MEASURE-THEORETIC LIE BRACKETS FOR NONSMOOTH VECTOR FIELDS

Giulia Cavagnari*

Department of Mathematical Sciences
Rutgers University - Camden
311 N. 5th Street, Camden, NJ 08102, USA

Antonio Marigonda

Department of Computer Science
University of Verona
Strada Le Grazie 15, I-37134 Verona, Italy

Abstract. In this paper we prove a generalization of the classical notion of commutators of vector fields in the framework of measure theory, providing an extension of the set-valued Lie bracket introduced by Rampazzo-Sussmann for Lipschitz continuous vector fields. The study is motivated by some applications to control problems in the space of probability measures, modeling situations where the knowledge of the state is probabilistic, or in the framework of multi-agent systems, for which only a statistical description is available. Tools of optimal transportation theory are used.

1. **Introduction.** In [17] the authors give a generalization of the classical notion of *Lie bracket* (or *commutator*) of two smooth vector fields X, Y, in order to study the commutativity of the flows of two vector fields basically just assuming that the flows are well-defined (e.g., the two vector fields are locally Lipschitz continuous). In this framework, the classical Lie bracket $[X, Y](\cdot)$ appears to be defined only a.e. w.r.t. Lebesgue measure, moreover, as showed with many examples in [17], even at the points where it can be defined, it does not catch all the local features of the two flows.

By means of a suitable construction, in [17] the authors define an object, called *set-valued Lie bracket*, which associates to *every point* of the space a suitable *set* $[X, Y]_{\mathrm{set}}(\cdot)$, which in the classical smooth case is reduced to the usual Lie bracket, and turns out to be the convex hull of the upper Kuratowski limit of the classical Lie brackets (which are defined in a Lebesgue full measure subset, in particular in a dense subset).

They also prove that the basic properties enjoyed by the classical Lie bracket, have their natural counterparts. More precisely, if X, Y are locally Lipschitz vector fields on a manifold M of class C^2, denoted by ϕ_t^X and ϕ_t^Y their flows at time t, it is proved that

2010 *Mathematics Subject Classification.* Primary: 34A60, 49J15; Secondary: 46G10.

Key words and phrases. Lie bracket, optimal transport, set-valued analysis.

The authors have been supported by INdAM-GNAMPA Project 2016: *Stochastic Partial Differential Equations and Stochastic Optimal Transport with Applications to Mathematical Finance*.

* Corresponding author: Giulia Cavagnari.

1. *Asymptotic formula*: for all $q \in M$ we have
$$\lim_{\substack{t,s\to 0\\ t\neq 0,\, s\neq 0}} \operatorname{dist}\left(\frac{\phi^Y_{-t}\circ\phi^X_{-s}\circ\phi^Y_t\circ\phi^X_s(q)-q}{ts}, [X,Y]_{\mathrm{set}}(q)\right)=0.$$
2. *Commutativity of the flows*: $\phi^X_t\circ\phi^Y_s(q)=\phi^Y_s\circ\phi^X_t(q)$ iff $[X,Y]_{\mathrm{set}}(q)=\{0\}$ for all $q\in M$.
3. *Simultaneous flow-box theorem*: if X_i, $i=1,\dots,d$ are locally Lipschitz vector fields on a d-dimensional manifold M of class C^2 satisfying $[X_i,X_j]_{\mathrm{set}}(q)=\{0\}$ for all $q\in M$ $i,j=1,\dots,d$ then, around every point $q\in M$ where the vector fields are independent, there exists a Lipschitz change of coordinates with Lipschitz inverse, sending X_i to the i-th element of the canonical basis e_i of T_qM.

The main ingredient to prove the results of [17] is an *exact* integral formula expressing the difference $\phi^Y_{-t}\circ\phi^X_{-s}\circ\phi^Y_t\circ\phi^X_s(q)-q$ (proved in Lemma 4.5 of [17]). In this context, the term *exact* is used in opposition to *asymptotic*. This integral formula turns out very useful to be handled, and, together with a regularization argument, yields all the main results of the paper.

In [16], these results are applied to give a nonsmooth version of the Frobenius theorem for Lipschitz distributions of vector fields on a manifold. The generalization of the construction of [17] to higher order Lie bracket is not straigthforward, as pointed out in Section 7 of [17], and has been recently proved in the two papers [10], which generalized the exact formula for the single Lie bracket to general nested brackets, and the recent [11].

To make the computations, the authors in [17] make extensively use of the *Agrachev - Gamkrelidze formalism* (AGF), introduced by Agrachev and Gamkrelidze in the '70s. The main idea of this formalism is to embed all the main objects of the flow analysis in a convenient subspace of the space of distributions, using the linear structure of the latter to perform all the computations. In this setting, each point q is represented by a Dirac delta δ_q, a vector field is seen as a differential operator on C^∞_c functions, and the flow of vector fields is seen as the push forward operator. With the AGF formalism, all the computations turn out to be shortned and simplified (see e.g. Section 2 of [17] for an outline of the AGF formalism, containing also a rigorous justification and some examples).

It is well known that, in the classical framework, the vector space $\mathrm{Lie}(\mathscr{F})$ generated by all the vector fields built from a given set $\mathscr{F}$ of vector fields by means of possibly nested Lie bracket, is deeply related to *controllability* properties of the finite-dimensional driftless control-affine systems where the controlled vector fields are the element of $\mathscr{F}$. Roughly speaking, Lie bracket operations enlarge the set of admissible *displacements* that a particle can reach in a given amount of time by following the admissible trajectories of the system, even if, in general, a Lie bracket does not give an admissible *direction* for the system.

Hence, the study of *higher order conditions for attainability* plays an important role. In the classical finite-dimensional setting, Petrov's condition represents a first order requirement on the trajectory and can be interpreted as the request that for each point sufficiently near to the target there exists an admissible trajectory which points *sufficiently towards* the target at the first order, indeed it involves the first order term of at least one admissible trajectory, i.e. an admissible velocity. Since it is a strong condition to be satisfied, it is natural to look for higher order conditions when the first one does not hold, by involving higher order terms of the expansion of the trajectory. It has been studied (see [12]) that these conditions involve Lie bracket of admissible vector fields and they can be viewed as Petrov's conditions of higher order.

In recent papers [4], [5], [6], [7], [8], [9], some control problems in the space of probability measures on $\mathbb{R}^d$ are studied, as a natural generalization of control problems in finite-dimensional space when the initial state is known only up to some uncertainty, or to model situations where the number of agents is so huge to make viable only a statistical (macroscopic) description of the system. In the first case, the time-evolving measure represents our probabilistic knowledge about the state of the particle, while, in the second case, it represents the statistical distribution of the agents.

In all these problems, the dynamics is given by a *controlled* continuity equation, to be satisfied in the sense of distributions, where the current density (which is the control in the problem) is chosen among the Borel selections of a given set-valued map, which can be seen as the underlying microscopic classical dynamics, i.e., the dynamics followed by each agent. In order to study controllability problems in this framework, it turns out to be a natural problem to define some correspondent quantity for the Lie bracket in a measure-theoretic setting by using tools of transport theory. The study of controllability conditions involving measure-theoretic Lie bracket is still an open problem in this setting. We refer the reader to [13, 14] for the study of sufficient conditions granting *small time-local attainability* in finite-dimension.

