

SUPERPOSITION PRINCIPLE AND SCHEMES FOR MEASURE DIFFERENTIAL EQUATIONS

FABIO CAMILLI, GIULIA CAVAGNARI *, RAUL DE MAIO, AND BENEDETTO PICCOLI

ABSTRACT. Measure Differential Equations (MDE) describe the evolution of probability measures driven by probability velocity fields, i.e. probability measures on the tangent bundle. They are, on one side, a measure-theoretic generalization of ordinary differential equations; on the other side, they allow to describe concentration and diffusion phenomena typical of kinetic equations. In this paper, we analyze some properties of this class of differential equations, especially highlighting their link with nonlocal continuity equations. We prove a representation result in the spirit of the Superposition Principle by Ambrosio-Gigli-Savaré, and we provide alternative schemes converging to a solution of the MDE, with a particular view to uniqueness/non-uniqueness phenomena.

1. INTRODUCTION

The theory of Measure Differential Equations (MDE in brief) has been recently introduced in [13]. A Cauchy problem for a MDE is given by

$$(1.1) \quad \begin{cases} \dot{\mu}_t = V[\mu_t], \\ \mu_{t=0} = \mu_0, \end{cases}$$

where $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$, the space of probability measures on \mathbb{R}^d , and V is a probability vector field (PVF in brief), i.e. a map assigning to a probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ a (Young) probability measure $V[\mu]$ on the tangent bundle $T\mathbb{R}^d$ such that the first projection of $V[\mu]$ is the base measure μ itself.

MDEs can be seen as a generalization of Ordinary Differential Equations (ODEs) to the space of measures, compatible with the natural map $x \rightarrow \delta_x$ associating to every point of \mathbb{R}^d (more generally, a manifold) the Dirac delta centered at the same point. Lipschitz continuity of the ODE guarantees existence and uniqueness of a Lipschitz semigroup of solutions to the corresponding MDE, which is also linked to existence and uniqueness of (distributional) solutions of the corresponding transport/continuity equation. More

2010 *Mathematics Subject Classification.* 35S99, 35F20, 35F25, 28A50.

Key words and phrases. Measure differential equations, superposition principle, measure-valued solution, probability vector fields.

* Corresponding author: Giulia Cavagnari, giulia.cavagnari@polimi.it.

precisely, the above mentioned link with the continuity equation framework is given as follows. If $V[\mu]$ is concentrated on a graph, i.e.

$$(1.2) \quad V[\mu] = \mu \otimes \delta_{v(x)}$$

for a given Borel measurable vector field $v : \mathbb{R}^d \rightarrow \mathbb{R}^d$, then (1.1) reduces to the continuity equation driven by v ,

$$\begin{cases} \partial_t \mu_t + \operatorname{div}(v \mu_t) = 0. \\ \mu_{t=0} = \mu_0, \end{cases}$$

As proved in [13], the map from vector fields v to probability vector fields V as in (1.2) is a monoid morphism if we endow the latter with fiber-convolution operation.

The study of linear and nonlinear transport equations, in the framework of weak measure solutions, has received a lot of attention in the recent time (see [1, 3–8]). This theory is indeed relatively flexible to describe a large variety of phenomena, as a continuum model for interacting particle systems (see e.g. [10, 11] for the derivation of mean-field equations as limit of interacting N -particles dynamics). The MDE approach can be seen as a further generalization of this technique, when the uncertainty affects not only the position of the particles, but also the law governing their evolution. As will be proved in Section 3, MDEs are in fact an alternative language to describe phenomena that fits into the study of nonlocal continuity equations, where the macroscopic evolution of the state of the system is ruled by a vector field that depends on the evolving state itself. In fact, in [13] it is shown how MDEs can provide a unified model to study phenomena such as finite-speed diffusion and concentration (the latter using one-sided Lipschitz conditions), including conservation laws with discontinuous fluxes which may generate delta waves [16]. Finally MDE allows to prove mean-field type results for multi-particle systems. For all these reasons, models aimed at describing collective motion [2, 5, 18], such as pedestrian traffic [8], vehicular traffic [4] or general multi-particle systems are natural applications of this study. Moreover, a further development of the theory on MDEs in presence of a source term has been recently provided in [14].

We recall that existence of weak measure solutions to (1.1) has been proved in [13] by means of an approximation scheme, called Lattice Approximate Solutions (LAS in brief). The scheme is obtained by discretizing the equation in space, time and velocity and moving convex combinations of Dirac masses through the resulting discrete dynamical system.

Aim of this paper is to provide a further analysis of (1.1) to better understand certain properties regarding the solutions of the problem. The first result is an equivalence relation with a nonlocal continuity equation dynamics, stated in Proposition 3.1, then we give an extension of the Superposition Principle by Ambrosio-Gigli-Savaré in the context of MDEs. We will provide a representation result for a solution of a MDE, similarly to what occurs for continuity equations with a local vector field (see [1, Theorem 8.2.1]), characterizing a (possibly not unique) solution of (1.1) with a superposition of integral curves coming from a suitable underlying particle system. In the same spirit, we also provide a consistent probabilistic representation for the LAS scheme in [13].

In the second part of the paper, we consider alternative schemes converging to a weak solution of the MDE. We first define a semi-discrete in time Lagrangian scheme for (1.1) and we prove that, up to subsequences, it converges to the same limit of the LAS scheme. Moreover, we introduce another semi-discrete in time scheme obtained by taking the barycenter of the PVF at each time step, before moving the mass. We show with an example that this mean velocity scheme may converge to a different weak solution of (1.1) with respect to the LAS/Lagrangian schemes. This fact highlights the weak framework of the MDE theory, in what concerns uniqueness of solutions which, in general, is not expected (see Remark 3.2 and the examples provided in Section 6). However, up to restrict the study to the class of solutions that can be obtained as limits of LASs (or with the semi-discrete in time Lagrangian scheme, see Corollary 4.3), in [13, Section 5] the author prove the uniqueness of a Lipschitz semigroup associated to (1.1) by prescribing the evolution of convex combinations of Dirac measures for a small initial time.

The paper is organized as follows: in Section 2 we give some preliminaries on optimal transport and measure theory, recalling the MDE setting and the definition of the LAS scheme introduced in [13]; in Section 3 we highlight the correspondences between weak solutions of MDE and (distributional) solutions of a suitable continuity equation, we thus exploit a Superposition Principle for MDEs and a probabilistic representation construction for the LASs; in Section 4 we provide a Lagrangian approximation scheme; in Section 5, we present another approximating scheme converging to a different solution of (1.1) and finally, in Section 6, we discuss some clarifying examples, where we make use of the previously defined schemes and of what analyzed in Section 3.

2. PRELIMINARIES AND FIRST RESULTS

We recall some preliminary definitions and results (we address the reader to [1, 17, 19] as relevant resources regarding optimal transport and measure theory). Given a complete separable metric space X , we denote by $\mathcal{P}(X)$ the set of Borel probability measures on X , by $\mathcal{P}_p(X)$ the subset of $\mathcal{P}(X)$ whose elements have finite p -moment and by $\mathcal{P}_c(X)$ the subset of $\mathcal{P}_p(X)$ whose elements have compact support. We endow the set $\mathcal{P}_p(X)$ with the p -Wasserstein distance W_p^X , and we consider the metric W_1^X on $\mathcal{P}_c(X)$. In the case $p = 1$, we recall a special duality formula, called *the Kantorovich-Rubinstein duality*

$$W^X(\mu, \nu) = \sup \left\{ \int_X f d(\mu - \nu) : f : X \rightarrow \mathbb{R}, \text{Lip}(f) \leq 1 \right\}.$$

Proposition 2.1. $\mathcal{P}_p(X)$ endowed with the p -Wasserstein metric, W_p^X , is a complete separable metric space. Moreover, given a sequence $\{\mu_n\}_{n \in \mathbb{N}} \subseteq \mathcal{P}_p(X)$ and $\mu \in \mathcal{P}_p(X)$, we have that the following are equivalent

- (1) $\lim_{n \rightarrow \infty} W_p^X(\mu_n, \mu) = 0$,
- (2) $\mu_n \rightharpoonup^* \mu$ and $\{\mu_n\}_{n \in \mathbb{N}}$ has uniformly integrable p -moments.

In the following, we recall the definition of push-forward (see [1, Section 5.2]) and a disintegration result (see [1, Theorem 5.3.1]).

Definition 2.2 (Pushforward). Let X, Y be two separable metric spaces, $\mu \in \mathcal{P}(X)$, and $r : X \rightarrow Y$ be a Borel map. We define the *pushforward measure* $r\#\mu \in \mathcal{P}(Y)$ by $r\#\mu(B) := \mu(r^{-1}(B))$, for any Borel set $B \subseteq Y$. An equivalent definition is as follows,

$$\langle r\#\mu, f \rangle := \int_X f(r(x)) d\mu(x),$$

for every bounded (or $r\#\mu$ -integrable) Borel function $f : Y \rightarrow \mathbb{R}$.

Theorem 2.3 (Disintegration). Let \mathbf{X}, X be complete separable metric spaces, $\boldsymbol{\mu} \in \mathcal{P}(\mathbf{X})$ and $r : \mathbf{X} \rightarrow X$ be a Borel map. Then there exists a $r\#\boldsymbol{\mu}$ -a.e. uniquely determined Borel family of probability measures $\{\mu_x\}_{x \in X} \subset \mathcal{P}(\mathbf{X})$ such that $\mu_x(\mathbf{X} \setminus r^{-1}(x)) = 0$ for $r\#\boldsymbol{\mu}$ -a.e. $x \in X$. Furthermore

$$\int_{\mathbf{X}} f(z) d\boldsymbol{\mu}(z) = \int_X \int_{r^{-1}(x)} f(z) d\mu_x(z) d(r\#\boldsymbol{\mu})(x),$$

for any bounded Borel map $f : \mathbf{X} \rightarrow \mathbb{R}$. We will write $\boldsymbol{\mu} = (r\#\boldsymbol{\mu}) \otimes \mu_x$.

Remark 2.4. As pointed out in [1, Section 5.3], if $\mathbf{X} = X \times Y$ and $r^{-1}(x) \subseteq \{x\} \times Y$ for all $x \in X$, then we can identify each measure $\mu_x \in \mathcal{P}(X \times Y)$ with a measure defined on Y . We will make a strong use of this result throughout the paper.

We recall now the definition of convolution between measures and product with a coefficient $a \in \mathbb{R}$. We denote with χ_A the characteristic function of $A \subseteq \mathbb{R}^d$.

Definition 2.5 (Convolution). We define *the convolution operator* $\oplus : \mathcal{P}(\mathbb{R}^d) \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{P}(\mathbb{R}^d)$ by $(\mu \oplus \nu)(B) := \int_{\mathbb{R}^d \times \mathbb{R}^d} \chi_B(x+y) d\nu(y) d\mu(x)$, for any Borel set $B \subseteq \mathbb{R}^d$. Equivalently we may define

$$\langle \mu \oplus \nu, f \rangle := \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x+y) d\nu(y) d\mu(x),$$

for any $\mu \oplus \nu$ -integrable Borel function $f : \mathbb{R}^d \rightarrow \mathbb{R}$.

We define *the product operator* $\cdot : \mathbb{R} \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{P}(\mathbb{R}^d)$, $(a, \mu) \mapsto a \cdot \mu$, by

$$(2.1) \quad (a \cdot \mu)(B) := \int_{\mathbb{R}^d} \chi_B(ax) d\mu(x),$$

for any Borel set $B \subseteq \mathbb{R}^d$.

Remark 2.6. Observe that $\mathcal{P}_c(\mathbb{R}^d)$ is closed w.r.t. convolution and product operators. In particular, as pointed out in [13, Section 6.1] we have that the operation \oplus defines a monoid structure over $\mathcal{P}_c(\mathbb{R}^d)$.

Before setting the theory to study Measure Differential Equations, let us recall briefly the definition of solution for a continuity equation. These two notions will be compared in Section 3.

Definition 2.7 (Continuity equation). Given $T > 0$, a Borel family of probability measures $\boldsymbol{\mu} = \{\mu_t\}_{t \in [0, T]} \subseteq \mathcal{P}(\mathbb{R}^d)$ and a Borel map $w : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $w_t(x) = w(t, x)$, such that $\int_{\mathbb{R}^d} |w_t(x)| d\mu_t(x)$ belongs to $L^1([0, T])$, we say that $\boldsymbol{\mu}$ solves the *continuity equation*

$$(2.2) \quad \partial_t \mu_t + \operatorname{div}(w_t \mu_t) = 0,$$

if for every $f \in C_c^\infty(\mathbb{R}^d)$ there holds

$$(2.3) \quad \frac{d}{dt} \langle \mu_t, f \rangle = \int_{\mathbb{R}^d} (\nabla f(x) \cdot w_t(x)) d\mu_t(x),$$

in the sense of distributions on $(0, T)$.

1 According to Lemma 8.1.2 in [1], there exists a curve $\tilde{\mu} : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^d)$ which is
 2 continuous with respect to narrow convergence and such that $\tilde{\mu}(t) = \mu_t$ for \mathcal{L}^1 -a.e. $t \in$
 3 $(0, T)$, i.e. each solution of the continuity equation admits a unique narrowly continuous
 4 representative.

