Arch. Rational Mech. Anal. 233 (2019) 1289–1317
Digital Object Identifier (DOI) https://doi.org/10.1007/s00205-019-01379-4

Measure Differential Equations

Benedetto Piccoli

Communicated by C. De Lellis

Abstract

A new type of differential equations for probability measures on Euclidean spaces, called measure differential equations (briefly MDEs), is introduced. MDEs correspond to probability vector fields, which map measures on an Euclidean space to measures on its tangent bundle. Solutions are intended in weak sense and existence, uniqueness and continuous dependence results are proved under suitable conditions. The latter are expressed in terms of the Wasserstein metric on the base and fiber of the tangent bundle. MDEs represent a natural measure-theoretic generalization of ordinary differential equations via a monoid morphism mapping sums of vector fields to fiber convolution of the corresponding probability vector Fields. Various examples, including finite-speed diffusion and concentration, are shown, together with relationships to partial differential equations. Finally, MDEs are also natural mean-field limits of multi-particle systems, with convergence results extending the classical Dobrushin approach.

1. Introduction

The evolution of many physical and biological systems can be modeled by ordinary or partial differential equations. To include a representation of uncertainties, the state of the system can be modeled by a probability distribution or a random variable rather than a point of a Euclidean space (or a manifold.) Stochastic differential equations (SDEs) [13] offer a well-developed and successful tool to describe the evolution of random variables.

We define a new type of differential equation for probability measures. The point of view is that of optimal transport, thus we endow the space of probability measures (on an Euclidean space $\mathbb{R}^n$) with the Wasserstein metric. The latter is defined in terms of solutions to the optimal transport problem, first proposed by Monge in 1781 and then extended by Kantorovich in 1942; see [17] for more complete historical perspectives. We first introduce the concept of a Probability Vector Field (briefly PVF), which is a map assigning to every probability measure μ on $\mathbb{R}^n$ a probability measure $V[\mu]$ on $T\mathbb{R}^n$ (the tangent bundle), whose marginal on the base is μ itself. Simply put words, the fiber values of $V[\mu]$ provide the velocities along which the mass of μ is spread. Given a PVF V, the corresponding Measure Differential Equation (briefly MDE) reads $\dot{\mu} = V[\mu]$ and a solution is defined in the usual weak sense. This concept can be seen as a natural generalization of transport equations based on Differential Inclusions ([7]) and the approach of DiPerna–Lions ([9]) to kinetic equations, see Remark 1.

If V is sublinear (for the size of measures' support) and continuous with respect to the Wasserstein metrics on $\mathbb{R}^n$ and $T\mathbb{R}^n$, then we obtain a solution using approximation and compactness. More precisely, by discretizing in space, time and velocities we construct approximate solutions consisting of finite sums of Dirac deltas moving on a lattice of $\mathbb{R}^n$ which are called Lattice Approximate Solutions (briefly LASs.) LASs can be seen as generalizations of probabilistic Cellular Automata, defined using V.

To address continuous dependence from initial data, it is not enough to ask for Lipschitz continuity of V for the Wasserstein metrics. This is due to the fact that the fiber marginal of $V[\mu]$ has a meaning of an infinitesimal displacement, opposed to the base marginal. Therefore we introduce a different quantity, which computes the Wasserstein distance over the fiber restricted to transference plans which are optimal over the base, see (21). This allows as to obtain the existence of a Lipschitz semigroup of solutions, obtained as limit of LASs, from Lipschitz-type assumptions. Weak solutions to Cauchy Problems for MDEs are not expected to be unique, thus we address the question of uniqueness at the level of a semigroup. For this purpose, we introduce the concept of a Dirac germ, which consists of a small-time evolution for finite sums of Dirac deltas. Then we show uniqueness of a Lipschitz semigroup, compatible with a given Dirac germ. Therefore uniqueness questions can be addressed by looking for unique limits to LASs with finite sums of Dirac deltas as initial data.

We then explore various connections of MDEs with classical approaches. First, we show that MDEs represent a natural measure-theoretic generalization of Ordinary Differential Equations (briefly ODEs). An MDE is naturally associated to an ODE by moving masses along the ODE solutions. Lipschitz continuity of the ODE implies the existence of a Lipschitz semigroup for the corresponding MDE (which is the only one compatible with ODE solutions). The correspondence of ODEs-MDEs defines a map, which is a monoid morphism between the space of vector fields, endowed with the usual sum, and the space of PVFs, endowed with a fiber-convolution operation. Moreover, the map sends the multiplication by a scalar to the natural counterpart of scalar multiplication over the fiber, see Proposition 6.3.

MDEs can model both diffusion and concentration phenomena. We first show that a PVF V, which is constant on the fiber component, gives rise to a simple translation (because of the Law of Large Numbers.) On the other side, it is possible to define PVFs which depend on the global properties of the measures, providing finite speed diffusion. For MDEs representing concentration, uniqueness is obtained by one-sided Lipschitz-type conditions, mimicking the one-sided Lipschitz conditions for ODEs. Moreover, MDEs extend are theory of conservation laws with discontinuous fluxes.

Finally, kinetic models are considered. Dobrushin's approach ([10]) is recovered as a special case of MDEs in the sense that, given a multi-particle system whose dynamics given by ODEs, one can define a corresponding MDE under appropriate conditions (for example indistinguishibility of particles and uniform Lipschitz estimates). Moreover, the MDE enjoys well-posedness properties and compatibility with the empirical probability distributions defined by the multi-particle system.

The paper is organized as follows: in Section 2 we define PVFs, MDEs and solutions to MDEs. Then, in Section 3, we prove the existence of solutions to Cauchy Problems for MDEs under continuity assumption, and, in Section 4, the existence of a Lipschitz semigroup of solutions under appropriate Lipschitz-type assumptions. Uniqueness of Lipschitz semigroups is addressed in Section 5 using the concept of Dirac germ (Definition 5.1). The relationship of MDEs with ODEs is explored in Section 6, while examples of finite-speed diffusion and concentration phenomena are given in Section 7. Finally, results for mean-field limits of multi-particle systems, seen as special cases of MDEs, are provided in Section 8.

2. Basic Definitions

For simplicity we restrict ourselves to $\mathbb{R}^n$, but a local theory can be easily developed for manifolds admitting a partition of unity and, for every $R > 0$, $B(0, R)$ for the ball of radius R centered at the origin. We use the symbol $T\mathbb{R}^n$ for the tangent bundle of $\mathbb{R}^n$, and $\pi_1 : T\mathbb{R}^n \to \mathbb{R}^n$ for the projection to the base $\mathbb{R}^n$, that is $\pi_1(x, v) = x$. We use the identification $T\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$ and endow $T\mathbb{R}^n$ with the Euclidean norm of $\mathbb{R}^{2n}$. We also define $\pi_{13} : (T\mathbb{R}^n)^2 \mapsto (\mathbb{R}^n)^2$ by $\pi_{13}(x, v, y, w) = (x, y)$ (that is the projection on the bases for both components). For every $A \subset \mathbb{R}^n$, χ_A indicates the characteristic function of the set A and $\mathcal{C}_c^\infty(\mathbb{R}^n)$ indicates the space of smooth functions with compact support. Given a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ ($\mathbb{N}$ indicates natural numbers), we set $|\alpha| = \sum_i \alpha_i$, $\partial^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}$ and, for $k \in N$ and $A \subset \mathbb{R}^n$, we define $\|f\|_{\mathcal{C}^k(A)} = \sup_{x\in A, |\alpha|=k} |\partial^\alpha f(x)|$.

Given (X, d) Polish space (complete separable metric space), we indicate by $\mathcal{P}(X)$ the set of probability measures on X, that is positive Borel measures with total mass equal to one. Given $\mu \in \mathcal{P}(X)$ we indicate by $\mathrm{Supp}(\mu)$ its support and we define $\mathcal{P}_c(X)$ to be the set of probability measures with compact support. Given (X_1, d_1), (X_2, d_2) Polish spaces, $\mu \in \mathcal{P}(X_1)$ and $\phi : X_1 \to X_2$ measurable, we define the push forward $\phi\#\mu \in \mathcal{P}(X_2)$ by $\phi\#\mu(A) = \mu(\phi^{-1}(A)) = \mu(\{x \in X_1 : \phi(x) \in A\})$. Given $\mu \in \mathcal{P}(X_1)$ and $\nu_x \in \mathcal{P}(X_2)$, $x \in X_1$, we define $\mu \otimes \nu_x$ by $\int_{X_1\times X_2} \phi(x, v)\, \mathrm{d}(\mu \otimes \nu_x) = \int_{X_1} \int_{X_2} \phi(x, v)\mathrm{d}\nu_x(v)\, \mathrm{d}\mu(x)$.

Definition 2.1. A Probability Vector Field (briefly PVF) on $\mathcal{P}(\mathbb{R}^n)$ is a map $V : \mathcal{P}(\mathbb{R}^n) \to \mathcal{P}(T\mathbb{R}^n)$ such that $\pi_1\#V[\mu] = \mu$.

Given a PVF V, we define the corresponding Measure Differential Equation (MDE) by

$$\dot{\mu} = V[\mu]. \tag{1}$$

In simple words, $V[\mu]$ restricted to $T_x\mathbb{R}^n$ indicates the directions towards which the mass of μ at x is spread. For every $\mu_0 \in \mathcal{P}(\mathbb{R}^n)$ we define the Cauchy problem

$$\dot{\mu} = V[\mu], \qquad \mu(0) = \mu_0. \tag{2}$$

A solution to (2) in weak sense is defined as follows:

Definition 2.2. A solution to (2) is a map $\mu : [0, T] \to \mathcal{P}(\mathbb{R}^n)$ such that $\mu(0) = \mu_0$ and the following holds. For every $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$, the integral $\int_{T\mathbb{R}^n}(\nabla f(x) \cdot v)\, \mathrm{d}V[\mu(s)](x, v)$ is defined for almost every s, the map $s \to \int_{T\mathbb{R}^n}(\nabla f(x) \cdot v)\, \mathrm{d}V[\mu(s)](x, v)$ belongs to $L^1([0, T])$, and the map $t \to \int f \,\mathrm{d}\mu(t)$ is absolutely continuous and, for almost every $t \in [0, T]$, it satisfies

$$\frac{d}{\mathrm{d}t}\int_{\mathbb{R}^n} f(x)\, \mathrm{d}\mu(t)(x) = \int_{T\mathbb{R}^n} (\nabla f(x) \cdot v)\, \mathrm{d}V[\mu(t)](x, v). \tag{3}$$

Alternatively, we may ask the following condition to hold for every $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ and $t \in [0, T]$:

$$\int_{\mathbb{R}^n} f(x)\, \mathrm{d}\mu(t)(x) = \int_{\mathbb{R}^n} f(x)\, \mathrm{d}\mu_0(x) + \int_0^t \int_{T\mathbb{R}^n} (\nabla f(x) \cdot v)\, \mathrm{d}V[\mu(s)](x, v)\, \mathrm{d}s. \tag{4}$$

Remark 1. Given a Lipschitz vector field v, one can define an evolution of measures along the unique solutions of v. If Φ is the flow generated by v, this reads $\mu(t) = \Phi_t \# \mu(0)$, where $\mu(0)$ is the initial datum, and solves $\dot{\mu}(t) + L_v\mu(t) = 0$, where L_v is the Lie derivative along v so $L_v\mu = \nabla \cdot (v\mu)$. In the weak sense $\mu(\cdot)$ solves (3) with $V[\mu(t)] = \mu(t) \otimes \delta_{v(x)}$ (see also Section 6). This is the classical approach of DiPerna–Lions [9] to kinetic equations. One can generalize this approach by replacing $v(x)$ with a multifunction $V(x)$, see [7], or allowing $V(x)$ to be a probability measure on $T_x\mathbb{R}^n$. Definition 2.2 is a further generalization allowing V to depend on the global properties of the measure μ. Other approaches to define evolution for measures are present in the literature, most notably the Wasserstein gradient flows. In [1] (Sections 10.3 and 11.2) the authors exhibit λ-contracting semigroups (for measures with second moments) defined by subdifferentials of λ-convex functionals, the latter playing the role of PVFs.

Notice that (3) is a very weak concept of a solution, thus we do not expect to have a uniqueness, see Example 3 in Section 7.1. Even more, if we replace $V[\mu]$ with its average on the fiber the concept of solution does not change. More precisely, solutions to (1) coincide with solutions to the continuity equation $\partial_t\mu + \nabla \cdot (v\,\mu) = 0$ where $v(t, x)$ is the barycenter of $V[\mu(t)]$, that is $v(t, x) = \int_{\mathbb{R}^n} v\, \mathrm{d}\nu_{t,x}(v)$ where $V[\mu_t] = \mu(t) \otimes_x \nu_{t,x}$. The special case of constant $V[\cdot]$ is illustrated in Proposition 7.1. For instance, consider the PVF given by $\bar{V}[\mu] = \mu \otimes_x \frac{1}{2}(\delta_{-1} + \delta_1)$. It is easy to check that the constant solution $\mu(t) \equiv \delta_{x_0}$ solves the Cauchy problem (2) for $\mu_0 = \delta_{x_0}$. On the other side, Wasserstein gradient flows provide a more restrictive definition of solutions guaranteeing uniqueness and a refined characterization of the (right) time derivative of the semigroup, see Thorem 11.2.1 of [1].

