



**HAL**  
open science

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

Benedetto Piccoli, Francesco Rossi, Magali Tournus

► **To cite this version:**

Benedetto Piccoli, Francesco Rossi, Magali Tournus. A Wasserstein norm for signed measures, with application to non local transport equation with source term. 2019. hal-01665244v4

**HAL Id: hal-01665244**

**<https://hal.science/hal-01665244v4>**

Preprint submitted on 10 Oct 2019 (v4), last revised 10 Dec 2021 (v5)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Benedetto Piccoli <sup>\*</sup>      Francesco Rossi <sup>†</sup>      Magali Tournus <sup>‡</sup>

## Abstract

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

With the formulation and the new topological properties we obtain for this norm, we 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 [21, 22].

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, 10] 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 [11] define the distance  $Wb_2$  between measures with different masses

---

<sup>\*</sup>Department of Mathematical Sciences, Rutgers University Camden, Camden, NJ, USA. Email: piccoli@camden.rutgers.edu.

<sup>†</sup>Dipartimento di Matematica “Tullio Levi-Civita” Università degli Studi di Padova, Via Trieste 63, 35121 Padova, Italy. Email: francesco.rossi@math.unipd.it

<sup>‡</sup>Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France. Email: mtournus@math.cnrs.fr.

on a bounded domain. Further generalizations for positive measures with different masses, based on the Wasserstein distance, are introduced in [8, 14, 15].

Such generalizations still have a drawback: both measures need to be positive. The first contribution of this paper is then the definition of a norm on the space of signed Radon measures with finite mass on  $\mathbb{R}^d$ . Such norm, based on an optimal transport approach, induces a distance generalizing the Wasserstein distance to signed measures. We then prove that this norm corresponds to the extension of the so-called Kantorovich distance for finite signed Radon measures introduced in [12] in the dual form

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

The novelty then lies in the dual interpretation of this norm in the framework of optimal transportation. We also prove new topological properties and characterizations of this norm.

The second main contribution of the paper is to use this norm to guarantee well-posedness of the following non local transport equation with a source term being a signed measure. We study the following PDE

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

for  $x \in \mathbb{R}^d$  and  $\mu_0 \in \mathcal{M}^s(\mathbb{R}^d)$ , 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 development in biology, see a review in [17]. 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 the local topological degree) in domain occupied by a superconducting sample satisfies (2) with  $h[\mu_t] = 0$  and where  $v[\mu_t]$  is the magnetic field induced in the sample (see [3] and [16]).

Another motivation to study equation (2) in the framework of signed measure 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, 13], 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 need a space of signed measures equipped with a metric of Wasserstein type, see e.g. [4].

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  $W_1(\mu, \nu) = W_1(\mu^+ + \nu^-, \mu^- + \nu^+)$ . 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  $W_2(\mu, \nu) = 4$  whereas  $W_2(\mu, \eta) + W_2(\eta, \nu) = \sqrt{2} + \sqrt{2}$ .

We use the same trick from [3] to turn the generalized Wasserstein distance  $W_1^{a,b}$  into a distance for signed measures, by setting  $W_1^{a,b}(\mu, \nu)$  as  $W_1^{a,b}(\mu^+ + \nu^-, \mu^- + \nu^+)$ . For the same reason as mentioned above, this construction cannot be done for  $p \neq 1$ , in particular, no quadratic distance can be obtained with this construction. The space of signed measures being a vector space, we also define a norm  $\|\mu\|^{a,b} = W_1^{a,b}(\mu, 0)$ . Notice that to define the norm  $\|\cdot\|^{a,b}$ , we need to restrict ourselves to Radon measures with finite mass, since the generalized Wasserstein distance [19] may not be defined for Radon measures with infinite mass. Since the terms ‘‘Wasserstein’’ and ‘‘norm’’ are usually not used together, we emphasize that  $\|\cdot\|^{a,b}$  is a norm in the sense of linear vector spaces. We then use the norm  $\|\cdot\|^{a,b}$  to study existence and uniqueness of the solution to the equation (2).

The regularity assumptions made in this paper 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

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

(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

$$|v[\mu](x) - v[\mu](y)| \leq L|x - y|, \quad |v[\mu](x)| \leq M. \quad (4)$$

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

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

Hypothesis (H-1) guarantees that  $v[\mu_t]$  is continuous in space, and then the product  $v[\mu_t] \cdot \mu_t$  is well-defined. The main result about equation (2) is the following:

**Theorem 1** (Existence and uniqueness). *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 the space  $\mathcal{C}^0([0, 1], \mathcal{M}^s(\mathbb{R}^d))$  equipped with  $\|\mu_t\| = \sup_{t \in [0, 1]} \|\mu_t\|^{a,b}$ . 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:*

$$\|\mu_t - \nu_t\|^{a,b} \leq \|\mu_0 - \nu_0\|^{a,b} \exp(C_1 t), \quad C_1 = 2L + 2K(P + \min\{|\mu_0|, |\nu_0|\}) + Q,$$

the following estimates on the mass and support:

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

the solution is Lipschitz in time:

$$\|\mu_{t+\tau} - \mu_t\|^{a,b} \leq C_2 \tau, \quad C_2 = P + bM(P(t + \tau) + |\mu_0|), \quad \tau \geq 0.$$

A precise definition of measure-valued weak solution for equation (2) is provided at the beginning of Section 4.

**Remark 2.** *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 preliminary results which are needed for the rest of the paper. In Section 3, we define the generalized Wasserstein distance for signed measures, we show that it can be used to define a norm, and prove some topological properties. Section 4 is devoted to the use of the norm defined here to guarantee existence, uniqueness, and stability to initial condition for the transport equation (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 3.** *We say that*

- $\mu \ll \nu$  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  $\nu(E^c) = 0$

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

**Lemma 4.** *We consider  $\mu$  and  $\nu$  two measures in  $\mathcal{M}(\mathbb{R}^d)$ . Then, there exists a unique measure  $\mu \wedge \nu$  which satisfies*

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

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

$$\mu \wedge \nu = \min\{f, 1\}\nu. \quad (7)$$

*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

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

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

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

We define

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

If  $\nu(B) = 0$ , then  $f \leq 1$   $\nu$ -a.e., hence  $\eta(A) \leq \int_A \min\{f, 1\}d\nu$ . We then assume  $\nu(B) > 0$ . Then

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

which contradicts the fact that  $\eta \leq \nu$  and  $\eta \leq \mu$ . This implies that  $\eta$  satisfying (8) does not exist, and then (7) holds.  $\square$

**Lemma 5.** *Let  $\mu$  and  $\nu$  be two measures in  $\mathcal{M}(\mathbb{R}^d)$ . Then  $\eta \leq \mu + \nu$  implies  $\eta - (\mu \wedge \eta) \leq \nu$ .*

*Proof.* Take  $A$  a Borel set. We write  $\mu = f\eta + \eta_\perp$ , with  $\eta_\perp \perp \eta$ . Then  $\eta \wedge \mu = \min\{f, 1\}\eta$ , and we can write

$$\eta(A) - (\eta \wedge \mu)(A) = \int_A \left( \max\{1 - f, 0\} \right) d\eta.$$

Define  $B = A \cap \{f < 1\}$ , and  $E$  such that  $\eta(A \cap E) = \eta(A)$  and  $\eta_\perp(A \cap E^c) = \eta_\perp(A)$ . It then holds,

$$\eta(A) - (\eta \wedge \mu)(A) = \int_{B \cap E} (1 - f) d\eta(x) = \eta(B \cap E) + \eta_\perp(B \cap E) - \mu(B \cap E) \leq \nu(B \cap E) \leq \nu(A).$$

Since this estimate holds for any Borel set  $A$ , the statement is proved.  $\square$

## 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  $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$  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.** *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, weak and narrow convergences are equivalent. It is not the case for signed measure and we precise here what we call narrow 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,  $\mathcal{C}_c^0(\mathbb{R}^d; \mathbb{R})$  is the set of continuous functions with compact support on  $\mathbb{R}^d$ , and  $\mathcal{C}_0^0(\mathbb{R}^d; \mathbb{R})$  is the set of continuous functions on  $\mathbb{R}^d$  that vanish at infinity.

**Definition 8** (Narrow 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 for all  $\varphi \in \mathcal{C}_b^0(\mathbb{R}^d; \mathbb{R})$ ,  $\int_{\mathbb{R}^d} \varphi(x) d\mu_n(x) \rightarrow \int_{\mathbb{R}^d} \varphi(x) d\mu(x)$ .*

*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 for all  $\varphi \in \mathcal{C}_c^0(\mathbb{R}^d; \mathbb{R})$ ,  $\int_{\mathbb{R}^d} \varphi(x) d\mu_n(x) \rightarrow \int_{\mathbb{R}^d} \varphi(x) d\mu(x)$ .*

**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 [9, Theorem 1.41]. Notice that in [9], vague convergence is called weak convergence. In [12, 21] 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 remind key properties of the generalized Wasserstein distance. The usual Wasserstein distance  $W_p(\mu, \nu)$  was defined between two measures  $\mu$  and  $\nu$  of same mass  $|\mu| = |\nu|$ , see more details in [21].