5 *Remark 2.8.* We can give a meaning to a solution of (2.2) also when the vector field w is
 6 nonlocal depending on the solution itself, i.e. $w : [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}^d$, $w(t, x, \mu) =$
 7 $w_t[\mu](x)$. In this case, (2.2) reads

$$(2.4) \quad \partial_t \mu_t + \operatorname{div}(w_t[\mu_t] \mu_t) = 0,$$

8 to be understood in the distributional sense. Given $\mu_0 \in \mathcal{P}_p(\mathbb{R}^d)$, $p > 1$, we recall that
 9 if the vector field w is Lipschitz continuous in (x, μ) uniformly w.r.t. t , i.e. there exists
 10 $L > 0$ such that

$$(2.5) \quad |w(t, x, \mu^1) - w(t, y, \mu^2)| \leq L \left(|x - y| + W^{\mathbb{R}^d}(\mu^1, \mu^2) \right) \quad \forall x, y \in \mathbb{R}^d, \forall \mu^1, \mu^2 \in \mathcal{P}_1(\mathbb{R}^d),$$

11 then there exists a unique solution $\mu \in C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ to (2.4) with initial datum $\mu_{t=0} =$
 12 μ_0 . This result is proved and detailed for example in [12, Theorem 6.1]. Similar conditions
 13 are used in [15] to prove existence and uniqueness of solutions to (2.4) with initial datum
 14 μ_0 , in the class of measures that are absolutely continuous w.r.t. the Lebesgue measure.
 15 Under assumptions granting uniqueness, in [15] the authors construct the solution through
 16 converging numerical schemes.

17 **2.1. Recalls on Measure Differential Equations.** In this section we recall some basic
 18 definitions introduced in [13] that are at the base of the investigations proposed in this
 19 paper.

20 **Definition 2.9.** A *probability vector field* (PVF) is a map $V : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{P}(T\mathbb{R}^d)$ consistent
 21 with the projection on the first component, i.e. $\pi_1 \# V[\mu] = \mu$.

22 By Theorem 2.3 and Remark 2.4, we can write $V[\mu] = \mu \otimes \nu_x[\mu]$ for a μ -a.e. uniquely
 23 determined Borel family of probability measures $\{\nu_x[\mu]\}_{x \in \mathbb{R}^d} \subseteq \mathcal{P}(\mathbb{R}^d)$ defined on the fibers
 24 $T_x \mathbb{R}^d = \mathbb{R}^d$.

1 **Let $T > 0$.** Given $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ and a PVF V , we consider the following Cauchy problem

$$(2.6) \quad \begin{cases} \dot{\mu}_t = V[\mu_t], & t \in (0, T) \\ \mu_{t=0} = \mu_0, \end{cases}$$

2 where the nonlocal dynamics is called *Measure Differential Equation* (MDE). A solution
3 to this problem has to be interpreted as follows.

4 **Definition 2.10.** A solution of (2.6) is a map $\mu : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^d)$, $t \mapsto \mu_t$ such that
5 $\mu_{t=0} = \mu_0$ and that satisfies for a.e. $t \in [0, T]$

$$(2.7) \quad \frac{d}{dt} \langle \mu_t, f \rangle = \int_{T\mathbb{R}^d} (\nabla f(x) \cdot v) dV[\mu_t](x, v),$$

for any $f \in C_c^\infty(\mathbb{R}^d)$ such that the right-hand side is defined for a.e. t , the map $t \mapsto \int_{T\mathbb{R}^d} (\nabla f(x) \cdot v) dV[\mu_t](x, v)$ belongs to $L^1([0, T])$, and the map $t \mapsto \int_{\mathbb{R}^d} f d\mu_t$ is absolutely continuous. Equivalently,

$$\langle \mu_t - \mu_0, f \rangle = \int_0^t \int_{T\mathbb{R}^d} (\nabla f(x) \cdot v) dV[\mu_s](x, v) ds, \quad \forall f \in C_c^\infty(\mathbb{R}^d).$$

6 *Remark 2.11.* Notice that, in the trivial case when $V[\mu] = \mu \otimes \delta_{v(x)}$ for some Borel vector
7 field $v : \mathbb{R}^d \rightarrow \mathbb{R}^d$, the MDE in (2.6) reduces to the continuity equation $\partial_t \mu_t + \operatorname{div}(v \mu_t) = 0$
8 (see [13, Section 6]).

9 We stress that $V[\mu]$ is a probability measure on $T\mathbb{R}^d$ where the components of its ele-
10 ments (x, v) represent, respectively, the position and the infinitesimal displacement. We
11 recall another notion to measure distances between PVFs introduced in [13].

Definition 2.12. Given $V_i \in \mathcal{P}_c(T\mathbb{R}^d)$, $i = 1, 2$, and denoted by $\mu_i := \pi_{13\#} V_i$ the marginal
of V_i , we define

$$\mathcal{W}(V_1, V_2) = \inf \left\{ \int_{T\mathbb{R}^d \times T\mathbb{R}^d} |v - w| dT(x, v, y, w) : T \in \Pi(V_1, V_2), \pi_{13\#} T \in \Pi_{opt}(\mu_1, \mu_2) \right\},$$

12 where $\Pi(V_1, V_2)$ is the set of all the transference plans from V_1 to V_2 and $\Pi_{opt}(\mu_1, \mu_2)$
13 is the set of the optimal transference plans from μ_1 to μ_2 , and $\pi_{13} : (T\mathbb{R}^d)^2 \rightarrow (\mathbb{R}^d)^2$,
14 $(x, v, y, w) \mapsto (x, y)$.

The object \mathcal{W} computes the minimal displacements of the fiber components assuming
that marginals μ_i are transported in an optimal way. It is important to notice that \mathcal{W} is
not a metric since it can vanish for distinct elements in $\mathcal{P}_c(T\mathbb{R}^d)$. Moreover, it is easy to

verify that

$$W^{T\mathbb{R}^d}(V_1, V_2) \leq \mathcal{W}(V_1, V_2) + W^{\mathbb{R}^d}(\mu_1, \mu_2).$$

1 Considering the problem set in $\mathcal{P}_c(\mathbb{R}^d)$, we recall here the main assumptions required
 2 to have existence and convergence of approximation schemes for solutions of an MDE (see
 3 [13]).

4 **(H1) Sublinearity:** there exists a constant $C > 0$ such that for all $\mu \in \mathcal{P}_c(\mathbb{R}^d)$,

$$\sup_{(x,v) \in \text{supp} V[\mu]} |v| \leq C(1 + \sup_{x \in \text{supp} \mu} |x|);$$

5 **(H2) Continuity of PVF:** the map $V : \mathcal{P}_c(\mathbb{R}^d) \rightarrow \mathcal{P}_c(T\mathbb{R}^d)$ is continuous;

6 **(H3) Local lipschitzianity in μ -variable:** V is locally Lipschitz, in particular for
 7 every $R > 0$ there exists a constant $L = L(R) > 0$ such that

$$\mathcal{W}(V[\mu], V[\nu]) \leq L \cdot W^{\mathbb{R}^d}(\mu, \nu),$$

8 for every $\mu, \nu \in \mathcal{P}_c(\mathbb{R}^d)$ such that $\text{supp}(\mu), \text{supp}(\nu) \subset B(0, R)$, the open ball of
 9 radius R centered in $0 \in \mathbb{R}^d$.

10 In the following, we recall the scheme provided in [13] that has been used in order to
 11 prove existence of solutions to (2.6). Let us start introducing some notation. **Let $T > 0$.**
 12 For $N \in \mathbb{N} \setminus \{0\}$, set

$$\Delta_N = \frac{T}{N}, \quad \Delta_N^v = \frac{1}{N}, \quad \Delta_N^x = \Delta_N^v \Delta_N = \frac{T}{N^2}$$

13 be respectively the time, the velocity and the space-step sizes, noticing that, differently
 14 from [13], we set the time step size to T/N in place of $1/N$, for our convenience. Considering
 15 the corresponding grid in $[0, T] \times [-N, N]^d \times [-TN, TN]^d$, we denote by x_i the discretization
 16 points in space, and by v_i the discretization points for the space of velocities. We now
 17 build some objects aiming at providing a discrete approximation for $\mu \in \mathcal{P}_c(\mathbb{R}^d)$ and
 18 $V[\mu] \in \mathcal{P}_c(T\mathbb{R}^d)$ by concentrating the mass on the points of the grid. Denoting with
 19 $Q = ([0, \frac{T}{N^2}])^d$ and $Q' = ([0, \frac{1}{N}])^d$, we define

$$\mathcal{A}_N^x(\mu) := \sum_i m_i^x(\mu) \delta_{x_i}, \quad \mathcal{A}_N^v(V[\mu]) := \sum_i \sum_j m_{ij}^v(V[\mu]) \delta_{(x_i, v_j)},$$

- 1 where $m_i^x(\mu) := \mu(x_i + Q)$ and $m_{ij}^v(V[\mu]) := V[\mu](\{(x_i, v) : v \in v_j + Q'\})$. Notice that
 2 given $\mu \in \mathcal{P}_c(\mathbb{R}^d)$, for N sufficiently large we have

$$(2.8) \quad W^{\mathbb{R}^d}(\mathcal{A}_N^x(\mu), \mu) \leq \Delta_N^x, \quad W^{T\mathbb{R}^d}(\mathcal{A}_N^v(V[\mu]), V[\mu]) \leq \Delta_N^v.$$

Definition 2.13 (Lattice Approximate Solution (LAS)). Let V be a PVF satisfying (H1). Given $\mu_0 \in \mathcal{P}_c(\mathbb{R}^d)$, $T > 0$ and $N \in \mathbb{N}$, the *Lattice Approximate Solution (LAS)* $\mu^N : [0, T] \rightarrow \mathcal{P}_c(\mathbb{R}^d)$, $t \mapsto \mu_t^N$, is defined, by recursion, as follows

$$\begin{aligned} \mu_0^N &= \mathcal{A}_N^x(\mu_0) \\ \mu_{k+1}^N &= \mu^N((k+1)\Delta_N) = \sum_{ij} m_{ij}^v(V[\mu^N(k\Delta_N)]) \delta_{x_i + \Delta_N v_j}, \end{aligned}$$

notice that $\text{supp}(\mu_k^N)$ is contained on the space grid. By time-interpolation we can define μ^N for all times as

$$\mu^N(k\Delta_N + t) = \sum_{ij} m_{ij}^v(V[\mu^N(k\Delta_N)]) \delta_{x_i + t v_j}.$$

- 3 Notice that, thanks to the growth assumption (H1) on V and since $\mu_0 \in \mathcal{P}_c(\mathbb{R}^d)$, then
 4 the support of μ_t^N is uniformly bounded (see [13, Lemma 3.3]), i.e. there exists $R =$
 5 $R(T, C, \mu_0) > 0$ such that

$$(2.9) \quad \text{supp } \mu_t^N \subset B(0, R), \quad \forall t \in [0, T], \forall N \in \mathbb{N} \setminus \{0\}.$$

- 6 We address the reader to [13] for results granting the convergence of the LAS scheme to
 7 a solution of (2.6).

- 8 We recall here a definition, used in [13] to derive the uniqueness of solutions to (2.6) at
 9 the level of the semigroup, when restricting the analysis to a certain class of trajectories.
 10 This will be resumed later on in Section 4.

- 11 **Definition 2.14.** A Lipschitz semigroup for (2.6) is a map $S : [0, T] \times \mathcal{P}_c(\mathbb{R}^d) \rightarrow \mathcal{P}_c(\mathbb{R}^d)$
 12 such that for every $\mu, \eta \in \mathcal{P}_c(\mathbb{R}^d)$ and $t, s \in [0, T]$ the following holds:

- 13 (i) $S_0\mu = \mu$ and $S_t S_s \mu = S_{t+s}\mu$;
 14 (ii) the map $t \rightarrow S_t\mu$ is a solution of (2.6);

(iii) for every $R > 0$ there exists $C = C(R) > 0$ such that if $\text{supp}(\mu), \text{supp}(\eta) \subset B(0, R)$ then:

$$\begin{cases} \text{supp}(S_t\mu) \subset B(0, e^{Ct}(R+1)), \\ W^{\mathbb{R}^d}(S_t\mu, S_t\eta) \leq e^{Ct}W^{\mathbb{R}^d}(\mu, \eta), \\ W^{\mathbb{R}^d}(S_t\mu, S_s\mu) \leq C|t-s|. \end{cases}$$

1

3. A SUPERPOSITION PRINCIPLE FOR MDEs

2 We start this section by exploiting the definition of solutions for an MDE comparing it
3 with solutions of a continuity equation. We see that, by definition, in order for μ to be a
4 solution of an MDE, it is sufficient to solve a nonlocal continuity equation driven by the
5 barycenter of what $V[\mu]$ prescribes on the fibers $T_x\mathbb{R}^d$ along the solution. In this sense,
6 MDEs can be viewed as an alternative language to nonlocal continuity equations.