Our strategy can be summarized as follows: by exploiting the main idea of the AGF formalism, instead of considering Dirac deltas, we consider probability measures on $\mathbb{R}^d$, and define our object as limit (in a suitable topology) of an asymptotic formula like the one considered by Rampazzo-Sussmann, but instead of the *evaluation* at the point q, corresponding to the choice of δ_q, we consider the *push forward* of a probability measure μ along the flow. Under suitable assumptions, we are able to consider the convexified Kuratowski upper limit of this construction as in [17], thus defining a set-valued measure theoretic Lie bracket, which - by construction - satisfies the asymptotic formula and the commutativity property. We notice that this object, being a set of vector-valued measures absolutely continuous w.r.t. μ, has no longer a pointwise meaning, unless the starting measure is purely atomic.

We give also some representation formula, which allows to compare our results with the results of [17], showing that in the case of Dirac deltas, the two constructions agree and, slightly more generally, under the Lipschitz assumptions of [17], the density of each element w.r.t. a general probability measure μ is an L^p_μ-selection of the set-valued Lie bracket defined in [17].

The paper is structured as follows: in Section 2 we review some preliminaries of measure theory and of differential geometry, in Section 3 we introduce the main objects of our study and formulate the main results, in Section 4 we compare our result with the construction of Rampazzo-Sussmann. We conclude providing an example illustrating our construction in Section 5.

2. Preliminaries and notation.

2.1. Measure theory.

Our main reference for this part is [2].

Given $T > 0$, we set $\Gamma_T = C^0([0,T];\mathbb{R}^d)$, which is a separable Banach space when endowed with the usual sup norm. We denote by $e_t : \mathbb{R}^d \times \Gamma_T \to \mathbb{R}^d$ the evaluation operator, defined as $e_t(x,\gamma) = \gamma(t)$. Notice that $e_t(x,\gamma)$ does not depend on x.

Given a family of Banach spaces $\{X_i\}_{i\in I}$, we define the Borel maps $r_i : \prod_{j\in I} X_j \to X_i$, $r_i(x_I) = x_i$ for all $i \in I$. We denote with $\mathrm{Id}_{\mathbb{R}^d}$ the *identity map* on $\mathbb{R}^d$. Given a complete and separable metric space X, let us denote by $\mathscr{P}(X)$ the space of Radon probability measures on X. We have that $\mathscr{P}(X) = (C^0_b(X))'$ and that the w^*-topology on $\mathscr{P}(X)$ induced by this duality is metrizable (for instance by the Prokhorov metric). We will denote by $d_{\mathscr{P}}$ any metric on $\mathscr{P}(X)$ inducing the w^*-topology on $\mathscr{P}(X)$. We will denote by $\mathscr{M}(\mathbb{R}^d;\mathbb{R}^k)$ the set of vector-valued Radon measures on $\mathbb{R}^d$ with values in $\mathbb{R}^k$. If $f : X \to Y$ is a Borel map with X, Y separable metric spaces, the *push forward* of a measure $\mu \in \mathscr{P}(X)$ is the measure $f\sharp\mu \in \mathscr{P}(Y)$ defined by $f\sharp\mu(B) = \mu(f^{-1}(B))$ for every Borel set $B \subseteq Y$.

If X is a separable metric space, we will denote with $\mathrm{Bor}(X)$ the set of Borel maps from X to $\mathbb{R}$ and with $\mathrm{Bor}_b(X)$ the set of bounded Borel maps from X to $\mathbb{R}$.

For the following, let X be a separable Banach space.

Definition 2.1 (Wasserstein distance)**.** Given $\mu_1, \mu_2 \in \mathscr{P}(X)$, $p \geq 1$, we define the *p-Wasserstein distance* between μ_1 and μ_2 by setting

$$W_p(\mu_1,\mu_2) := \left(\inf\left\{\iint_{X\times X} |x_1 - x_2|^p\, d\pi(x_1,x_2) \,:\, \pi \in \Pi(\mu_1,\mu_2)\right\}\right)^{1/p}, \qquad (1)$$

where the set of *admissible transport plans* $\Pi(\mu_1,\mu_2)$ is defined by

$$\Pi(\mu_1,\mu_2) := \Bigg\{\pi \in \mathscr{P}(X\times X) : \begin{array}{l} \pi(A_1\times X) = \mu_1(A_1), \\ \pi(X\times A_2) = \mu_2(A_2), \end{array} \\ \text{for all } \mu_i\text{-measurable sets } A_i,\ i = 1,2\Bigg\}.$$

Definition 2.2 (p-moment)**.** Let $\mu \in \mathscr{P}(X)$, $p \geq 1$. We say that μ has finite p-moment if

$$\mathrm{m}_p(\mu) := \int_X |x|^p\, d\mu(x) < +\infty.$$

Equivalently, we have that μ has p-moment finite if and only if for every $x_0 \in X$ we have

$$\int_X |x - x_0|^p\, d\mu(x) < +\infty.$$

We denote by $\mathscr{P}_p(X)$ the subset of $\mathscr{P}(X)$ consisting of probability measures with finite p-moment.

Definition 2.3 (Uniform integrability)**.** Let $\mathscr{K} \subseteq \mathscr{P}(X)$, $g : X \to [0,+\infty]$ be a Borel function. We say that

1. g is *uniformly integrable with respect to* $\mathscr{K}$ if
$$\lim_{k\to\infty} \sup_{\mu\in\mathscr{K}} \int_{\{x\in X:\, g(x)>k\}} g(x)\, d\mu(x) = 0.$$
2. the set $\mathscr{K}$ *has uniformly integrable p-moments*, $p \geq 1$, if $|x|^p$ is uniformly integrable with respect to $\mathscr{K}$.
3. if $\mathscr{K} = \{\mu_n\}_{n\in\mathbb{N}} \subseteq \mathscr{P}(X)$, $p \geq 1$, $\mu_n \rightharpoonup^* \mu \in \mathscr{P}(X)$, the set $\mathscr{K}$ has uniformly integrable p-moments if and only if
$$\lim_{n\to\infty} \int_X f(x)\, d\mu_n(x) = \int_X f(x)\, d\mu(x),$$

for every continuous function $f : \mathbb{R}^d \to \mathbb{R}$ such that there exist $a, b \geq 0$ and $x_0 \in X$ with $|f(x)| \leq a + b|x - x_0|^p$ for every $x \in X$.

Proposition 1. *$\mathscr{P}_p(X)$ endowed with the p-Wasserstein metric $W_p(\cdot,\cdot)$ is a complete separable metric space. Moreover, given a sequence $\{\mu_n\}_{n\in\mathbb{N}} \subseteq \mathscr{P}_p(X)$ and $\mu \in \mathscr{P}_p(X)$, we have that the following are equivalent*

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

2.2. **Differential geometry.**

Definition 2.4 (Formal bracket)**.** We denote by $\mathrm{Diffeo}(\mathbb{R}^d)$ the set of all diffeomorphisms of $\mathbb{R}^d$. Let $\psi, \varphi \in \mathrm{Diffeo}(\mathbb{R}^d)$ be two diffeomorphisms. We define their *formal bracket* by setting:

$$[\psi, \varphi](x) := \psi \circ \varphi \circ \psi^{-1} \circ \varphi^{-1}(x).$$

Since for every $\psi, \varphi \in \mathrm{Diffeo}(\mathbb{R}^d)$ we have that $[\psi, \varphi] \in \mathrm{Diffeo}(\mathbb{R}^d)$, by iterating the procedure we can construct *formal bracket expressions* by nesting formal brackets of diffeomorphisms. Given a subset $\mathscr{S} \subseteq \mathrm{Diffeo}(\mathbb{R}^d)$, we define the *length* (also *order* or *depth*) of nested formal brackets of elements of $\mathscr{S}$ by induction. If $\varphi \in \mathscr{S}$ is a single diffeomorphism, then $\mathrm{ord}\,(\varphi) = 1$. Otherwise, if A and B are formal bracket expressions of elements of $\mathscr{S}$, we set $\mathrm{ord}\,[A, B] = \mathrm{ord}\,A + \mathrm{ord}\,B$.