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

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

We notice that

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

and in particular

$$W_p^{a,b} = \frac{b}{b'} W_p^{a',b'}, \quad \text{for } \frac{a}{b} = \frac{a'}{b'}. \quad (11)$$

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

**Lemma 12.** *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

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

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 13** (Kantorovitch Rubinstein duality). *For  $\mu, \nu$  in  $\mathcal{M}(\mathbb{R}^d)$ , it holds*

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

**Lemma 14** (Properties of the generalized Wasserstein distance). *Let  $\mu, \nu, \eta, \mu_1, \mu_2, \nu_1, \nu_2$  be some positive measures with finite mass on  $\mathbb{R}^d$ . The following properties hold*

1.  $W_p^{a,b}(\mu_1 + \mu_2, \nu_1 + \nu_2) \leq W_p^{a,b}(\mu_1, \nu_1) + W_p^{a,b}(\mu_2, \nu_2)$ .
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 13 for  $W^{1,1}$ . For general  $a > 0, b > 0$ , we proceed as follows. Let  $\mu, \nu$  be two measures. 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  $\Pi(\bar{\mu}, \bar{\nu})$ . Define  $D_\lambda : x \rightarrow \lambda x$  with  $\lambda > 0$  the dilation in  $\mathbb{R}^n$ . It holds

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

As a consequence, it holds

$$W^{a,b}(D_\lambda \# \mu, D_\lambda \# \nu) = W^{a,\lambda b}(\mu, \nu).$$

We now show that this implies  $W^{a,b}(\mu + \eta, \nu + \eta) = W^{a,b}(\mu, \nu)$ . Indeed, also applying Kantorovich-Rubinstein for  $W^{1,1}$  and (11) with  $a' = 1, b' = \lambda = \frac{b}{a}$ , it holds

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

□

**Definition 15** (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  $\eta \leq \mu$ , 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  $\eta' \leq \nu$ . Indeed, since  $\eta \leq \mu$ , it holds  $f \leq 1$ . Thus, for all Borel set  $B$  of  $\mathbb{R}^d$  we have

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

### 3 Generalized Wasserstein norm for signed measures

In this section, we define the generalized Wasserstein distance for signed measures and prove some of its properties. The idea is to follow what was already done in [3] for generalizing the classical Wasserstein distance.



**Definition 16** (Generalized Wasserstein distance extended to signed measures). *For  $\mu, \nu$  two signed measures with finite mass over  $\mathbb{R}^d$ , 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 17.** *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 point out 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 14 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  $\mu = \nu$ . As explained above, we can choose the Jordan decomposition for both  $\mu$  and  $\nu$ . Since  $W_1^{a,b}$  is a distance, we obtain  $\mu_+ + \nu_- = \mu_- + \nu_+$ . The orthogonality of  $\mu_+$  and  $\mu_-$  and of  $\nu_+$  and  $\nu_-$  implies that  $\mu_+ = \nu_+$  and  $\mu_- = \nu_-$ , and 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 14, 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 18.** *Let  $\mu, \nu, \eta, \mu_1, \mu_2, \nu_1, \nu_2$  in  $\mathcal{M}^s(\mathbb{R}^d)$  with finite mass on  $\mathbb{R}^d$ . 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 14 then equals  $W_1^{a,b}(\mu_+ + \nu_-, \mu_- + \nu_+) = \mathbb{W}_1^{a,b}(\mu, \nu)$ .

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 14. □

**Definition 19.** 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 20.** The space of signed measures  $(\mathcal{M}^s(\mathbb{R}^d), \|\cdot\|^{a,b})$  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 18, 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 13, we have

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

□

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

**Example 21.** Take  $\mu = \delta_x - \delta_y$ . Then

$$\|\mu\|^{a,b} = \mathbb{W}_1^{a,b}(\delta_x - \delta_y, 0) = W_1^{a,b}(\delta_x, \delta_y) = \inf_{\substack{\tilde{\mu}, \tilde{\nu} \in \mathcal{M}(\mathbb{R}^d) \\ |\tilde{\mu}| = |\tilde{\nu}|}} \{a(|\delta_x - \tilde{\mu}| + |\delta_y - \tilde{\nu}|) + bW_1(\tilde{\mu}, \tilde{\nu})\}.$$

Using Lemma 12, 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

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

The expression above depends on the distance between the Dirac masses  $\delta_x$  and  $\delta_y$ . For  $b|x - y| < 2a$ , then 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  $b|x - y| \geq 2a$ , then the minimum is attained for  $\epsilon = 0$  and  $\|\mu\|^{a,b} = 2a$ , and we say that all the mass is cancelled (or removed).

### 3.1 Topological properties

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

**Proposition 22.** For  $a > 0$ ,  $b > 0$ , the norm  $\|\cdot\|^{a,b}$  is equivalent to  $\|\cdot\|^{1,1}$ .

*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 12. 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 give now 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

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

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 duality for the norm  $\mathbb{W}_1^{1,1}$ . This shows that  $\mathbb{W}_1^{1,1}$  coincides with the bounded Lipschitz distance introduced in [12], also called Fortet Mourier distance in [22].

**Proposition 23** (Kantorovich duality). *The signed generalized Wasserstein distance  $\mathbb{W}_1^{1,1}$  coincides with the bounded Lipschitz distance: for  $\mu, \nu$  in  $\mathcal{M}^s(\mathbb{R}^d)$ , it holds*

$$\mathbb{W}_1^{1,1}(\mu, \nu) = \sup \left\{ \int_{\mathbb{R}^d} \varphi d(\mu - \nu); \varphi \in \mathcal{C}_b^{0,Lip}, \|\varphi\|_\infty \leq 1, \|\varphi\|_{Lip} \leq 1 \right\}$$

We emphasize that Proposition 23 does not coincide with Lemma 13, since it involves non-compactly supported test functions.

*Proof.* By using Lemma 13 we have

$$\begin{aligned} \mathbb{W}_1^{1,1}(\mu, \nu) &= W_1^{a,b}(\mu_+ + \nu_-, \nu_+ + \mu_-) \\ &= \sup \left\{ \int_{\mathbb{R}^d} \varphi d(\mu_+ - \mu_- - (\nu_+ - \nu_-)); \varphi \in \mathcal{C}_c^{0,Lip}, \|\varphi\|_\infty \leq 1, \|\varphi\|_{Lip} \leq 1 \right\} \\ &= \sup \left\{ \int_{\mathbb{R}^d} \varphi d(\mu - \nu); \varphi \in \mathcal{C}_c^{0,Lip}, \|\varphi\|_\infty \leq 1, \|\varphi\|_{Lip} \leq 1 \right\}. \end{aligned}$$

We denote by

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

First observe that  $S < +\infty$ . Indeed, it holds  $\int_{\mathbb{R}^d} \varphi d(\mu - \nu) \leq \|\varphi\|_\infty(|\mu| + |\nu|) < +\infty$ . Denote with  $\varphi_n$  a sequence of functions of  $\mathcal{C}_b^{0,Lip}$  such that  $\int_{\mathbb{R}^d} \varphi_n d(\mu - \nu) \rightarrow S$  as  $n \rightarrow \infty$ . 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  $\|\rho_n\|_\infty \leq 1$ . For the sequence  $\psi_n = \varphi_n \rho_n$  of functions of  $\mathcal{C}_c^{0,Lip}$ , it holds

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

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

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

and Proposition 23 is proved.  $\square$

**Remark 24.** We observe that a sequence  $\mu_n$  of  $\mathcal{M}^s(\mathbb{R})$  which satisfies  $\|\mu_n\|^{a,b} \xrightarrow{n \rightarrow \infty} 0$  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

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

See Example 21 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 21, depending on the sign of  $2a - \frac{2b}{n^2}$ , we either cancel the mass or transport it. For  $n$  large enough,  $2a \geq \frac{2b}{n^2}$ , so we transport the mass. Thus for  $n$  sufficiently large

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

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

**Remark 25.** Norm  $\|\cdot\|^{1,1}$  does not metrize narrow convergence, contrarily to what is stated in [12]. Indeed, take  $\mu_n = \delta_{\sqrt{2\pi n + \frac{\pi}{2}}} - \delta_{\sqrt{2\pi n + \frac{3\pi}{2}}}$ . We have

$$\|\mu_n\|^{1,1} \leq \left| \sqrt{2\pi n + \frac{\pi}{2}} - \sqrt{2\pi n + \frac{3\pi}{2}} \right| \xrightarrow{n \rightarrow \infty} 0,$$

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 26.** We have as a direct consequence of Proposition 23 that

$$\|\mu_n - \mu\|^{a,b} \xrightarrow{n \rightarrow \infty} 0 \quad \Rightarrow \quad \forall \varphi \in \mathcal{C}_b^{0,Lip}(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} \varphi d\mu_n \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}^d} \varphi d\mu. \quad (12)$$

However, the reciprocal statement of (12) is false: define

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

For

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

it is clear that

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

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 23,  $\|\mu_n - 0\| \geq \frac{\pi}{2}$  does not converge to zero. We now prove that, for each  $\varphi$  in  $\mathcal{C}_b^{0,Lip}(\mathbb{R})$ , it holds  $\int_{\mathbb{R}} \varphi d\mu_n \rightarrow 0$ . 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 we 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  $na_n \rightarrow 0$  for any  $2\pi$ -periodic Lipschitz function  $f$ , following the ideas of [23, 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,  $b'_n \rightarrow 0$ , and this implies  $na_n \rightarrow 0$ .