7 **Proposition 3.1.** *Let $T > 0$, $V : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{P}(T\mathbb{R}^d)$ be a PVF satisfying*

$$(3.1) \quad \int_0^T \int_{T\mathbb{R}^d} |v| dV[\mu_t](x, v) dt < +\infty.$$

8 *Then $\mu : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^d)$, $t \mapsto \mu_t$ is a solution of the MDE $\dot{\mu}_t = V[\mu_t]$ if and only if μ is a*
9 *solution of the nonlocal continuity equation (2.4) for a Borel vector field $w : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \rightarrow$*
10 *\mathbb{R}^d defined by*

$$(3.2) \quad w[\mu](x) := \int_{\pi_1^{-1}(x)} v d\nu_x[\mu](v), \quad \mu_t\text{-a.e. } x \in \mathbb{R}^d,$$

11 *for μ -a.e. $x \in \mathbb{R}^d$, where we denoted with $\nu_y[\mu]$ the disintegration of $V[\mu]$ w.r.t. π_1 .*

Proof. The proof follows by disintegration technique. Let us first notice that by (3.1)

$$\int_0^T \int_{\mathbb{R}^d} |w[\mu_t](x)| d\mu_t(x) dt \leq \int_0^T \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |v| d\nu_x[\mu_t](v) d\mu_t(x) dt < +\infty.$$

The conclusion comes straightforwardly noting that (2.7) can be written in terms of (2.3) as follows

$$\begin{aligned} \frac{d}{dt} \langle \mu_t, f \rangle &= \int_{T\mathbb{R}^d} (\nabla f(x) \cdot v) dV[\mu_t](x, v) \\ &= \int_{\mathbb{R}^d} \int_{\pi_1^{-1}(x)} (\nabla f(x) \cdot v) d\nu_x[\mu_t](v) d\mu_t(x) \\ &= \int_{\mathbb{R}^d} (\nabla f(x) \cdot w[\mu_t](x)) d\mu_t(x). \end{aligned}$$

Remark 3.2. As briefly discussed in Remark 2.8, we know that we can guarantee uniqueness of solutions to the nonlocal continuity equation (2.4) (with fixed initial datum) under the Lipschitz assumption (2.5) on the driving velocity field w . However, when we set the problem in the framework of MDEs, provide such condition is not an easy task, even having the equivalence just proved in Proposition 3.1. Indeed by Proposition 3.1, if we want to have (Lipschitz) continuity of w w.r.t. $x \in \mathbb{R}^d$, we should assume the disintegration $x \mapsto \nu_x[\mu]$ of the PVF $V[\mu]$ w.r.t. the projection to the base π_1 to be (Lipschitz) continuous, while in general it is just Borel measurable (see Theorem 2.3).

Thus, in general, uniqueness of a weak solution for (2.6) is not expected (see [13, Example 3]) even under assumptions (H1) – (H3). We will discuss further examples in Section 6.

Inspired by the previous result, we show how to construct a Superposition Principle (see [1, Theorem 8.2.1] for the local continuity equation dynamics) adapted to the language of MDEs. The procedure is similar to the one used in [7], where the authors provide a representation result for solutions of a continuity equation associated with Carathéodory solutions of a differential inclusion. This result, proved in [7], is exploited in [6], where the authors study optimal control problems in the space of probability measures with microscopic dynamics ruled, precisely, by a differential inclusion.

We split the statement into two parts. In the first part, we see that any measure $\eta \in \mathcal{P}(\Gamma_{[0,T]})$, concentrated on curves that follow a given PVF V in integral average, generates a solution of the MDE.

For $I \subseteq \mathbb{R}$ interval, we denote by Γ_I the set of continuous curves from I to \mathbb{R}^d and by e_t the evaluation operator $e_t : \Gamma_I \rightarrow \mathbb{R}^d$, $\gamma \mapsto \gamma(t)$, for $t \in I$, while $AC(I; \mathbb{R}^d)$ is the space of absolutely continuous curves from I to \mathbb{R}^d .

Theorem 3.3 (Superposition Principle for MDEs - Part I). *Let $T > 0$, $V : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{P}(T\mathbb{R}^d)$ be a PVF, $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$. Let $\eta \in \mathcal{P}(\Gamma_{[0,T]})$ be concentrated on the set of curves $\gamma \in \Gamma_{[0,T]}$ such that the following conditions hold*

- (i) $\gamma \in AC([0, T]; \mathbb{R}^d)$;
- (ii) for a.e. $t \in [0, T]$ and for $e_t\#\eta$ -a.e. $y \in \mathbb{R}^d$ we have

$$\int_{e_t^{-1}(y)} \dot{\gamma}(t) d\eta_{t,y}(\gamma) = \int_{\pi_1^{-1}(y)} v d\nu_y[e_t\#\eta](v),$$

- 1 where $\eta_{t,y}$ is the disintegration of $\boldsymbol{\eta}$ w.r.t. e_t , and $\nu_y[e_t\sharp\boldsymbol{\eta}]$ is the disintegration of
2 $V[e_t\sharp\boldsymbol{\eta}]$ w.r.t. the projection to the base π_1 ;
3 (iii) $\int_0^T \int_{\Gamma_{[0,T]}} |\dot{\gamma}(s)| d\boldsymbol{\eta}(\gamma) dt < +\infty$ and $\int_0^T \int_{T\mathbb{R}^d} |v| dV[e_t\sharp\boldsymbol{\eta}](x, v) dt < +\infty$.
- 4 Then, denoted with $\mu_t := e_t\sharp\boldsymbol{\eta}$, we have that $\boldsymbol{\mu} = \{\mu_t\}_{t \in [0,T]} \subseteq \mathcal{P}(\mathbb{R}^d)$ is a solution of the
5 MDE system (2.6).

Proof. Let us consider any $f \in C_c^\infty(\mathbb{R}^d)$. First, we check that $\int_{T\mathbb{R}^d} (\nabla f(x) \cdot v) dV[e_s\sharp\boldsymbol{\eta}](x, v)$ is defined for almost every $s \in [0, T]$. Indeed, immediately by hypothesis (iii),

$$\int_{T\mathbb{R}^d} (\nabla f(x) \cdot v) dV[e_s\sharp\boldsymbol{\eta}](x, v) \leq \|\nabla f\|_\infty \int_{T\mathbb{R}^d} |v| dV[e_s\sharp\boldsymbol{\eta}](x, v) < +\infty$$

- 6 for a.e. $s \in [0, T]$. Thus, we also have that $s \mapsto \int_{T\mathbb{R}^d} (\nabla f(x) \cdot v) dV[e_s\sharp\boldsymbol{\eta}](x, v)$ belongs to
7 $L^1([0, T])$.

Secondly, the map $t \mapsto \int_{\mathbb{R}^d} f d\mu_t$ is absolutely continuous. Indeed, for $0 \leq s < t \leq T$ we have

$$\begin{aligned} \left| \int_{\mathbb{R}^d} f d\mu_s - \int_{\mathbb{R}^d} f d\mu_t \right| &\leq \int_s^t \int_{\Gamma_{[0,T]}} |\nabla f(\gamma(\tau)) \cdot \dot{\gamma}(\tau)| d\boldsymbol{\eta}(\gamma) d\tau \\ &\leq \int_s^t \int_{\Gamma_{[0,T]}} |\nabla f(\gamma(\tau))| \cdot |\dot{\gamma}(\tau)| d\boldsymbol{\eta} d\tau \\ &\leq \|\nabla f\|_\infty \int_s^t \int_{\Gamma_{[0,T]}} |\dot{\gamma}(\tau)| d\boldsymbol{\eta} d\tau < +\infty, \end{aligned}$$

- 8 thanks to hypothesis (iii).

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

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} f(x) d\mu_t(x) &= \int_{\Gamma_{[0,T]}} \nabla f(\gamma(t)) \cdot \dot{\gamma}(t) d\boldsymbol{\eta}(\gamma) \\ &= \int_{\mathbb{R}^d} \nabla f(y) \int_{e_t^{-1}(y)} \dot{\gamma}(t) d\eta_{t,y}(\gamma) d\mu_t(y) \\ &= \int_{\mathbb{R}^d} \nabla f(y) \int_{\pi_1^{-1}(y)} v d\nu_y[\mu_t](v) d\mu_t(y) \\ &= \int_{T\mathbb{R}^d} \nabla f(x) \cdot v dV[\mu_t](x, v), \end{aligned}$$

- 9 where we used hypothesis (ii). □

1 *Remark 3.4.* We observe that the second request in item (iii) can be satisfied assuming
 2 hypothesis (H1) for the PVF V together with the hypothesis

$$\int_0^T \int_{\Gamma_{[0,T]}} \sup_{\gamma \in \text{supp } \boldsymbol{\eta}} |\dot{\gamma}(t)| d\boldsymbol{\eta}(\gamma) dt < +\infty.$$

3 Let us now pass to the other implication. In the second part of the statement that we
 4 are going to see, we want to prove the existence of a probabilistic representation $\boldsymbol{\eta}$ starting
 5 from a solution $\boldsymbol{\mu}$ of the MDE system (2.6) with given PVF V . In this general framework,
 6 this can be easily provided by disintegration technique and thanks to the Superposition
 7 Principle in [1, Theorem 8.2.1].

Theorem 3.5 (Superposition Principle for MDEs - Part II). *Let $T > 0$, $V : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{P}(T\mathbb{R}^d)$ be a PVF, $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$. Let $\boldsymbol{\mu} = \{\mu_t\}_{t \in [0,T]} \subseteq \mathcal{P}(\mathbb{R}^d)$ be an absolutely continuous solution of the MDE system (2.6) for a PVF $V : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{P}(T\mathbb{R}^d)$ satisfying*

$$\int_0^T \int_{T\mathbb{R}^d} |v|^p dV[\mu_t](x, v) dt < +\infty, \text{ for some } p > 1.$$

8 *Then there exists a probability measure $\boldsymbol{\eta} \in \mathcal{P}(\Gamma_{[0,T]})$ such that*

9 (i) $\boldsymbol{\eta}$ is concentrated on curves γ such that $\gamma \in AC([0, T]; \mathbb{R}^d)$ is a solution of the ODE

10 $\dot{\gamma}(t) = w_t(\gamma(t))$ for a.e. $t \in (0, T)$, where $w : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $(t, y) \mapsto w_t(y) =$
 11 $\int_{\pi_1^{-1}(y)} v d\nu_y[\mu_t](v)$ for a.e. t and μ_t -a.e. y ;

12 (ii) $\mu_t = e_t \# \boldsymbol{\eta}$ for any $t \in [0, T]$.

13 Straightforwardly, if $\hat{\boldsymbol{\eta}} \in \mathcal{P}(\Gamma_{[0,T]})$ realizes (i) and (ii), then for a.e. $t \in [0, T]$ and for
 14 μ_t -a.e. $y \in \mathbb{R}^d$ we have

$$\int_{e_t^{-1}(y)} \dot{\gamma}(t) d\hat{\eta}_{t,y}(\gamma) = \int_{\pi_1^{-1}(y)} v d\nu_y[\mu_t](v),$$

15 where $\hat{\eta}_{t,y}$ is the disintegration of $\hat{\boldsymbol{\eta}}$ w.r.t. e_t , and $\nu_y[\mu_t]$ is the disintegration of $V[\mu_t]$ w.r.t.
 16 the projection on the first component π_1 .

Proof. Let us take any $f \in C_c^\infty(\mathbb{R}^d)$. Let $\boldsymbol{\mu} = \{\mu_t\}_t$ be as in the statement. Then, by Definition 2.10 and by disintegrating V w.r.t. the projection map π_1 ,

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} f(y) d\mu_t(y) &= \int_{T\mathbb{R}^d} \nabla f(y) \cdot v dV[\mu_t](y, v) \\ &= \int_{\mathbb{R}^d} \nabla f(y) \cdot \int_{\pi_1^{-1}(y)} v d\nu_y[\mu_t](v) d\mu_t(y) \end{aligned}$$

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

2 By denoting with $w : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $w_t(y) := \int_{\pi_1^{-1}(y)} v d\nu_y[\mu_t](v)$ for a.e. t and μ_t -a.e.
 3 y , we notice that w is a Borel vector field by disintegration property (Theorem 2.3) and
 4 that μ is a solution of the continuity equation $\partial_t \mu_t + \operatorname{div}(w_t \mu_t) = 0$.

Moreover, by Jensen's inequality, we get

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} |w_t(y)|^p d\mu_t(y) dt &= \int_0^T \int_{\mathbb{R}^d} \left| \int_{\pi_1^{-1}(y)} v d\nu_y[\mu_t](v) \right|^p d\mu_t(y) dt \\ &\leq \int_0^T \int_{T\mathbb{R}^d} |v|^p dV[\mu_t](y, v) dt < +\infty, \end{aligned}$$

5 thanks to the hypothesis on V . We can thus apply the classical Superposition Principle in
 6 [1, Theorem 8.2.1] to get (i – ii).

Let now $\hat{\eta}$ be as in the statement, hence $\dot{\gamma}(t) = w_t(\gamma(t))$ for a.e. $t \in (0, T)$ and for $\hat{\eta}$ -a.e.
 γ . Last property is straightforward by disintegration and definition of w , indeed for any
 test function $\varphi \in C_c^\infty([0, T] \times \mathbb{R}^d)$,

$$\begin{aligned} \int_0^T \int_{\Gamma_{[0, T]}} \varphi(t, y) \dot{\gamma}(t) d\hat{\eta}(\gamma) dt &= \int_0^T \int_{\mathbb{R}^d} \varphi(t, y) \int_{e_t^{-1}(y)} \dot{\gamma}(t) d\hat{\eta}_{t, y}(\gamma) d\mu_t(y) dt \\ &= \int_0^T \int_{\mathbb{R}^d} \varphi(t, y) \int_{e_t^{-1}(y)} w_t(y) d\hat{\eta}_{t, y}(\gamma) d\mu_t(y) dt \\ &= \int_0^T \int_{\mathbb{R}^d} \varphi(t, y) \int_{\pi_1^{-1}(y)} v d\nu_y[\mu_t](v) d\mu_t(y) dt. \end{aligned}$$

7

□

8 We now complete the analysis concerning the connection between Superposition Prin-
 9 ciple and MDEs by giving an example of an explicit and consistent construction for a
 10 probabilistic representation of the LAS scheme.

11 Let $\mu \subseteq \mathcal{P}(\mathbb{R}^d)$ be a solution of the MDE system (2.6) obtained as uniform-in-time limit
 12 of LASs $\{\mu^N\}_{N \in \mathbb{N}}$. We now construct a probabilistic representation for μ that is concen-
 13 trated on uniform limits of the trajectories $\gamma_{i, j} : [0, T] \rightarrow \mathbb{R}^d$, $\gamma_{i, j}(t) = x_i + tv_j$, where the
 14 LASs μ^N are concentrated.