Definition 2.5. Let $X : \mathbb{R}^d \to \mathbb{R}^d$ be a locally Lipschitz vector field. Given $x \in \mathbb{R}^d$, we denote by $\phi^X_t(x)$ or $\phi^X(t, x)$ the *flow of X starting from x*, i.e. the (unique) solution of $\dot{x}(s) = X(x(s))$, $x(0) = x$ evaluated at $s = t$. We have $\phi^X(0, x) = x$ and $\dfrac{\partial}{\partial t}\phi^X(t, x) = X(\phi^X(t, x))$.

For t sufficiently small, it is well known that $\phi^X_t(\cdot)$ is a diffeomorphism. Given two C^1-smooth vector fields X, Y, we have that

$$\begin{cases} \dfrac{d}{dt}[\phi^X_t, \phi^Y_t](x)_{|t=0} = 0, \\ \\ \dfrac{d^2}{dt^2}[\phi^X_t, \phi^Y_t](x)_{|t=0} = 2[X, Y](x), \end{cases}$$

where on the right hand side we have the usual *Lie bracket* of vector fields defined in local coordinates by:

$$[X, Y](x) = \langle \nabla Y(x), X(x)\rangle - \langle \nabla X(x), Y(x)\rangle.$$

The correspondence between the first nonvanishing derivative at 0 of flows generating the bracket and the order of the Lie bracket is explained in the following classical result (see e.g., Theorem 1 in [15]).

Theorem 2.6. *Let $k \in \mathbb{N}\setminus\{0, 1\}$, M be a manifold of class C^k, and for $i = 1, \ldots, k$ let $\phi^i : \mathbb{R} \times M \supset U_{\phi^i} \to M$ be a smooth map of class C^k such that*

1. *U_{ϕ^i} is an open neighborhood of $\{0\} \times M$ in $\mathbb{R} \times M$,*
2. *ϕ^i_t is a diffeomorphism of class C^k on its domain,*
3. $\phi^i_0 = \mathrm{Id}_M$ *and* $\dfrac{\partial}{\partial t}\phi^i_t\Big|_{t=0} = X_i \in \mathrm{Vec}_{k-1}(M)$,

where $\mathrm{Vec}_k(M)$ *is the set of vector fields on* M *of class* C^k*. Then for each formal bracket expression* B *of order* k *(w.r.t.* $\mathscr{S}=\{\phi^i : i=1,\dots,k\}$*) we have*

$$\begin{aligned} \frac{\partial^j}{\partial t^j} B(\phi^1_t,\dots,\phi^k_t)\Big|_{t=0} &= 0 \quad \forall 1\le j<k,\\ \frac{1}{k!}\cdot\frac{\partial^k}{\partial t^k} B(\phi^1_t,\dots,\phi^k_t)\Big|_{t=0} &= B(X_1,\dots,X_k), \end{aligned}$$

where the last expression is computed substituting each ϕ^i_t *with* X_i *in* $B(\phi^1_t,\dots,\phi^k_t)$*, and then computing the nested Lie brackets of vector fields.*

3. Measure-theoretic Lie bracket.

In this section we introduce the basic objects of our analysis, proving also the main results of the paper.

Definition 3.1 (Measures associated to a family of transformations)**.** Let $T>0$, $\mathcal{K}\subseteq\mathscr{P}(\mathbb{R}^d)$, $\mu\in\mathrm{cl}_{d_{\mathscr{P}}}\mathcal{K}$ and let $\mathbf{\Psi}_{\mathcal{K}}=\{\Psi_t(\cdot)\}_{t\in[0,T]}$ be a family of maps such that

(D_1) $\Psi_t:\mathbb{R}^d\to\mathbb{R}^d$ is a Borel map for all $t\in[0,T]$;
(D_2) $t\mapsto\Psi_t(x)$ is continuous from $[0,T]$ to $\mathbb{R}^d$;
(D_3) $\Psi_0=\mathrm{Id}_{\mathbb{R}^d}$;
(D_4) $\Psi_t\sharp\mu\in\mathcal{K}$ for all $t\in]0,T]$,

where $\mathrm{cl}_{d_{\mathscr{P}}}$ denotes the closure in the w^*-topology. If $\mathcal{K}=\mathscr{P}(\mathbb{R}^d)$ we will omit the subscript $\mathcal{K}$.

Define the measures $\boldsymbol{\eta}^{\mathbf{\Psi}_{\mathcal{K}}}_\mu\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$ and $\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu,t}\in\mathscr{P}(\mathbb{R}^d\times\mathbb{R}^d)$ by setting for any $t\in]0,T]$, $m\in\mathbb{N}\setminus\{0\}$, $\varphi\in\mathrm{Bor}_b(\mathbb{R}^d\times\Gamma_T)$, $\psi\in\mathrm{Bor}_b(\mathbb{R}^d\times\mathbb{R}^d)$

$$\begin{aligned} \iint_{\mathbb{R}^d\times\Gamma_T}\varphi(x,\gamma)\,d\boldsymbol{\eta}^{\mathbf{\Psi}_{\mathcal{K}}}_\mu(x,\gamma) &:= \int_{\mathbb{R}^d}\varphi(x,\gamma_x)\,d\mu(x),\\ \iint_{\mathbb{R}^d\times\mathbb{R}^d}\psi(x,y)\,d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu,t}(x,y) &:= \int_{\mathbb{R}^d}\psi\left(x,\frac{\Psi_t(x)-x}{t^m}\right)d\mu(x), \end{aligned}$$

where $\gamma_x(\cdot)\in\Gamma_T$ is defined by $\gamma_x(t)=\Psi_t(x)$. Notice that for $\boldsymbol{\eta}^{\mathbf{\Psi}_{\mathcal{K}}}_\mu$-a.e. $(x,\gamma)\in\mathbb{R}^d\times\Gamma_T$ we have $e_0(x,\gamma)=\gamma(0)=x$.

Defined the map $Q^m_t:\mathbb{R}^d\times\Gamma_T\to\mathbb{R}^d$ by

$$Q^m_t(x,\gamma) := \frac{e_t(x,\gamma)-e_0(x,\gamma)}{t^m},$$

we have $\boldsymbol{\eta}^{\mathbf{\Psi}_{\mathcal{K}}}_\mu=\mu\otimes\delta_{\gamma_x}$, $\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu,t}=(e_0\times Q^m_t)\sharp\boldsymbol{\eta}^{\mathbf{\Psi}_{\mathcal{K}}}_\mu=\left(\mathrm{Id}_{\mathbb{R}^d},\dfrac{\Psi_t-\mathrm{Id}_{\mathbb{R}^d}}{t^m}\right)\sharp\mu$, where for $t\neq 0$ the map $e_0\times Q^m_t:\mathbb{R}^d\times\Gamma_T\to\mathbb{R}^d\times\mathbb{R}^d$ is defined as

$$(e_0\times Q^m_t)(x,\gamma)=\left(\gamma(0),\frac{\gamma(t)-\gamma(0)}{t^m}\right).$$

Remark 1. The main motivation for considering a general subset $\mathcal{K}$ of $\mathscr{P}(\mathbb{R}^d)$ comes from applications, where for example we are able to measure only averaged quantities w.r.t. Lebesgue's measure.