To isolate unique solutions to MDEs we introduce the concept of Lattice Approximate Solutions (briefly LAS), see Definition 3.1, and look at uniqueness at the level of Lipschitz semigroup, see Theorem 5.2. For a good illustration of this phenomenon consider again $\bar{V}[\mu] = \mu \otimes_x \frac{1}{2}(\delta_{-1} + \delta_1)$ and V defined in Example 1 in Section 7.1. Solutions obtained as limit to LAS to (1) for $\bar{V}$ are constant in time, while those for V diffuse mass from the barycenter of the initial datum. However, for the Cauchy problem with $\mu_0 = \delta_0$ the constant solutions for $\bar{V}$ are also solutions for V. Notice that LAS are defined for simplicity considering the lattice $\mathbb{Z}^n/(N^2)$, $N \in \mathbb{N}$, however the main results still hold for other lattices, for instance obtained changing origin or frame.

3. Existence of Solutions to Cauchy Problems for MDEs

For simplicity we will focus on the set $\mathcal{P}_c(\mathbb{R}^n)$ of probability measures with compact support, but other sets with compactness properties may be used, for instance based on bounds on the moments. First we need to introduce some concepts from optimal transport theory. We refer the reader to [1,15,17,18] for a complete perspective.

Given (X, d) Polish space and given $\mu, \nu \in \mathcal{P}(X)$ we indicate by $P(\mu, \nu)$ the set of transference plans from μ to ν, that is the set of probability measures on $X \times X$ with marginals equal to μ and ν respectively. Given $\tau \in P(\mu, \nu)$, let $J(\tau)$ be its transportation cost

$$J(\tau) = \int_{X^2} \mathrm{d}(x, y)\, \mathrm{d}\tau(x, y).$$

The Monge–Kantorovich or optimal transport problem amounts to find τ that minimizes $J(\tau)$ and the Wasserstein metric is defined by

$$W^X(\mu, \nu) = \inf_{\tau \in P(\mu,\nu)} J(\tau).$$

For simplicity of notation we drop the superscript if $X = \mathbb{R}^n$.

Remark 2. In general one can define the family of Wasserstein metrics W_p, $p \geqq 1$, by setting

$$J(\tau) = \left(\int_{X^2} \mathrm{d}(x, y)^p\, \mathrm{d}\tau(x, y) \right)^{\frac{1}{p}}.$$

Here we focus, for simplicity, on the case $p = 1$ but will provide comments for the general case.

We indicate by $P^{\mathrm{opt}}(\mu, \nu)$ the set of optimal transference plans, that is minimizing $J(\tau)$. If μ, ν have finite first moments, that is $\int |x|\, \mathrm{d}\mu(x) < +\infty$, $\int |x|\, \mathrm{d}\nu(x) < +\infty$, then $P^{\mathrm{opt}}(\mu, \nu)$ is not empty. Thus this holds true if μ, ν have compact support. We always endow $\mathcal{P}(X)$ with the Wasserstein metric and the relative topology. Let us recall the Kantorovich–Rubinstein duality

$$W^X(\mu, \nu) = \sup \left\{ \int_X f\, \mathrm{d}(\mu - \nu) \ : f : X \to \mathbb{R},\ Lip(f) \leqq 1 \right\}, \tag{5}$$

where $Lip(f)$ indicates the Lipschitz constant of f. We have the following:

Lemma 3.1. *Consider a sequence* $\mu_N \subset \mathcal{P}_c(\mathbb{R}^n)$ *and assume there exists* $R > 0$ *such that* $\mathrm{Supp}(\mu_N) \subset B(0, R)$ *for every* N. *Then there exists* $\mu \in \mathcal{P}_c(\mathbb{R}^n)$ *and a subsequence, still indicated by* μ_N, *such that* $W(\mu_N, \mu) \to 0$.

The proof of the Lemma 3.1 is standard and we postpone it to the Appendix.

Our assumptions to prove the existence of solutions are the following:

(H:bound) V is support sublinear, that is there exists $C > 0$ such that for every $\mu \in \mathcal{P}_c(X)$ it holds that

$$\sup_{(x,v)\in \mathrm{Supp}(V[\mu])} |v| \leqq C \left(1 + \sup_{x\in \mathrm{Supp}(\mu)} |x|\right).$$

(H:cont) Given $R > 0$ denote by $\mathcal{P}_c^R(T\mathbb{R}^n)$ the set of probability measures with support contained in $B(0, R)$. For every $R > 0$ the map $V : \mathcal{P}_c^R(\mathbb{R}^n) \to \mathcal{P}_c(T\mathbb{R}^n)$ (restriction of V) is continuous (for the topology given by the Wasserstein metrics $W^{\mathbb{R}^n}$ and $W^{T\mathbb{R}^n}$).

Remark 3. Notice that on the set $\mathcal{P}_c^R(T\mathbb{R}^n)$ all Wasserstein metrics W_p, $p \geqq 1$, gives rise to the same topology (weak topology), thus condition (H:cont) is equivalent to continuity with respect to any of the W_p.

A stronger continuity condition would be asking for the map $V : \mathcal{P}_c(\mathbb{R}^n) \to \mathcal{P}_c(T\mathbb{R}^n)$ to be continuous (for the topology given by the Wasserstein metrics $W^{\mathbb{R}^n}$ and $W^{T\mathbb{R}^n}$). This would require as to consider sequences of measures with mass going to infinity.

To prove the existence of solutions to a Cauchy problem (2), we define a sequence of approximate solutions using a scheme of Euler type. We first introduce some more notation.

For $N \in \mathbb{N}$ let $\Delta_N = \frac{1}{N}$ be the time step size, $\Delta_N^v = \frac{1}{N}$ the velocity step size and $\Delta_N^x = \Delta_N^v \Delta_N = \frac{1}{N^2}$ the space step size. We also define x_i to be the $(2N^3 + 1)^n$ equispaced discretization points of $\mathbb{Z}^n/(N^2) \cap [-N, N]^n$ and v_j to be the $(2N^2 + 1)^n$ equispaced discretization points of $\mathbb{Z}^n/N \cap [-N, N]^n$. Given $\mu \in \mathcal{P}_c(\mathbb{R}^n)$ we define the following operator providing an approximation by finite sums of Dirac deltas:

$$\mathcal{A}_N^x(\mu) = \sum_i m_i^x(\mu)\delta_{x_i}, \tag{6}$$

where

$$m_i^x(\mu) = \mu(x_i + Q) \tag{7}$$

with $Q = ([0, \frac{1}{N^2}[)^n$. Similarly, given $\mu \in \mathcal{P}_c(\mathbb{R}^n)$, whose support is contained in the set $\mathbb{Z}^n/(N^2) \cap [-N, N]^n$, we set

$$\mathcal{A}_N^v(V[\mu]) = \sum_i \sum_j m_{ij}^v(V[\mu])\, \delta_{(x_i, v_j)}, \tag{8}$$

where

$$m_{ij}^{v}(V[\mu]) = V[\mu](\{(x_i, v) : v \in v_j + Q'\}), \tag{9}$$

with $Q' = ([0, \frac{1}{N}[)^n$. For every $\mu \in \mathcal{P}_c(\mathbb{R}^n)$ there exists N such that $\mathrm{Supp}(\mu_0) \subset [-N, N]^n$, thus from the definition of $\mathcal{A}_N^x$ and $\mathcal{A}_N^v$, we easily get

Lemma 3.2. *Given $\mu \in \mathcal{P}_c(\mathbb{R}^n)$, for N sufficiently big the following holds:*

$$W(\mathcal{A}_N^x(\mu), \mu) \leqq \sqrt{n}\,\Delta_N^x, \qquad W^{T\mathbb{R}^n}(\mathcal{A}_N^v(V[\mu]), V[\mu]) \leqq \sqrt{n}\,\Delta_N^v.$$

We are now ready to define a sequence of approximate solutions.

Definition 3.1. Consider V satisfying (H:bound). Given the Cauchy Problem (2), $T > 0$ and $N \in \mathbb{N}$ such that $e^{C_N T}(R_N + 1) < N$ (see Lemma 3.3 for definition of C_N and R_N), we define the Lattice Approximate Solution (LAS) $\mu^N : [0, T] \to \mathcal{P}_c(\mathbb{R}^n)$ as follows:

Recalling (6)–(9), we set $\mu_0^N = \mathcal{A}_N^x(\mu_0)$ and, by recursion, define

$$\mu_{\ell+1}^N = \mu^N((\ell+1)\Delta_N) = \sum_i \sum_j m_{ij}^v(V[\mu^N(\ell\Delta_N)])\,\delta_{x_i + \Delta_N v_j}. \tag{10}$$

By definition of Δ_N, Δ_N^v, Δ_N^x and (10), $\mathrm{Supp}(\mu_\ell^N)$ is contained in the set $\mathbb{Z}^n/(N^2) \cap [-N, N]^n$, thus we can write $\mu_\ell^N = \sum_i m_i^{N,\ell} \delta_{x_i}$ for some $m_i^{N,\ell} \geqq 0$. Finally μ^N is defined for all times by time-interpolation:

$$\mu^N(\ell\Delta_N + t) = \sum_{ij} m_{ij}^v(V[\mu^N(\ell\Delta_N)])\,\delta_{x_i + t\, v_j}. \tag{11}$$

In other words, to define $\mu_{\ell+1}^N$, we approximate $V[\mu_\ell^N]$ by $\mathcal{A}_N^v(V[\mu_\ell^N])$ and use the corresponding velocities to move the Dirac deltas of μ_ℓ^N. Fix now $T > 0$. Because of assumption (H:bound), the support of μ_ℓ^N keeps uniformly bounded on the time interval $[0, T]$, as detailed in the next Lemma.

Lemma 3.3. *Given a PVF V satisfying (H:bound), μ_0 with* $\mathrm{Supp}(\mu_0) \subset B(0, R)$ *and ℓ such that $\ell\Delta_N \leqq T$, the following holds true:*

$$\mathrm{Supp}(\mu_\ell^N) \subset B\left(0, e^{C_N T}(R_N + 1) - 1\right), \tag{12}$$

where $C_N = C + \frac{\sqrt{n}}{N}$ and $R_N = R + \frac{\sqrt{n}}{N^2}$.

Proof. Set $s_i^N = \sup_{x \in \mathrm{Supp}(\mu_i^N)} |x|$. Since the support of μ_0^N is contained in the ball $B(0, R)$ by definition of $\mathcal{A}_N^x$, we have $s_0^N \leqq R_N$. Notice that $\mathcal{A}_N^v(V[\cdot])$ satisfies (H:bound) with C replaced by C_N. Then $s_{i+1}^N \leqq C_N \Delta_N + (1 + C_N \Delta_N) s_i^N$, thus $s_\ell^N \leqq C_N \Delta_N \left(\sum_{i=0}^{\ell-1} (1 + C_N\Delta_N)^i\right) + (1 + C_N \Delta_N)^\ell s_0^N$. The first addendum is estimated by the integral $\int_0^{\ell\Delta_N} C_N e^{C_N \tau} \mathrm{d}\tau$, while the second with $e^{C_N T} s_0^N$ thus we conclude. □

Because of Lemma 3.3, the definition of $V[\cdot]$ and the choice of N in Definition 3.1, we have $\sum_j m^v_{ij}(V[\mu^N_\ell]) = m^{N,\ell}_i$, thus the mass is conserved. Notice that the support of $\mu^N(t)$ is not, in general, contained in $\mathbb{Z}^n/(N^2) \cap [-N, N]^n$. We can now state the main result of this Section.

Theorem 3.1. *Given a PVF V satisfying (H:bound) and (H:cont), for every $T > 0$ and $\mu_0 \in \mathcal{P}_c(\mathbb{R}^n)$ there exists a solution $\mu : [0, T] \to \mathcal{P}_c(\mathbb{R}^n)$ to the Cauchy problem (2) obtained as uniform-in-time limit of LASs for the Wasserstein metric.*

Moreover, if $\mathrm{Supp}(\mu_0) \subset B(0, R)$*, then*

$$W(\mu(t), \mu(s)) \leqq C\, e^{CT}\, (R + 1)\, |t - s|. \tag{13}$$

Proof. We have

$$W(\mu^N_{\ell+1}, \mu^N_\ell) = W\left(\sum_{i,j} m^v_{ij}(V[\mu^N_\ell])\, \delta_{x_i+\Delta_N v_j}, \sum_i m^{N,\ell}_i \delta_{x_i}\right).$$

Since $\sum_j m^v_{ij}(V[\mu^N_\ell]) = m^{N,\ell}_i$, we can define a transference plan from μ^N_ℓ to $\mu^N_{\ell+1}$ by moving the mass $m^{N,\ell}_i \delta_{x_i}$ to $\sum_j m^v_{ij}(V[\mu^N_\ell])\, \delta_{x_i+\Delta_N v_j}$. Thus we obtain

$$W(\mu^N_{\ell+1}, \mu^N_\ell) \leqq \sum_{i,j} m^v_{ij}(V[\mu^N_\ell])\, |x_i + \Delta_N\, v_j - x_i| = \Delta_N \sum_{i,j} m^v_{ij}(V[\mu^N_\ell])\, |v_j|.$$

Let $R > 0$ be such that $\mathrm{Supp}(\mu_0) \subset B(0, R)$. Using Lemma 3.3 and (H:bound), we deduce that $m^v_{ij}(V[\mu^N_\ell]) \neq 0$ only if $|v_j| \leqq C_N\, e^{C_N T}\, (R_N + 1)$. Thus we get

$$W(\mu^N_{\ell+1}, \mu^N_\ell) \leqq C_N\, e^{C_N T}\, (R_N + 1)\, \Delta_N.$$

Repeating the same reasoning for $\mu^N(t)$ (see (11)), we get

$$W(\mu^N(t), \mu^N(s)) \leqq C_N\, e^{C_N T}\, (R_N + 1)\, |t - s|. \tag{14}$$

Therefore, the sequence $\mu^N : [0, T] \to \mathcal{P}_c(\mathbb{R}^n)$ is uniformly Lipschiz for the Wasserstein metric. By Ascoli-Arzelá Theorem, there exists a subsequence, still indicated by μ^N, which converges uniformly to a Lipschitz curve $\mu : [0, T] \to \mathcal{P}_c(\mathbb{R}^n)$. Since $C_N \to C$ and $R_N \to R$, we have that $\mu(\cdot)$ satisfies (13).