15 Let us start by fixing some notation. Given $I_1, I_2 \subset \mathbb{R}$ nonempty and compact intervals,
 16 with $\max I_1 = \min I_2$, we define

(1) the *set of compatible trajectories*

$$\mathcal{D}_{I_1, I_2} := \{(\gamma_1, \gamma_2) \in \Gamma_{I_1} \times \Gamma_{I_2} : \gamma_1(\max I_1) = \gamma_2(\min I_2)\};$$

(2) the *concatenation* $\gamma_1 \star \gamma_2 \in \Gamma_{I_1 \cup I_2}$ of curves $\gamma_1 \in \Gamma_{I_1}$, $\gamma_2 \in \Gamma_{I_2}$, with $\gamma_1(\max I_1) = \gamma_2(\min I_2)$, is a map from $I_1 \cup I_2$ to \mathbb{R}^d defined as follows

$$\gamma_1 \star \gamma_2(t) = \begin{cases} \gamma_1(t), & \text{if } t \in I_1, \\ \gamma_2(t), & \text{if } t \in I_2; \end{cases}$$

(3) the *merge map* $M_{I_1, I_2} : \mathcal{D}_{I_1, I_2} \rightarrow \Gamma_{I_1 \cup I_2}$, $(\gamma_1, \gamma_2) \mapsto \gamma_1 \star \gamma_2$. We will omit the subscripts I_1, I_2 when clear.

Definition 3.6. Let $T > 0$, and $\mu^N = \{\mu_t^N\}_{t \in [0, T]} \subseteq \mathcal{P}_c(\mathbb{R}^d)$ be the LAS defined in Definition 2.13. Denote with $I_a^b := [a\Delta_N, b\Delta_N]$, for $a, b \in \mathbb{N}$, $a \leq b$. We define

(1) the following measure in $\mathcal{P}(\Gamma_{I_\ell^{\ell+1}})$

$$\eta_{I_\ell^{\ell+1}}^N := \sum_{i,j} m_{i,j}^v(V[\mu_{\ell\Delta_N}^N]) \delta_{\gamma_{i,j}^\ell} = \int_{T\mathbb{R}^d} \delta_{\gamma_{x,v}^\ell} d\mathcal{A}_N^v(V[\mu_{\ell\Delta_N}^N])(x, v),$$

where $\gamma_{i,j}^\ell = \gamma_{x_i, v_j}^\ell$ are the solutions of the LAS characteristic system defined on $I_\ell^{\ell+1}$, thus $\gamma_{i,j}^\ell(t) = x_i + (t - \ell\Delta_N)v_j$, for $t \in I_\ell^{\ell+1}$;

(2) $\eta_{I_0^{h+1}}^N := \mu_{h\Delta_N}^N \otimes M_{I_0^h, I_h^{h+1}} \# (\eta_{I_0^h, x}^N \otimes \eta_{I_h^{h+1}, x}^N) \in \mathcal{P}(\Gamma_{[0, (h+1)\Delta_N]})$, defined by recursion for $h = 1, \dots, N-1$, where $\eta_{I_0^h, x}^N$ and $\eta_{I_h^{h+1}, x}^N$ are respectively the disintegrations of $\eta_{I_0^h}^N$ and $\eta_{I_h^{h+1}}^N$ w.r.t. $e_{h\Delta_N}$. That is

$$\int_{\Gamma_{I_0^{h+1}}} \varphi(\gamma) d\eta_{I_0^{h+1}}^N(\gamma) := \int_{\mathbb{R}^d} \left(\int_{e_{h\Delta_N}^{-1}(x)} \varphi(\gamma) d \left[M_{I_0^h, I_h^{h+1}} \# (\eta_{I_0^h, x}^N \otimes \eta_{I_h^{h+1}, x}^N) \right] (\gamma) \right) d\mu_{h\Delta_N}^N(x),$$

for any bounded Borel function $\varphi : \Gamma_{I_0^{h+1}} \rightarrow \mathbb{R}$.

(3) $\eta^N := \eta_{I_0^N}^N \in \mathcal{P}(\Gamma_{[0, T]})$.

Proposition 3.7. Let V be a PVF satisfying (H1) and (H2), $\mu_0 \in \mathcal{P}_c(\mathbb{R}^d)$, and μ be a solution of the MDE system (2.6) obtained as uniform-in-time limit of LASs $\{\mu^N\}_N$ for the Wasserstein metric. Let η^N be as in Definition 3.6. Then

(1) η^N is a probabilistic representation for the LAS μ^N , i.e. $\mu_t^N = e_t \# \eta^N$ for all $t \in [0, T]$;

(2) $\eta^N \rightharpoonup^* \eta$ up to subsequences, and η is a probabilistic representation for μ .

Proof of (1). First we prove that $\mu_t^N = e_t \# \boldsymbol{\eta}_{I_\ell^{\ell+1}}^N$ for all $\ell = 0, \dots, N-1$, and $t \in I_\ell^{\ell+1}$. By Definition 2.13, $\mu_t^N := \sum_{i,j} m_{i,j}^v(V([\mu_{\ell\Delta_N}^N]) \delta_{x_i+(t-\ell\Delta_N)v_j})$. Thus, for all $\varphi \in C_b^0(\mathbb{R}^d)$,

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(x) d\mu_t^N &= \sum_{i,j} \varphi(x_i + (t - \ell\Delta_N)v_j) \cdot m_{i,j}^v(V[\mu_{\ell\Delta_N}^N]) \\ &= \sum_{i,j} \varphi(\gamma_{i,j}^\ell(t)) \cdot m_{i,j}^v(V[\mu_{\ell\Delta_N}^N]) = \int_{\Gamma_{I_\ell^{\ell+1}}} \varphi(\gamma(t)) d\boldsymbol{\eta}_{I_\ell^{\ell+1}}^N. \end{aligned}$$

Let us now conclude the proof by showing that $e_t \# \boldsymbol{\eta}_{I_0^{h+1}}^N = \mu_t^N$ for all $h = 1, \dots, N-1$, and $t \in I_0^{h+1}$. Indeed, for all $\varphi \in C_b^0(\mathbb{R}^d)$,

$$\begin{aligned} \int \varphi(\gamma(t)) d\boldsymbol{\eta}_{I_0^{h+1}}^N(\gamma) &= \int \varphi(\gamma(t)) d \left[\mu_{h\Delta_N}^N \otimes M \# (\eta_{I_0^h, y}^N \otimes \eta_{I_h^{h+1}, y}^N) \right] (\gamma) \\ &= \int \varphi(\gamma(t)) d \left[M \# (\eta_{I_0^h, y}^N \otimes \eta_{I_h^{h+1}, y}^N) \right] (\gamma) d\mu_{h\Delta_N}^N(y) \\ &= \int \varphi(\gamma_1 \star \gamma_2(t)) d\eta_{I_0^h, y}^N(\gamma_1) d\eta_{I_h^{h+1}, y}^N(\gamma_2) d\mu_{h\Delta_N}^N(y) \\ &= \int \varphi(\gamma_2(t)) d\eta_{I_h^{h+1}, y}^N(\gamma_2) d\mu_{h\Delta_N}^N(y) \\ &= \int \varphi(\gamma_2(t)) d\boldsymbol{\eta}_{I_h^{h+1}}^N(\gamma_2) = \int \varphi(x) d\mu_t^N(x), \end{aligned}$$

- 1 where in the fourth equality we assumed, without loss of generality, $t \in I_h^{h+1}$, otherwise we
- 2 iterate the same procedure for $\boldsymbol{\eta}_{I_0^h}^N$. In the last two passages we used what proved before,
- 3 i.e. $e_t \# \boldsymbol{\eta}_{I_h^{h+1}}^N = \mu_t^N$.

- 4 *Proof of (2).* First, let us prove that the family $\{\boldsymbol{\eta}^N\}_N$ is tight, thus there exists $\boldsymbol{\eta} \in$
- 5 $\mathcal{P}(\Gamma_{[0,T]})$ such that $\boldsymbol{\eta}^N \rightharpoonup^* \boldsymbol{\eta}$, up to a non-relabeled subsequence. We proceed in the same
- 6 way as in [1, Theorem 8.2.1]. Indeed, we use [1, Lemma 5.2.2] with

$$\boldsymbol{r}^1 := e_0 : \gamma \mapsto \gamma(0) \in \mathbb{R}^d, \quad \boldsymbol{r}^2 : \gamma \mapsto \gamma - \gamma(0) \in \Gamma_{[0,T]},$$

and we notice that $\boldsymbol{r}^1 \times \boldsymbol{r}^2$ is proper, and by the previous item we have that the family $\{\boldsymbol{r}^1 \# \boldsymbol{\eta}^N\}_N$ is given by the first marginals $\{\mu_0^N\}$ which is tight (indeed, it narrowly converges to μ_0 by assumption), furthermore $\beta^N := \boldsymbol{r}^2 \# \boldsymbol{\eta}^N \in \Gamma_{[0,T]}$ satisfy for $p > 1$,

$$\begin{aligned} \int_{\Gamma_{[0,T]}} \int_0^T |\dot{\gamma}(t)|^p dt d\beta^N(\gamma) &= \int_{\Gamma_{[0,T]}} \int_0^T \left| \frac{d}{dt} (\gamma - \gamma(0))(t) \right|^p dt d\boldsymbol{\eta}^N(\gamma) \\ &= \int_0^T \int_{\Gamma_{[0,T]}} |\dot{\gamma}(t)|^p d\boldsymbol{\eta}^N(\gamma) dt \end{aligned}$$

$$\leq T \cdot \sup_{t \in [0, T]} \sup_{(x, v) \in \text{supp } \mathcal{A}_N^v(V[\mu_t^N])} |v|^p$$

1 which is uniformly bounded thanks to assumption (H1) on V and recalling (2.9). Hence,
 2 the tightness of the family β^N follows by [1, Remark 5.1.5], since for $p > 1$ the functional
 3 $\gamma \mapsto \int_0^T |\dot{\gamma}|^p dt$ (set to $+\infty$ if $\gamma \notin AC^p([0, T]; \mathbb{R}^d)$ or if $\gamma(0) \neq 0$) has compact sublevels in
 4 $\Gamma_{[0, T]}$. Thus, the family $\{\eta^N\}_N$ is tight.
 5 By *weak**-convergence of μ_t^N to μ_t for all $t \in [0, T]$ and of η^N to some η up to sub-
 6 sequences, and since from item (1), $\mu_t^N = e_t \# \eta^N$, then we immediately have that η is a
 7 probabilistic representation for μ , i.e. $e_t \# \eta = \mu_t$. By construction (see [1, Theorem 5.1.8]),
 8 we have that η is supported on the curves $\gamma \in \Gamma_{[0, T]}$, where γ are the uniform limits of the
 9 LASs characteristics where η^N is supported. \square

4. A SEMI-DISCRETE LAGRANGIAN SCHEME FOR MDE

10
 11 In this section, we first define a semi-discrete in time Lagrangian scheme for (2.6) and
 12 compare it to the LAS scheme in Definition 2.13, showing that they converge to the same
 13 limit. Fixed $T > 0$, for $N \in \mathbb{N}$ we set $\Delta t^N = T/N$ and we define a partition of $[0, T]$ by

$$(4.1) \quad D_N = \{t_k^N = k\Delta t^N, k = 0, \dots, N\}.$$

14 To simplify the notation, we omit the index N in t_k^N and in Δt^N if there is no ambiguity.
 15 Given $\mu_0 \in \mathcal{P}_c(\mathbb{R}^d)$ and a PVF V , we set

$$(4.2) \quad \begin{cases} \bar{\mu}_{t_0}^N := \mu_0; \\ \bar{\mu}_{t_{k+1}}^N := \int_{T\mathbb{R}^d} \delta_{x+v\Delta t^N} d\nu_x[\bar{\mu}_{t_k}^N](v) d\bar{\mu}_{t_k}^N(x) = \bar{\mu}_{t_k}^N \oplus \Delta t \cdot \nu_x[\bar{\mu}_{t_k}^N], \end{cases}$$

16 by Bochner integration formula (see [9]). Applying iteratively the previous definition, we
 17 get

$$(4.3) \quad \begin{aligned} \bar{\mu}_{t_{k+1}}^N &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^{(k+1)d}} \delta_{x_{k+1}^N} d\nu_{x_k^N}[\bar{\mu}_{t_k}^N](v_k^N) \dots d\nu_x[\bar{\mu}_{t_0}^N](v_0) \right) d\mu_0(x) \\ &= \mu_0 \oplus \left(\bigoplus_{i=0}^k \Delta t \cdot \nu_{x_i^N}[\bar{\mu}_{t_i}^N] \right), \end{aligned}$$

- 1 where $x_{k+1}^N := x + \sum_{j=0}^k v_j^N \Delta t^N$, for $x \in \text{supp}(\mu_0)$ and $v_j^N \in \text{supp}(\nu_{x_j^N}[\bar{\mu}_{t_j}^N])$. We extend
 2 $\bar{\mu}^N$ to the interval $[0, T]$ by setting for $t \in (t_k, t_{k+1}]$

$$\bar{\mu}_t^N := \int_{T\mathbb{R}^d} \delta_{x+v(t-t_k)} d\nu_x[\bar{\mu}_{t_k}^N](v) d\bar{\mu}_{t_k}^N(x),$$