We will now provide some estimates on the p-moments of the measures $\boldsymbol{\eta}^{\mathbf{\Psi}_{\mathcal{K}}}_\mu$ and $\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu,t}$ associated to $\mathbf{\Psi}_{\mathcal{K}}$.

Lemma 3.2 (Estimates on moments)**.** *Let* $T>0$*,* $p\ge 1$*,* $\mathcal{K}\subseteq\mathscr{P}_p(\mathbb{R}^d)$*,* $\mu\in\mathrm{cl}_{d_{\mathscr{P}}}\mathcal{K}$ *and let* $\mathbf{\Psi}_{\mathcal{K}}=\{\Psi_t(\cdot)\}_{t\in[0,T]}$ *be a family of maps satisfying assumptions* (D_1)*,* (D_2)*,* (D_3)*,* (D_4)*.*

1. If $\dfrac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m} \in L^p_\mu(\mathbb{R}^d)$*, we have*

$$\mathrm{m}_p(\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu,t}) \le \left(\left\| \frac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m} \right\|_{L^p_\mu} + \mathrm{m}_p^{1/p}(\mu) \right)^p .$$

2. If there exists a Borel map $f : \mathbb{R}^d \to [0, +\infty]$ *with* $|\Psi_t(x) - x| \le f(x)$ *for all* $t \in [0, T]$ *and* $x \in \mathbb{R}^d$*, we have*

$$\mathrm{m}_p(\boldsymbol{\eta}^{\boldsymbol{\Psi}_{\mathcal{K}}}_\mu) \le \mathrm{m}_p(\mu) + \left(\|f\|_{L^p_\mu} + \mathrm{m}_p^{1/p}(\mu) \right)^p .$$

Proof.

1. If $\dfrac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m} \in L^p_\mu(\mathbb{R}^d)$, we have

$$\begin{aligned}
\mathrm{m}_p(\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu,t}) &\le \iint_{\mathbb{R}^d \times \mathbb{R}^d} (|x| + |y|)^p \, d\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu,t}(x,y) \\
&\le \left(\left(\iint_{\mathbb{R}^d \times \mathbb{R}^d} |x|^p \, d\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu,t}(x,y) \right)^{1/p} + \left(\iint_{\mathbb{R}^d \times \mathbb{R}^d} |y|^p \, d\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu,t}(x,y) \right)^{1/p} \right)^p \\
&= \left(\left\| \frac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m} \right\|_{L^p_\mu} + \mathrm{m}_p^{1/p}(\mu) \right)^p .
\end{aligned}$$

2. If there exists a Borel map $f : \mathbb{R}^d \to [0, +\infty]$ with $|\Psi_t(x) - x| \le f(x)$ for all $t \in [0, T]$ and $x \in \mathbb{R}^d$, we have by Monotone Convergence Theorem

$$\begin{aligned}
\mathrm{m}_p(\boldsymbol{\eta}^{\boldsymbol{\Psi}_{\mathcal{K}}}_\mu) &= \iint_{\mathbb{R}^d \times \Gamma_T} (|x|^p + \|\gamma\|^p_\infty) \, d\boldsymbol{\eta}^{\boldsymbol{\Psi}_{\mathcal{K}}}_\mu(x,\gamma) = \int_{\mathbb{R}^d} (|x|^p + \|\gamma_x\|^p_\infty) \, d\mu(x) \\
&\le \mathrm{m}_p(\mu) + \int_{\mathbb{R}^d} (\|\gamma_x - x\|_\infty + |x|)^p \, d\mu(x) \\
&\le \mathrm{m}_p(\mu) + \left(\|f\|_{L^p_\mu} + \mathrm{m}_p^{1/p}(\mu) \right)^p .
\end{aligned}$$

□

We define now a measure-theoretic object related to the limit of $\dfrac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m}$ as $t \to 0^+$.

Definition 3.3 (Measure-theoretic expansion)**.** Let $T > 0$, $m \in \mathbb{N}$, $m \ge 1$, $p \ge 1$, $\mathcal{K} \subseteq \mathscr{P}_p(\mathbb{R}^d)$, $\mu \in \mathrm{cl}_{d_{\mathscr{P}}} \mathcal{K}$ and let $\boldsymbol{\Psi}_{\mathcal{K}} = \{\Psi_t(\cdot)\}_{t \in [0,T]}$ be a family of maps satisfying assumptions (D_1), (D_2), (D_3), (D_4). Define the following set

$$P^p_m(\mu, \boldsymbol{\Psi}_{\mathcal{K}}) := \bigcap_{\substack{\delta > 0 \\ 0 < \sigma < T}} \mathrm{cl}_{W_p} \left\{ \pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu',t} \in \mathscr{P}_p(\mathbb{R}^d \times \mathbb{R}^d) : \begin{array}{l} 0 < t \le \sigma,\ 0 < d_{\mathscr{P}}(\mu', \mu) \le \delta, \\ \mu' \in \mathcal{K} \end{array} \right\},$$

where cl_{W_p} denotes the closure in the W_p-topology, and $\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu',t}$ is defined as in Definition 3.1.

We notice that

1. $P^p_m(\mu, \boldsymbol{\Psi}_{\mathcal{K}})$ is W_p-closed.
2. $\pi \in P^p_m(\mu, \boldsymbol{\Psi}_{\mathcal{K}})$ if and only if there exist $\{t_i\}_{i \in \mathbb{N}} \subseteq]0, T]$ and $\{\mu^{(i)}\}_{i \in \mathbb{N}} \subseteq \mathcal{K}$ such that $t_i \to 0$, $\mu^{(i)} \rightharpoonup^* \mu$, and $W_p\left(\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}, \pi\right) \to 0$ as $i \to +\infty$.

3. For any $\pi \in P^p_m(\mu, \boldsymbol{\Psi}_{\mathcal{K}})$ we have that $r_1\sharp\pi = \mu$, indeed, given $t_i \to 0^+$, $\{\mu^{(i)}\}_{i\in\mathbb{N}} \subseteq \mathcal{K}$, $\mu^{(i)} \rightharpoonup^* \mu$ such that $W_p(\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}, \pi) \to 0$, we have in particular $r_1\sharp\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i} \rightharpoonup^* r_1\sharp\pi$, since convergence in W_p implies w^*-convergence, and $r_1\sharp\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i} = \mu^{(i)} \rightharpoonup^* \mu$.

We can disintegrate each element $\pi \in P^p_m(\mu, \boldsymbol{\Psi}_{\mathcal{K}})$ with respect to r_1 obtaining a family of probability measures $\{\sigma^\pi_x\}_{x\in\mathbb{R}^d}$ which is μ-a.e. uniquely defined and satisfies $\pi = \mu \otimes \sigma^\pi_x$. Thus we can define the set

$$V^p_m(\mu, \boldsymbol{\Psi}_{\mathcal{K}}) := \left\{ V \in L^p_\mu(\mathbb{R}^d; \mathbb{R}^d) : V(x) = \int_{\mathbb{R}^d} y\, d\sigma^\pi_x(y),\ \pi = \mu \otimes \sigma^\pi_x \in P^p_m(\mu, \boldsymbol{\Psi}_{\mathcal{K}}) \right\}.$$

Remark 2. Roughly speaking, the second marginal of each element $\pi \in P^p_m(\mu, \boldsymbol{\Psi}_{\mathcal{K}})$ represents a limit point of the vector valued measure $\dfrac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m}\mu'$ for $\mu' \in \mathcal{K}$ converging to μ and $t \to 0^+$. To recover an object defined pointwise μ-a.e., we take its barycenter, obtaining the map V.