We now prove that the limit $\mu(t)$ satisfies (4). Set $m^{N,\ell}_{ij} = m^v_{ij}(V[\mu^N_\ell])$, thus $\sum_j m^{N,\ell}_{ij} = m^{N,\ell}_i$. Given $f \in \mathcal{C}^\infty_c(\mathbb{R}^n)$ and $\bar{\ell}$, we compute

$$\begin{aligned}
\int_{\mathbb{R}^n} f\, \mathrm{d}(\mu^N_{\bar{\ell}} - \mu^N_0) &= \int_{\mathbb{R}^n} f\, \mathrm{d}\left(\sum_{\ell=0}^{\bar{\ell}-1} \left(\mu^N_{\ell+1} - \mu^N_\ell\right)\right) \\
&= \sum_\ell \int_{\mathbb{R}^n} f\, \mathrm{d}\left(\sum_{ij} m^{N,\ell}_{ij}\, \delta_{x_i+\Delta_N v_j} - \sum_i m^{N,\ell}_i\, \delta_{x_i}\right) \\
&= \sum_\ell \sum_{ij} m^{N,\ell}_{ij} (f(x_i + \Delta_N\, v_j) - f(x_i))
\end{aligned}$$

$$= \sum_\ell \sum_{ij} m_{ij}^{N,\ell} \Big[\Delta_N \, (\nabla f(x_i) \cdot v_j) + \|f\|_{\mathcal{C}^2(B(0,C'))} O(\Delta_N^2) \Big],$$

where $C' = e^{CT}(R+1)$, so that $\mathrm{Supp}(\mu_\ell^N) \subset B(0, C')$, and $\sup_N O(\Delta_N^k)\, N^k < +\infty$, $k \in \mathbb{N}$, thus

$$= \sum_\ell \int_{\ell\Delta_N}^{(\ell+1)\Delta_N} \int_{T\mathbb{R}^n} (\nabla f(x) \cdot v) \, \mathrm{d}(\mathcal{A}_N^v(V[\mu_\ell^N]))(x, v) \, \mathrm{d}t + O(\Delta_N).$$

Notice that $\mathrm{Supp}(V[\mu_\ell^N]) \subset B(0, C') \times B(0, C(1 + C'))$. If $x, y \in B(0, C')$ and $v, w \in B(0, C(1 + C'))$, we can estimate using triangular and Cauchy-Schwarz inequality:

$$|\nabla f(x) \cdot v - \nabla f(y) \cdot w| \leqq |\nabla f(x) - \nabla f(y)| \, |v| + |\nabla f(y)| \, |v - w|.$$

For the first addendum by Mean Value Theorem we have: $|(\nabla f)_i(x) - (\nabla f)_i(y)| \leqq |\nabla(\nabla f)_i((1-\lambda_i)x + \lambda_i y)| \, |x - y|$ for some $\lambda_i \in [0, 1]$, thus $|\nabla f(x) - \nabla f(y)| \leqq \sqrt{n}\, \sup_i |\nabla(\nabla f)_i((1-\lambda_i)x + \lambda_i y)| \, |x - y| \leqq n\, |x - y| \, \|f\|_{\mathcal{C}^2(B(0,C'))}$. Finally, we get

$$n\|f\|_{\mathcal{C}^2(B(0,C'))} \, |x - y| \, C(1 + C') + \sqrt{n}\|f\|_{\mathcal{C}^1(B(0,C'))} \, |v - w| \leqq L(f)|(x, v) - (y, w)|,$$

where $L(f) = \sqrt{2}\max\{n\|f\|_{\mathcal{C}^2(B(0,C'))}\, C(1+C'), \sqrt{n}\|f\|_{\mathcal{C}^1(B(0,C'))}\}$. Define $g = g(x, v)$ to be the function $\nabla f(x) \cdot v$ restricted to the set $B(0, C') \times B(0, C(1+C'))$, thus $L(f)$ is a Lipschitz constant for the function g. Using the Kirszbraun theorem (see [16] Thorem 1.31), we can extend g to the whole tangent bundle $T\mathbb{R}^n$ with the same Lipschitz constant $L(f)$. Set $\xi = \chi_{B(0,C') \times B(0,C(1+C'))}$ (the characteristic function of the set $B(0, C') \times B(0, C(1 + C'))$, then using the Kantorovich–Rubinstein duality (5) for $X = T\mathbb{R}^n$ and Lemma 3.2, we get

$$\begin{aligned} L(f) &\left| \int_{T\mathbb{R}^n} \frac{\nabla f(x) \cdot v}{L(f)} \, \mathrm{d}(\mathcal{A}_N^v(V[\mu_\ell^N]) - V[\mu_\ell^N])(x, v) \right| \\ &= L(f) \left| \int_{T\mathbb{R}^n} \frac{(\nabla f(x) \cdot v)\xi(x, v)}{L(f)} \, \mathrm{d}(\mathcal{A}_N^v(V[\mu_\ell^N]) - V[\mu_\ell^N])(x, v) \right| \\ &= L(f) \left| \int_{T\mathbb{R}^n} \frac{g(x, v)}{L(f)} \, \mathrm{d}(\mathcal{A}_N^v(V[\mu_\ell^N]) - V[\mu_\ell^N])(x, v) \right| \\ &\leqq L(f)\, W^{T\mathbb{R}^n}(\mathcal{A}_N^v(V[\mu_\ell^N]), V[\mu_\ell^N]) \leqq \sqrt{n}\, L(f)\, \Delta_N^v. \end{aligned} \tag{15}$$

For every $t \in [0, T]$ and $N \in \mathbb{N}$, set $\ell_N(t) = \lfloor \frac{t}{N} \rfloor$, where $\lfloor s \rfloor = \max\{n \in \mathbb{N} : n \leqq s\}$. Thus we have proved

$$\begin{aligned} &\int_{\mathbb{R}^n} f \, \mathrm{d}(\mu_{\ell_N(t)}^N - \mu_0^N) \\ &\quad = \int_0^{\ell_N(t)\Delta_N} \int_{T\mathbb{R}^n} (\nabla f(x) \cdot v) \, \mathrm{d}(V[\mu_{\ell_N(s)}^N])(x, v) \, \mathrm{d}s + O(\Delta_N). \end{aligned} \tag{16}$$

We will now pass to the limit in the left and right-hand side of equation (16). Since μ^N converges uniformly-in-time to μ for the Wasserstein metric, there exists ϵ_N, with $\epsilon_N \to 0$ as $N \to +\infty$, such that $W(\mu_N(t), \mu(t)) \leqq \epsilon_N$. Then, for every $t \in [0, T]$, one has

$$W(\mu^N_{\ell_N(t)}, \mu(t)) \leqq W(\mu^N_{\ell_N(t)}, \mu^N(t)) + W(\mu_N(t), \mu(t)) \leqq O(\Delta_N) + \epsilon_N, \quad (17)$$

where we used (14) and $|\ell_N(t) - t| \leqq \Delta_N$ to estimate the first term. From (17) we get

$$\begin{aligned} \left| \int_{\mathbb{R}^n} f \, \mathrm{d}(\mu^N_{\ell_N(t)} - \mu(t)) \right| &\leqq \sqrt{n} \|f\|_{\mathcal{C}^0(B(0,C'))} W(\mu^N_{\ell_N(t)}, \mu(t)) \\ &\leqq O(\Delta_N) + \sqrt{n} \|f\|_{\mathcal{C}^0(B(0,C'))} \epsilon_N, \end{aligned} \quad (18)$$

while, from Lemma 3.2,

$$\left| \int_{\mathbb{R}^n} f \, \mathrm{d}(\mu^N_0 - \mu_0) \right| = O(\Delta_N). \quad (19)$$

Set $\psi_N(s) = \int_{T\mathbb{R}^n} (\nabla f(x) \cdot v) \, \mathrm{d}(V[\mu^N_{\ell_N(s)}] - V[\mu(s)])$, then using the Kantorovich–Rubinstein duality (5) as in (15), we have

$$|\psi_N(s)| \leqq \sqrt{n} \, L(f) W^{T\mathbb{R}^n}(V[\mu^N_{\ell_N(s)}], V[\mu(s)]).$$

From (17) and (H:cont), we have that $W^{T\mathbb{R}^n}(V[\mu^N_{\ell_N(s)}], V[\mu(s)])$ converges pointwise to zero for $s \in [0, T]$ as $N \to +\infty$, thus the same is true for $\psi_N(s)$. From Lemma 3.3 and (H:bound), we have that $\psi_N(s)$ is uniformly bounded for $N \in \mathbb{N}$ and $s \in [0, T]$, thus by Lebesgue dominated convergence, ψ_N converges to 0 in $L^1(]0, T[)$. Then, again using Lemma 3.3 and (H:bound), we deduce that for every $t \in [0, T]$,

$$\begin{aligned} \lim_{N \to +\infty} \Bigg| \int_0^{\ell_N(t)\Delta_N} & \int_{T\mathbb{R}^n} (\nabla f(x) \cdot v) \, \mathrm{d}(V[\mu^N_{\ell_N(s)}]) \, \mathrm{d}s \\ & - \int_0^t \int_{T\mathbb{R}^n} (\nabla f(x) \cdot v) \; V[\mu(s)]) \, \mathrm{d}s \Bigg| = 0. \end{aligned} \quad (20)$$

Passing to the limit in (16), using (18)–(20), we conclude. □

4. Lipschitz Semigroup of Solutions to MDEs

We now investigate continuous dependence from initial data. More precisely, we provide a new condition ensuring the existence of a Lipschitz semigroup of solutions obtained as limit of LASs.

Notice that $V[\mu]$ is supported on $T\mathbb{R}^n$ but the two components (x, v) have different meanings, indeed v represents a tangent vector thus an infinitesimal displacement. For this reason, instead of using $W^{T\mathbb{R}^n}$, we are going to introduce another concept to measure distances among elements of $\mathcal{P}(T\mathbb{R}^N)$.

Definition 4.1. Consider $V_i \in \mathcal{P}_c(T\mathbb{R}^n)$, $i = 1, 2$, and denote by μ_i the marginal over the base, that is $\pi_1 \# V_i = \mu_i$. We define the following quantity:

$$\mathcal{W}(V_1, V_2) = \inf \left\{ \int_{T\mathbb{R}^n \times T\mathbb{R}^n} |v - w| \, \mathrm{d}T(x, v, y, w) \ : \right. \\ \left. T \in P(V_1, V_2),\ \pi_{13}\#T \in P^{\mathrm{opt}}(\mu_1, \mu_2) \right\}. \tag{21}$$

The condition $\pi_{13}\#T \in P^{\mathrm{opt}}(\mu_1, \mu_2)$ tells us that T acts optimally on the base transporting μ_1 to μ_2. Therefore $\mathcal{W}$ gives the optimal transport distance of the fiber components based on optimal ways to transport the marginals on the base.

Remark 4. Notice that $\mathcal{W}$ is not a metric since it can vanish for distinct elements of $\mathcal{P}(T\mathbb{R}^N)$. It would be tempting to add the term $|x - y|$ to the integrand in (21) (or a norm of (x, v, y, w)) but we would not obtain a metric, because the triangular inequality does not hold. A simple example is obtained by setting in $T\mathbb{R}^2$: $V_1 = \frac{1}{2}\delta_{((0,0),(1,0))} + \frac{1}{2}\delta_{((1,0),(3,0))}$, $V_2 = \frac{1}{2}\delta_{((0,1),(1,0))} + \frac{1}{2}\delta_{((1,-1),(3,0))}$ and $V_3 = \frac{1}{2}\delta_{((1,1),(1,0))} + \frac{1}{2}\delta_{((0,-1),(3,0))}$. There exists a unique optimal transference plan from $\mu_1 = \pi_1\#V_1$ to $\mu_2 = \pi_1\#V_2$, moving the mass from $(0, 0)$ to $(0, 1)$ and from $(1, 0)$ to $(1, -1)$, a unique optimal transference plan from μ_2 to $\mu_3 = \pi_1\#V_3$, moving the mass from $(0, 1)$ to $(1, 1)$ and from $(1, -1)$ to $(0, -1)$, and a unique optimal transference plan from μ_1 to μ_3, moving the mass from $(0, 0)$ to $(0, -1)$ and from $(1, 0)$ to $(1, 1)$. Then, the set $\mathcal{T}_{12} = \{T \in P(V_1, V_2) : \pi_{13}\#T \in P^{opt}(\mu_1, \mu_2)\}$ has a unique element and $\inf_{T\in\mathcal{T}_{12}} \int(|x - y| + |v - w|)\mathrm{d}T = 1$. Defining similarly $\mathcal{T}_{23}$ and $\mathcal{T}_{13}$, we get $\inf_{T\in\mathcal{T}_{23}} \int(|x - y| + |v - w|)\mathrm{d}T = 1$ and $\inf_{T\in\mathcal{T}_{13}} \int(|x - y| + |v - w|)\mathrm{d}T = 3$.