- 3 and we denote $\bar{\mu}^N = \{\bar{\mu}_t^N\}_{t \in [0, T]}$. Due to the assumptions on V , we have that $\bar{\mu}_t^N \in \mathcal{P}(\mathbb{R}^d)$
 4 for all $t \in [0, T]$. Moreover, since for some $R > 0$, $\text{supp}(\mu_0) \subset B(0, R)$, then by (H1) and
 5 arguing as in [13, Lemma 3.3] it follows that

$$(4.4) \quad \text{supp}(\bar{\mu}_t^N) \subset B(0, e^{CT}(R+1)), \quad \forall t \in [0, T].$$

- 6 **Theorem 4.1.** *Let $\mu_0 \in \mathcal{P}_c(\mathbb{R}^d)$. Then, the scheme (4.2) converges, up to a subsequence,*
 7 *to a solution of (2.6).*
 8 *Moreover, assume that there exists a sequence $\{N_k\}_{k \in \mathbb{N}}$ such that both the scheme (4.2) and*
 9 *the LAS schemes in Definition 2.13 converge. Then, they converge to the same solution of*
 10 *(2.6).*

Proof. We first show that sequence (4.2) is equi-Lipschitz continuous in time. For $f \in \text{Lip}_1(\mathbb{R}^d)$, by (H1) and (4.4) we have

$$\begin{aligned} |\langle \bar{\mu}_{t_{k+1}}^N - \bar{\mu}_{t_k}^N, f \rangle| &= \left| \int_{T\mathbb{R}^d} [f(x + v\Delta t) - f(x)] d\nu_x[\bar{\mu}_{t_k}^N](v) d\bar{\mu}_{t_k}^N(x) \right| \\ &\leq \Delta t \int_{T\mathbb{R}^d} |v| d\nu_x[\bar{\mu}_{t_k}^N] d\bar{\mu}_{t_k}^N(x) \leq \Delta t C(1 + e^{CT}(R+1)). \end{aligned}$$

- 11 It follows that

$$W^{\mathbb{R}^d}(\bar{\mu}_t^N, \bar{\mu}_s^N) \leq K|t - s|, \quad \forall t, s \in [0, T],$$

for $K = K(R, T) > 0$. By Ascoli-Arzelà Theorem, the sequence $\{\bar{\mu}^N\}_{N \in \mathbb{N}}$ admits at least a subsequence, still denoted by $\bar{\mu}^N$, which converges to a measure map $\bar{\mu} \in \text{Lip}_K([0, T], \mathcal{P}_c(\mathbb{R}^d))$ such that $\bar{\mu}_{t=0} = \mu_0$.

We now prove that $\bar{\mu}$ is a solution of (2.6). For simplicity we index with N the converging subsequence. Given $t \in (t_k^N, t_{k+1}^N]$ and $f \in C^\infty(\mathbb{R}^d) \cap \text{Lip}_1(\mathbb{R}^d)$ such that $\|f\|_{C^2(\mathbb{R}^d)} \leq 1$,

we have

$$\begin{aligned}
& \langle \bar{\mu}_t - \mu_0, f \rangle - \int_0^t \int_{T\mathbb{R}^d} \nabla f(x) \cdot v \, dV[\bar{\mu}_s](x, v) \, ds \\
&= \langle \bar{\mu}_t - \bar{\mu}_t^N, f \rangle + \langle \bar{\mu}_t^N - \mu_0, f \rangle - \int_0^t \int_{T\mathbb{R}^d} \nabla f(x) \cdot v \, dV[\bar{\mu}_s](x, v) \, ds \\
(4.5) \quad &= \langle \bar{\mu}_t - \bar{\mu}_t^N, f \rangle + \langle \bar{\mu}_t^N - \bar{\mu}_{t_k}^N, f \rangle + \\
&\quad + \sum_{i=0}^{k-1} \left[\langle \bar{\mu}_{t_{i+1}}^N - \bar{\mu}_{t_i}^N, f \rangle - \int_{t_i}^{t_{i+1}} \int_{T\mathbb{R}^d} \nabla f(x) \cdot v \, dV[\bar{\mu}_s](x, v) \, ds \right] + \\
&\quad - \int_{t_k}^t \int_{T\mathbb{R}^d} \nabla f(x) \cdot v \, dV[\bar{\mu}_s](x, v) \, ds
\end{aligned}$$

Recalling that $\bar{\mu}_{t_{i+1}}^N = \int_{T\mathbb{R}^d} \delta_{x+v\Delta t} d\nu_x[\bar{\mu}_{t_i}^N] d\bar{\mu}_{t_i}^N(x)$, we have

$$\begin{aligned}
& \langle \bar{\mu}_{t_{i+1}}^N - \bar{\mu}_{t_i}^N, f \rangle - \int_{t_i}^{t_{i+1}} \int_{T\mathbb{R}^d} \nabla f(x) \cdot v \, dV[\bar{\mu}_s](x, v) \, ds \\
&= \int_{T\mathbb{R}^d} [f(x + v\Delta t) - f(x)] \, dV[\bar{\mu}_{t_i}^N](x, v) - \int_{t_i}^{t_{i+1}} \int_{T\mathbb{R}^d} \nabla f(x) \cdot v \, dV[\bar{\mu}_s](x, v) \, ds \\
&= \int_{t_i}^{t_{i+1}} \int_{T\mathbb{R}^d} \frac{d}{ds} f(x + (s - t_i)v) \, dV[\bar{\mu}_{t_i}^N](x, v) \, ds - \int_{t_i}^{t_{i+1}} \int_{T\mathbb{R}^d} \nabla f(x) \cdot v \, dV[\bar{\mu}_s](x, v) \, ds \\
&= \int_{t_i}^{t_{i+1}} \int_{T\mathbb{R}^d} \nabla f(x + (s - t_i)v) \cdot v \, dV[\bar{\mu}_{t_i}^N](x, v) \, ds - \int_{t_i}^{t_{i+1}} \int_{T\mathbb{R}^d} \nabla f(x) \cdot v \, dV[\bar{\mu}_s](x, v) \, ds \\
&= \int_{t_i}^{t_{i+1}} \int_{T\mathbb{R}^d} (\nabla f(x + (s - t_i)v) - \nabla f(x)) \cdot v \, dV[\bar{\mu}_{t_i}^N](x, v) \, ds + \\
&\quad + \int_{t_i}^{t_{i+1}} \int_{T\mathbb{R}^d} \nabla f(x) \cdot v \, d(V[\bar{\mu}_{t_i}^N] - V[\bar{\mu}_s])(x, v) \, ds \\
&= I_{1,i} + I_{2,i}.
\end{aligned}$$

We now estimate $I_{1,i}$ and $I_{2,i}$. By (H1) and (4.4), we have

$$\begin{aligned}
|I_{1,i}| &\leq \int_{t_i}^{t_{i+1}} \int_{T\mathbb{R}^d} |\nabla f(x + (s - t_i)v) - \nabla f(x)| \cdot |v| \, dV[\bar{\mu}_{t_i}^N](x, v) \, ds \\
&\leq \|f\|_{C^2(\mathbb{R}^d)} \int_{t_i}^{t_{i+1}} \int_{T\mathbb{R}^d} |s - t_i| |v|^2 \, dV[\bar{\mu}_{t_i}^N](x, v) \, ds \\
&\leq \|f\|_{C^2(\mathbb{R}^d)} C^2 (1 + e^{CT} (R + 1))^2 \Delta t^2.
\end{aligned}$$

By the Kantorovich-Rubinstein duality, (H3) and the triangular inequality, we get

$$\begin{aligned}
I_{2,i} &\leq \int_{t_i}^{t_{i+1}} W^{T\mathbb{R}^d}(V[\bar{\mu}_{t_i}^N], V[\bar{\mu}_s]) \, ds \leq \int_{t_i}^{t_{i+1}} L \cdot W^{\mathbb{R}^d}(\bar{\mu}_{t_i}^N, \bar{\mu}_s) \, ds \\
&\leq \int_{t_i}^{t_{i+1}} L(W^{\mathbb{R}^d}(\bar{\mu}_{t_i}^N, \bar{\mu}_s^N) + W^{\mathbb{R}^d}(\bar{\mu}_s^N, \bar{\mu}_s)) \, ds \leq \int_{t_i}^{t_{i+1}} L(K(s - t_i) + W^{\mathbb{R}^d}(\bar{\mu}_s^N, \bar{\mu}_s)) \, ds
\end{aligned}$$

$$\leq LK \left(\Delta t^2 + \int_{t_i}^{t_{i+1}} W^{\mathbb{R}^d}(\bar{\mu}_s^N, \bar{\mu}_s) ds \right).$$

Replacing the previous estimates in (4.5), we get

$$\begin{aligned} & \langle \bar{\mu}_t - \mu_0, f \rangle - \int_0^t \int_{T\mathbb{R}^d} \nabla f(x) \cdot v dV[\bar{\mu}_s](x, v) ds \\ & \leq \langle \bar{\mu}_t - \bar{\mu}_t^N, f \rangle + \sum_{i=0}^{k-1} (I_{1,i} + I_{2,i}) + \langle \bar{\mu}_t^N - \bar{\mu}_{t_k}^N, f \rangle - \int_{t_k}^t \int_{T\mathbb{R}^d} \nabla f(x) \cdot v dV[\bar{\mu}_s](x, v) ds \\ & \leq \langle \bar{\mu}_t - \bar{\mu}_t^N, f \rangle + LK \int_0^t W^{\mathbb{R}^d}(\bar{\mu}_s^N, \bar{\mu}_s) ds + 2K' T \Delta t, \end{aligned}$$

- 1 where $K' = \max\{LK, C^2(1 + e^{CT}(R+1))^2\}$. Passing to the limit for $N \rightarrow \infty$ in the
- 2 previous inequality and recalling that $W^{\mathbb{R}^d}(\bar{\mu}_t^N, \bar{\mu}_t) \rightarrow 0$, we finally get that

$$\langle \bar{\mu}_t - \mu_0, f \rangle - \int_0^t \int_{T\mathbb{R}^d} \nabla f(x) \cdot v dV[\bar{\mu}_s](x, v) ds = 0$$

- 3 for any f and therefore $\bar{\mu}$ is a solution of (2.6).

Let us now prove the second part of the theorem. Let $\{\mu^N\}$ be a convergent (sub-) sequence generated by the LAS scheme in Definition 2.13, and $\{\bar{\mu}^N\}$ be a convergent one generated by the scheme (4.2). Let us denote by $\mu, \bar{\mu}$ the corresponding limits. Then

$$W^{\mathbb{R}^d}(\mu_t, \bar{\mu}_t) \leq W^{\mathbb{R}^d}(\mu_t, \mu_t^N) + W^{\mathbb{R}^d}(\bar{\mu}_t^N, \bar{\mu}_t) + W^{\mathbb{R}^d}(\mu_t^N, \bar{\mu}_t^N).$$

Since the first two terms on the right-hand side of the last inequality converge to 0 for $N \rightarrow +\infty$, we have to study only the convergence of the last term. Let $f \in \text{Lip}_1(\mathbb{R}^d)$, $t \in (t_k, t_{k+1}]$, then

$$\begin{aligned} (4.6) \quad \langle \mu_t^N - \bar{\mu}_t^N, f \rangle &= \int_{T\mathbb{R}^d} f(x + (t - t_k)v) d(V[\mu_{t_k}^N] - V[\bar{\mu}_{t_k}^N]) \\ &+ \int_{T\mathbb{R}^d} f(x + (t - t_k)v) d(\mathcal{A}_N^v(V[\mu_{t_k}^N]) - V[\mu_{t_k}^N]). \end{aligned}$$

Notice that this computation holds thanks to the common time-grid shared by the two schemes. For the first term, we have

$$\int_{T\mathbb{R}^d} f(x + (t - t_k)v) d(V[\mu_{t_k}^N] - V[\bar{\mu}_{t_k}^N]) = \int_{T\mathbb{R}^d} f(x + w) d(V^{\Delta t^N}[\mu_{t_k}^N] - V^{\Delta t^N}[\bar{\mu}_{t_k}^N]),$$

where we have denoted $dV^{\Delta t^N}[\eta] = d((t - t_k) \cdot \nu_x[\eta])d\eta$, referring to the notation in (2.1). Then, we can observe that the map $\psi : (x, w) \rightarrow x + w$ belongs to $\text{Lip}_1(T\mathbb{R}^d, \mathbb{R}^d)$. Since $f \in \text{Lip}_1(\mathbb{R}^d, \mathbb{R})$, we have $f \circ \psi \in \text{Lip}_1(T\mathbb{R}^d, \mathbb{R})$. Then, from the previous inequality and

the Kantorovich-Rubinstein duality it follows that

$$\begin{aligned} \int_{T\mathbb{R}^d} f(x + (t - t_k)v) d(V[\mu_{t_k}^N] - V[\bar{\mu}_{t_k}^N]) &\leq W^{T\mathbb{R}^d}(V^{\Delta t^N}[\mu_{t_k}^N], V^{\Delta t^N}[\bar{\mu}_{t_k}^N]) \\ &\leq \Delta t^N \mathcal{W}(V[\mu_{t_k}^N], V[\bar{\mu}_{t_k}^N]) + W^{\mathbb{R}^d}(\mu_{t_k}^N, \bar{\mu}_{t_k}^N) \leq (L\Delta t^N + 1)W^{\mathbb{R}^d}(\mu_{t_k}^N, \bar{\mu}_{t_k}^N), \end{aligned}$$

where the last inequality is a consequence of (H3). For the second term in (4.6), by the same argument and (2.8), we found it is bounded by $\frac{1}{N^2}$. Then

$$\begin{aligned} \langle \mu_t^N - \bar{\mu}_t^N, f \rangle &\leq \frac{1}{N^2} + (1 + L\Delta t^N)W^{\mathbb{R}^d}(\mu_{t_k}^N, \bar{\mu}_{t_k}^N) \\ &\leq (1 + L\Delta t^N)^{k+1} o\left(\frac{1}{N}\right) + \frac{\sum_{l=0}^k (1 + L\Delta t^N)^l}{N^2} \\ &\leq e^{\frac{LT(k+1)}{N}} \cdot o\left(\frac{1}{N}\right) + \frac{e^{\frac{LT(k+1)}{N}} - 1}{NLT} \end{aligned}$$

1 and therefore the two schemes converge to the same limit, up to subsequences. \square

2 Before giving a further consideration coming as a consequence of the previous theorem,
3 we recall the following uniqueness result proved in [13, Theorem 5.2].