The set of vector-valued measures $\{V\mu : V \in V^p_m(\mu, \boldsymbol{\Psi}_{\mathcal{K}})\}$ will be the object generalizing the asymptotic behaviour of the vector-valued measure $\dfrac{\Psi_t - \mathrm{Id}_{\mathbb{R}^d}}{t^m}\mu'$, in the sense precised below.

Lemma 3.4 (Interpretation)**.** *Let $T > 0$, $m \in \mathbb{N}$, $m \geq 1$, $p \geq 2$, $\mathcal{K} \subseteq \mathscr{P}_p(\mathbb{R}^d)$, $\mu \in \mathrm{cl}_{d_{\mathscr{P}}}\mathcal{K}$ and let $\boldsymbol{\Psi}_{\mathcal{K}} = \{\Psi_t(\cdot)\}_{t\in[0,T]}$ be a family of maps satisfying assumptions (D_1), (D_2), (D_3), (D_4). Then if $V \in V^p_m(\mu, \boldsymbol{\Psi}_{\mathcal{K}})$ there exist $\{\mu^{(i)}\}_{i\in\mathbb{N}} \subseteq \mathcal{K}$ and $\{t_i\}_{i\in\mathbb{N}} \subseteq]0, T]$ such that $\mu^{(i)} \rightharpoonup^* \mu$, $t_i \to 0^+$ and*

$$\lim_{i\to+\infty} \frac{\Psi_{t_i}\sharp\mu^{(i)} - \mu^{(i)}}{t_i^m} = -\mathrm{div}(V\mu),$$

in the sense of distributions.

Proof. Let $V \in V^p_m(\mu, \boldsymbol{\Psi}_{\mathcal{K}})$. There exist sequences $\{t_i\}_{i\in\mathbb{N}} \subseteq]0, T]$ and $\{\mu^{(i)}\}_{i\in\mathbb{N}} \subseteq \mathcal{K}$, and a family of probability measures $\{\sigma_x\}_{x\in\mathbb{R}^d}$ uniquely defined for μ-a.e. $x \in \mathbb{R}^d$ such that $\mu^{(i)} \rightharpoonup^* \mu$, $t_i \to 0^+$ and, set $\pi := \mu \otimes \sigma_x$, we have $W_p(\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}, \pi) \to 0^+$ and

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} \varphi(x) y\, d\pi(x,y) = \int_{\mathbb{R}^d} \varphi(x) V(x)\, d\mu,$$

for any $\varphi \in C^\infty_c(\mathbb{R}^d)$.

For any $\varphi \in C^\infty_c(\mathbb{R}^d)$ we set $R_\varphi : [0, +\infty[\times\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$,

$$R_\varphi(t, x, y) := \varphi(x + t^m y) - \varphi(x) - \langle \nabla\varphi(x), t^m y \rangle,$$

and, recalling the smoothness of φ, we have

$$\frac{|R_\varphi(t, x, y)|}{t^m} \leq t^m \|D^2\varphi\|_\infty |y|^2 \chi_{\mathrm{supp}\,\varphi}(x).$$

In particular, for i sufficiently large we obtain

$$\begin{aligned}\iint_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|R_\varphi(t_i, x, y)|}{t_i^m}\, d\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y) &\leq t_i^m \|D^2\varphi\|_\infty \iint_{\mathbb{R}^d\times\mathbb{R}^d} |y|^2\, d\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y)\\ &\leq t_i^m \|D^2\varphi\|_\infty \mathrm{m}_2(\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i})\\ &\leq t_i^m \|D^2\varphi\|_\infty (1 + \mathrm{m}_p(\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i})).\end{aligned}$$

Hence we have

$$\begin{aligned}\langle\varphi,\frac{\Psi_{t_i}\sharp\mu^{(i)}-\mu^{(i)}}{t_i^m}\rangle &= \frac{1}{t_i^m}\left[\int_{\mathbb{R}^d}\varphi(x)\,d\Psi_{t_i}\sharp\mu^{(i)}(x)-\int_{\mathbb{R}^d}\varphi(x)\,d\mu^{(i)}(x)\right]\\ &= \frac{1}{t_i^m}\int_{\mathbb{R}^d}\left[\varphi\left(x+t_i^m\frac{\Psi_{t_i}(x)-x}{t_i^m}\right)-\varphi(x)\right]\,d\mu^{(i)}(x)\\ &= \frac{1}{t_i^m}\iint_{\mathbb{R}^d\times\mathbb{R}^d}\left[\varphi\left(x+t_i^m y\right)-\varphi(x)\right]\,d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y)\\ &= \iint_{\mathbb{R}^d\times\mathbb{R}^d}\langle\nabla\varphi(x),y\rangle\,d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y)+\iint_{\mathbb{R}^d\times\mathbb{R}^d}\frac{R_\varphi(t_i,x,y)}{t_i^m}\,d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y)\\ &\le \iint_{\mathbb{R}^d\times\mathbb{R}^d}\langle\nabla\varphi(x),y\rangle\,d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y)+t_i^m\|D^2\varphi\|_\infty(1+\mathrm{m}_p(\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i})).\end{aligned}$$

Taking the limit for $i\to+\infty$, and recalling that $\mathrm{m}_p(\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{t_i,\mu^{(i)}})$ is uniformly bounded since $W_p(\pi,\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i})\to 0$, we have

$$\lim_{i\to+\infty}\langle\varphi,\frac{\Psi_{t_i}\sharp\mu^{(i)}-\mu^{(i)}}{t_i^m}\rangle\le\int_{\mathbb{R}^d}\langle\nabla\varphi(x),V(x)\rangle\,d\mu(x)=-\langle\varphi,\mathrm{div}(V\mu)\rangle,$$

which concludes the proof by the arbitrariness of $\varphi\in C_c^\infty(\mathbb{R}^d)$. □

Corollary 1. *In the same assumptions of Lemma 3.4, assume that*

$$\lim_{t\to 0}\left\|\frac{\Psi_t-\mathrm{Id}_{\mathbb{R}^d}}{t^m}\right\|_{L^p_\mu}=0.$$

Then

1. $\displaystyle\lim_{t\to 0}\frac{W_p(\Psi_t\sharp\mu,\mu)}{t^m}=0$;
2. *for every* $\varphi\in\mathrm{Lip}(\mathbb{R}^d)$ *we have*

$$\lim_{t\to 0}\int_{\mathbb{R}^d}\frac{\varphi\circ\Psi_t(x)-\varphi(x)}{t^m}\,d\mu(x)=0.$$

Proof. The result comes immediately, since we have

$$\begin{aligned}\left(\frac{W_p(\Psi_t\sharp\mu,\mu)}{t^m}\right)^p &\le \iint_{\mathbb{R}^d\times\mathbb{R}^d}\frac{|\Psi_t(x)-x|^p}{t^{pm}}\,d\mu(x)\\ &= \iint_{\mathbb{R}^d\times\mathbb{R}^d}|y|^p\,d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu,t}(x,y),\\ \left|\int_{\mathbb{R}^d}\frac{\varphi\circ\Psi_t(x)-\varphi(x)}{t^m}\,d\mu(x)\right|^p &= \left|\iint_{\mathbb{R}^d\times\mathbb{R}^d}\frac{\varphi(x+t^m y)-\varphi(x)}{t^m}\,d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu,t}(x,y)\right|^p\\ &\le \mathrm{Lip}^p(\varphi)\iint_{\mathbb{R}^d\times\mathbb{R}^d}|y|^p\;d\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu,t}(x,y),\end{aligned}$$

and in both cases the right hand side tends to 0 by assumption. □

We are going to provide now a sufficent condition ensuring that the above defined sets are nonempty.