We are now ready to state a new assumption, which is a local Lipschitz-type condition on the map $\mu \to V[\mu]$ for $\mathcal{W}$. We require that

(H:lip) for every $R > 0$ there exists $K = K(R) > 0$ such that if Supp(μ), Supp $(\nu) \subset B(0, R)$ then

$$\mathcal{W}(V[\mu], V[\nu]) \leqq K\ W(\mu, \nu). \tag{22}$$

Remark 5. We notice that also here we could use any of the Wasserstein distances W_p, $p \geqq 1$, as follows. First define $\mathcal{W}_p$ using the term $(\int_{T\mathbb{R}^n \times T\mathbb{R}^n} |v - w|^p \ \mathrm{d}T(x, v, y, w))^{1/p}$ in (21). Then the proof of Theorem 4.1 is modified as follows. In (29) we estimate $W_p^p(\mu^N_{\ell+1}, \nu^N_{\ell+1})$. Notice that for $a, b \in \mathbb{R}^n$, we have $\frac{d}{\mathrm{d}\epsilon}|a + \epsilon b|^p\big|_{\epsilon=0} = p\,|a|^{p-1}\frac{\langle b,a\rangle}{|a|}$. Therefore we can write $|(x + \Delta_N\, v) - (y + \Delta_N\, w)| = |x - y|^p + \Delta_N\, p|x - y|^{p-2}\langle x - y, v - w\rangle + o(\Delta_N) \leqq |x - y|^p + \Delta_N\, p|x - y|^{p-1}|v - w| + o(\Delta_N)$ (where we used Cauchy-Schwarz inequality). We define I_1 and I_2 in the same way and for I_2 we use Hölder inequality with exponents p and $(p - 1)/p$.

The quantity $\mathcal{W}$ in general can not compare to $W^{T\mathbb{R}^n}$, which weights in the same way the base and the fiber. However, we have the following:

Lemma 4.1. *Given μ, $\nu \in \mathcal{P}(\mathbb{R}^n)$, it holds that*

$$W^{T\mathbb{R}^n}(V[\mu], V[\nu]) \leqq \mathcal{W}(V[\mu], V[\nu]) + W(\mu, \nu). \tag{23}$$

In particular, (H:lip) implies the local Lipschitz continuity of V with respect to $W^{T\mathbb{R}^n}$.

Proof. By definition, we have

$$\begin{aligned}
&W^{T\mathbb{R}^n}(V[\mu], V[\nu]) \\
&= \inf_{T\in P(V[\mu],V[\nu])} \int_{(T\mathbb{R}^n)^2} |(x, v) - (y, w)| \, \mathrm{d}T(x, v, y, w) \\
&\leqq \inf_{T\in P(V[\mu],V[\nu])} \int_{(T\mathbb{R}^n)^2} (|x - y| + |v - w|) \, \mathrm{d}T(x, v, y, w) \\
&\leqq \inf_{T\in P(V[\mu],V[\nu]),\pi_{13}\#T\in P^{opt}(\mu,\nu)} \int_{(T\mathbb{R}^n)^2} (|x - y| + |v - w|) \, \mathrm{d}T(x, v, y, w) \\
&= W(\mu, \nu) + \mathcal{W}(V[\mu], V[\nu]).
\end{aligned}$$

□

Remark 6. The converse of Lemma 4.1 does not hold true, since $\mathcal{W}$ can not be estimated in terms of W and $W^{T\mathbb{R}^n}$. To see this consider $n = 1$, define $\varphi(x) = \sin(\frac{1}{x})$ for $x \neq 0$ and $\varphi(0) = 0$, and set $V[\mu] = \mu \otimes_x \delta_{\varphi(x)}$. Consider $\mu = \frac{1}{2}(\delta_{2k\pi+\frac{1}{2}\pi} + \delta_{k\pi+\frac{3}{2}\pi})$ and $\mu = \frac{1}{2}(\delta_{2k\pi+\frac{3}{2}\pi} + \delta_{k\pi+\frac{1}{2}\pi})$ with k large enough. The only optimal transference plan between μ and μ' moves $\delta_{2k\pi+\frac{1}{2}\pi}$ to $\delta_{2k\pi+\frac{3}{2}\pi}$ and one can easily compute $W(\mu, \mu') \leqq \frac{C}{k^2}$ for some $C > 0$. On the other side, the optimal transference plan on $T\mathbb{R}^n$ between $V[\mu]$ and $V'[\mu]$ sends $\delta_{(2k\pi+\frac{1}{2}\pi,1)}$ to $\delta_{(k\pi+\frac{1}{2}\pi}, 1)$ thus we get $W^{T\mathbb{R}^n}(V[\mu], V[\mu']) \leqq \frac{C}{k}$ for some $C > 0$. Finally $\mathcal{W}$ is computed using the only plan that projects onto the optimal transference plan between μ and μ', thus we get $\mathcal{W}(V[\mu], V[\mu']) = 2$.

We are now ready to prove the existence of semigroups of solutions. First we give the following:

Definition 4.2. Let V be a PVF satisfying (H:bound) and $T > 0$. A Lipschitz semigroup for (1) is a map $S : [0, T] \times \mathcal{P}_c(\mathbb{R}^n) \to \mathcal{P}_c(\mathbb{R}^n)$ such that for every $\mu, \nu \in \mathcal{P}_c(\mathbb{R}^n)$ and $t, s \in [0, T]$ the following holds:

(i) $S_0\mu = \mu$ and $S_t S_s \mu = S_{t+s} \mu$;
(ii) the map $t \mapsto S_t\mu$ is a solution to (1);
(iii) for every $R > 0$ there exists $C(R) > 0$ such that if $\mathrm{Supp}(\mu), \mathrm{Supp}(\nu) \subset B(0, R)$, then

$$\mathrm{Supp}(S_t\mu) \subset B(0, e^{Ct}(R + 1)), \tag{24}$$
$$W(S_t\mu, S_t\nu) \leqq e^{C(R)t} W(\mu, \nu), \tag{25}$$
$$W(S_t\mu, S_s\mu) \leqq C(R) \, |t - s|. \tag{26}$$

In Definition 4.2 it is reasonable to ask the constant $C(R)$ to be uniform for supports contained in $B(0, R)$ because the semigroup is defined on the compact interval $[0, T]$ and because of the uniform bounds provided by Lemma 3.3.

Next, Theorem provides the existence of a Lipschitz semigroup of solutions to an MDE, obtained via limit of LAS.

Theorem 4.1. *Given V satisfying (H:bound) and (H:lip), and $T > 0$, there exists a Lipschitz semigroup of solutions to* (1)*, obtained passing to the limit in LASs.*

Proof. We first prove Wasserstein estimates on LASs for different initial data. Fix $\mu_0, \nu_0 \in \mathcal{P}_c(\mathbb{R}^n)$ and call μ^N, respectively ν^N, the LAS defined using μ_0, respectively ν_0, as initial datum. First, from Lemma 3.2 we get

$$W(\mu_0^N, \nu_0^N) = W(\mathcal{A}_N^x(\mu_0), \mathcal{A}_N^x(\nu_0)) \leqq W(\mu_0, \nu_0) + 2\,\Delta_N^x. \tag{27}$$

Let us now estimate the Wasserstein distance between μ_ℓ^N and ν_ℓ^N by recursion:

$$W(\mu_{\ell+1}^N, \nu_{\ell+1}^N) = W\left(\sum_{i,j} m_{ij}^v(V[\mu_\ell^N])\,\delta_{x_i+\Delta_N v_j}, \sum_{i.j} m_{ij}^v(V[\nu_\ell^N])\,\delta_{x_i+\Delta_N v_j}\right).$$

Let $R > 0$ be such that $\mathrm{Supp}(\mu_0), \mathrm{Supp}(\nu_0) \subset B(0, R)$. By Lemma 3.3, the supports of μ^N and ν^N are uniformly contained in $B(0, e^{C_N T}(R+1))$, thus, by assumption (H:lip), there exists $K = K(R)$, depending on R, such that for N sufficiently big,

$$\mathcal{W}(V[\mu_\ell^N], V[\nu_\ell^N]) \leqq K\,W(\mu_\ell^N, \nu_\ell^N),$$

thus there exists $T \in P(V[\mu_\ell^N], V[\nu_\ell^N])$ such that

$$\int_{T\mathbb{R}^n \times T\mathbb{R}^n} |v - w|\,\mathrm{d}T(x, v, y, w) \leqq K\;W(\mu_\ell^N, \nu_\ell^N) + \Delta_N, \tag{28}$$

and $\pi_{13}\#T \in P^{opt}(\mu_\ell^N, \nu_\ell^N)$. We will now construct a transference plan from $\mu_{\ell+1}^N$ to $\nu_{\ell+1}^N$ by moving masses using the plan T. More precisely, define $\tau_{ij} \in \mathcal{P}_c((\mathbb{R}^n)^2)$ by $\tau_{ij}(A, B) = T(\{(x_i, v, x_j, w) : x_i + \Delta_N v \in A, x_j + \Delta_N w \in B\})$. In other words if T moves a mass from $\delta_{(x_i,v)}$ to $\delta_{(x_j,w)}$ then τ_{ij} moves the same mass from $\delta_{x_i+\Delta_N v}$ to $\delta_{x_j+\Delta_N w}$. We then define $\tau = \sum_{ij} \tau_{ij} \in P(\mu_{\ell+1}^N, \nu_{\ell+1}^N)$. Notice that setting $Y(x, v; y, w) := (x + \Delta_N v, y + \Delta_N w)$ it holds $\tau = Y\#T$. We have the following:

$$\begin{aligned}
&W(\mu_{\ell+1}^N, \nu_{\ell+1}^N)\\
&\quad\leqq \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|\,\mathrm{d}\tau(x, y)\\
&\quad= \int_{T\mathbb{R}^n \times T\mathbb{R}^n} |(x + \Delta_N v) - (y + \Delta_N w)|\,\mathrm{d}T(x, v, y, w)\\
&\quad\leqq \int_{T\mathbb{R}^n \times T\mathbb{R}^n} |x - y|\,\mathrm{d}T(x, v, y, w)
\end{aligned}$$

$$+\int_{T\mathbb{R}^n\times T\mathbb{R}^n} \Delta_N\,|v-w|\,\mathrm{d}T(x,v,y,w) \doteq I_1 + I_2. \tag{29}$$

Since $\pi_{13}\#T \in P^{opt}(\mu_\ell^N, \nu_\ell^N)$, we have

$$I_1 = W(\mu_\ell^N, \nu_\ell^N),$$

while by (28), we get

$$I_2 \leqq \Delta_N \left(K\ W(\mu_\ell^N, \nu_\ell^N) + \Delta_N\right).$$

Finally it holds that

$$W(\mu_{\ell+1}^N, \nu_{\ell+1}^N) \leqq (1 + K\,\Delta_N)\ W(\mu_\ell^N, \nu_\ell^N) + \Delta_N^2. \tag{30}$$

Combining (27) and (30) we get

$$\begin{aligned} W(\mu_\ell^N, \nu_\ell^N) &\leqq (1+K\,\Delta_N)^\ell\ (W(\mu_0,\nu_0) + 2\Delta_N^x) + \sum_{k=0}^{\ell-1}(1+K\,\Delta_N)^k\ \Delta_N^2 \\ &\leqq e^{K\,\ell\Delta_N}\ (W(\mu_0,\nu_0) + 2\Delta_N^x) + \frac{e^{K\,(\ell-1)\Delta_N}-1}{K}\ \Delta_N \\ &\leqq e^{K\,\ell\Delta_N}\left(W(\mu_0,\nu_0) + 2\Delta_N^x + \frac{\Delta_N}{K}\right). \end{aligned} \tag{31}$$

Now, define the countable set $\mathcal{D}^q = \{\mu_0 \in \mathcal{P}_c(\mathbb{R}^n) : \mu_0 = \sum_{i=1}^N m_i\delta_{x_i}, N \in \mathbb{N}, 0 < m_i \in \mathbb{Q}, \sum_i m_i = 1, x_i \in \mathbb{Q}^n\}$. By Lemma 4.1, hypothesis (H:lip) implies (H:cont), thus for every $\mu_0 \in \mathcal{D}^q$, we can apply Theorem 3.1 and find a subsequence of μ^N which converges uniformly on $[0, T]$ for the Wasserstein metric to a solution satisfying (13). Using a diagonal argument we find a subsequence, still indicated by μ^N, which converges uniformly on $[0, T]$ for every $\mu_0 \in \mathcal{D}^q$ to a solution $S_t\mu_0$. Moreover, given $\mu, \nu \in \mathcal{D}^q$, with $\mathrm{Supp}(\mu), \mathrm{Supp}(\nu) \subset B(0, R)$, passing to the limit in (31) and using (13), we have for $K = K(e^{CT}(R+1))$ that

$$W(S_t\mu, S_t\nu) \leqq e^{Kt}W(\mu,\nu). \tag{32}$$

By (32) and the density of $\mathcal{D}^q$ in $\mathcal{P}_c(\mathbb{R}^n)$, we can uniquely extend the map S to the whole set $\mathcal{P}_c(\mathbb{R}^n)$ by approximation.