**Proposition 27.** Assume that  $\|\mu_n\|^{a,b} \xrightarrow{n \rightarrow \infty} 0$ , then  $\Delta m_n := |\mu_n^+| - |\mu_n^-| \xrightarrow{n \rightarrow \infty} 0$ .

*Proof.* We have by definition  $\|\mu_n\|^{a,b} = W_1^{a,b}(\mu_n^+, \mu_n^-)$ . We denote by  $\bar{\mu}_n^+, \bar{\mu}_n^-$  the minimizers in the right hand side of (9) realizing the distance  $W_1^{a,b}(\mu_n^+, \mu_n^-)$ . We have

$$\|\mu_n\|^{a,b} = a (|\mu_n^+ - \bar{\mu}_n^+| + |\mu_n^- - \bar{\mu}_n^-|) + b W_1(\bar{\mu}_n^+, \bar{\mu}_n^-), \quad |\bar{\mu}_n^+| = |\bar{\mu}_n^-|.$$

Since  $\|\mu_n\|^{a,b} \xrightarrow{n \rightarrow \infty} 0$ , each of the three terms converges to zero as well. Since by Lemma 12 we can assume  $\bar{\mu}_n^+ \leq \mu_n^+$  and  $\bar{\mu}_n^- \leq \mu_n^-$ , we have

$$\begin{aligned} \left| |\mu_n^+| - |\mu_n^-| \right| &= \left| |\mu_n^+ - \bar{\mu}_n^+ + \bar{\mu}_n^+| - |\mu_n^- - \bar{\mu}_n^- + \bar{\mu}_n^-| \right| \\ &= \left| |\mu_n^+ - \bar{\mu}_n^+| + |\bar{\mu}_n^+| - |\mu_n^- - \bar{\mu}_n^-| - |\bar{\mu}_n^-| \right| \\ &= \left| |\mu_n^+ - \bar{\mu}_n^+| - |\mu_n^- - \bar{\mu}_n^-| \right| \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

□

We remind 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  $(\mathcal{M}^s(\mathbb{R}^d), \|\cdot\|^{a,b})$  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  $(\mathcal{M}^s(\mathbb{R}^d), \|\cdot\|^{a,b})$ : 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*

$$W_1^{a,b}(\mu_n, \mu_{n+k}) \leq 2b \sum_{i=n+1}^{n+k} \frac{1}{2^i} \leq 2b \sum_{i=n+1}^{+\infty} \frac{1}{2^i} \xrightarrow{n \rightarrow \infty} 0.$$

*However, the sequence  $(\mu_n)_n$  does not converge in  $(\mathcal{M}^s(\mathbb{R}^d), \|\cdot\|^{a,b})$ . As seen in Remark 26, the convergence for the norm  $\|\cdot\|^{a,b}$  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*

$$\langle \mu_n - \mu, \varphi \rangle = \sum_{i=n+1}^{+\infty} \left( \varphi\left(i + \frac{1}{2^i}\right) - \varphi\left(i - \frac{1}{2^i}\right) \right) \xrightarrow{n \rightarrow \infty} 0.$$

*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  $(\mathcal{M}^s(\mathbb{R}^d), \|\cdot\|^{a,b})$ . If  $\mu_n$  is tight and has uniformly bounded mass, then it converges in  $(\mathcal{M}^s(\mathbb{R}^d), \|\cdot\|^{a,b})$ .*

*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,  $\mu_{\varphi_1(n)}^+ \xrightarrow{n \rightarrow \infty} \mu^+$  vaguely. Then,  $|\mu_n^-|$  being uniformly bounded, there exists  $\varphi_2$  non decreasing such that for  $\varphi = \varphi_1 \circ \varphi_2$  it holds

$$\mu_{\varphi(n)}^- \xrightarrow{n \rightarrow \infty} \mu^- \text{ vaguely.}$$

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  $W_1^{a,b}(\mu_{\varphi(n)}^+, \mu^+) \xrightarrow{n \rightarrow \infty} 0$  and  $W_1^{a,b}(\mu_{\varphi(n)}^-, \mu^-) \xrightarrow{n \rightarrow \infty} 0$  (see [19, Theorem 13]). Then, we have

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

since  $(\mu_n)_n$  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  $n$ .

**Theorem 30.** *The two following statements are equivalent:*

$$(i) \quad \|\mu_n - \mu\|^{a,b} \xrightarrow{n \rightarrow \infty} 0.$$

(ii) *There exists four positive measures  $z_n^+, z_n^-, m_n^+, m_n^- \in \mathcal{M}(\mathbb{R}^d)$  such that*

$$\begin{aligned} & W_1^{a,b}(z_n^+, z_n^-) \xrightarrow{n \rightarrow \infty} 0, \\ \mu_n^+ = z_n^+ + m_n^+, & \quad \text{with} \quad W_1^{a,b}(m_n^+, \mu^+) \xrightarrow{n \rightarrow \infty} 0, \\ \mu_n^- = z_n^- + m_n^-, & \quad W_1^{a,b}(m_n^-, \mu^-) \xrightarrow{n \rightarrow \infty} 0, \\ & \quad \{m_n^+\}_n \text{ and } \{m_n^-\}_n \text{ are tight and bounded in mass,} \end{aligned}$$

where  $\mu = \mu^+ - \mu^-$  is the Jordan decomposition, and  $\mu_n = \mu_n^+ - \mu_n^-$  is any decomposition.

*Proof.* We start by proving (i)  $\Rightarrow$  (ii). We have  $\|\mu_n - \mu\|^{a,b} = \mathbb{W}_1^{a,b}(\mu_n, \mu) = W_1^{a,b}(\mu_n^+ + \mu^-, \mu_n^- + \mu^+) \xrightarrow{n \rightarrow \infty} 0$ . We denote by  $a_n \leq (\mu_n^+ + \mu^-)$  and  $b_n \leq (\mu_n^- + \mu^+)$  a choice of minimizers realizing  $W_1^{a,b}(\mu_n^+ + \mu^-, \mu_n^- + \mu^+)$ , as well as  $\pi_n$  being a minimizing transference plan from  $a_n$  to  $b_n$ . We have  $|a_n| = |b_n|$ . Following the wording of Example 21, the measures  $a_n$  and  $b_n$  are the transported mass, and the measures  $(\mu_n^+ - \mu^- - a_n)$  and  $(\mu_n^- - \mu^+ - b_n)$  are the cancelled mass.

**Step 1. The cancelled mass.** We define by  $a_n^+$  and  $b_n^-$  the largest transported mass which is respectively below  $\mu_n^+$  and  $\mu_n^-$

$$\begin{aligned} a_n^+ &= \mu_n^+ \wedge a_n, & b_n^- &= \mu_n^- \wedge b_n, \\ a_n^- &= a_n - a_n^+, & b_n^+ &= b_n - b_n^-. \end{aligned}$$

The mass which is cancelled is then  $r_n = r_n^+ + r_n^- := (\mu_n^+ - a_n^+) + (\mu_n^- - a_n^-)$  and  $r_n^* = r_n^{*, -} + r_n^{*, +} := (\mu_n^- - b_n^-) + (\mu_n^+ - b_n^+)$ . The cancelled mass  $r_n$  and  $r_n^*$  are expressed here as the sum of two positive measures. Indeed, it is clear by definition that  $a_n^+ \leq \mu_n^+$ , and since  $a_n \leq \mu_n^+ + \mu^-$ , Lemma 5, gives that  $a_n^- = a_n - a_n^+ \wedge \mu_n^+ \leq \mu^-$ . We reason the same way for  $r_n^*$ . Then, we have  $W_1^{a,b}(\mu_n^+ + \mu^-, \mu_n^- + \mu^+) = a(|\mu_n^+ - a_n^+| + |\mu_n^- - a_n^-| + |\mu_n^- - b_n^-| + |\mu_n^+ - b_n^+|) + bW_1(a_n, b_n)$ . Since  $W_1^{a,b}(\mu_n^+ + \mu^-, \mu_n^- + \mu^+)$  goes to zero, each of the five terms of the above decomposition goes to zero, and in particular,  $|\mu_n^+ - a_n^+| \xrightarrow{n \rightarrow \infty} 0$  and  $|\mu_n^- - b_n^-| \xrightarrow{n \rightarrow \infty} 0$  which implies that

$$W_1^{a,b}(\mu_n^+ - a_n^+, 0) \xrightarrow{n \rightarrow \infty} 0, \quad W_1^{a,b}(\mu_n^- - b_n^-, 0) \xrightarrow{n \rightarrow \infty} 0. \quad (13)$$

**Step 2. The transported mass.** The mass  $a_n^+$  is split into two pieces:  $\nu_n$  is sent to  $\mu_n^-$ , and  $\xi_n$  is sent to  $\mu^+$ . Denote by  $\bar{a}_n^+$  the image of  $a_n^+$  under  $\pi_n$  (using Definition 15), then we define  $\nu_n^* = \bar{a}_n^+ \wedge \mu_n^-$ . (Still using Definition 15), we denote by  $\nu_n$  the image of  $\nu_n^*$  under  $\pi_n$ . Then, we define  $\xi_n$  such that  $a_n^+ = \nu_n + \xi_n$ , and we denote by  $\xi_n^*$  the image of  $\xi_n$  under  $\pi_n$ . By definition, we have

$$W_1(a_n, b_n) = W_1(\nu_n, \nu_n^*) + W_1(\xi_n, \xi_n^*) + W_1(w_n, w_n^*) + W_1(\alpha_n, \alpha_n^*), \quad (14)$$

with  $a_n^+ = \nu_n + \xi_n$ ,  $w_n^*$  is defined so that  $b_n^- = \nu_n^* + w_n^*$ ,  $w_n$  is the image of  $w_n^*$  under  $\pi_n$ ,  $\alpha_n$  is defined so that  $\mu^- = w_n + \alpha_n$ ,  $\alpha_n^*$  is the image of  $\alpha_n$  under  $\pi_n$ , and it can be checked that  $\mu^+ = \xi_n^* + \alpha_n^*$ . Since  $W_1(a_n, b_n) \xrightarrow{n \rightarrow \infty} 0$ , each of the four term of the sum (14) is going to zero.