Lemma 3.5 (Nontriviality)**.** *Let $T>0$, $m\in\mathbb{N}$, $m\ge 1$, $p\ge 1$, $\mathcal{K}\subseteq\mathscr{P}_p(\mathbb{R}^d)$, $\mu\in\mathrm{cl}_{d_{\mathscr{P}}}\mathcal{K}$ and let $\mathbf{\Psi}_{\mathcal{K}}=\{\Psi_t(\cdot)\}_{t\in[0,T]}$ be a family of maps satisfying assumptions (D_1), (D_2), (D_3), (D_4).*

1. *$P^p_m(\mu,\boldsymbol{\Psi}_\mathcal{K})\neq\emptyset$ if and only if $V^p_m(\mu,\boldsymbol{\Psi}_\mathcal{K})\neq\emptyset$. More precisely, if $\pi=\mu\otimes\sigma^\pi_x\in P^p_m(\mu,\boldsymbol{\Psi}_\mathcal{K})$ then the map defined as*

$$V(x)=\int_{\mathbb{R}^d} y\,d\sigma^\pi_x(y)$$

belongs to $L^p_\mu(\mathbb{R}^d;\mathbb{R}^d)$.

2. *Assume that*

$$\liminf_{\substack{W_p(\mu',\mu)\to 0\\ \mu'\in\mathcal{K}\\ t\to 0^+}}\frac{\|\Psi_t-\mathrm{Id}_{\mathbb{R}^d}\|_{L^p_{\mu'}}}{t^m}=:C<+\infty,$$

then $P^p_m(\mu,\boldsymbol{\Psi}_\mathcal{K})\neq\emptyset$, which implies also $V^p_m(\mu,\boldsymbol{\Psi}_\mathcal{K})\neq\emptyset$.

Proof.

1. Given $\pi\in P^p_m(\mu,\boldsymbol{\Psi}_\mathcal{K})$ as in the statement, we estimate the L^p_μ-norm of $V(\cdot)$ by applying Jensen's inequality

$$\begin{aligned}\|V\|^p_{L^p_\mu}&=\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d} y\,d\sigma^\pi_x(y)\right|^p d\mu(x)\le\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|y|^p\,d\sigma^\pi_x(y)\right)d\mu(x)\\&=\iint_{\mathbb{R}^d\times\mathbb{R}^d}|y|^p\,d\pi(x,y)\le\mathrm{m}_p(\pi)<+\infty.\end{aligned}$$

Then we have that $V\in V^p_m(\mu,\boldsymbol{\Psi}_\mathcal{K})$, which turns out to be nonempty. The converse is trivial.

2. Let $\{\mu^{(i)}\}_{i\in\mathbb{N}}$ be a sequence in $\mathcal{K}$, $\{t_i\}_{i\in\mathbb{N}}\subseteq]0,T]$ be such that

$$W_p(\mu^{(i)},\mu)\to 0,\qquad t_i\to 0^+,\qquad \lim_{i\to+\infty}\frac{\|\Psi_{t_i}-\mathrm{Id}_{\mathbb{R}^d}\|_{L^p_{\mu^{(i)}}}}{t_i^m}=C.$$

Since $W_p(\mu^{(i)},\mu)\to 0$, we have that there exists $C'>0$ such that $\mathrm{m}_p^{1/p}(\mu^{(i)})\le C'$ for all $i\in\mathbb{N}$. Define $\pi^{\boldsymbol{\Psi}_\mathcal{K},m}_{\mu^{(i)},t_i}$ as in Definition 3.1, and notice that, by assumption, for i sufficiently large we have $\left\|\dfrac{\Psi_{t_i}-\mathrm{Id}_{\mathbb{R}^d}}{t_i^m}\right\|_{L^p_{\mu^{(i)}}}\le C+1$. Thus, according to Lemma 3.2 item (1),

$$\mathrm{m}_p(\pi^{\boldsymbol{\Psi}_\mathcal{K},m}_{\mu^{(i)},t_i})\le\left(\left\|\frac{\Psi_{t_i}-\mathrm{Id}_{\mathbb{R}^d}}{t_i^m}\right\|_{L^p_{\mu^{(i)}}}+\mathrm{m}_p^{1/p}(\mu^{(i)})\right)^p\le(C+C'+1)^p.$$

In particular, according to Remark 5.1.5 in [2], up to passing to a subsequence, we can assume that there exists $\pi_\infty\in\mathscr{P}_p(\mathbb{R}^d\times\mathbb{R}^d)$ such that $W_p(\pi^{\boldsymbol{\Psi}_\mathcal{K},m}_{\mu^{(i)},t_i},\pi_\infty)\to 0$, yielding $\pi_\infty\in P^p_m(\mu,\boldsymbol{\Psi}_\mathcal{K})$ and $\mathrm{m}_p(\pi_\infty)\le(C+C'+1)^p$. To conclude, it is enough to apply the previous item. □

The following localization result allows us to restrict our attention in the computation of $P^p_m(\mu,\boldsymbol{\Psi}_\mathcal{K})$ just on the measures supported in a neighborhood of $\operatorname{supp}\mu$.

Lemma 3.6 (Localization). *Let $T>0$, $m\in\mathbb{N}$, $m\ge 1$, $p\ge 1$, $\mathcal{K}\subseteq\mathscr{P}_p(\mathbb{R}^d)$ such that if $\mu_1\in\mathcal{K}$ and $\mu_2\ll\mu_1$, then also $\mu_2\in\mathcal{K}$. Let $\mu\in\mathrm{cl}_{d_\mathscr{P}}\mathcal{K}$ and $\boldsymbol{\Psi}_\mathcal{K}=\{\Psi_t(\cdot)\}_{t\in[0,T]}$ be a family of maps satisfying assumptions (D_1), (D_2), (D_3), (D_4). Then we have*

$$
\begin{aligned}
&P^p_m(\mu, \boldsymbol{\Psi}_{\mathcal{K}})\\
&= \bigcap_{\substack{0<\delta<T\\ W\subseteq\mathbb{R}^d \text{ open}\\ \operatorname{supp}\mu\subseteq W}} \mathrm{cl}_{W_p}\left\{\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu',t} \in \mathscr{P}_p(\mathbb{R}^d\times\mathbb{R}^d):\ \begin{array}{l} 0<d_{\mathscr{P}}(\mu',\mu)\le\delta,\ \mu'\in\mathcal{K}\\ 0<t\le\delta,\ \operatorname{supp}\mu'\subseteq\overline{W}\end{array}\right\},
\end{aligned}
$$