We now show that, following the proof of Theorem 3.1, we can conclude that $S_t\mu$ is a solution for every $\mu \in \mathcal{P}_c(\mathbb{R}^n)$. We use the notation $\mu_\ell^N(\mu)$ to indicate the subsequence of LASs, with μ as initial datum, converging to the trajectory $S_t\mu$. Consider a sequence $\mu^\nu \in \mathcal{D}^q$ such that $W(\mu^\nu, \mu) \to 0$ as $\nu \to +\infty$. Then, because of (32), there exists a $\eta_\nu > 0$, $\eta_\nu \to 0$ as $\nu \to +\infty$, such that $W(S_t\mu^\nu, S_t\mu) \leqq \eta_\nu$ and $W(\mu_{\ell_N(t)}^N(\mu^\nu), \mu_{\ell_N(t)}^N(\mu)) \leqq \eta_\nu$ for every $t \in [0, T]$. Notice that (16) holds replacing μ_0 with μ^ν. Moreover, using (17), we can estimate

$$W(\mu_{\ell_N(t)}^N(\mu^\nu), S_t\mu) \leqq O(\Delta_N) + \epsilon_N + \eta_\nu. \tag{33}$$

Then we define the function $\psi_{N,\nu}(s) = \int_{T\mathbb{R}^n}(\nabla f(x)\cdot v)\ \mathrm{d}(V[\mu_{\ell_N(s)}^N(\mu^\nu)] - V[\mu_{\ell_N(s)}^N(\mu)]])$. Since the supports of $\mu_{\ell_N(t)}^N(\mu^\nu)$ are uniformly bounded, we can apply Lebesgue dominated convergence and conclude from (33) that $\psi_{N,\nu}$ converges to zero in $L^1(]0,T[)$ as $\nu \to +\infty$. Passing to the limit in ν and N in (16) we conclude that (ii) of Definition 4.2 holds true for S on the whole set $\mathcal{P}_c(\mathbb{R}^n)$. Moreover, again by approximation, we get that (13) and (32) hold on the whole set $\mathcal{P}_c(\mathbb{R}^n)$, thus S satisfies also (iii).

Let us now prove i) of Definition 4.2. From (27), we get $S_0\mu = \mu$. Consider $\mu \in \mathcal{P}_c(\mathbb{R}^n)$ and $t, s \in [0,T]$. We use again the notation $\mu^N_\ell(\mu)$ to indicate the subsequence of LASs defined having μ as initial datum and converging to the trajectory $S_t\mu$. Then, for every ϵ there exists N such that if $\ell = \lfloor Ns \rfloor$ (where $\lfloor s \rfloor = \sup\{n \in \mathbb{N} : n \leqq s\}$ is the usual floor function) and $\ell' = \lfloor Nt \rfloor$, then

$$W(S_s\mu, \mu^N_\ell(\mu)) \leqq \epsilon, \quad W(S_t\mu^N_\ell(\mu), \mu^N_{\ell'}(\mu^N_\ell(\mu))) \leqq \epsilon. \tag{34}$$

Notice that $|\ell + \ell' - \lfloor N(t+s) \rfloor| \leqq 1$, thus, possibly changing N, we also get

$$W(\mu^N_{\ell+\ell'}(\mu), S_{t+s}\mu) \leqq \epsilon. \tag{35}$$

By definition, $\mu^N_{\ell'}(\mu^N_\ell(\mu)) = \mu^N_{\ell+\ell'}(\mu)$, thus, using (32), (34) and (35), we estimate,

$$\begin{aligned} W(S_tS_s\mu, S_{t+s}\mu) &\leqq W(S_tS_s\mu, S_t\mu^N_\ell(\mu)) \\ &\quad + W(S_t\mu^N_\ell(\mu), \mu^N_{\ell'}(\mu^N_\ell(\mu))) \\ &\quad + W(\mu^N_{\ell'}(\mu^N_\ell(\mu)), \mu^N_{\ell+\ell'}(\mu)) + W(\mu^N_{\ell+\ell'}(\mu), S_{t+s}\mu) \\ &\leqq e^{Kt} W(S_s\mu, \mu^N_\ell(\mu)) + 2\epsilon \leqq e^{Kt}\epsilon + 2\epsilon. \end{aligned}$$

For the arbitrariness of ϵ, (i) also holds true for S, and the proof is complete. □

5. Uniqueness of Solutions Semigroup to MDEs

Definition 2.2 is not expected to guarantee uniqueness in general, see Example 3 in Section 7.1. However we can obtain the uniqueness of a Lipschitz semigroup prescribing the small-time evolution of finite sums of Dirac deltas.

We first define the concept of a Dirac germ.

Definition 5.1. Consider V a PVF satisfying (H:bound) and define $\mathcal{D} = \{\mu \in \mathcal{P}_c(\mathbb{R}^n) : \mu = \sum_{i=1}^N m_i\delta_{x_i}, N \in \mathbb{N}, 0 < m_i, \sum_i m_i = 1, x_i \in \mathbb{R}^n\}$. A Dirac germ γ is a map assigning to every $\mu \in \mathcal{D}$ a Lipschitz curve $\gamma_\mu : [0, \epsilon(\mu)] \to \mathcal{P}_c(\mathbb{R}^n)$, $\epsilon(\mu) > 0$ uniformly positive for uniformly bounded supports, such that $\gamma_\mu(0) = \mu$ and γ_μ is a solution to (1).

Roughly speaking, a Dirac germ is a prescribed evolution of solutions for finite sums of Dirac deltas (for sufficiently small times).

Definition 5.2. Consider V a PVF satisfying (H:bound), $T > 0$ and a Dirac germ γ. A Lipschitz semigroup, compatible with the Dirac germ γ, is a Lipschitz semigroup $S : [0,T] \times \mathcal{P}_c(\mathbb{R}^n) \to \mathcal{P}_c(\mathbb{R}^n)$ for (1) such that the following holds: for every $R > 0$, denoting $\mathcal{D}_R = \{\mu \in \mathcal{D} : \mathrm{Supp}(\mu) \subset B(0,R)\}$, there exists $C(R) > 0$ such that for every $t \in [0, \inf_{\mu\in\mathcal{D}_R} \epsilon(\mu)]$ it holds that

$$\sup_{\mu\in\mathcal{D}_R} W(S_t\mu, \gamma_\mu(t)) \leqq C(R)\, t^2. \tag{36}$$

In other words, S is a Lipschitz semigroup whose trajectories are well approximated by the Dirac germ. To prove the uniqueness of a Dirac semigroup (compatible with a given Dirac germ), we use the following Lemma:

Lemma 5.1. *Let S be a Lipschitz semigroup and $\mu : [0, T] \to \mathcal{P}_c(\mathbb{R}^n)$ a Lipschitz continuous curve, then we have*

$$W(S_t\mu(0), \mu(t)) \leqq e^{Kt} \int_0^t \liminf_{h\to 0+} \frac{1}{h} W(S_h\mu(s), \mu(s+h))\, ds.$$

Lemma 5.1 was proved in [5] (Theorem 2.9) for semigroups on Banach spaces, but is valid also for metric spaces. For the reader's convenience we detail the proof in the Appendix.

We are now ready to prove the following:

Theorem 5.1. *Let V be a PVF satisfying (H:bound), $T > 0$ and a Dirac germ γ. There exists at most one Lipschitz semigroup compatible with γ.*

Proof. Let S^1, S^2 be two Lipschitz semigroups compatible with γ. By Lemma 5.1, we have for every $\mu \in \mathcal{P}_c(\mathbb{R}^n)$ that

$$W(S^1_t\mu, S^2_t\mu) \leqq e^{Kt} \int_0^t \liminf_{h\to 0+} \frac{1}{h} W(S^1_h S^2_s\mu, S^2_{s+h}\mu)\, \mathrm{d}s.$$

There exists $R > 0$ such that $\mathrm{Supp}(\mu) \subset B(0, R)$, thus from (24) of Definition 4.2, $\mathrm{Supp}(S^i_t\mu) \subset B(0, e^{CT}(R+1))$ for every $t \in [0, T]$. Moreover, there exists $C_i = C_i(e^{CT}(R+1)) > 0$, $i = 1, 2$, such that (25) and (26) hold true.

Now, fix s. For every ϵ there exists $\mu_s \in \mathcal{D}$ such that $W(\mu_s, S^2_s\mu) \leqq \epsilon$ and, from Definition 5.2, there exists C, depending on R and T, such that

$$W(S^1_h\mu_s, \gamma_{\mu_s}(h)) \leqq C\, h^2, \qquad W(S^2_h\mu_s, \gamma_{\mu_s}(h)) \leqq C\, h^2.$$

Then

$$\begin{aligned} W(S^1_h S^2_s\mu, S^2_{s+h}\mu) &= W(S^1_h S^2_s\mu, S^2_h S^2_s\mu) \\ &\leqq W(S^1_h S^2_s\mu, S^1_h\mu_s) + W(S^1_h\mu_s, \gamma_{\mu_s}(h)) \\ &\quad + W(\gamma_{\mu_s}(h), S^2_h\mu_s) + W(S^2_h\mu_s, S^2_h S^2_s\mu) \\ &\leqq 2\,(e^{\tilde{C}h}\epsilon + C\,h^2), \end{aligned}$$

where $\tilde{C} = \max\{C_1, C_2\}$. Since ϵ is arbitrary and does not depend on h, we conclude that

$$\liminf_{h\to 0+} \frac{1}{h} W(S^1_h S^2_s\mu, S^2_{s+h}\mu) = 0,$$

which gives $W(S^1_t\mu, S^2_t\mu) = 0$. □

We are now ready to state our main uniqueness result. We remark that there may exist solutions which are not the limit of LAS.

Theorem 5.2. *Let V be a PVF satisfying (H:bound) and (H:lip). If for every $\mu \in \mathcal{D}$ (finite sum of Dirac deltas) the sequence of LASs μ^N converges to a unique limit, then there exists a unique Lipschitz semigroup whose trajectories are limits of LASs.*

Proof. By Theorem 4.1 there exists a Lipschitz semigroup S obtained via limits of LASs. Such a semigroup is unique because of the density of finite sums of Dirac masses in $\mathcal{P}_c(\mathbb{R}^n)$. □

Theorem 5.2 allows as to reduce the question of uniqueness of a Lipschitz semigroup (compatible with LAS limits) to that of understanding the uniqueness of LAS limits for finite sums of Dirac deltas. The latter question is much simpler than the general uniqueness of Lipschitz semigroups and, in the next sections, we provide examples of PVFs for which we can apply Theorem 5.2.

6. Ordinary Differential Equations and MDEs

In this section we show natural connections between Ordinary Differential Equations (briefly ODEs) and MDEs. We start with the following definition:

Definition 6.1. Consider an ODE $\dot{x} = v(x)$, $v : \mathbb{R}^n \to \mathbb{R}^n$. We define a PVF V^v by

$$V^v[\mu] = \mu \otimes \delta_{v(x)}.$$

The main question is if V^v satisfies hypothesis (H:bound) and (H:cont) or (H:lip).

Notice that (H:bound) easily follows from a sublinear growth requirement on v, that is there exists $C > 0$ such that $|v(x)| \leqq C(1 + |x|)$.

(H:cont) holds if v is continuous. Indeed, fix $R > 0$ and $f \in \mathcal{C}_c^\infty(T\mathbb{R}^n)$, then the map $x \to f(x, v(x))$ is continuous and bounded on $B(0, R)$. Since $\int_{T\mathbb{R}^n} f(x, v)\mathrm{d}(V^v[\mu] - V^v[\nu]) = \int_{R^n} f(x, v(x))\mathrm{d}(\mu - \nu)$, we conclude.