**Step 3. Conclusion.**

Let us write

$$z_n^+ = \nu_n + (\mu_n^+ - a_n^+), \quad z_n^- = \nu_n^* + (\mu_n^- - b_n^-), \quad m_n^+ = \xi_n, \quad m_n^- = w_n^*.$$

We show here that the sequences defined hereinabove satisfy the conditions stated in (ii). First, we have  $z_n^+ + m_n^+ = \nu_n^+ + (\mu_n^+ - a_n^+) + \xi_n = \mu_n^+$  and similarly,  $z_n^- + m_n^- = \nu_n^* + (\mu_n^- - b_n^-) + w_n^* = \mu_n^-$ .

Then, we have

$$\begin{aligned} W_1^{a,b}(z_n^+, z_n^-) &= W_1^{a,b}(\nu_n + (\mu_n^+ - a_n^+), \nu_n^* + (\mu_n^- - b_n^-)) \\ &\leq W_1^{a,b}(\nu_n, \nu_n^*) + W_1^{a,b}(\mu_n^+ - a_n^+, \mu_n^- - b_n^-) \quad \text{using Lemma 14} \\ &\leq bW_1(\nu_n, \nu_n^*) + W_1^{a,b}(\mu_n^+ - a_n^+, 0) + W_1^{a,b}(0, \mu_n^- - b_n^-) \\ &\xrightarrow{n \rightarrow \infty} 0, \quad \text{using (13) and (14)}. \end{aligned}$$

Here, we also used that for  $|\mu| = |\nu|$ ,  $W_1^{a,b}(\mu, \nu) \leq bW_1(\mu, \nu)$ . This is trivial with the definition of  $W_1^{a,b}$ . Now, we also have

$$\begin{aligned} W_1^{a,b}(m_n^+, \mu^+) &= W_1^{a,b}(\xi_n, \mu^+) \leq W_1^{a,b}(\xi_n, \xi_n^*) + W_1^{a,b}(\xi_n^*, b_n^+) + W_1^{a,b}(b_n^+, \mu^+) \quad (\text{triangular inequality}) \\ &= W_1^{a,b}(\xi_n, \xi_n^*) + W_1^{a,b}(\alpha_n^*, 0) + W_1^{a,b}(\mu^+ - b_n^+, 0) \end{aligned}$$

since  $\alpha_n^* + \xi_n^* = b_n^+$ . We know that  $W_1^{a,b}(\xi_n, \xi_n^*) \leq bW_1(\xi_n, \xi_n^*) \xrightarrow{n \rightarrow \infty} 0$  using (14), and that  $W_1^{a,b}(\mu^+ - b_n^+, 0) \xrightarrow{n \rightarrow \infty} 0$  using (13). Let us explain now why  $W_1^{a,b}(\alpha_n^*, 0) \xrightarrow{n \rightarrow \infty} 0$ . We remind that  $W_1(\alpha_n, \alpha_n^*) \xrightarrow{n \rightarrow \infty} 0$ ,  $\alpha_n \leq a_n^- \leq \mu^-$ ,  $\alpha_n^* \leq b_n^+ \leq \mu^+$ . Since  $(\alpha_n)_n$  is uniformly bounded in mass, then there exists  $\alpha \in \mathcal{M}(\mathbb{R}^d)$  such that  $\alpha_{\varphi(n)} \xrightarrow{n \rightarrow \infty} \alpha$  vaguely (see Lemma 9). We have also that  $(\alpha_{\varphi(n)})_n$  is tight, since  $\alpha_{\varphi(n)} \leq \mu^-$  which has a finite mass. Using Theorem 13 of [18], we deduce that  $W_1^{a,b}(\alpha_{\varphi(n)}, \alpha) \xrightarrow{n \rightarrow \infty} 0$ . Then,  $W_1^{a,b}(\alpha_{\varphi(n)}^*, \alpha) \leq W_1^{a,b}(\alpha_{\varphi(n)}^*, \alpha_{\varphi(n)}) + W_1^{a,b}(\alpha_{\varphi(n)}, \alpha) \leq W_1(\alpha_{\varphi(n)}^*, \alpha_{\varphi(n)}) + W_1^{a,b}(\alpha_{\varphi(n)}, \alpha) \xrightarrow{n \rightarrow \infty} 0$ . Then, using again Theorem 13 of [18], we deduce that  $\alpha_{\varphi(n)}^* \xrightarrow{n \rightarrow \infty} \alpha$  vaguely. Since  $\alpha_n \leq \mu^-$ , we have  $\alpha \leq \mu^-$ . Likewise,  $\alpha_n^* \leq \mu^+$  implies  $\alpha \leq \mu^+$ . Since  $\mu^- \perp \mu^+$ , we have  $\alpha = 0$ . We have  $W_1^{a,b}(\alpha_{\varphi(n)}, 0) \xrightarrow{n \rightarrow \infty} 0$  and  $W_1^{a,b}(\alpha_{\varphi(n)}^*, 0) \xrightarrow{n \rightarrow \infty} 0$ . The sequence  $(\alpha_n)_n$  satisfies the following property: each of its subsequences admits a subsequence converging to zero. Thus, we have that the whole sequence is converging to zero, i.e.  $W_1^{a,b}(\alpha_n, 0) \xrightarrow{n \rightarrow \infty} 0$  and  $W_1^{a,b}(\alpha_n^*, 0) \xrightarrow{n \rightarrow \infty} 0$ . Lastly, the tightness of  $(m_n^+)_n$  and  $(m_n^-)_n$  is given again by Theorem 13 of [18], since  $W_1^{a,b}(m_n^\pm, \mu^\pm) \xrightarrow{n \rightarrow \infty} 0$ .

We prove now that (ii)  $\Rightarrow$  (i). Let us assume (ii). We have

$$\begin{aligned} \|\mu_n - \mu\|^{a,b} &= W_1^{a,b}(\mu_n^+ + \mu^-, \mu_n^- + \mu^+) = W_1^{a,b}(z_n^+ + m_n^+ + \mu^-, z_n^- + m_n^- + \mu^+) \\ &\leq W_1^{a,b}(z_n^+, z_n^-) + W_1^{a,b}(m_n^+, \mu^+) + W_1^{a,b}(\mu^-, m_n^-) \xrightarrow{n \rightarrow \infty} 0, \end{aligned}$$

where the last inequality comes from Lemma 14. This proves which is (i).  $\square$



## 4 Application to the transport equation with source term

This section is devoted to the use of the norm defined in Definition 19 to guarantee existence, uniqueness, and stability with respect to initial condition for the transport equation (2).

**Definition 31** (Measure-valued weak solution). *A measure-valued weak solution to (2) is a map  $\mu \in \mathcal{C}^0([0, 1]; \mathcal{M}^s(\mathbb{R}^d))$  such that  $\mu_{t=0} = \mu_0$  and for all  $\Phi \in \mathcal{D}(\mathbb{R}^d)$  it holds*

$$\frac{d}{dt} \langle \Phi, \mu_t \rangle = \langle v[\mu_t] \cdot \nabla \Phi, \mu_t \rangle + \langle h[\mu_t], \Phi \rangle, \quad (15)$$

where

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

### 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 variation of  $\|\mu - \nu\|^{a,b}$  after the action of a flow on  $\mu$  and  $\nu$ . 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  $|T\#\mu| \leq |\mu|$ , 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 32.** *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*

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

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

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

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

*Proof.* The proof is a direct consequence of [21, Theorem 5.34] combined with [6, Theorem 2.1.1].  $\square$

**Lemma 33.** *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:*

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

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

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

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

$$\begin{aligned}
\|\phi_t^v \# \mu - \phi_t^v \# \nu\|^{a,b} &= \mathbb{W}_1^{a,b}(\phi_t^v \# \mu, \phi_t^v \# \nu) = \mathbb{W}_1^{a,b}(\phi_t^v \# \mu^+ - \phi_t^v \# \mu^-, \phi_t^v \# \nu^+ - \phi_t^v \# \nu^-) \\
&= W_1^{a,b}(\phi_t^v \# (\mu^+ + \nu^-), \phi_t^v \# (\mu^- + \nu^+)) \\
&\leq e^{Lt} W_1^{a,b}(\mu^+ + \nu^-, \mu^- + \nu^+) \quad \text{by [20, Prop. 10]} \\
&= e^{Lt} \|\mu - \nu\|^{a,b}.
\end{aligned}$$