Proof. The inclusion $\supseteq$ holds trivially true. We prove the converse inclusion. Let $\pi \in P^p_m(\mu, \boldsymbol{\Psi}_{\mathcal{K}})$, in particular there exists $\{\mu^{(i)}\}_{i\in\mathbb{N}} \subseteq \mathcal{K}$, $\mu^{(i)} \rightharpoonup^* \mu$, $t_i \to 0^+$ such that $W_p(\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}, \pi) \to 0$. Let $W \subseteq \mathbb{R}^d$ be open and such that $\operatorname{supp}\mu \subseteq W$. Define $\varphi_W \in C^\infty_c(\mathbb{R}^d)$ such that $0 \le \varphi_W(\mathbb{R}^d) \le 1$, $\varphi_W(x) \equiv 1$ for all $x \in \operatorname{supp}(\mu)$ and $\operatorname{supp}\varphi_W \subseteq W$. Set

$$
\mu^{(i)}_W := \frac{\varphi_W \mu^{(i)}}{\displaystyle\int_{\mathbb{R}^d} \varphi_W(x)\, d\mu^{(i)}(x)} \in \mathcal{K},
$$

by hypothesis. Let $\psi \in C^0_b(\mathbb{R}^d)$. Then, since $\psi\varphi_W \in C^0_b(\mathbb{R}^d)$, we have

$$
\begin{aligned}
\lim_{i\to+\infty} \int_{\mathbb{R}^d} \psi(x)\, d\mu^{(i)}_W(x) &= \lim_{i\to+\infty} \frac{\displaystyle\int_{\mathbb{R}^d} \psi(x)\varphi_W(x)\, d\mu^{(i)}(x)}{\displaystyle\int_{\mathbb{R}^d} \varphi_W(x)\, d\mu^{(i)}(x)} = \frac{\displaystyle\int_{\mathbb{R}^d} \psi(x)\varphi_W(x)\, d\mu(x)}{\displaystyle\int_{\mathbb{R}^d} \varphi_W(x)\, d\mu(x)}\\
&= \int_{\mathbb{R}^d} \psi(x)\, d\mu(x),
\end{aligned}
$$

since $\varphi_W \equiv 1$ on $\operatorname{supp}\mu$. Thus we have $\mu^{(i)}_W \rightharpoonup^* \mu$ for all $0<\delta<T$. For any $0<\delta<T$ we have

$$
\lim_{i\to+\infty} \int_{\mathbb{R}^d} \varphi_W(x)\, d\mu^{(i)}(x) = 1,
$$

thus there exists $i_\delta \in \mathbb{N}$ such that $\displaystyle\int_{\mathbb{R}^d} \varphi_W(x)\, d\mu^{(i)}(x) \ge \frac{1}{2}$, for all $i \ge i_\delta$.

This implies $\mathrm{m}_p(\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)}_W,t_i}) \le 2\mathrm{m}_p(\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i})$, for all $i \ge i_\delta$, by Monotone Convergence Theorem. Since by assumption $W_p(\pi, \pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}) \to 0$, we have that $\mathrm{m}_p(\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i})$ is uniformly bounded, and so, up to passing to a non relabeled subsequence, we have that there exists $\pi' \in \mathscr{P}_p(\mathbb{R}^d\times\mathbb{R}^d)$ such that $W_p(\pi', \pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)}_W,t_i}) \to 0$ as $i \to +\infty$. To prove that $\pi = \pi'$, which will conclude the proof by the arbitrariness of W and δ, it is enough to show that $d_{\mathscr{P}}(\pi, \pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)}_W,t_i}) \to 0$. Indeed, for any $\psi \in C^0_b(\mathbb{R}^d\times\mathbb{R}^d)$ we have

$$
\begin{aligned}
\lim_{i\to+\infty} \iint_{\mathbb{R}^d\times\mathbb{R}^d} \psi(x,y)\, d\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)}_W,t_i}(x,y) &= \lim_{i\to+\infty} \int_{\mathbb{R}^d} \psi\left(x, \frac{\Psi_{t_i}(x)-x}{t_i^m}\right) d\mu^{(i)}_W(x)\\
&= \lim_{i\to+\infty} \frac{\displaystyle\int_{\mathbb{R}^d} \varphi_W(x)\psi\left(x, \frac{\Psi_{t_i}(x)-x}{t_i^m}\right) d\mu^{(i)}(x)}{\displaystyle\int_{\mathbb{R}^d} \varphi_W(x)\, d\mu^{(i)}}\\
&= \lim_{i\to+\infty} \int_{\mathbb{R}^d} \varphi_W(x)\psi\left(x,y\right)\, d\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y)\\
&= \int_{\mathbb{R}^d} \varphi_W(x)\psi\left(x,y\right)\, d\pi(x,y)
\end{aligned}
$$

$$= \int_{\mathbb{R}^d} \psi(x,y)\, d\pi(x,y),$$

and so $W_p(\pi, \pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{t_i,\mu^{(i)}_W}) \to 0$ as $i \to +\infty$, $\{\mu^{(i)}\}_{i\in\mathbb{N}} \subseteq \mathcal{K}$ and $\operatorname{supp}\mu^{(i)}_W \subseteq \overline{W}$ for all $i \in \mathbb{N}$. □

We will now provide some representation formulas for the function on $V^p_m(\mu, \mathbf{\Psi}_{\mathcal{K}})$, proving also some refinement under additional assumptions. These will be used to establish a comparison with the set-valued Lie brackets defined by Rampazzo-Sussmann in [17].

Definition 3.7. Let $T > 0$, $m \in \mathbb{N}$, $m \geq 1$, $\mathcal{K} \subseteq \mathscr{P}(\mathbb{R}^d)$, $\mu \in \mathrm{cl}_{d_{\mathscr{P}}}\mathcal{K}$, $D \subseteq \mathbb{R}^d$, and let $\mathbf{\Psi}_{\mathcal{K}} = \{\Psi_t(\cdot)\}_{t\in[0,T]}$ be a family of maps satisfying assumptions (D_1), (D_2), (D_3), (D_4). For every $\delta > 0$, $0 < \sigma < T$, and $z \in \mathbb{R}^d$, define the sets

$$S^{\sigma,\delta}_{m,D}(z) := \left\{ \frac{\Psi_t(y) - y}{t^m} : 0 < t < \sigma,\, y \in B(z,\delta) \cap D \right\},$$

$$K^{\sigma,\delta}_{m,D}(z) := \begin{cases} \overline{\mathrm{co}}\, S^{\sigma,\delta}_{m,D}(z), & \text{if } S^{\sigma,\delta}_{m,D}(z) \neq \emptyset, \\ \\ \emptyset, & \text{otherwise,} \end{cases}$$

$$E_{m,D} := \{z \in D : \text{ there exists } \sigma_z, \delta_z > 0 \text{ such that } S^{\sigma_z,\delta_z}_{m,D}(z) \text{ is bounded}\}.$$

If $D = \mathbb{R}^d$ we will write $S^{\sigma,\delta}_m(z)$, $K^{\sigma,\delta}_m(z)$, thus omitting D.

Theorem 3.8 (Representation formula)**.** *Let $T > 0$, $m \in \mathbb{N}$, $m \geq 1$, $p \geq 1$, $\mathcal{K} \subseteq \mathscr{P}_p(\mathbb{R}^d)$, $\mu \in \mathrm{cl}_{d_{\mathscr{P}}}\mathcal{K}$ and let $\mathbf{\Psi}_{\mathcal{K}} = \{\Psi_t(\cdot)\}_{t\in[0,T]}$ be a family of maps satisfying assumptions (D_1), (D_2), (D_3), (D_4). Let $D \subseteq \mathbb{R}^d$ and assume that the following condition holds*

(H_1) $\mu'(D) = 1$ for all $\mu' \in \mathcal{K}$.