4 **Theorem 4.2.** *Let V be a PVF satisfying (H1) and (H3). Assume that, for every μ_0*
5 *obtained as convex combination of Dirac deltas, the sequence of LASs converges to a unique*
6 *limit. Then there exists a unique Lipschitz semigroup whose trajectories are limits of LASs.*

7 Then, as a corollary of Theorem 4.1, we are able to get a uniqueness result for a Lipschitz
8 semigroup of MDEs obtained as limit of the semi-discrete Lagrangian scheme, as follows.

9 **Corollary 4.3.** *Under the assumptions of Theorem 4.2, there exists a unique Lipschitz*
10 *semigroup generated by the semi-discrete Lagrangian scheme (4.2) and it coincides with*
11 *that generated by LASs.*

12 5. A MEAN VELOCITY SCHEME FOR MDE

13 In this section, inspired by the discussion developed in Section 3, we provide another
14 approximation scheme for the problem (2.6). By the result proved in Proposition 3.1, this
15 scheme is analogous to the semi-discrete in time Lagrangian scheme studied in [15, Scheme
16 1] for a nonlocal continuity equation, indeed it is based on the choice of a barycentric
17 velocity field. As our Lagrangian scheme in Section 4, also this scheme is semi-discrete

- 1 in time but, due to a different choice of the velocity field, it may converge to a different
 2 solution of the MDE (see also Remark 3.2 for a discussion about general lack of uniqueness).
 3 This fact will be exploited in Section 6 with clarifying examples.

4 We define Δt^N and t_k^N as in Section 4. Given $\mu_0 \in \mathcal{P}_c(\mathbb{R}^d)$ and a PVF V satisfying
 5 (H1) – (H3), the new approximation scheme is given iteratively by

$$(5.1) \quad \begin{cases} \hat{\mu}_{t=0}^N = \mu_0; \\ \bar{v}_{t_j}(x) := \int_{\mathbb{R}^d} v d\nu_x[\hat{\mu}_{t_j}^N](v), \\ \hat{\mu}_{t_{j+1}}^N = \hat{\mu}_{t_j}^N \oplus \Delta t^N \cdot \delta_{\bar{v}_{t_j}(x)}, \end{cases}$$

6 for $j \in \{0, \dots, N-1\}$. The scheme (5.1) transports the measure distribution $\hat{\mu}_{t_j}^N$ by a
 7 velocity field obtained as the barycenter of the velocity measure ν_x at $\hat{\mu}_{t_j}^N$.

8 **Theorem 5.1.** *Let $\mu_0 \in \mathcal{P}_c(\mathbb{R}^d)$. Then, the scheme (5.1) converges, up to a subsequence,*
 9 *to a solution of (2.6).*

10 *Proof.* We first prove that, given $R > 0$ such that $\text{supp}(\mu_0) \subset B(0, R)$, there exists $K =$
 11 $K(R, T) > 0$ such that for N sufficiently large

$$(5.2) \quad \text{supp}(\hat{\mu}_{t_j}^N) \subset B(0, K), \quad \forall j \in \{0, \dots, N\}.$$

Indeed

$$\begin{aligned} \sup_{\text{supp}(\hat{\mu}_{t_{j+1}}^N)} |x| &= \sup_{\text{supp}(\hat{\mu}_{t_j}^N)} |x + \Delta t \bar{v}_{t_j}(x)| \leq \sup_{\text{supp}(\hat{\mu}_{t_j}^N)} |x| + \Delta t \sup_{\text{supp}(V[\hat{\mu}_{t_j}^N])} |v| \\ &\leq \Delta t C + (1 + \Delta t C) \sup_{\text{supp}(\hat{\mu}_{t_j}^N)} |x| \leq \dots \leq \Delta t C \sum_{k=0}^j (1 + \Delta t C)^k + (1 + \Delta t C)^{j+1} R \\ &\leq e^{Cj\Delta t} - 1 + Re^{TC} \leq e^{TC}(1 + R) - 1, \end{aligned}$$

where we used (H1). We now prove the equicontinuity in time of the scheme: $t, s \in [0, T]$
 and $N \gg 1$ such that $t - s > \Delta t^N$; then there exists $\{t_j^N\} = \{t_j\}$ such that $s < t_j <$
 $\dots t_{j+k} < t$. Let $f \in \text{Lip}_1(\mathbb{R}^d)$, then

$$\langle \hat{\mu}_t^N - \hat{\mu}_s^N, f \rangle = \langle \hat{\mu}_t^N - \hat{\mu}_{t_{j+k}}^N, f \rangle + \sum_{i=1}^k \langle \hat{\mu}_{t_{j+i}}^N - \hat{\mu}_{t_{j+i-1}}^N, f \rangle + \langle \hat{\mu}_{t_j}^N - \hat{\mu}_s^N, f \rangle.$$

By construction,

$$\begin{aligned}\langle \hat{\mu}_{t_{j+1}}^N - \hat{\mu}_{t_j}^N, f \rangle &= \int_{T\mathbb{R}^d} (f(x + \Delta t^N \bar{v}_j) - f(x)) d\hat{\mu}_{t_j}^N \\ &\leq \Delta t^N \int_{\mathbb{R}^d} |\bar{v}_j(x)| d\hat{\mu}_{t_j}^N(x) \leq \Delta t^N \int_{T\mathbb{R}^d} |v| dV[\hat{\mu}_{t_j}^N].\end{aligned}$$

Hence by (H1) and the equi-boundedness of supports, it follows

$$\langle \hat{\mu}_{t_{j+1}}^N - \hat{\mu}_{t_j}^N, f \rangle \leq \Delta t^N \int_{\mathbb{R}^d} C(1 + |x|) d\hat{\mu}_{t_j}^N(x) \leq \Delta t^N C(1 + K).$$

Analogously $\langle \hat{\mu}_t^N - \hat{\mu}_{t_{j+k}}^N, f \rangle \leq |t - t_{j+k}|C(1 + K)$ and $\langle \hat{\mu}_{t_j}^N - \hat{\mu}_s^N, f \rangle \leq |t_j - s|C(1 + K)$.

Hence, taking the supremum for $f \in \text{Lip}_1(\mathbb{R}^d)$, we have

$$W^{\mathbb{R}^d}(\hat{\mu}_t^N, \hat{\mu}_s^N) \leq |t - s|C(1 + K).$$

- 1 Since the support of $\hat{\mu}_{t_j}^N$ is bounded, uniformly in N , it immediately follows that the
- 2 sequence $\{\hat{\mu}^N\}_{N \in \mathbb{N}}$ have bounded first and second momentum and therefore there exists
- 3 $\mu \in C([0, T]; \mathcal{P}_c(\mathbb{R}^d))$ such that, up to a subsequence,

$$(5.3) \quad \sup_{t \in [0, T]} W^{\mathbb{R}^d}(\hat{\mu}_t^N, \mu_t) \rightarrow 0, \text{ for } N \rightarrow +\infty.$$

- 4 We now prove that μ is a solution of (2.6). Given $f \in C_c^\infty(\mathbb{R}^d)$ such that $\|f\|_{C^2(\mathbb{R}^d)} \leq 1$,
- 5 we rewrite

$$\begin{aligned}(5.4) \quad \langle \hat{\mu}_{t_k}^N - \hat{\mu}_0^N, f \rangle &= \sum_{j=1}^k \langle \hat{\mu}_{t_j}^N - \hat{\mu}_{t_{j-1}}^N, f \rangle, \\ &= \sum_{j=1}^k \int_{\mathbb{R}^d} (f(x + \Delta t \bar{v}_{t_{j-1}}(x)) - f(x)) d\hat{\mu}_{t_{j-1}}^N \\ &= \sum_{j=1}^k \int_{\mathbb{R}^d} \int_{t_{j-1}}^{t_j} \frac{d}{ds} f(x + (s - t_{j-1})\bar{v}_{t_{j-1}}(x)) ds d\hat{\mu}_{t_{j-1}}^N(x) \\ &= \sum_{j=1}^k \int_{t_{j-1}}^{t_j} \int_{\mathbb{R}^d} \bar{v}_{t_{j-1}}(x) \cdot \nabla f(x + (s - t_{j-1})\bar{v}_{t_{j-1}}(x)) d\hat{\mu}_{t_{j-1}}^N(x) ds.\end{aligned}$$

We estimate

$$\begin{aligned}
& \left| \int_{t_{j-1}}^{t_j} \left(\int_{\mathbb{R}^d} \bar{v}_{t_{j-1}}(x) \cdot \nabla f(x + (s - t_{j-1})\bar{v}_{t_{j-1}}(x)) d\hat{\mu}_{t_{j-1}}^N(x) - \int_{T\mathbb{R}^d} v \cdot \nabla f(x) dV[\mu_s] \right) ds \right| \\
&= \left| \int_{t_{j-1}}^{t_j} \left(\int_{T\mathbb{R}^d} v \cdot \nabla f(x + (s - t_{j-1})\bar{v}_{t_{j-1}}(x)) dV[\hat{\mu}_{t_{j-1}}^N] - \int_{T\mathbb{R}^d} v \cdot \nabla f(x) dV[\mu_s] \right) ds \right| \\
&\leq \int_{t_{j-1}}^{t_j} \left| \int_{T\mathbb{R}^d} v \cdot \nabla (f(x + (s - t_{j-1})\bar{v}_{t_{j-1}}(x)) - f(x)) dV[\hat{\mu}_{t_{j-1}}^N] \right| ds \\
&+ \int_{t_{j-1}}^{t_j} \left| \int_{T\mathbb{R}^d} v \cdot \nabla f(x) d(V[\hat{\mu}_{t_{j-1}}^N] - V[\mu_s]) \right| ds \\
&\leq \int_{t_{j-1}}^{t_j} (s - t_{j-1}) \int_{\mathbb{R}^d} \|D^2 f\|_{C^2} |\bar{v}_{t_{j-1}}(x)|^2 d\hat{\mu}_{t_{j-1}}^N(x) + \int_{t_{j-1}}^{t_j} W^{T\mathbb{R}^d}(V[\hat{\mu}_{t_{j-1}}^N], V[\mu_s]) ds \\
&\leq \int_{t_{j-1}}^{t_j} (s - t_{j-1}) \int_{T\mathbb{R}^d} |v|^2 dV[\hat{\mu}_{t_{j-1}}^N](x) ds + \Delta t \sup_{t \in [t_{j-1}, t_j]} W^{T\mathbb{R}^d}(V[\hat{\mu}_{t_{j-1}}^N], V[\mu_t]).
\end{aligned}$$

Therefore, by (H1) and (5.2), we have

$$\begin{aligned}
& \left| \sum_{j=1}^k \int_{t_{j-1}}^{t_j} \int_{\mathbb{R}^d} \bar{v}_{t_{j-1}}(x) \nabla f(x + (s - t_{j-1})\bar{v}_{t_{j-1}}(x)) d\hat{\mu}_{t_{j-1}}^N(x) ds - \int_0^T \int_{T\mathbb{R}^d} v \cdot \nabla f(x) dV[\mu_s](x, v) ds \right| \\
&\leq \sum_{j=1}^k \left(\int_{t_{j-1}}^{t_j} (s - t_{j-1}) \int_{T\mathbb{R}^d} |v|^2 dV[\hat{\mu}_{t_{j-1}}^N](x) ds + \Delta t \sup_{t \in [t_{j-1}, t_j]} W^{T\mathbb{R}^d}(V[\hat{\mu}_{t_{j-1}}^N], V[\mu_t]) \right) \\
&\leq \sum_{j=1}^k \frac{\Delta t^2}{2} C' + T \sup_{1 \leq j \leq k} \sup_{t \in [t_{j-1}, t_j]} W^{T\mathbb{R}^d}(V[\hat{\mu}_{t_{j-1}}^N], V[\mu_t]) \\
&\leq \Delta t T C' + T \sup_{1 \leq j \leq k} \sup_{t \in [t_{j-1}, t_j]} W^{T\mathbb{R}^d}(V[\hat{\mu}_{t_{j-1}}^N], V[\mu_t]).
\end{aligned}$$

- 1 By (H2), V is uniformly continuous on $B(0, K)$, hence we conclude that the right hand-
- 2 side vanishes as $N \rightarrow +\infty$, thanks to (5.2) and (5.3). If $t_k^N \rightarrow t$ for $N \rightarrow \infty$, since the term
- 3 on the left side in (5.4) converges to $\langle \mu_t - \mu_0, f \rangle$ by construction, the previous estimate
- 4 implies that μ is a weak solution to (2.6). \square

5 *Remark 5.2.* For sake of completeness, similarly to Definition 3.6, we can provide an explicit
6 formula to construct a probabilistic representation for the scheme introduced in Section 4
7 and for the mean velocity one, as follows. Let $I_a^b := [a\Delta t^N, b\Delta t^N]$, for $a, b \in \mathbb{N}$, $a \leq b$,
8 with Δt^N and t_k as in Section 4. Denote respectively with $\bar{\mu}^N = \{\bar{\mu}_t^N\}_{t \in [0, T]}$ and $\hat{\mu}^N =$
9 $\{\hat{\mu}_t^N\}_{t \in [0, T]}$ the schemes defined in (4.2) and (5.1). For $k = 0, \dots, N - 1$, we define

(A) $\bar{\eta}_{I_k^{k+1}}^N := \bar{\mu}_{t_k}^N \oplus (\cdot - t_k) \cdot \nu_x[\bar{\mu}_{t_k}^N] \in \mathcal{P}(\Gamma_{I_k^{k+1}})$, i.e.,

$$\bar{\eta}_{I_k^{k+1}}^N = \int_{\mathbb{R}^d \times \mathbb{R}^d} \delta_{x+(\cdot-t_k)v} d\nu_x[\bar{\mu}_{t_k}^N](v) d\bar{\mu}_{t_k}^N(x) = \int_{T\mathbb{R}^d} \delta_{x+(\cdot-t_k)v} dV[\bar{\mu}_{t_k}^N](x, v);$$

(B) $\hat{\eta}_{I_k^{k+1}}^N := \hat{\mu}_{t_k}^N \oplus (\cdot - t_k) \cdot \delta_{\bar{v}_{t_k}(x)} \in \mathcal{P}(\Gamma_{I_k^{k+1}})$, i.e., $\hat{\eta}_{I_k^{k+1}}^N = \int_{\mathbb{R}^d} \delta_{x+(\cdot-t_k)\bar{v}_{t_k}(x)} d\hat{\mu}_{t_k}^N(x)$.