For the second inequality,

$$\begin{aligned}
\mathbb{W}_1^{a,b}(\mu, \phi_t^v \# \mu) &= W_1^{a,b}(\mu^+ + \phi_t^v \# \mu^-, \mu^- + \phi_t^v \# \mu^+) \\
&\leq W_1^{a,b}(\mu^+, \phi_t^v \# \mu^+) + W_1^{a,b}(\mu^-, \phi_t^v \# \mu^-) \quad (\text{Lemma 14}) \\
&\leq b t \|v\|_{\mathcal{C}^0} (|\mu^+| + |\mu^-|) \quad \text{by [20, Prop. 10]} \\
&= b t \|v\|_{L^\infty(0,1; \mathcal{C}^0(\mathbb{R}))} |\mu| \quad \text{since } \mu = \mu^+ - \mu^- \text{ is the Jordan decomposition.}
\end{aligned}$$

The third inequality is given by

$$\begin{aligned}
\|\phi_t^v \# \mu - \phi_t^w \# \mu\|^{a,b} &= \mathbb{W}_1^{a,b}(\phi_t^v \# \mu^+ + \phi_t^w \# \mu^-, \phi_t^w \# \mu^+ + \phi_t^v \# \mu^-) \\
&\leq W_1^{a,b}(\phi_t^v \# \mu^+, \phi_t^w \# \mu^+) + W_1^{a,b}(\phi_t^w \# \mu^-, \phi_t^v \# \mu^-) \\
&\leq b W_1(\phi_t^v \# \mu^+, \phi_t^w \# \mu^+) + W_1(\phi_t^w \# \mu^-, \phi_t^v \# \mu^-) \\
&\leq (|\mu^+| + |\mu^-|) \frac{(e^{Lt} - 1)}{L} \|v - w\|_{L^\infty(0,1; \mathcal{C}^0(\mathbb{R}))} \quad \text{using [20, Prop. 10] with } \mu = \nu.
\end{aligned}$$

The last inequality is deduced from the first and the third one using triangular inequality.  $\square$

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

In this section, we build a solution to (2) as the limit of a sequence of approximated solutions defined in the following scheme. We then prove that (2) admits a unique solution.

Consider  $\mu_0 \in \mathcal{M}^s(\mathbb{R}^d)$  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  $h \in \mathcal{C}^{0,Lip}(\mathcal{M}^s(\mathbb{R}^d), \mathcal{M}^s(\mathbb{R}^d))$  satisfying (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  $[0, 1]$ .

SCHEME

**Initialization.** Fix  $k \in \mathbb{N}$ . Define  $\Delta t = \frac{1}{2^k}$ . Set  $\mu_0^k = \mu_0$ .

**Induction.** Given  $\mu_{i\Delta t}^k$  for  $i \in \{0, 1, \dots, 2^k - 1\}$ , define  $v_{i\Delta t}^k := v[\mu_{i\Delta t}^k]$  and

$$\mu_t^k = \Phi_{t-i\Delta t}^{v_{i\Delta t}^k} \# \mu_{i\Delta t}^k + (t - i\Delta t)h[\mu_{i\Delta t}^k], \quad t \in [i\Delta t, (i+1)\Delta t]. \quad (16)$$

The scheme is a natural operator splitting: the flow  $\Phi_{t-i\Delta t}$  encodes the transport part  $\partial_t \mu + \text{div}(v\mu) = 0$  while  $(t - i\Delta t)h$  encodes the reaction  $\partial_t \mu = h$ .

**Proposition 34.** *The sequence  $(\mu_t^k)_k$  defined in the scheme above is a Cauchy sequence in the space  $\mathcal{C}^0([0, 1], \mathcal{M}^s(\mathbb{R}^d), \|\cdot\|)$  with*

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

Moreover, it is uniformly bounded in mass and compactly supported, i.e.

$$|\mu_t^k| < Pt + |\mu_0|, \quad \text{supp}\{\mu_t\} \subset B(0, R' + tM), \quad 0 \leq t \leq 1, \quad (17)$$

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

Let us mention that the estimate (17) is expected at the discrete level from the PDE (2) with the assumptions (H-1), (H-2), (H-3). Indeed, the transport part should preserve mass (more precisely  $|T\#\mu| \leq |\mu|$  as discussed in subsection 4.1, while the reaction term  $|h| \leq P$  gives a mass growth that is at most linear  $Pt$ . 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 faster than  $|v| \leq M$ .

*Proof.* Let  $L$  be the Lipschitz constant in (H-2). We assume to have  $k$  sufficiently large to have  $e^{Lt} \leq 1 + 2Lt$  for all  $t \leq [0, \Delta t]$ . This holds e.g. for  $L\Delta t \leq 1$ , hence  $k \geq \log_2(L)$ .

We also notice that the sequence built by the scheme satisfies

$$|\mu_t^k| \leq Pt + |\mu_0|, \quad t \in [0, 1], \quad (18)$$

where  $P$  is such that  $|h[\mu]| \leq P$  by (H-3). Indeed, it holds for  $t \in [i\Delta t, (i+1)\Delta t]$

$$|\mu_t^k| \leq |\Phi_t^{v_{i\Delta t}} \# \mu_{i\Delta t}^k| + \Delta t |h[\mu_{i\Delta t}^k]| \leq |\mu_{i\Delta t}^k| + \Delta t P,$$

and the result follows by induction on  $i$  (for  $k$  fixed). This proves (17). The sequence  $(\mu_t^k)_{k \in \mathbb{N}}$  also has uniformly bounded support. Indeed, first observe that  $\text{supp}\{\mu\} = \text{supp}\{\mu^+\} \cup \text{supp}\{\mu^-\}$ , where  $(\mu^+, \mu^-)$  is the Jordan decomposition of  $\mu$ . Choose  $\mathcal{K}$  such that  $\text{supp}\{\mu_0\} \subset \mathcal{K}$  and use (16) 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  $\mathcal{K}_{t,M,R} \subset B(0, R' + tM)$ . Since such set does not depend on  $t$  for  $0 \leq t \leq 1$ , while  $M, R$  are fixed, then  $\mu_t^k$  have uniformly bounded support.

We now follow the notations of [18] and define  $m_j^k := \mu_{\frac{j}{2^k}}^k$ ,  $v_j^k := v[m_j^k]$  and the corresponding flow  $f_t^{j,k} := \phi_t^{v_j^k}$ . Fix  $k \in \mathbb{N}$  and  $t \in [0, 1]$ . Define  $j \in \{0, 1, \dots, 2^k\}$  such that  $t \in \left] \frac{j}{2^k}, \frac{j+1}{2^k} \right]$ . The following inequalities rely on Lemma 33 and (H-1), (H-2), (H-3).

**First case.** If  $t \in \left] \frac{j}{2^k}, \frac{2j+1}{2^{k+1}} \right]$ .

We call  $t' = t - \frac{j}{2^k} \leq \frac{1}{2^{k+1}}$  and we obtain

$$\begin{aligned} \mathbb{W}_1^{a,b}(\mu_t^k, \mu_t^{k+1}) &= \mathbb{W}_1^{a,b}(f_{t'}^{j,k} \# m_j^k + t' h[m_j^k], f_{t'}^{2j,k+1} \# m_{2j}^{k+1} + t' h[m_{2j}^{k+1}]) \\ &\leq \mathbb{W}_1^{a,b}(f_{t'}^{j,k} \# m_j^k, f_{t'}^{2j,k+1} \# m_{2j}^{k+1}) + \mathbb{W}_1^{a,b}(t' h[m_j^k], t' h[m_{2j}^{k+1}]) \\ &\leq e^{Lt'} \mathbb{W}_1^{a,b}(m_j^k, m_{2j}^{k+1}) + |m_j^k| \frac{(e^{Lt'} - 1)}{L} \|v_j^k - v_{2j}^{k+1}\|_{C^0(\mathbb{R}^d)} + t' Q \mathbb{W}_1^{a,b}(m_j^k, m_{2j}^{k+1}) \\ &\leq \mathbb{W}_1^{a,b}(m_j^k, m_{2j}^{k+1}) \left( e^{Lt'} + (P + |\mu^0|) \frac{1}{L} (e^{Lt'} - 1) + t' Q \right) \end{aligned}$$

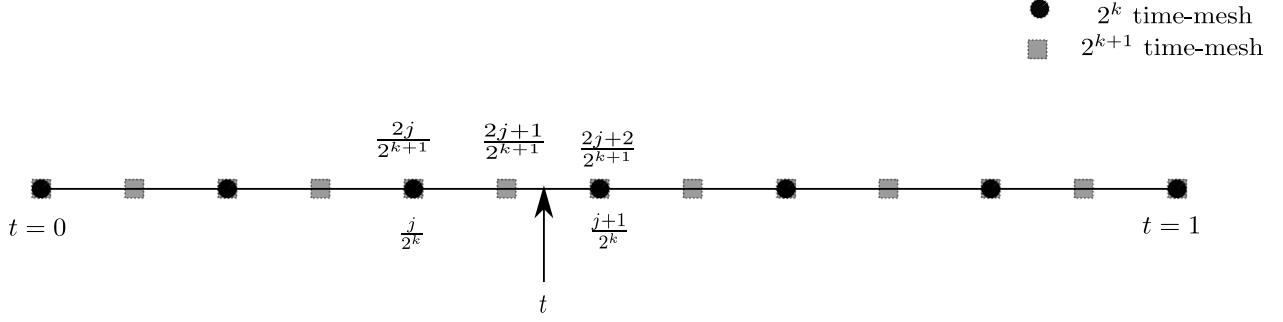
Since it holds

$$e^{Lt'} \leq 1 + 2Lt' \leq 1 + 2L2^{-(k+1)}, \quad \frac{(e^{Lt'} - 1)}{L} \leq 2 \cdot 2^{-(k+1)},$$

we have

$$\|\mu_t^k - \mu_t^{k+1}\|^{a,b} \leq \|m_j^k - m_{2j}^{k+1}\|^{a,b} \left(1 + 2^{-(k+1)} (2L + 2(P + |\mu^0|) + Q)\right), \quad t \in \left[\frac{j}{2^k}, \frac{2j+1}{2^{k+1}}\right]. \quad (19)$$