We first prove the following:

Proposition 6.1. *V^v satisfies (H:lip) for finite sums of Dirac deltas if and only if v is locally Lipschitz continuous.*

Proof. Assume first v to be locally Lipschitz, fix $R > 0$ and let $L(v, R)$ be the Lipschitz constant of v on $B(0, R)$. Consider two probability measures $\mu_i = \sum_{j=1}^{N_i} m_j^i\, \delta_{x_j^i}$, $i = 1, 2$. As above, we can assume $N_1 = N_2 = N$ and that there exists a map $\sigma : \{1, \dots, N\} \to \{1, \dots, N\}$ so that an optimal transference plan between μ_1 and μ_2 moves the mass $m_j^1 = m_{\sigma(j)}^2$ from x_j^1 to $x_{\sigma(j)}^2$. Then it holds that

$$\begin{aligned}\mathcal{W}(V^v[\mu_1], V^v[\mu_2]) &= \mathcal{W}\left(V^v\left[\sum_{j=1}^N m_j^1\, \delta_{x_j^1}\right], V^v\left[\sum_{j=1}^N m_j^2\, \delta_{x_j^2}\right]\right)\\ &\leqq \sum_{j=1}^N m_j^1\, |v(x_j^1) - v(x_{\sigma(j)}^2)|\end{aligned}$$

$$\leqq L(v, R) \sum_{j=1}^{N} m_j^1 |x_j^1 - x_{\sigma(j)}^2| = L(v, R)\, W(\mu_1, \mu_2).$$

Conversely, assume V^v to satisfy (H:lip). Taking two points, $x, y \in B(0, R)$, we have

$$|v(x) - v(y)| = \mathcal{W}(\delta_{(x,v(x))}, \delta_{(y,v(y))}) \leqq K\, W(\delta_x, \delta_y) = K\, |x - y|,$$

thus we conclude that v is locally Lipchitz continuous. □

From the uniqueness of solutions for locally Lipschitz vector fields we obtain

Theorem 6.1. *Consider a locally Lipschitz vector field $v : \mathbb{R}^n \to \mathbb{R}^n$ with sublinear growth (that is there exists $C > 0$ such that $|v(x)| \leqq C(1 + |x|)$), then V^v satisfies (H:lip) and the MDE, associated to V^v, admits a unique Lipschitz semigroup obtained as limit of LASs.*

Proof. Property (H:bound) for V^v follows from the sublinear growth of v. Consider now $\mu, \nu \in \mathcal{P}_c(R^n)$ and let $R > 0$ be such that $\mathrm{Supp}(\mu), \mathrm{Supp}(\nu) \subset B(0, R)$. Define the map $Z(x, y) = (x, v(x), y, v(y))$. If $\tau \in P^{opt}(\mu, \nu)$, then $T = Z\#\tau$ satisfies

$$\begin{aligned}\int_{T\mathbb{R}^n} |v - w| \,\mathrm{d}T &= \int_{\mathbb{R}^n} |v(x) - v(y)| \,\mathrm{d}\tau(x, y)\\ &\leqq L(v, R) \int_{\mathbb{R}^n} |x - y| \,\mathrm{d}\tau(x, y) = L(v, R)\, W(\mu, \nu),\end{aligned}$$

where $L(v, R)$ indicates a Lipschitz constant for v on the set $B(0, R)$. Thus (H:lip) holds true.

Now, consider a finite sum of Dirac deltas $\mu = \sum_i m_i \delta_{x_i}$ and indicate by $x_i(\cdot)$ the unique solution to the Cauchy problem $\dot{x} = v(x)$, $x(0) = x_i$. We define the Euler approximate trajectory x_i^N by recursion: $x_i^N(0) = x_i$ and $x_i^N(t) = x_i^N(\lfloor \frac{t}{\Delta_N} \rfloor \Delta_N) + (t - \lfloor \frac{t}{\Delta_N} \rfloor) f(x_i^N(\lfloor \frac{t}{\Delta_N} \rfloor \Delta_N))$ (where $\lfloor \cdot \rfloor$ is the floor function). Notice that the sequence of LAS is exactly given by $\mu^N(t) = \sum_i m_i \delta_{x_i^N(t)}$. Because of the sublinear growth of v, given $T > 0$, the Euler approximate solutions x_i^N and exact solutions x_i are all contained in a ball $B(0, R)$ over the time interval $[0, T]$. The error function $z_i(t) = |x_i^N(t) - x_i(t)|^2$ satisfies

$$\begin{aligned}\dot{z}(t) &= 2\left\langle f\left(x_i^N\left(\left\lfloor \frac{t}{\Delta_N} \right\rfloor \Delta_N\right)\right) - f(x_i(t)), x_i^N(t) - x_i(t)\right\rangle\\ &\leqq 2Lz(t) + 2\left|\left\langle f(x_i^N(t)) - f\left(x_i^N\left(\left\lfloor \frac{t}{\Delta_N} \right\rfloor \Delta_N\right)\right), x_i^N(t) - x_i(t)\right\rangle\right|\\ &\leqq 2Lz(t) + 2LC\Delta_N \cdot 2R,\end{aligned}$$

where $L = L(v, R)$ is a Lipschitz constant for v on $B(0, R)$ and $C = C(v, R)$ a bound for v on $B(0, R)$. By Gronwall's Lemma we have $z(t) \leqq 4LCR\, \Delta_N\, (\frac{e^{2LT-1}}{2L})$. Therefore, given $T > 0$, $x_i^N(\cdot)$ converges uniformly on $[0, T]$ to $x_i(\cdot)$, thus μ^N converges in Wasserstein distance to $\mu(t) = \sum_i m_i \delta_{x_i(t)}$ uniformly on $[0, T]$. We conclude by applying Theorem 5.2. □

We also have the following:

Proposition 6.2. *Assume v is locally Lipschitz continuous with sublinear growth, then the solution to the Cauchy problem for $\dot{\mu} = V^v[\mu]$ with initial datum μ_0, with compact support, is the unique solution to the transport equation*

$$\mu_t + div(v\,\mu) = 0, \quad \mu(0) = \mu_0.$$

The proof follows immediately from uniqueness of weak solutions to the transport equation, see [17].

6.1. A Natural Monoid Structure for PVFs

We now describe a natural monoid structure and scalar product on the set of PVFs, built upon the connections between vector fields and PVFs.

First we define a fiber convolution for measures on $T\mathbb{R}^n$ with same marginal on the base. More precisely, given $\mu_b \otimes \mu_x, \mu_b \otimes \nu_x \in \mathcal{P}(T\mathbb{R}^n)$, for every $B \subset T\mathbb{R}^n$ we define

$$\begin{aligned}&(\mu_b \otimes \mu_x) *_f (\mu_b \otimes \nu_x)(B)\\&= \int_{\mathbb{R}^n} \left(\int_{(\mathbb{R}^n)^2} \chi_B(x, v+w)\, \mathrm{d}\mu_x(v)\mathrm{d}\nu_x(w) \right) \mathrm{d}\mu_b(x).\end{aligned}$$

Given two PVFs V_1, V_2, we denote $V_i[\mu] = \mu \otimes \nu_x^i$ ($\nu_x^i = \nu_x^i[\mu]$) the disintegration of V_i on base-fiber of $T\mathbb{R}^n$, then we can define

$$(V^1 \oplus_f V^2)(\mu) = (\mu \otimes \nu_x^1) *_f (\mu \otimes \nu_x^2).$$

We have the following:

Proposition 6.3. *The operation $\oplus_f$ defines an abelian monoid structure over the set of PVFs.*

Proof. Commutativity and associativity follows from the same property of convolution of measures (and linearity of the integration over the base). The neutral element is given by $V[\mu] = \mu \otimes \delta_0$. □

Notice that every PVF V^v is invertible and its inverse is V^{-v}, but other elements are not invertible, thus $\oplus_f$ does not define a group structure. However, the sum of two vector fields v_i is mapped to the fiber-convolution of their PVFs, indeed,

$$V^{v_1+v_2}(\mu) = \mu \otimes \delta_{v_1(x)+v_2(x)} = (\mu \otimes \delta_{v_1(x)}) *_f (\mu \otimes \delta_{v_2(x)}).$$

For every $\lambda \in \mathbb{R}$ and $B \subset T\mathbb{R}^n$, set

$$(\lambda \cdot_f V_i)[\mu](B) = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \chi_B(x, \lambda v)\, \mathrm{d}\nu_x^i[\mu](v) \right) \mathrm{d}\mu(x).$$

Let us denote by $Vec(\mathbb{R}^n)$ the set of locally Lipschitz vector fields with sublinear growth endowed with the usual vector space structure, and by $PVec(\mathbb{R}^n)$ the set of PVFs satisfying (H:bound) and (H:lip) endowed with the operations $\oplus_f$ and $\cdot_f$, then we have

Proposition 6.4. *The map $v \to V^v$ is a monoid isomorphism from $Vec(\mathbb{R}^n)$ to $PVec(\mathbb{R}^n)$. Moreover $V^{\lambda v} = \lambda \cdot_f V^v$.*

7. Finite Speed Diffusion and Concentration

In this section, we show examples of MDEs which reproduce diffusion and concentration phenomena. The former can be obtained using PVF which depend on global quantities and satisfy condition (H:lip) (while we also show that diffusion can not be obtained by constant PVFs). In particular, we are able to model diffusions with uniformly bounded speed. Concentration is achieved by PVFs violating (H:lip), but still guaranteeing convergence of LASs to unique limits and the existence of the Lipschitz semigroup.

7.1. Diffusion

Let us start proving

Proposition 7.1. *If a PVF V does not depend on μ, that is $V[\mu] = \mu \otimes \bar{V}$ for some $\bar{V} \in \mathcal{P}_c(\mathbb{R}^n)$, then the solution to (2) obtained as limit LAS is given by constant translation at speed $\bar{v} = \int v\, d\bar{V}$, that is for every Borel set A one has $\mu(t)(A) = \mu_0(A - t\bar{v})$.*

Proof. As pointed out in Remark 1 of Section 2 we can replace $\bar{V}$ with $\bar{v}$ without changing the definition of solution. Therefore the conclusion follows from the uniqueness of solutions for the PVF $V^{\bar{v}}$. □

We now provide a first example of finite speed diffusion obtained by letting the PVF V depend on global properties of μ. This example is related to the Wasserstein gradient flow with interaction energy $\Phi(\mu) := -\int_{\mathbb{R}\times\mathbb{R}} |x-y| \mathrm{d}\mu(x)\mathrm{d}\mu(y)$, see [4].

Example 1. For every $\mu \in \mathcal{P}_c(\mathbb{R})$ define

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

Notice that we have $\mu(]-\infty, B(\mu)[) \leqq \frac{1}{2} \leqq \mu(]-\infty, B(\mu)])$, then we set $\eta = \mu(]-\infty, B(\mu)]) - \frac{1}{2}$ so $\mu(\{B(\mu)\}) = \eta + \frac{1}{2} - \mu(]-\infty, B(\mu)[)$. We define $V[\mu] = \mu \otimes \nu_x$, with

$$\nu_x = \begin{cases} \delta_{-1} & \text{if } x < B(\mu) \\ \delta_1 & \text{if } x > B(\mu) \\ \frac{1}{\mu(\{B(\mu)\})}\left(\eta\delta_1 + \left(\frac{1}{2} - \mu(]-\infty, B(\mu)[)\right)\delta_{-1}\right) & \text{if } x = B(\mu),\ \mu(\{B(\mu)\}) > 0. \end{cases} \tag{37}$$

We then have the following:

Proposition 7.2. *Let V be the PVF defined in (37). V satisfies (H:bound) and (H:lip) and LASs admit a unique limit, thus the conclusions of Theorem 5.2 holds true. Moreover, the solution to (2), obtained as the limit LASs μ^N, satisfies*

$$\mu(t)(A) = \mu_0((A\cap]-\infty, B(\mu)-t[)+t) + \mu_0((A\cap]B(\mu)+t, +\infty[)-t) \\ + \frac{1}{\mu_0(\{B(\mu_0)\})}\big(\eta\delta_{B(\mu_0)+t}(A) \\ + \left(\frac{1}{2} - \mu_0(]-\infty, B(\mu_0)[)\right)\delta_{B(\mu_0)-t}(A)\bigg).$$

In particular, we have that:

(i) *The solution to* (2) *with* $\mu_0 = \delta_{x_0}$ *is given by* $\mu(t) = \frac{1}{2}\delta_{x_0+t} + \frac{1}{2}\delta_{x_0-t}$*;*
(ii) *The solution to* (2) *with* $\mu_0 = \chi_{[a,b]}\,\lambda$ *(where* χ *is the characteristic function and* λ *is the Lebesgue measure) is given by* $\mu(t) = \chi_{[a-t,\frac{a+b}{2}-t]}\lambda + \chi_{[\frac{a+b}{2}+t,b+t]}\lambda$.

Proof. The PVF V satisfies (H:bound) by definition. Consider two measures μ, $\nu \in \mathcal{P}_c(\mathbb{R})$ and assume first $\mu(\{B(\mu)\}) = \nu(\{B(\nu)\}) = 0$, that is the barycenters are not atoms. Then any optimal plan between μ and ν moves the mass of μ to the left, respectively right, of $B(\mu)$ to the mass of ν the left, respectively right, of $B(\nu)$ (see Theorem 2.18 and Remark 2.19 (ii) in [17]). Therefore $\mathcal{W}(V[\mu], V[\nu]) = 0$ and (H:lip) is trivially satisfied.

If μ or ν do have an atom at the barycenter, then we can still construct a $T \in P(V[\mu], V[\nu]))$ whose support is contained in the set $\{(x, 1, y, 1) : x, y \in \mathbb{R}^n\} \cup \{(x, -1, y, -1) : x, y \in \mathbb{R}^n\}$, thus again we conclude $\mathcal{W}(V[\mu], V[\nu]) = 0$, and (H:lip) is trivially satisfied.

The other claims follow by direct computations. □

Example 1 can be generalized as follows:

Example 2. Consider an increasing map $\varphi : [0, 1] \to \mathbb{R}$ and define $V_\varphi[\mu] = \mu \otimes_x J_\varphi(x)$, where

$$J_\varphi(x) = \begin{cases} \delta_{\varphi(F_\mu(x))} & \text{if } F_\mu(x^-) = F_\mu(x), \\ \frac{\varphi\#\left(\chi_{[F_\mu(x^-),F_\mu(x)]}\lambda\right)}{F_\mu(x)-F_\mu(x^-)} & \text{otherwise,} \end{cases}$$

where $F_\mu(x) = \mu(]-\infty, x])$, the cumulative distribution of μ, and λ the Lebesgue measure. Simply put, $V[\mu]$ moves the ordered masses with speed prescribed by φ.

