laws. *J. Hyperbolic Differ. Equ.*, 2(1):91–107, 2005.

[5] Giovanni A. Bonaschi, José A. Carrillo, Marco Di Francesco, and Mark A. Peletier. Equivalence of gradient flows and entropy solutions for singular nonlocal interaction equations in 1D. *ESAIM Control Optim. Calc. Var.*, 21(2):414–441, 2015.

[6] Alberto Bressan and Benedetto Piccoli. *Introduction to the mathematical theory of control*, volume 2 of *AIMS Series on Applied Mathematics*. AIMS, 2007.

[7] Luis A. Caffarelli and Robert J. McCann. Free boundaries in optimal transport and Monge-Ampère obstacle problems. *Ann. of Math. (2)*, 171(2):673–730, 2010.

[8] Lénaïc Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. Unbalanced optimal transport: dynamic and Kantorovich formulations. *J. Funct. Anal.*, 274(11):3090–3123, 2018.

[9] Lenaic Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. An interpolating distance between optimal transport and Fisher–Rao metrics. *Foundations of Computational Mathematics*, 18(1):1–44, 2018.

[10] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015.

[11] Alessio Figalli. The optimal partial transport problem. *Arch. Ration. Mech. Anal.*, 195(2):533–560, 2010.

[12] Alessio Figalli and Nicola Gigli. A new transportation distance between non-negative measures, with applications to gradients flows with Dirichlet boundary conditions. *J. Math. Pures Appl. (9)*, 94(2):107–130, 2010.

[13] Leonid G. Hanin. An extension of the Kantorovich norm. In *Monge Ampère equation: applications to geometry and optimization (Deerfield Beach, FL, 1997)*, volume 226 of *Contemp. Math.*, pages 113–130. Amer. Math. Soc., Providence, RI, 1999.

[14] François James and Nicolas Vauchelet. Equivalence between duality and gradient flow solutions for one-dimensional aggregation equations. *Discrete Contin. Dyn. Syst.*, 36(3):1355–1382, 2016.

[15] Stanislav Kondratyev, Léonard Monsaingeon, and Dmitry Vorotnikov. A new optimal transport distance on the space of finite Radon measures. *Adv. Differential Equations*, 21(11-12):1117–1164, 2016.

[16] Matthias Liero, Alexander Mielke, and Giuseppe Savaré. Optimal transport in competition with reaction: the Hellinger-Kantorovich distance and geodesic curves. *SIAM J. Math. Anal.*, 48(4):2869–2911, 2016.

[17] Edoardo Mainini. A description of transport cost for signed measures. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 390:147–181, 308–309, 2011.

[18] Benedetto Piccoli and Francesco Rossi. Transport equation with nonlocal velocity in Wasserstein spaces: convergence of numerical schemes. *Acta Appl. Math.*, 124:73–105, 2013.

[19] Benedetto Piccoli and Francesco Rossi. Generalized Wasserstein distance and its application to transport equations with source. *Arch. Ration. Mech. Anal.*, 211(1):335–358, 2014.

[20] Benedetto Piccoli and Francesco Rossi. On properties of the generalized Wasserstein distance. *Arch. Ration. Mech. Anal.*, 222(3):1339–1365, 2016.

[21] Benedetto Piccoli and Francesco Rossi. Measure-theoretic models for crowd dynamics. In *Crowd Dynamics, Volume 1*, pages 137–165. Springer, 2018.

[22] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[23] Cédric Villani. *Optimal transport, Old and New*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2009.

[24] A. Zygmund. *Trigonometric series. Vol. I, II.* Cambridge University Press, 3rd edition, 2002.