**Second case.** If  $t \in \left[\frac{2j+1}{2^{k+1}}, \frac{j+1}{2^k}\right]$ .



We call  $t' = t - \frac{2j+1}{2^{k+1}} \leq \frac{1}{2^{k+1}}$  and we obtain

$$\begin{aligned} \mu_t^k &= f_{t'+\frac{1}{2^{k+1}}}^{j,k} \# m_j^k + \left(t' + \frac{1}{2^{k+1}}\right) h[m_j^k] = f_{t'}^{j,k} \# f_{\frac{1}{2^{k+1}}}^{j,k} \# m_j^k + t' h[m_j^k] + \frac{1}{2^{k+1}} h[m_j^k], \\ \mu_t^{k+1} &= f_{t'}^{2j+1,k+1} \# \left(f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1} + \frac{1}{2^{k+1}} h[m_{2j}^{k+1}]\right) + t' h \left[f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1} + \frac{1}{2^{k+1}} h[m_{2j}^{k+1}]\right] \\ &= f_{t'}^{2j+1,k+1} \# f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1} + \frac{1}{2^{k+1}} f_{t'}^{2j+1,k+1} \# h[m_{2j}^{k+1}] + t' h \left[f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1} + \frac{1}{2^{k+1}} h[m_{2j}^{k+1}]\right]. \end{aligned}$$

It then holds

$$\begin{aligned} \|\mu_t^k - \mu_t^{k+1}\|^{a,b} &\leq \mathbb{W}_1^{a,b} \left( f_{t'}^{j,k} \# f_{\frac{1}{2^{k+1}}}^{j,k} \# m_j^k, f_{t'}^{2j+1,k+1} \# f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1} \right) \\ &\quad + \frac{1}{2^{k+1}} \mathbb{W}_1^{a,b} \left( h[m_j^k], f_{t'}^{2j+1,k+1} \# h[m_{2j}^{k+1}] \right) \\ &\quad + t' \mathbb{W}_1^{a,b} \left( h[m_j^k], h \left[ f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1} + \frac{1}{2^{k+1}} h[m_{2j}^{k+1}] \right] \right). \end{aligned} \quad (20)$$

Use now Lemma 33 to prove the following estimate

$$\begin{aligned} &\mathbb{W}_1^{a,b} \left( f_{t'}^{j,k} \# f_{\frac{1}{2^{k+1}}}^{j,k} \# m_j^k, f_{t'}^{2j+1,k+1} \# f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1} \right) \\ &\leq (1 + 2L2^{-(k+1)}) \mathbb{W}_1^{a,b} \left( f_{\frac{1}{2^{k+1}}}^{j,k} \# m_j^k, f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1} \right) + 2^{-(k+1)} 2P \|v_j^k - v_{2j+1}^{k+1}\|_{C^0(\mathbb{R}^d)}. \end{aligned}$$

Since, according to the first case, it holds both

$$\mathbb{W}_1^{a,b} \left( f_{\frac{1}{2^{k+1}}}^{j,k} \# m_j^k, f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1} \right) \leq \|m_j^k - m_{2j}^{k+1}\|^{a,b} \left(1 + 2^{-(k+1)} (2L + 2(P + |\mu^0|))\right)$$

and

$$\begin{aligned}
\|v_j^k - v_{2j+1}^{k+1}\|_{C^0(\mathbb{R}^d)} &\leq K\mathbb{W}_1^{a,b}(m_j^k, m_{2j+1}^{k+1}) \leq K\mathbb{W}_1^{a,b}(m_j^k, m_{2j}^{k+1}) + K\mathbb{W}_1^{a,b}(m_{2j}^{k+1}, m_{2j+1}^{k+1}) \\
&\leq K\mathbb{W}_1^{a,b}(m_j^k, m_{2j}^{k+1}) + K\mathbb{W}_1^{a,b}(m_{2j}^{k+1}, m_{2j+1}^{k+1}) \\
&= K\mathbb{W}_1^{a,b}(m_j^k, m_{2j}^{k+1}) + K\mathbb{W}_1^{a,b}(m_{2j}^{k+1}, f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1}) \\
&= K\mathbb{W}_1^{a,b}(m_j^k, m_{2j}^{k+1}) + KM2^{-(k+1)},
\end{aligned}$$

we have

$$\begin{aligned}
&\mathbb{W}_1^{a,b} \left( f_{t'}^{j,k} \# f_{\frac{1}{2^{k+1}}}^{j,k} \# m_j^k, f_{t'}^{2j+1,k+1} \# f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1} \right) \\
&\leq \|m_j^k - m_{2j}^{k+1}\|^{a,b} \left( 1 + 2^{-(k+1)} (4L + 2(P + |\mu^0|)(1 + L) + 2KP) \right) + 2^{-2(k+1)} 2PKM.
\end{aligned} \tag{21}$$

Moreover, it also holds both

$$\begin{aligned}
&\mathbb{W}_1^{a,b} \left( h[m_j^k], f_{t'}^{2j+1,k+1} \# h[m_{2j}^{k+1}] \right) \\
&\leq \mathbb{W}_1^{a,b} \left( h[m_j^k], f_{t'}^{2j+1,k+1} \# h[m_j^k] \right) + \mathbb{W}_1^{a,b} \left( f_{t'}^{2j+1,k+1} \# h[m_j^k], f_{t'}^{2j+1,k+1} \# h[m_{2j}^{k+1}] \right) \\
&\leq t'MP + e^{Lt'} Q \|m_j^k - m_{2j}^{k+1}\|^{a,b} \leq +MP2^{-(k+1)} + (1 + 2L2^{-(k+1)}) \|m_j^k - m_{2j}^{k+1}\|^{a,b},
\end{aligned} \tag{22}$$

and

$$\begin{aligned}
&\mathbb{W}_1^{a,b} \left( m_j^k, f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1} + \frac{1}{2^{k+1}} h[m_{2j}^{k+1}] \right) \\
&\leq \mathbb{W}_1^{a,b} \left( m_j^k, f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1} \right) + 2^{-(k+1)} \mathbb{W}_1^{a,b} \left( 0, h[m_{2j}^{k+1}] \right) \\
&\leq \mathbb{W}_1^{a,b} \left( m_j^k, m_{2j}^{k+1} \right) + \mathbb{W}_1^{a,b} \left( m_{2j}^{k+1}, f_{\frac{1}{2^{k+1}}}^{2j,k+1} \# m_{2j}^{k+1} \right) + 2^{-(k+1)} aP \\
&\leq \|m_j^k - m_{2j}^{k+1}\|^{a,b} + 2^{-(k+1)} (|\mu^0| + P(1 + a)).
\end{aligned} \tag{23}$$

Plugging (21), (22) and (23) into (20), and combining it with (19) we find **in both cases**

$$\|\mu_t^k - \mu_t^{k+1}\|^{a,b} \leq (1 + 2^{-k} C_1) \|m_j^k - m_{2j}^{k+1}\|^{a,b} + C_2 2^{-2k}, \quad t \in \left] \frac{j}{2^k}, \frac{j+1}{2^k} \right], \tag{24}$$

with

$$C_1 = (1 + 3L + (P + |\mu^0|)(1 + L) + KP + Q), \quad C_2 = \frac{1}{4} (MP(1 + 2K) + |\mu^0| + P(1 + a)).$$

In particular, plugging  $t = (j + 1)/2^k$  in (24), we get

$$\|m_{j+1}^k - m_{2(j+1)}^{k+1}\|^{a,b} \leq (1 + 2^{-k} C_1) \|m_j^k - m_{2j}^{k+1}\|^{a,b} + C_2 2^{-2k},$$

and by induction on  $j$  (for  $k$  fixed), we obtain

$$\|m_{2^k}^k - m_{2^{k+1}}^{k+1}\|^{a,b} \leq \sum_{j=0}^{2^k-1} (1 + 2^{-k} C_1)^j 2^{-2k} C_2 \leq \frac{C_2}{C_1} (e^{C_1} - 1) 2^{-k}.$$

From (24) it holds

$$\|\mu_t^k - \mu_t^{k+1}\| \leq \|m_{2^k}^k - m_{2^{k+1}}^{k+1}\|^{a,b},$$

and then we conclude

$$\|\mu_t^k - \mu_t^{k+1}\| \leq \frac{C_2}{C_1} (e^{C_1} - 1) 2^{-k}.$$

Since the right hand side is the term of a convergent series, then  $(\mu_t^k)_k$  is a Cauchy sequence.  $\square$