Now, we can build $\bar{\eta}^N$ and $\hat{\eta}^N$ by applying items (2 – 3) of Definition 3.6 and replacing item (1) respectively with (A) and (B).

Following the same line as in Proposition 3.7, we can prove that an analogous result holds also for the semi-discrete in time Lagrangian scheme (or the mean velocity one) by replacing the LAS scheme μ^N with the scheme $\bar{\mu}^N$ (or $\hat{\mu}^N$) and using the representation $\bar{\eta}^N$ (or $\hat{\eta}^N$) just provided.

6. EXAMPLES

In this section we present some examples in $d = 1$ aimed at clarifying the work of the various proposed schemes, in particular we show that the LAS scheme and the mean velocity one in (5.1) may converge to different solutions of a MDE. We also exploit the result stated in Proposition 3.1. For simplicity of computations and without loss of generality, let us set $\Delta_N = \Delta t^N = 1/N$ as a time-step size for all the schemes. We denote with $\mathcal{L} = \mathcal{L}^1$ the 1-dimensional Lebesgue measure.

Example 6.1 (Splitting particle). For every $\mu \in \mathcal{P}_c(\mathbb{R})$ define:

$$B(\mu) = \sup \left\{ x : \mu([-\infty, x]) \leq \frac{1}{2} \right\}.$$

Set $\eta(\mu) = \mu([-\infty, B(\mu)]) - \frac{1}{2}$ so $\mu(\{B(\mu)\}) = \eta(\mu) + \frac{1}{2} - \mu([-\infty, B(\mu)[[$. We define

$V[\mu] = \mu \otimes \nu_x[\mu]$, with

$$(6.1) \quad \nu_x[\mu] = \begin{cases} \delta_{-1} & \text{if } x < B(\mu) \\ \delta_1 & \text{if } x > B(\mu) \\ \frac{1}{\mu(\{B(\mu)\})} (\eta\delta_1 + (\frac{1}{2} - \mu([-\infty, B(\mu)[[) \delta_{-1}) & \text{if } x = B(\mu), \mu(\{B(\mu)\}) > 0. \end{cases}$$

The solution obtained as limit of LASs, satisfies:

$$\mu_t(A) = \mu_0((A \cap]-\infty, B(\mu_0) - t[) + t) + \mu_0((A \cap]B(\mu_0) + t, +\infty[) - t)$$

$$+ \frac{1}{\mu_0(\{B(\mu_0)\})} \left(\eta\delta_{B(\mu_0)+t}(A) + (\frac{1}{2} - \mu_0([-\infty, B(\mu_0)[[) \delta_{B(\mu_0)-t}(A) \right).$$

1 In particular:

2 i) The solution with $\mu_0 = \delta_{x_0}$ is given by $\mu_t = \frac{1}{2}\delta_{x_0+t} + \frac{1}{2}\delta_{x_0-t}$, as illustrated in

3 **Figure 1**;

4 ii) The solution with $\mu_0 = \chi_{[a,b]} \lambda$ (where χ is the characteristic function and $\lambda =$
5 $\frac{1}{b-a} \mathcal{L}_{\perp[a,b]}$) is given by $\mu_t = \chi_{[a-t, \frac{a+b}{2}-t]} \lambda + \chi_{[\frac{a+b}{2}+t, b+t]} \lambda$.

6 The same behavior is valid for the scheme (4.2) (see Theorem 4.1). Moreover, it can be
7 verified that the stationary solution $\{\delta_{x_0}\}_t$ is the unique limit of the mean velocity scheme
8 (5.1) when $\mu_0 = \delta_{x_0}$, while this scheme has the same behavior of the LASs one when
9 $\mu_0 = \chi_{[a,b]} \lambda$. Hence the limit solution depends, in general, on the given approximation
10 scheme.

11 Let us read these results in light of Proposition 3.1 and Remark 3.2, in the easiest case
12 when we start from $\mu_0 = \delta_{x_0}$. First, we notice that the maps $x \mapsto \nu_x[\mu]$ and thus also

$$(6.2) \quad x \mapsto w[\mu](x) := \int_{\pi_1^{-1}(x)} v \, d\nu_x[\mu](v)$$

13 are not even continuous, thus uniqueness of solutions to the continuity equation $\partial_t \mu_t +$
14 $\operatorname{div}(w[\mu_t] \mu_t) = 0$, or equivalently to the MDE $\dot{\mu}_t = V[\mu_t]$, with $\mu|_{t=0} = \mu_0$, is not
15 guaranteed. We can verify that both the curves $\boldsymbol{\mu}^1 = \{\delta_{x_0}\}_t$ and $\boldsymbol{\mu}^2 = \{\mu_t\}_t$, with
16 $\mu_t = \frac{1}{2}\delta_{x_0+t} + \frac{1}{2}\delta_{x_0-t}$, selected by the schemes (4.2) and (5.1) are in fact solutions. In-
17 deed, along $\boldsymbol{\mu}^1$, we have $B(\delta_{x_0}) = x_0$ and $\eta(\delta_{x_0}) = \frac{1}{2}$, thus $\nu_{x_0}[\delta_{x_0}] = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$ and
18 $w[\delta_{x_0}](x_0) = 0$. Hence straightforwardly, (2.4) with w as in (6.2) and so the MDE are
19 satisfied for $\boldsymbol{\mu}^1$.

Concerning $\boldsymbol{\mu}^2$, for $t > 0$ we have $B(\mu_t) = x_0 + t$, $\eta(\mu_t) = \frac{1}{2}$, thus $\nu_{x_0-t}[\mu_t] = \delta_{-1}$ and
 $\nu_{x_0+t}[\mu_t] = \delta_1$. Hence,

$$\begin{aligned} \int_{T\mathbb{R}^d} (\nabla f(x) \cdot v) \, dV[\mu_t](x, v) &= \int_{\mathbb{R}^d} \nabla f(x) \cdot \left[\int_{\pi_1^{-1}(x)} v \, d\nu_x[\mu_t](v) \right] d\mu_t(x) \\ &= \frac{1}{2} \nabla f(x_0 + t) - \frac{1}{2} \nabla f(x_0 - t) \\ &= \frac{d}{dt} \int_{\mathbb{R}^d} f(x) \, d\mu_t(x), \end{aligned}$$

20 thus also $\boldsymbol{\mu}^2$ is a solution to the same MDE (or equivalently to (2.4) with w as in (6.2)).

21 *Example 6.2.* Let $V : \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(T\mathbb{R})$ a PVF defined by $V[\mu] := \mu \otimes \omega$, where $\omega :=$
22 $\frac{1}{2}(\delta_{-1} + \delta_1)$, and let $\mu_0 = \delta_0$. Then, both the LAS and (4.2) schemes give a binomial
23 distribution at every time (see **Figure 2**) while, as in the previous example, the mean

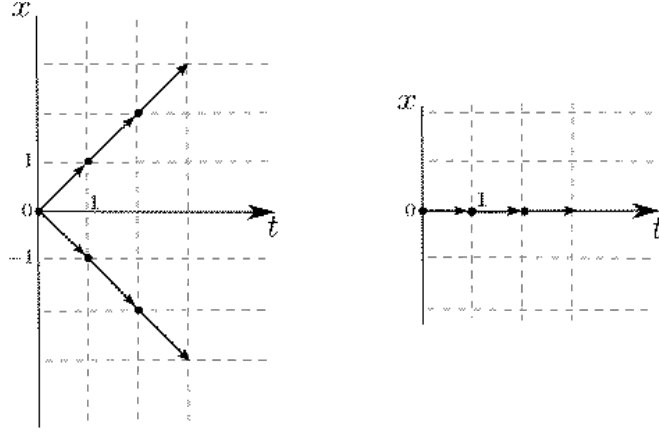


FIGURE 1. LAS and semi-discrete Lagrangian schemes on the left for $N = 1$. Mean-velocity scheme on the right.

- 1 velocity scheme remains stationary. However by the Law of Large Numbers, as $N \rightarrow +\infty$
- 2 all the three schemes univocally converge to the constant solution $\mu = \{\delta_0\}_{t \in [0, T]}$. We refer
- 3 to [13, Proposition 7.1] for a formal proof.
- 4 This limit behavior is in line with Proposition 3.1, indeed the MDE under consideration is
- 5 equivalent to the continuity equation driven by

$$w[\mu](x) = \int_{\pi_1^{-1}(x)} v d\omega(v) = 0,$$

- 6 for μ -a.e. $x \in \mathbb{R}^d$. By standard theory, the continuity equation driven by such a vector
- 7 field, $\partial_t \mu_t + \operatorname{div}(w\mu_t) = 0$, admits a unique solution. Hence, also the MDE as a unique
- solution, that is the stationary one.

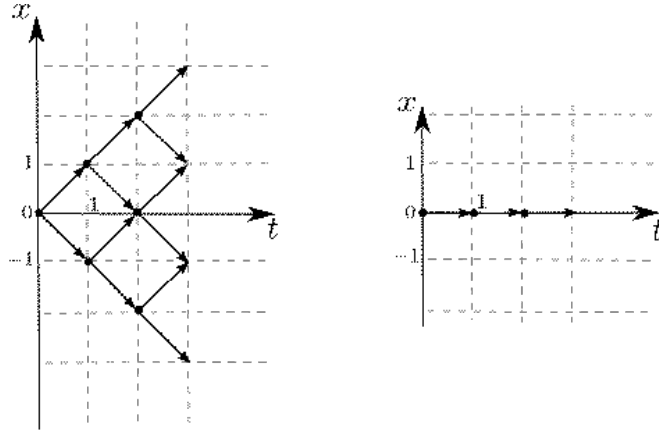


FIGURE 2. LAS and semi-discrete Lagrangian schemes on the left. Mean-velocity scheme on the right.

8

- 9 *Example 6.3.* Let $V : \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(T\mathbb{R})$ a PVF defined by $V[\mu] := \mu \otimes \omega$, where $\omega :=$
- 10 $\frac{1}{2}\chi_{[-1,1]}\mathcal{L} = \frac{1}{2}\mathcal{L}_{\llbracket -1,1 \rrbracket}$. Let $\mu_0 = \delta_0$. Considering the LAS scheme, as illustrated in **Figure**

1 **3**, we notice that for $N = 1$, the points v_j in the discretized space of velocities such that
 2 $m_{ij}^v(V[\mu]) \neq 0$ are $v_0 = -1$ and $v_1 = 0$, with equal weight. For $N = 2$, we get $v_0 = -1$,
 3 $v_1 = -1/2$, $v_2 = 0$ and $v_3 = 1/2$, hence we start to give mass also to positive $x \in \mathbb{R}$, thus
 obtaining $\mu_{|t=1/2}^2 = \frac{1}{4} \sum_{i=0}^3 \delta_{(-1/2+i/4)}$ and $\mu_{|t=1}^2 = \sum_{i=-2}^4 \frac{1}{16} (4 - |i-1|) \delta_{(-1/2+i/4)}$.

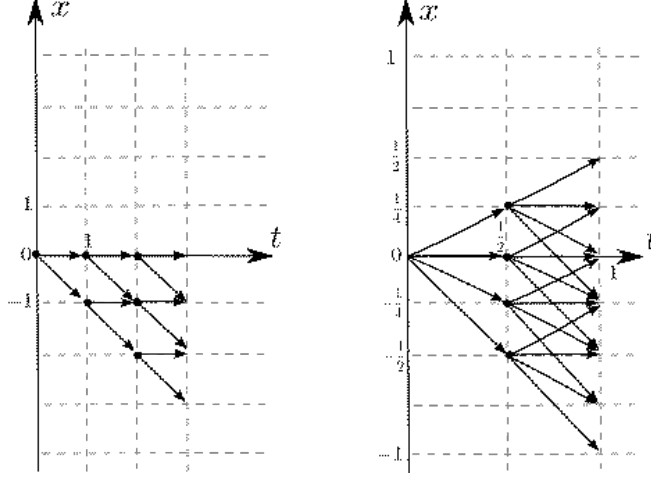


FIGURE 3. LAS scheme: for $N = 1$ (left) and $N=2$ (right).