Following the same proof of Proposition 7.2, we have that V_φ satisfies (H:bound) and (H:lip) if φ is bounded. If φ is a diffeomorphism, the conclusion of Theorem 5.2 holds true and the solution from δ_0 is given by $g(t, x)\lambda$ with

$$g(t, x) = \frac{1}{t\varphi'\left(\varphi^{-1}\left(\frac{x}{t}\right)\right)} = \frac{(\varphi^{-1})'\left(\frac{x}{t}\right)}{t}.$$

For example, if $\varphi(\alpha) = \alpha - \frac{1}{2}$, then $g(t, x) = \frac{1}{t}\chi_{[-\frac{t}{2},\frac{t}{2}]}$, so we get uniformly distributed masses. For $\varphi(x) = 4\,\text{sgn}(\alpha - \frac{1}{2})\,(\alpha - \frac{1}{2})^2$ we get $g(t, x) = \frac{1}{4\sqrt{tx}}$, which is unbounded at 0.

Example 3. Let us go back to the question of the uniqueness of the solutions. Consider the PVF V_1 defined in Example 1 and let $V_2[\mu] = \mu \otimes (\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1})$. From Proposition 7.2, the solution to $\dot{\mu} = V_1[\mu]$, $\mu(0) = \delta_0$ is given by $\mu_1(t) = \frac{1}{2}\delta_t + \frac{1}{2}\delta_{-t}$. From Proposition 7.1, the solution to $\dot{\mu} = V_2[\mu]$, $\mu(0) = \delta_0$ is given by $\mu_2(t) = \delta_0$. It is easy to check that $V_1[\delta_0] = V_2[\delta_0]$ and that μ_2 satisfies (3) both for V_1 and V_2.

It is also interesting to notice that the LASs μ^N for μ_1 coincides with μ_1 for every N. On the other side, given $f \in \mathcal{C}_c^\infty$, $\int f \,\mathrm{d}\mu_2(t) \equiv f(0)$ so $\frac{d}{\mathrm{d}t}\int f\,\mathrm{d}\mu_2(t) \equiv 0$. Thus μ_1 is trivially approximated by LASs, while μ_2 gives the trivial solution to (3).

7.2. *Concentration*

It is well known that, to achieve the existence and uniqueness of solutions to an ODE $\dot{x} = v(x)$, the locally Lipschitz condition on the vector field v can be relaxed to a one-sided locally Lipschitz condition

$$\langle v(x) - v(y), x - y\rangle \leqq L\,|x-y|^2, \tag{38}$$

where $\langle\cdot,\cdot\rangle$ indicates the scalar product of $\mathbb{R}^n$. See [12] for general results and Section 1 of Chapter 3 in [2] for a concise presentation. Similarly we can relax condition (H:lip) as follows: define

$$\mathcal{W}'(V_1, V_2) = \inf\left\{\int_{T\mathbb{R}^n\times T\mathbb{R}^n} \frac{\langle v-w, x-y\rangle}{|x-y|}\,\mathrm{d}T(x,v,y,w)\ :\ T \in P(V_1,V_2),\ \pi_{13}\#T \in P^{opt}(\mu_1,\mu_2)\right\}, \tag{39}$$

then assume that

(H':lip) for every $R > 0$ there exists $K = K(R) > 0$ such that if $\mathrm{Supp}(\mu)$, $\mathrm{Supp}(\nu) \subset B(0,R)$, then

$$\mathcal{W}'(V[\mu], V[\nu]) \leqq K\ W(\mu,\nu). \tag{40}$$

Remark 7. Also for (H':lip) we may consider W_p, $p \geqq 1$; the analogous condition for W_2 is satisfied by measure-valued subgradients of λ-convex functionals see formula (10.3.21) of [1].

We have the following:

Theorem 7.1. *Given V satisfying (H:bound) and (H':lip), and $T > 0$, passing to the limit in LASs we can define a Lipschitz semigroup of solutions to* (2).

Proof. The proof of Theorem 4.1 can be modified as follows: to estimate the Wasserstein distance between μ_ℓ^N and ν_ℓ^N, we first notice that for $a, b \in \mathbb{R}^n$ and $\epsilon > 0$ we have $\frac{d}{\mathrm{d}\epsilon}|a+\epsilon b|\big|_{\epsilon=0} = \frac{\langle b,a\rangle}{|a|}$, thus $|a+\epsilon b| = |a| + \epsilon\frac{\langle b,a\rangle}{|a|} + o(\epsilon)$. Setting $x - y = a$ and $v - w = b$ we can then write

$$|(x-y) + \Delta_N(v-w)| \leqq |x-y| + \Delta_N\left\langle v-w, \frac{x-y}{|x-y|}\right\rangle + o(\Delta_N).$$

Now assumption (H':lip) guarantees that (30) is still true and we can conclude in the same way as for Theorem 4.1. □

Examples of concentration are obtained easily, as follows:

Example 4. Consider an ODE $\dot{x} = v(x)$ with v satisfying (38) (with L bounded on compact sets). Then condition (H':lip) holds true and we can apply Theorem 7.1.

In $\mathbb{R}$ define $v(x) = \pm 1$ if $\pm x < 0$ and $v(0) = 0$. If we start with a uniform mass distributed on the interval $[-1, 1]$, the solution converges in time $t = 1$ to δ_0.

Example 5. Consider a scalar conservation law

$$u_t + \nabla \cdot (a(t, x)\, u) = 0,$$

with a satisfying $\langle a(t, x) - a(t, y), x - y \rangle \leqq L\, |x - y|^2$ uniformly in t and on compact sets. Then the conclusions of Theorem 6.1 hold true for the ODE $\dot{x} = a(t, x)$. One can thus recover the conclusions of Theorem 3.3 of [14] (the latter being more generally based on uniqueness of Filippov characteristics).

8. Mean-Field Limits for Multi-Particle Systems

A typical example of a multi-particle system is given by the system of ODEs:

$$\dot{x}_i = \frac{1}{m} \sum_{j=1}^{m} \phi(x_j - x_i), \tag{41}$$

where $x_i \in \mathbb{R}^n$, $i = 1, \ldots, m$, and ϕ is locally Lipschitz continuous and uniformly bounded. For every m and $x(\cdot) = (x_1(\cdot), \ldots, x_m(\cdot)) \in \mathbb{R}^{nm}$, solution to (41), consider the empirical probability measure of m particles:

$$\mu_m(t) = \frac{1}{m} \sum_{i=1}^{m} \delta_{x_i(t)}. \tag{42}$$

A typical problem is to understand the limit of μ_m as $m \to \infty$ (see for instance [11]); applications include problems from biology, crowd dynamics and other fields, see [6,8]. Dobrushin (see [10]) proved convergence, for the Wasserstein metric topology, of the empirical probability measures to solutions of the mean field equation

$$\mu_t + \nabla_x \cdot \left(\left(\int \phi(x - y) \mathrm{d}\mu(y) \right) \mu \right) = 0.$$

Let us consider the more general model

$$\dot{x}_i = v_i^m(x), \tag{43}$$

where $x = (x_1, \ldots, x_m)$, $x_i \in \mathbb{R}^n$ and v_i^m is locally Lipschitz continuous and uniformly bounded. We assume a condition of indistinguibility of particles. For every m, we indicate by Σ_m the set of permutations σ over the set $\{1, \ldots, m\}$. Given $\sigma \in \Sigma_m$ and $x = (x_1, \ldots, x_m)$, $x_i \in \mathbb{R}^n$, we define $x_\sigma = (x_{\sigma(1)}, \ldots, x_{\sigma(m)})$. Then we assume that

(IP) for every $x = (x_1, \dots, x_m)$, $x_i \in \mathbb{R}^n$, and $\sigma \in \Sigma_m$, it holds $v^m_{\sigma(i)}(x) = v^m_i(x_\sigma)$.

Notice that, given two empirical measures $\mu_j = \frac{1}{m}\sum_{i=1}^m \delta_{x_i^j}$, $j = 1, 2$, the Wasserstein distance between them is given by

$$W(\mu_1, \mu_2) = \inf_{\sigma\in\Sigma_m} \frac{1}{m}\sum_i |x_i^1 - x^2_{\sigma(i)}|,$$

then, setting $x^j = (x_1^j, \dots, x_m^j) \in \mathbb{R}^{mn}$, $j = 1, 2$, we can estimate

$$\begin{aligned}\frac{1}{m}\sum_i |v_i^m(x^1) - v^m_{\sigma(i)}(x^2)| &= \frac{1}{m}\sum_i |v_i^m(x^1) - v_i^m(x^2_\sigma)| \\ &\leqq \frac{1}{m}\sum_i L_i^m \, |x^1 - x^2_\sigma| \\ &= \frac{1}{m}\left(\sum_i L_i^m\right)\left(\sum_k |x_k^1 - x^2_{\sigma(k)}|^2\right)^{\frac{1}{2}} \\ &\leqq \left(\sum_i L_i^m\right)\cdot \frac{1}{m}\sum_k |x_k^1 - x^2_{\sigma(k)}|,\end{aligned}$$

where L_i^m is a (local) Lipschitz constant of v_i^m. Let $\Sigma_m(\mu_1, \mu_2)$ be the set of $\sigma \in \Sigma_m$ which realizes the Wasserstein distance $W(\mu_1, \mu_2)$. Then

$$\begin{aligned}&\inf_{\sigma\in\Sigma_m(\mu_1,\mu_2)} \frac{1}{m}\sum_i |v_i^m(x^1) - v^m_{\sigma(i)}(x^2)| \\ &\leqq \inf_{\sigma\in\Sigma_m(\mu_1,\mu_2)} \left(\sum_i L_i^m\right)\cdot\frac{1}{m}\sum_k |x_k^1 - x^2_{\sigma(k)}| \\ &= \left(\sum_i L_i^m\right)\cdot W(\mu_1, \mu_2).\end{aligned} \tag{44}$$

The left-hand side of (44) is precisely the term appearing in the definition of $\mathcal{W}$ (see (21)) if V is a PVF corresponding to the system (43). Assume there exist uniform bounds on the Lipschitz constants of v_i^m and a PVF obtained as limit as $m \to \infty$ in the following sense:

(A) Denote by $L_i^m(R)$ the Lipschitz constant of the vector field v_i^m over the set $B(0, R) \subset \mathbb{R}^{mn}$, then for every R it holds $\sup_m \sum_i L_i^m(R) < +\infty$. Moreover, there exists a PVF V such that for every sequence $x^N = (x_1^N, \dots, x^N_{m(N)})$, $N \in \mathbb{N}$, $m(N) \in \mathbb{N}$, $x_i^N \in \mathbb{R}^n$, define $\mu_N = \frac{1}{m(N)}\sum_{i=1}^{m(N)} \delta_{x_i^N}$, and it there exists $R > 0$ such that $|x_i^N| \leqq R$ and $\mu \in \mathcal{P}_c(\mathbb{R}^n)$ such that $\lim_{N\to\infty} W(\mu_N, \mu) = 0$, then

$$\lim_{N\to\infty} \mathcal{W}\left(\frac{1}{m(N)}\sum_{i=1}^{m(N)} \delta_{(x_i^N, v_i^{m(N)}(x^N))}, V[\mu]\right) = 0. \tag{45}$$

An example where assumption (A) is satisfied is given below, in Corollary 8.1. We have

Theorem 8.1. *Consider the system (43), assume v_i^m locally Lipschitz and globally bounded, (IP) and (A) hold true, and denote by V the PVF given by (A). Then V satisfies (H:lip) and the empirical probability measures (42), where $x = (x_1, \dots, x_m)(t)$ solves (43), are solutions to the MDE $\dot{\mu} = V[\mu]$. Moreover, there exists a unique Lipschitz semigroup for the MDE whose trajectories coincide with the empirical probability measures for finite sums of Dirac deltas.*

Proof. Let $x = (x_1, \dots, x_m)$, $x_i \in \mathbb{R}^n$, and $\mu = \frac{1}{m}\sum_{i=1}^m \delta_{x_i}$, then by taking the constant sequence in (A) $\mu_N \equiv \mu$, $m(N) = m$, we deduce $\mathcal{W}(\frac{1}{m}\sum_{i=1}^m \delta_{(x_i, v_i^m(x))}, V[\mu]) = 0$, thus, by Lemma 4.1, $V[\mu] = \frac{1}{m}\sum_{i=1}^m \delta_{(x_i, v_i^m(x))}$. Now, given $\mu \in \mathcal{P}_c(\mathbb{R}^n)$, let μ_N be a sequence of finite sums of Dirac deltas as in (A) with $W(\mu_N, \mu) \to 0$. From (45) and Lemma 4.1, we deduce $W^{T\mathbb{R}^n}(V[\mu], V[\mu_N]) \to 0$. Therefore we can uniquely define V on the whole $\mathcal{P}_c(\mathbb{R}^n)$ by approximation.