### 4.3 Proof of Theorem 1

In this section, we prove Theorem 1, 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.** Observe that the sequence given by the scheme  $(\mu_t^k)_k$  is a Cauchy sequence (Proposition 34) in the space  $(\mathcal{C}^0[0, 1], \mathcal{M}^s(\mathbb{R}^d))$  is uniformly bounded in mass and tight (see Proposition 34). Then, by using Theorem 29, we define

$$\mu_t := \lim_{k \rightarrow \infty} \mu_t^k,$$

where the convergence holds in the space  $\mathcal{C}^0([0, 1], \mathcal{M}^s(\mathbb{R}^d))$ . Denote the following for  $\varphi \in \mathcal{D}((0, 1) \times \mathbb{R}^d)$ :

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

The goal is to prove that for all  $\varphi \in \mathcal{D}((0, 1) \times \mathbb{R}^d)$ , we have

$$\int_0^1 dt (\langle \mu_t, \partial_t \varphi(t, x) + v[\mu_t] \cdot \nabla \varphi(t, x) \rangle + \langle h[\mu_t], \varphi(t, x) \rangle) = 0.$$

This implies (it is equivalent) that for all  $\phi \in \mathcal{D}((0, 1) \times \mathbb{R}^d)$ , (15) holds (see [2, chapter 8]). We first notice that

$$\sum_{j=0}^{2^k-1} \int_{j\Delta t}^{(j+1)\Delta t} dt \left( \langle \mu_t^k, \partial_t \varphi(t, x) + v[\mu_{j\Delta t}^k] \cdot \nabla \varphi(t, x) \rangle + \langle h[\mu_{j\Delta t}^k], \varphi(t, x) \rangle \right) \xrightarrow{k \rightarrow \infty} 0$$

Indeed,  $\nu_t := \phi_t^v \# \nu_0$  is a weak solution of  $\frac{\partial}{\partial t} \nu_t + \nabla \cdot (v(x) \nu_t) = 0$  with  $v$  a fixed vector field, and  $\eta_t = \eta_0 + th$  is a weak solution of  $\frac{\partial}{\partial t} \eta_t = h$ , with  $h$  a fixed measure. We apply this to  $\mu^k$  piecewise on each time interval. It then holds, using that  $\mu_t^k$  satisfies (16)

$$\begin{aligned} & \left| \sum_{j=0}^{2^k-1} \int_{j\Delta t}^{(j+1)\Delta t} dt \left( \langle \mu_t^k, \partial_t \varphi(t, x) + v[\mu_{j\Delta t}^k] \cdot \nabla \varphi(t, x) \rangle + \langle h[\mu_{j\Delta t}^k], \varphi(t, x) \rangle \right) \right| \\ &= \left| \sum_{j=0}^{2^k-1} \int_{j\Delta t}^{(j+1)\Delta t} dt \langle (t - j\Delta t) h[\mu_{j\Delta t}^k], v[\mu_{j\Delta t}^k] \cdot \nabla \varphi(t, x) \rangle \right| \\ &\leq MP \|\nabla \varphi\|_{\infty} 2^{-(k+1)} \xrightarrow{k \rightarrow \infty} 0. \end{aligned}$$

Now, to guarantee (15), it is enough to prove that

$$\begin{aligned} & \lim_{k \rightarrow \infty} \left| \int_0^1 dt (\langle \mu_t, \partial_t \varphi(t, x) + v[\mu_t] \cdot \nabla \varphi(t, x) \rangle + \langle h[\mu_t], \varphi(t, x) \rangle) \right. \\ & \quad \left. - \sum_{j=0}^{2^k-1} \int_{j\Delta t}^{(j+1)\Delta t} dt \left( \langle \mu_t^k, \partial_t \varphi(t, x) + v[\mu_{j\Delta t}^k] \cdot \nabla \varphi(t, x) \rangle + \langle h[\mu_{j\Delta t}^k], \varphi(t, x) \rangle \right) \right| = 0 \end{aligned}$$

We have

$$\left| \int_0^1 dt (\langle \mu_t, \partial_t \varphi(t, x) \rangle) - \sum_{j=0}^{2^k-1} \int_{j\Delta t}^{(j+1)\Delta t} dt \left( \langle \mu_t^k, \partial_t \varphi(t, x) \rangle \right) \right| \leq \|\partial_t \varphi\|_{\infty} \|\mu_t - \mu_t^k\| \xrightarrow{k \rightarrow \infty} 0,$$

$$\left| \int_0^1 dt \langle h[\mu_t], \varphi(t, x) \rangle - \sum_{j=0}^{2^k-1} \int_{j\Delta t}^{(j+1)\Delta t} dt \langle h[\mu_t^k], \varphi(t, x) \rangle \right| \leq Q \|\varphi\|_\infty \|\mu_t - \mu_t^k\| \xrightarrow[k \rightarrow \infty]{} 0,$$

and

$$\begin{aligned} & \left| \int_0^1 dt \langle \mu_t, v[\mu_t] \cdot \nabla \varphi(t, x) \rangle - \sum_{j=0}^{2^k-1} \int_{j\Delta t}^{(j+1)\Delta t} dt \langle \mu_t^k, v[\mu_{j\Delta t}^k] \cdot \nabla \varphi(t, x) \rangle \right| \\ & \leq \left| \sum_{j=0}^{2^k-1} \int_{j\Delta t}^{(j+1)\Delta t} dt \langle \mu_t^k - \mu_t, v[\mu_{j\Delta t}^k] \cdot \nabla \varphi(t, x) \rangle \right| + \left| \sum_{j=0}^{2^k-1} \int_{j\Delta t}^{(j+1)\Delta t} dt \langle \mu_t^k, (v[\mu_{j\Delta t}^k] - v[\mu_t^k]) \cdot \nabla \varphi(t, x) \rangle \right| \\ & + \left| \sum_{j=0}^{2^k-1} \int_{j\Delta t}^{(j+1)\Delta t} dt \langle \mu_t^k, (v[\mu_t] - v[\mu_t^k]) \cdot \nabla \varphi(t, x) \rangle \right| \\ & \leq \|\nabla \varphi\|_\infty \left( M \|\mu_t - \mu_t^k\| + LM(P + |\mu_0|)2^{-(k+1)} + (P + |\mu_0|)L \|\mu_t - \mu_t^k\| \right) \xrightarrow[k \rightarrow \infty]{} 0. \end{aligned}$$

This proves (15).

**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 31 to the transport equation (2) is Lipschitz with respect to time, since it satisfies

$$\|\mu_{t+\tau} - \mu_t\|^{a,b} \leq L_1 \tau, \quad t \geq 0, \tau \geq 0, \quad (25)$$

with  $L_1 = P + bM(P(t + \tau) + |\mu_0|)$ . 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

$$\partial_t \mu_t(x) + \operatorname{div} \cdot (w(t, x) \mu_t(x)) = b_t(x), \quad \mu_{|t=0}(x) = \mu_0(x). \quad (26)$$

Uniqueness of the solution of the linear equation (26) is a direct consequence of Lemma 32. Moreover, the scheme presented in Section 4.2 can be rewritten for the vector field  $w$  in which dependence with respect to time is added and dependence with respect to the measure is dropped. Thus, the unique solution  $\mu_t$  to (26) can be obtained as the limit of this scheme. We have for  $k \geq 0$  the following estimate

$$\|\mu_{t+\tau} - \mu_t\|^{a,b} \leq \|\mu_t - \mu_t^k\|^{a,b} + \|\mu_t^k - \mu_{t+\tau}^k\|^{a,b} + \|\mu_{t+\tau}^k - \mu_{t+\tau}\|^{a,b},$$

where  $\mu_t^k$  is given by the scheme. The first and third terms can be rendered as small as desired for  $k \geq k_0$  large enough, independent on  $t, \tau$ . For  $\ell := \min\{i \in \{1, \dots, 2^k\}, t \leq \frac{i}{2^k}\}$ ,  $j := \min\{i \in \{1, \dots, 2^k\}, t + \tau \leq \frac{i}{2^k}\}$  with the notations of the scheme, it holds

$$\begin{aligned} \|\mu_{t+\tau}^k - \mu_t^k\|^{a,b} &= \|m_j^k - m_\ell^k\|^{a,b} = \left\| \sum_{i=\ell}^{j-1} (m_{i+1}^k - m_i^k) \right\|^{a,b} = \left\| \sum_{i=\ell}^{j-1} (\phi_{\Delta t}^{v[m_i^k]} \# m_i^k + \Delta t h[m_i^k] - m_i^k) \right\|^{a,b} \\ &\leq \sum_{i=\ell}^{j-1} \|\phi_{\Delta t}^{v[m_i^k]} \# m_i^k - m_i^k\|^{a,b} + \Delta t \left\| \sum_{i=\ell}^{j-1} h[m_i^k] \right\|^{a,b}. \end{aligned}$$

Using Lemma 33 and (18), it holds

$$\sum_{i=\ell}^{j-1} \|\phi_{\Delta t}^{v[m_i^k]} \# m_i^k - m_i^k\|^{a,b} \leq \frac{j-\ell}{2^k} bM(P(t+\tau) + |\mu_0|) \leq bM(P(t+\tau) + |\mu_0|)\tau + \frac{bM(P(t+\tau) + |\mu_0|)}{2^k}. \quad (27)$$

Using (H-3), we have

$$\Delta t \left\| \sum_{i=\ell}^{j-1} h[m_i^k] \right\|^{a,b} \leq \frac{j-\ell}{2^k} P \leq P(t+\tau) \left( \tau + \frac{1}{2^k} \right), \quad (28)$$

Merging (27)-(28) and letting  $k \rightarrow \infty$ , we recover (25).