4

5 Coming to the semi-discrete Lagrangian scheme (4.2), at the first time-step we get the
 6 uniform distribution on $[-1, 1]$, while afterwards we obtain a normal distribution on $[-t, t]$
 7 (see **Figure 4**). Reasoning in the same way as in the previous example, by the Law of
 8 Large Numbers, the LAS scheme and so also the semidiscrete Lagrangian one converge to
 9 the constant solution as $N \rightarrow \infty$ (see [13, Proposition 7.1]). Trivially, the mean-velocity
 10 scheme shares the same behavior.

11 We can get this conclusion also through the very same discussion analyzed in Example 6.2,
 12 indeed by Proposition 3.1, we get uniqueness of solutions for this MDE, being equivalent
 13 to a continuity equation driven by $w = 0$.

14 *Example 6.4.* Let $V : \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(T\mathbb{R})$ a PVF defined by $V[\mu] := \mu \otimes \delta_{v(x)}$, where
 15 $v(x) := 2\sqrt{|x|}$, and $\mu_{|t=0} = \delta_{-1}$. Recalling Definition 2.10, by the atomic nature of the
 16 PVF V over the fibers $T_x\mathbb{R}$, we deduce that the set of solutions to the MDE coincides
 17 with the set of distributional solutions of the continuity equation driven by the vector field
 18 $v(\cdot)$. We can thus use the classical Superposition Principle [1, Theorem 8.2.1] to build the
 19 trajectories μ by considering the integral solutions of the underlying ODE $\dot{x}(t) = 2\sqrt{|x(t)|}$,
 20 with initial condition $x(0) = -1$. By classical theory we know that this system admits
 21 infinite solutions (called Peano's brush), such as the trivial one $x_\infty(t) = -(t-1)^2$ for $t \leq 1$,

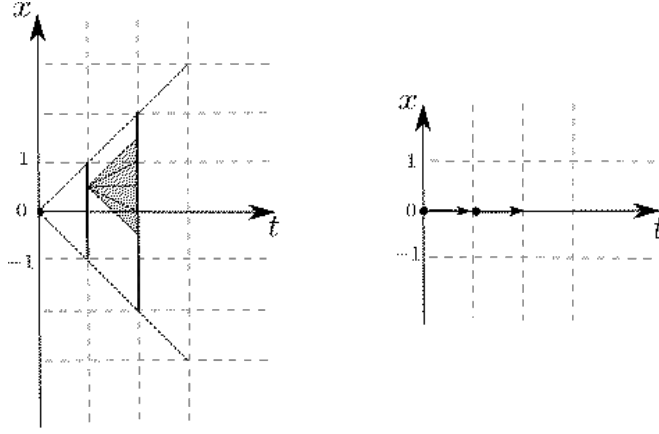


FIGURE 4. Semi-discrete Lagrangian scheme on the left. Mean-velocity scheme on the right.

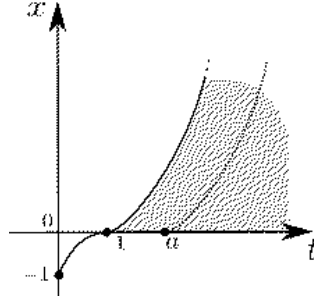


FIGURE 5. Peano's brush referred to Example 4.

1 $x_\infty(t) = 0$ for $t \geq 1$, but also the trajectories given by

$$(6.3) \quad x_a(t) = \begin{cases} -(t-1)^2, & \text{if } t \leq 1, \\ 0, & \text{if } 1 \leq t \leq a, \\ (t-a)^2, & \text{if } t \geq a, \end{cases}$$

2 as a varies in $[1, +\infty[$ (see **Figure 5**). In particular, among the infinite solutions of the
 3 MDE, we have $\boldsymbol{\mu}^1 = \{\mu_t^1\}_{t \in [0, T]}$, with $\mu_t^1 = \delta_{x_1(t)}$, and $\boldsymbol{\mu}^2 = \{\mu_t^2\}_{t \in [0, T]}$, with $\mu_t^2 = \delta_{x_\infty(t)}$.

4 Computing the LAS scheme for $N = 1$ we get:

- 5 (1) $\mu_0 = \delta_{-1}$, hence $v(-1) = 2$ which belongs to the velocity grid;
- 6 (2) so we get $\mu_{|t=1}^{N=1} = \delta_{(-1+2)} = \delta_1$, hence $v(1) = 2$ which belongs to the velocity grid;
- 7 (3) so we get $\mu_{|t=2}^{N=1} = \delta_{(1+2)} = \delta_3$, hence $v(3) = 2\sqrt{3}$ which does not belong to the
 8 velocity grid. Since $3 < 2\sqrt{3} < 4$, then the point in the discretized space of
 9 velocities for $N = 1$ such that $m_{ij}^v(V[\mu]) \neq 0$ is $v_j = 3$;
- 10 (4) so we get $\mu_{|t=3}^{N=1} = \delta_{3+3} = \delta_6$, and so on.

1 For $N = 2$ and $N = 3$, by performing similar computations we obtain the trajectories
 2 as represented in **Figure 6**. We can show that the LAS scheme converges to μ^2 , and
 3 thus so does the semidiscrete Lagrangian scheme, up to subsequences. Moreover we notice
 4 that, due to the atomic nature of V over the fibers $T_x\mathbb{R}$, the mean velocity scheme (5.1)
 5 coincides with the semidiscrete Lagrangian one (4.2). Thus, all the three schemes converge,
 6 up to subsequences to the same solution μ^2 . Finally we point out that the semidiscrete
 7 Lagrangian scheme corresponds to the Euler method for the underlying ODE. We also
 8 notice that in our case, for all $N \in \mathbb{N}$ the grid intersects the critical point $(t, x) = (1, 0)$
 9 where we loose local Lipschitzianity of the vector field. If we perform a perturbation of
 10 the grid, shifting it w.r.t. the critical point, then the schemes will converge to μ^1 , up to
 subsequences.

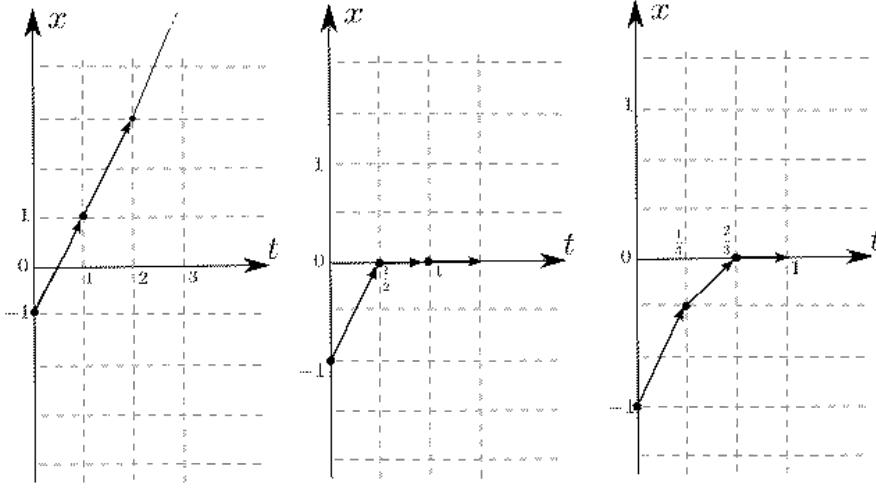


FIGURE 6. LAS scheme for $N= 1, 2, 3$.

11
 12 The lack of uniqueness for the notion of weak solution given in Definition 2.10 and
 13 exploited in the examples is not surprising, as already observed in Remark 3.2. On the
 14 other side, by Proposition 3.1, if the mean velocity field is enough regular, the theory in
 15 [1] would grant us the uniqueness of a solution of the MDE as push-forward of the initial
 16 condition.

17 **Acknowledgements:** G.C. thanks SBAI, “Sapienza” Università di Roma for its valu-
 18 able hospitality during the preparation of the paper. G.C. is also indebted with University
 19 of Pavia where this research has been partially carried out, in particular G.C. has been sup-
 20 ported by Cariplo Foundation and Regione Lombardia via project *Variational Evolution*

1 *Problems and Optimal Transport*, and by MIUR PRIN 2015 project *Calculus of Variations*,
 2 together with FAR funds of the Department of Mathematics of the University of Pavia.
 3 G.C. thanks also the support of the INdAM-GNAMPA Project 2019 *Optimal transport for*
 4 *dynamics with interaction* (“*Trasporto ottimo per dinamiche con interazione*”).

REFERENCES

- 6 [1] Ambrosio, L., Gigli, N. and Savaré, G.: *Gradient flows in metric spaces and in the space of probability*
 7 *measures*. Lectures in Mathematics ETH Zürich, 2nd ed., Birkhäuser Verlag, Basel, 2008.
- 8 [2] Bongini, M. and Buttazzo, G.: *Optimal control problems in transport dynamics*. Mathematical Models
 9 and Methods in Applied Sciences, Vol. 27, n. 3, pp. 427–451, 2017.
- 10 [3] Camilli, F., De Maio, R. and Tosin, A.: Measure-valued solutions to nonlocal transport equations on
 11 networks. J. Differential Equations, 264 (2018), pp. 7213–7241.
- 12 [4] Camilli, F., De Maio, R. and Tosin, A.: Transport of measures on networks. Networks & Heterogeneous
 13 Media, 2017, 12 (2), pp. 191–215.
- 14 [5] Cañizo, J. A., Carrillo J. A. and Rosado, J.: A well-posedness theory in measures for some kinetic
 15 models of collective motion. Math. Models Methods Appl. Sci., 21 (2011), pp. 515–539.
- 16 [6] Cavagnari, G., Marigonda, A. and Piccoli, B.: Generalized dynamic programming principle and sparse
 17 mean-field control problems. Journal of Mathematical Analysis and Applications, vol. 481, n. 1 (2020).
- 18 [7] Cavagnari, G., Marigonda, A. and Piccoli, B.: Superposition principle for differential inclusions. I.
 19 Lirkov, S. Margenov (Eds.). Large-Scale Scientific Computing, LSSC 2017, Lecture Notes in Computer
 20 Science, vol. 10665, pp. 201–209. Springer, Cham (2018).
- 21 [8] Cristiani, E., Piccoli, B. and Tosin, A.: *Multiscale Modeling of Pedestrian Dynamics*. Springer Inter-
 22 national Publishing, 2014, MS&A: Modeling, Simulation and Applications, Vol. 12
- 23 [9] Diestel, J. and Uhl, J.J.: *Vector Measures*. Amer. Math. Soc., Providence, 1977
- 24 [10] Jabin, P.-E.: *A review of the mean field limits for Vlasov equations*. Kinetic & Related Models, Vol.
 25 7, 2014.
- 26 [11] Golse, F.: *The mean-field limit for the dynamics of large particle systems*. Journées équations aux
 27 dérivées partielles, 2003.
- 28 [12] Orrieri, C.: *Large deviations for interacting particle systems: joint mean-field and small-noise limit*.
 29 arXiv:1810.12636.
- 30 [13] Piccoli, B.: Measure differential equations. Arch Rational Mech Anal, vol. 233, pp. 1289–1317 (2019).
- 31 [14] Piccoli, B. and Rossi, F.: *Measure dynamics with Probability Vector Fields and sources*. Discrete &
 32 Continuous Dynamical Systems - A, Vol. 39, 2019.
- 33 [15] Piccoli, B. and Rossi, F.: *Transport Equation with Nonlocal Velocity in Wasserstein Spaces: Conver-*
 34 *gence of Numerical Schemes*. Acta Applicandae Mathematicae, Vol. 124, pp. 73–105, 2011.
- 35 [16] Poupaud, F. and Rascle, M.: *Measure solutions to the linear multi-dimensional transport equation*
 36 *with non-smooth coefficients*. Communications in Partial Differential Equations, Vol. 22, n. 1-2, pp.
 37 225–267, 1997.

- 1 [17] Santambrogio, F.: *Optimal Transport for Applied Mathematicians*. Progress in Nonlinear Differential
2 Equations and Their Applications, Birkhäuser Basel, vol. 87, ed. 1 (2015).
3 [18] Vicsek, T. and Zafeiris, A.: *Collective motion*. Physics Reports, Vol. 517, n. 3, pp. 71–140, 2012.
4 [19] Villani, C.: *Topics in Optimal Transportation*. American Mathematical Society, Graduate Studies in
5 Mathematics, Vol. 58, 2003

FABIO CAMILLI: “SAPIENZA” UNIVERSITÀ DI ROMA,

6 DIPARTIMENTO DI SCIENZE DI BASE E APPLICATE PER L'INGEGNERIA

VIA SCARPA 16, I-00161 ROME, ITALY.

7 *Email address:* `camilli@sbai.uniroma1.it`

GIULIA CAVAGNARI: POLITECNICO DI MILANO,

8 DIPARTIMENTO DI MATEMATICA “F. BRIOSCHI”

PIAZZA LEONARDO DA VINCI, 32, I-20133 MILANO, ITALY.

9 *Email address:* `giulia.cavagnari@polimi.it`

RAUL DE MAIO: “SAPIENZA” UNIVERSITÀ DI ROMA,

10 DIPARTIMENTO DI SCIENZE DI BASE E APPLICATE PER L'INGEGNERIA

VIA SCARPA 16, I-00161 ROME, ITALY.

11 *Email address:* `raul.demaio@uniroma1.it`

BENEDETTO PICCOLI: RUTGERS UNIVERSITY - CAMDEN,

12 DEPARTMENT OF MATHEMATICAL SCIENCES

311 N. 5TH STREET CAMDEN, NJ 08102, USA.

13 *Email address:* `piccoli@camden.rutgers.edu`