Property (H:bound) for V follows from the boundedness of v_i^m. To prove (H:lip), consider $\mu_i \in \mathcal{P}_c(\mathbb{R}^n)$, $i = 1, 2$ and let μ_N^i be sequences as in (A) such that $\lim_{N\to\infty} W(\mu_N^i, \mu_i) = 0$. For every N and $\epsilon_N > 0$, there exists $T_N \in P(V[\mu_N^1], V[\mu_N^2])$, with $\pi_{13}\#T_N \in P^{opt}(\mu_N^1, \mu_N^2)$, $T_N^1 \in P(V[\mu_1], V[\mu_N^1])$, with $\pi_{13}\#T_N^1 \in P^{opt}(\mu_1, \mu_N^1)$, and $T_N^2 \in P(V[\mu_N^2], V[\mu_2])$, with $\pi_{13}\#T_N^2 \in P^{opt}(\mu_N^2, \mu_2)$, such that

$$\int_{T\mathbb{R}^n \times T\mathbb{R}^n} |v - w| \, \mathrm{d}T_N(x, v, y, w) \leqq \mathcal{W}(V[\mu_N^1], V[\mu_N^2]) + \epsilon_N,$$
$$\int_{T\mathbb{R}^n \times T\mathbb{R}^n} |v - w| \, \mathrm{d}T_N^i(x, v, y, w) \leqq \mathcal{W}(V[\mu_N^i], V[\mu_i]) + \epsilon_N, \qquad i = 1, 2.$$

We can compose the transference plans T_N^1, T_N and T_N^2 (see Lemma 5.3.2, remark 5.3.3 and Section 7.1 of [1]) thus there exists $\tilde{T}_N$ such that $\tilde{\pi}_{12}\#\tilde{T}_N = T_N^1$, $\tilde{\pi}_{23}\#\tilde{T}_N = T_N$, $\tilde{\pi}_{34}\#\tilde{T}_N = T_N^2$, and $\tilde{\pi}_{14}\#\tilde{T}_N \in P(V[\mu_1], V[\mu_2])$, where $\tilde{\pi}_{ij}$ is the projection on the i-th and j-th components of the Cartesian product $(T\mathbb{R}^n)^4$. Moreover, we have

$$\begin{aligned}
&\int_{(T\mathbb{R}^n)^2} |v - w| \, \mathrm{d}(\tilde{\pi}_{14}\#\tilde{T}_N)(x, v, y, w) \\
&\quad \leqq \int_{(T\mathbb{R}^n)^2} |v - w| \, \mathrm{d}(T_N^1 + T_N + T_N^2)(x, v, y, w) \\
&\quad = \mathcal{W}(V[\mu_1], V[\mu_N^1]) + \mathcal{W}(V[\mu_N^1], V[\mu_N^2]) + \mathcal{W}(V[\mu_N^2], V[\mu_2]) + 3\epsilon_N,
\end{aligned}$$

and

$$\begin{aligned}
&\int_{(T\mathbb{R}^n)^2} |x - y| \, \mathrm{d}(\tilde{\pi}_{14}\#\tilde{T}_N)(x, v, y, w) \\
&\quad \leqq \int_{(T\mathbb{R}^n)^2} |x - y| \, \mathrm{d}(T_N^1 + T_N + T_N^2)(x, v, y, w)
\end{aligned}$$

$$= W(\mu_1, \mu_N^1) + W(\mu_N^1, \mu_N^2) + W(\mu_N^2, \mu_2).$$

The sequence $\tilde{\pi}_{14}\#\tilde{T}_N$ is tight, thus narrowly relatively compact (Lemma 5.2.2 and Theorem 5.1.3 of [1]), so there exists a subsequence converging to $\tilde{T} \in \mathcal{P}((T\mathbb{R}^n)^2)$. The transportation costs are narrowly lower semicontinuous (Proposition 7.13 of [1]), thus we have that $\tilde{T} \in P(V[\mu_1], V[\mu_2])$ and

$$\int_{(T\mathbb{R}^n)^2} |v - w| \, \mathrm{d}\tilde{T}(x, v, y, w) \leqq \liminf_{N\to\infty} \int_{(T\mathbb{R}^n)^2} |v - w| \, \mathrm{d}(\tilde{\pi}_{14}\#\tilde{T}_N)(x, v, y, w).$$

Moreover,

$$\begin{aligned}\int_{(T\mathbb{R}^n)^2} |x - y| \, \mathrm{d}\tilde{T}(x, v, y, w) &\leqq \liminf_{N\to\infty} \int_{(T\mathbb{R}^n)^2} |x - y| \, \mathrm{d}(\tilde{\pi}_{14}\#\tilde{T}_N)(x, v, y, w) \\ &= W(\mu_1, \mu_2),\end{aligned}$$

thus $\pi_{13}\#\tilde{T} \in P^{opt}(\mu_1, \mu_2)$. Then

$$\begin{aligned}\mathcal{W}(V[\mu_1], V[\mu_2]) &\leqq \int_{(T\mathbb{R}^n)^2} |v - w| \, \mathrm{d}\tilde{T}(x, v, y, w) \\ &\leqq \liminf_{N\to\infty} \int_{(T\mathbb{R}^n)^2} |v - w| \, \mathrm{d}(\tilde{\pi}_{14}\#\tilde{T}_N)(x, v, y, w) \\ &\leqq \liminf_{N\to\infty} \Big(\mathcal{W}(V[\mu_1], V[\mu_N^1]) + \mathcal{W}(V[\mu_N^1], V[\mu_N^2]) \\ &\quad + \mathcal{W}(V[\mu_N^2], V[\mu_2]) + 3\epsilon_N \Big).\end{aligned}$$

Now we can choose $\epsilon_N \to 0$ as $N \to \infty$ and by (45) the first and third addendum in parenthesis tend to zero. By (44), the second addendum can be bounded by $\sup_m \sum_i L_i^m(R) \, W(\mu_1, \mu_2)$, thus it follows that

$$\mathcal{W}(V[\mu_1], V[\mu_2]) \leqq \left(\sup_m \sum_i L_i^m(R) \right) W(\mu_1, \mu_2),$$

then, by (A), V satisfies (H:lip).

From Theorem 4.1 there exists a Lipschitz semigroup of solutions to the MDE $\dot{\mu} = V[\mu]$, obtained as the limit of LASs. Moreover, using the local Lipschitz continuity of v_i^m, we can define a Dirac germ coinciding with the empirical probability measures (which, in turn, coincide with the unique limit of LASs). Then we conclude by applying Theorem 5.1. □

We easily obtain the following:

Corollary 8.1. *Consider* (41) *with ϕ bounded and locally Lipschitz. Then the conclusions of Theorem* 8.1 *hold true.*

Proof. The system (41) can be written as (43) with $v_i^m(x) = \frac{1}{m}\sum_{j=1}^m \phi(x_j - x_i)$. The uniform boundedness of v_i^m follows from the boundedness of ϕ. Moreover, if $L_\phi(R)$ is the Lipschitz constant of ϕ on $B(0, R)$, then $L_i^m(R) = \frac{1}{m}L_\phi(R)$ and $\sup_m \sum_i L_i^m(R) = L_\phi(R)$. Finally, defining $V[\mu] = \mu \otimes \int_{\mathbb{R}^n} \phi(x - y)\mathrm{d}\mu(y)$, (A) follows from the local Lipschitz continuity of ϕ. We conclude by applying Theorem 8.1. □

Remark 8. Kinetic models with concentration phenomena were studied in a number of papers, see for instance [3]. These models are expected to verify neither condition (H:lip) nor even (H':lip), however they exhibit the uniqueness of forward trajectories for empirical measures. It would be natural to apply the MDE theory to prolong solutions past blow-up times.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix

Proof of of Lemma 3.1. The sequence μ_N is tight, that is for every $\epsilon > 0$ there exists a compact set $K_\epsilon \subset \mathbb{R}^n$ such that for all N it holds $\mu_N(\mathbb{R}^n \setminus K_\epsilon) \leqq \epsilon$. This is trivially satisfied taking $K_\epsilon = B(0, R)$. Then, by the Prokhorov Theorem (see Theorem 5.1.3 of [1]), there exists a subsequence converging narrowly to $\mu \in \mathcal{P}_c(\mathbb{R}^n)$, that is $\int f\, \mathrm{d}\mu_N \to \int f\, \mathrm{d}\mu$ for every $f : \mathbb{R}^n \to \mathbb{R}$ continuous and bounded. Since the moments $\int |x| \mathrm{d}\mu_N$ are uniformly bounded, $W(\mu_N, \mu) \to 0$ (see Proposition 7.1.5 of [1]). □

Proof of of Lemma 5.1. The function

$$\liminf_{h\to 0+} \frac{1}{h} W(S_h\mu(s), \mu(s+h))$$

is measurable and bounded. Measurability follows from observing that the incremental ratios are continuous for fixed h and taking the infimum over $h \in \mathbb{Q}$, while boundedness from Lipschitz continuity of the semigroup trajectories and of $\mu(\cdot)$. Define

$$\psi(s) = W(S_{t-s}\mu(s), S_t\mu(0)),$$
$$x(s) = \psi(s) - e^{Kt} \int_0^s \liminf_{h\to 0+} \frac{1}{h} W(S_h\mu(r), \mu(r+h))\, dr.$$

Notice that $\psi(0) = x(0) = 0$ and ψ and x are Lipschitz continuous. Therefore for Rademacher Theorem $\dot{\psi}(s)$ and $\dot{x}(s)$ are defined for almost every s. Moreover, by the Lebesgue Theorem, ψ is approximately continuous for almost every s. Therefore, for almost every s, we have

$$\dot{x}(s) = \dot{\psi}(s) - e^{Kt} \liminf_{h\to 0+} \frac{1}{h} W(S_h\mu(s), \mu(s+h)).$$

Moreover,

$$\begin{aligned}\psi(s+h)-\psi(s) &= W(S_{t-(s+h)}\mu(s+h), S_t\mu(0)) - W(S_{t-s}\mu(s), S_t\mu(0))\\ &\leqq W(S_{t-(s+h)}\mu(s+h), S_{t-s}\mu(s))\\ &= W(S_{t-(s+h)}\mu(s+h), S_{t-(s+h)}S_h\mu(s))\\ &\leqq e^{Kt}W(\mu(s+h), S_h\mu(s)),\end{aligned}$$

which implies

$$\dot{\psi}(s) \leqq e^{Kt}\liminf_{h\to 0+}\frac{1}{h}W(\mu(s+h), S_h\mu(s)).$$

thus $\dot{x}(s) \leqq 0$ for almost every s. Finally, $x(t) \leqq 0$, which proves the Lemma. □

References

1. Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures. 2nd ed., Lectures in Mathematics ETH Zürich. Birkhäuser, Basel (2008)
2. Bardi, M., Capuzzo-Dolcetta, I.: Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA (1997)
3. Bertozzi, A.L., Laurent, T., Rosado, J.: L^p theory for the multidimensional aggregation equation. Commun. Pure Appl. Math. **64**(1), 45–83 (2011)
4. Bonaschi, G. A., Carrillo, J. A., Di Francesco, M., Peletier, M. A.: Equivalence of gradient flows and entropy solutions for singular nonlocal interaction equations in 1D. ESAIM Control Optim. Calc. Var. **21**(2), 414–441 (2015)
5. Bressan, A.: Hyperbolic Systems of Conservation Laws, Volume 20 of Oxford Lecture Series in Mathematics and Its Applications. Oxford University Press, Oxford (2000)
6. Carrillo, J. A., Fornasier, M., Rosado, J., Toscani, G.: Asymptotic flocking dynamics for the kinetic Cucker–Smale model. SIAM J. Math. Anal. **42**(1), 218–236 (2010)
7. Cavagnari, G., Marigonda, A., Piccoli, B.: Optimal synchronization problem for a multi-agent system. Netw. Heterog. Media **12**(2), 277–295 (2017)
8. Cristiani, E., Piccoli, B., Tosin, A.: Modeling self-organization in pedestrians and animal groups from macroscopic and microscopic viewpoints. In: Naldi, G., Pareschi, L., Toscani, G. (eds.) Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences, Modeling and Simulation in Science, Engineering and Technology, pp. 337–364. Birkhäuser, Boston (2010)
9. DiPerna, R.J., Lions, P.-L.: On the Cauchy problem for Boltzmann equations: global existence and weak stability. Ann. Math. **130**(2), 321–366 (1989)
10. Dobrušin, R.L.: Vlasov equations. Funktsional. Anal. i Prilozhen. **13**(2), 48–58 (1979)
11. Golse, F.: The mean-field limit for the dynamics of large particle systems. In: Journées "Équations aux Dérivées Partielles", pp. Exp. No. IX, 47. Univ. Nantes, Nantes (2003)
12. Martin, R.H. Jr.: Nonlinear operators and differential equations in Banach spaces. Wiley-Interscience, New York (1976)
13. Øksendal, B.: Stochastic Differential Equations. Universitext. 6th edn., Springer, Berlin (2003)
14. Poupaud, F., Rascle, M.: Measure solutions to the linear multi-dimensional transport equation with non-smooth coefficients. Commun. Partial Differ. Equ. **22**(1-2), 337–358 (1997)

15. Santambrogio, F.: Optimal Transport for Applied Mathematicians, Volume 87 of Progress in Nonlinear Differential Equations and Their Applications. Birkhäuser, Cham (2015)
16. Schwartz, J.T.: Nonlinear functional analysis. Gordon and Breach Science Publishers, New York (1969). Notes by H. Fattorini, R. Nirenberg and H. Porta, with an additional chapter by Hermann Karcher, Notes on Mathematics and its Applications
17. Villani, C.: Topics in Optimal Transportation, Volume 58 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI (2003)
18. Villani, C.: Optimal Transport. Springer, Berlin (2009)

Benedetto Piccoli
Rutgers University Camden,
Camden,
NJ, USA.
e-mail: piccoli@camden.rutgers.edu

(*Received October 17, 2017 / Accepted March 26, 2019*)
Published online March 30, 2019