**Step 3. Any weak solution to (2) satisfies the operator splitting estimate:**

$$\|\mu_{t+\tau} - (\phi_{\tau}^{v[\mu_t]} \# \mu_t + \tau h[\mu_t])\|^{a,b} \leq K_1 \tau^2, \quad (29)$$

for some  $K_1 > 0$ . Indeed, let us consider a solution  $\mu_t$  to (2). As in the previous step,  $\mu_t$  is the unique solution to (26), and thus it can be obtained as the limit of the sequence provided by the scheme. With the notations used in Step 2 and using Lemma 33

$$\begin{aligned} \|\mu_{t+\tau} - (\phi_{\tau}^{v[\mu_t]} \# \mu_t + \tau h[\mu_t])\|^{a,b} &\leq \|\mu_{t+\tau} - \mu_{t+\tau}^k\|^{a,b} + \|\mu_{t+\tau}^k - (\phi_{\tau}^{v[\mu_t^k]} \# \mu_t^k + \tau h[\mu_t^k])\|^{a,b} \\ &\quad + \tau \|h[\mu_t^k] - h[\mu_t]\|^{a,b} + \|\phi_{\tau}^{v[\mu_t]} \# \mu_t - \phi_{\tau}^{v[\mu_t^k]} \# \mu_t^k\|^{a,b}. \end{aligned}$$

The first, third and fourth terms can be rendered as small as needed for  $k$  sufficiently large, independently on  $\tau$ . We focus then on the second term. Assume for simplicity that  $t = \ell\Delta t$  and  $t + \tau = (\ell + n)\Delta t$ , we have

$$\|\mu_{t+\tau}^k - (\phi_{\tau}^{v[\mu_t^k]} \# \mu_t^k + \tau h[\mu_t^k])\|^{a,b} = \|m_{\ell+n}^k - (\phi_{n\Delta t}^{v[m_{\ell}^k]} \# m_{\ell}^k + n\Delta t h[m_{\ell}^k])\|^{a,b}.$$

For  $n = 2$ , we have

$$\begin{aligned} \|m_{\ell+2}^k - (\phi_{2\Delta t}^{v[m_{\ell}^k]} \# m_{\ell}^k + 2\Delta t h[m_{\ell}^k])\|^{a,b} &= \|\phi_{\Delta t}^{v[m_{\ell+1}^k]} \# m_{\ell+1}^k + \Delta t h[m_{\ell+1}^k] - \phi_{\Delta t}^{v[m_{\ell}^k]} \# \phi_{\Delta t}^{v[m_{\ell}^k]} \# m_{\ell}^k - 2\Delta t h[m_{\ell}^k]\|^{a,b} \\ &= \|\phi_{\Delta t}^{v[m_{\ell+1}^k]} \# \left( \phi_{\Delta t}^{v[m_{\ell}^k]} \# m_{\ell}^k + \Delta t h[m_{\ell}^k] \right) + \Delta t h[m_{\ell+1}^k] - \phi_{\Delta t}^{v[m_{\ell}^k]} \# \phi_{\Delta t}^{v[m_{\ell}^k]} \# m_{\ell}^k - 2\Delta t h[m_{\ell}^k]\|^{a,b} \\ &= \|\phi_{\Delta t}^{v[m_{\ell+1}^k]} \# \phi_{\Delta t}^{v[m_{\ell}^k]} \# m_{\ell}^k + \Delta t \phi_{\Delta t}^{v[m_{\ell+1}^k]} \# h[m_{\ell}^k] + \Delta t h[m_{\ell+1}^k] - \phi_{\Delta t}^{v[m_{\ell}^k]} \# \phi_{\Delta t}^{v[m_{\ell}^k]} \# m_{\ell}^k - 2\Delta t h[m_{\ell}^k]\|^{a,b} \\ &\leq \|\phi_{\Delta t}^{v[m_{\ell+1}^k]} \# \phi_{\Delta t}^{v[m_{\ell}^k]} \# m_{\ell}^k - \phi_{\Delta t}^{v[m_{\ell}^k]} \# \phi_{\Delta t}^{v[m_{\ell}^k]} \# m_{\ell}^k\|^{a,b} + \Delta t \|\phi_{\Delta t}^{v[m_{\ell+1}^k]} \# h[m_{\ell}^k] + h[m_{\ell+1}^k] - 2h[m_{\ell}^k]\|^{a,b} \end{aligned}$$

Using Step 2, we have  $\|m_{\ell+n}^k - m_{\ell}^k\| \leq L_1 n \Delta t$ . Then, using Lemma 33

$$\|m_{\ell+2}^k - (\phi_{2\Delta t}^{v[m_{\ell}^k]} \# m_{\ell}^k + 2\Delta t h[m_{\ell}^k])\|^{a,b} \leq C \Delta t^2$$

By induction on  $i = 1 \dots n$ , it then holds

$$\|m_{\ell+n}^k - (\phi_{n\Delta t}^{v[m_{\ell}^k]} \# m_{\ell}^k + n\Delta t h[m_{\ell}^k])\|^{a,b} \leq C(n\Delta t)^2,$$

and (29) follows.

**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  $\varepsilon(t) := \|\mu_t - \nu_t\|^{a,b}$ . We denote



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

Using Lemma 33 and Step 3, and  $e^{L\tau} \leq 1 + 2L\tau$  for  $0 \leq L\tau \leq \ln(2)$ , we have that  $\varepsilon(t)$  is Lipschitz and it satisfies

$$\begin{aligned} \varepsilon(t + \tau) &= \|\mu_{t+\tau} - \nu_{t+\tau}\|^{a,b} = \|\phi_\tau^{v[\mu_t]} \# \mu_t + \tau h[\mu_t] + R_\mu(t, \tau) - \phi_\tau^{v[\nu_t]} \# \nu_t - \tau h[\nu_t] - R_\nu(t, \tau)\|^{a,b} \\ &\leq \|\phi_\tau^{v[\mu_t]} \# \mu_t - \phi_\tau^{v[\nu_t]} \# \nu_t\|^{a,b} + \tau \|h[\mu_t] - h[\nu_t]\|^{a,b} + \|R_\mu(t, \tau)\|^{a,b} + \|R_\nu(t, \tau)\|^{a,b} \\ &\leq e^{L\tau} \|\mu_t - \nu_t\|^{a,b} + b(P + |\mu_0|) \frac{e^{L\tau} - 1}{L} \|v[\mu_t] - v[\nu_t]\|_{C^0} + \tau Q \|\mu_t - \nu_t\|^{a,b} + 2K_1\tau^2 \\ &\leq (e^{L\tau} + b(P + \min\{|\mu_0|, |\nu_0|\})2\tau K + \tau Q) \|\mu_t - \nu_t\|^{a,b} + 2K_1\tau^2 \\ &\leq (1 + \tau(2L + 2bK(P + \min\{|\mu_0|, |\nu_0|\}) + Q)) \|\mu_t - \nu_t\|^{a,b} + 2K_1\tau^2, \end{aligned}$$

which is

$$\frac{\varepsilon(t + \tau) - \varepsilon(t)}{\tau} \leq M\varepsilon(t) + 2K_1\tau, \quad t > 0, \tau \leq \frac{\ln(2)}{L}, \quad M = 2L + 2K(P + \min\{|\mu_0|, |\nu_0|\}) + Q. \quad (30)$$

Letting  $\tau$  go to zero, we deduce  $\varepsilon'(t) \leq M\varepsilon(t)$  almost everywhere. Then,  $\varepsilon(t) \leq \varepsilon(0) \exp(Mt)$ , that is 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  $\|\cdot\|^{a,b}$  is a norm, this implies  $\mu_t = \nu_t$  for all  $t$ , that is 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: A Journal Issued by the Courant Institute of Mathematical Sciences*, 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*. American Institute of Mathematical Sciences (AIMS), Springfield, MO, 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. An interpolating distance between optimal transport and fisher–rao metrics. *Foundations of Computational Mathematics*, Oct 2016.
- [9] 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.
- [10] Alessio Figalli. The optimal partial transport problem. *Arch. Ration. Mech. Anal.*, 195(2):533–560, 2010.
- [11] Alessio Figalli and Nicola Gigli. A new transportation distance between non-negative measures, with applications to gradient flows with Dirichlet boundary conditions. *J. Math. Pures Appl. (9)*, 94(2):107–130, 2010.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] Edoardo Mainini. A description of transport cost for signed measures. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 390(Teoriya Predstavleniĭ, Dinamicheskie Sistemy, Kombinatornye Metody. XX):147–181, 308–309, 2011.
- [17] Benedetto Piccoli and Francesco Rossi. Measure-theoretic models for crowd dynamics. in *Crowd Dynamics Volume 1 - Theory, Models, and Safety Problems*, N. Bellomo and L. Gibelli Eds, Birkhäuser, to appear.
- [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] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [22] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

- [23] A. Zygmund. *Trigonometric series. Vol. I, II.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2002. With a foreword by Robert A. Fefferman.