Then if $V \in V^p_m(\mu, \mathbf{\Psi}_{\mathcal{K}})$ we have

$$V(z) \in \bigcap_{\sigma,\delta>0} K^{\sigma,\delta}_{m,D}(z), \qquad \textit{for } \mu\textit{-a.e. } z \in \mathbb{R}^d, \tag{2}$$

$$V(z) \in \mathrm{co} \bigcap_{\sigma,\delta>0} \overline{S^{\sigma,\delta}_{m,D}(z)}, \qquad \textit{for } \mu\textit{-a.e. } z \in E_{m,D}. \tag{3}$$

Proof. Let $V \in V^p_m(\mu, \mathbf{\Psi}_{\mathcal{K}})$. There exist sequences $\{t_i\}_{i\in\mathbb{N}} \subseteq]0,T]$, $t_i \to 0^+$ and $\{\mu^{(i)}\}_{i\in\mathbb{N}} \subseteq \mathcal{K}$, $\mu^{(i)} \rightharpoonup^* \mu$, and a family of probability measures $\{\xi_x\}_{x\in\mathbb{R}^d}$ uniquely defined for μ-a.e. $x \in \mathbb{R}^d$ such that denoted by $\pi := \mu \otimes \xi_x$, we have $W_p(\pi^{\mathbf{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}, \pi) \to 0$ and $V(x) = \displaystyle\int_{\mathbb{R}^d} y\,d\xi_x(y)$ for μ-a.e. $x \in \mathbb{R}^d$.

For any $\sigma \in]0,T]$ we define a set-valued map $G_\sigma : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ by taking

$$G_\sigma(x) := \bigcap_{\delta>0} K^{\sigma,\delta}_{m,D}(x).$$

Notice that $\operatorname{dom} G_\sigma \supseteq D$. This set-valued map has closed graph, indeed, let $\{x_n\}_{n\in\mathbb{N}}$, $\{y_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^d$, $x, y \in \mathbb{R}^d$ be such that $x_n \to x$, $y_n \to y$, $y_n \in G_\sigma(x_n)$ for all $n \in \mathbb{N}$. Fix $\delta > 0$ and let $n_\delta > 0$ be such that $|x_n - x| < \delta$ for all $n \geq n_\delta$. For every $\delta' > 0$ and $n \geq n_\delta$ we have that

$$y_n \in \overline{\mathrm{co}}\, S^{\sigma,\delta'}_{m,D}(x_n) \subseteq \overline{\mathrm{co}}\, S^{\sigma,\delta'+|x_n-x|}_{m,D}(x) \subseteq \overline{\mathrm{co}}\, S^{\sigma,\delta'+\delta}_{m,D}(x).$$

By passing to the limit as $n \to +\infty$ we have $y \in \overline{\mathrm{co}}\, S^{\sigma,\delta'+\delta}_{m,D}(x)$ for all $\delta', \delta > 0$, and then by taking the intersection on $\delta, \delta' > 0$ we have $y \in G_\sigma(x)$.

Since G_σ has closed graph, the map $g_\sigma(x,y) := I_{G_\sigma(x)}(y)$ is l.s.c. and nonnegative (set $I_\emptyset \equiv +\infty$), moreover $g_\sigma(x,\cdot)$ is convex for all $x \in \mathbb{R}^d$.

By Jensen's inequality we have

$$\begin{aligned}\int_{\mathbb{R}^d} g_\sigma(x, V(x))\, d\mu(x) &= \int_{\mathbb{R}^d} g_\sigma\left(x, \int_{\mathbb{R}^d} y\, d\xi_x(y)\right) d\mu(x)\\ &\le \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} g_\sigma(x,y)\, d\xi_x(x)\, d\mu(x)\\ &= \iint_{\mathbb{R}^d\times\mathbb{R}^d} g_\sigma(x,y)\, d\pi(x,y).\end{aligned}$$

Recalling Lemma 5.1.7 in [2], by l.s.c. of $g_\sigma(\cdot,\cdot)$ we have

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} g_\sigma\,(x,y)\; d\pi(x,y) \le \liminf_{i\to+\infty} \iint_{\mathbb{R}^d\times\mathbb{R}^d} g_\sigma(x,y)\, d\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y).$$

We obtain

$$\begin{aligned}\int_{\mathbb{R}^d} g_\sigma(x, V(x))\, d\mu(x) &\le \iint_{\mathbb{R}^d\times\mathbb{R}^d} g_\sigma\,(x,y)\; d\pi(x,y)\\ &\le \liminf_{i\to+\infty} \iint_{\mathbb{R}^d\times\mathbb{R}^d} g_\sigma(x,y)\, d\pi^{\boldsymbol{\Psi}_{\mathcal{K}},m}_{\mu^{(i)},t_i}(x,y)\\ &= \liminf_{i\to+\infty} \int_{\mathbb{R}^d} g_\sigma\left(x, \frac{\Psi_{t_i}(x)-x}{t_i^m}\right) d\mu^{(i)}(x).\end{aligned}$$

Since there exists $i_\sigma \ge 0$ such that $t_i \le \sigma$ for all $i \ge i_\sigma$, then for any $x \in D$ we have

$$g_\sigma\left(x, \frac{\Psi_{t_i}(x)-x}{t_i^m}\right) = 0, \qquad \text{for all } i \ge i_\sigma. \tag{4}$$

This implies

$$\int_{\mathbb{R}^d} g_\sigma(x, V(x))\, d\mu(x) \le \liminf_{i\to+\infty} \int_{\mathbb{R}^d\setminus D} g_\sigma\left(x, \frac{\Psi_{t_i}(x)-x}{t_i^m}\right) d\mu^{(i)}(x).$$

Thus, since by hypothesis $\mu^{(i)}(D) = 1$ for all $i \in \mathbb{N}$, we have $g_\sigma(x, V(x)) = 0$ for μ-a.e. $x \in \mathbb{R}^d$. Recalling the arbitrariness of $\sigma > 0$, for μ-a.e. $x \in \mathbb{R}^d$

$$V(x) \in \bigcap_{\sigma>0}\bigcap_{\delta>0} K^{\sigma,\delta}_{m,D}(x) = \bigcap_{\sigma,\delta>0} K^{\sigma,\delta}_{m,D}(x),$$

which proves (2).

Since

$$\bigcap_{\sigma,\delta>0} K^{\sigma,\delta}_{m,D}(z) \supseteq \mathrm{co} \bigcap_{\sigma,\delta>0} \overline{S^{\sigma,\delta}_{m,D}(z)}$$

for all $z \in \mathbb{R}^d$, to prove (3) we must show that equality holds when $z \in E_{m,D}$. By definition of $E_{m,D}$, there exist $\delta_z > 0$ and $0 < \sigma_z < T$ such that $S^{\sigma,\delta}_{m,D}(z)$ is bounded for all $0 < \sigma < \sigma_z$ and $0 < \delta < \delta_z$, so we can find a sequence $t_i \to 0^+$, a sequence $y_i \to z$, and a vector $\xi(z) \in \mathbb{R}^d$ such that

$$\lim_{i\to\infty} \frac{\Psi_{t_i}(y_i) - y_i}{t_i^m} = \xi(z),$$

and, by construction, we have $\xi(z) \in \overline{S^{\sigma,\delta}_{m,D}(z)}$ for all $\sigma, \delta > 0$.

Thus $\xi(z) \in \bigcap_{\sigma,\delta>0} \overline{S^{\sigma,\delta}_{m,D}(z)}$, and so the set $\mathrm{co} \bigcap_{\sigma,\delta>0} \overline{S^{\sigma,\delta}_{m,D}(z)}$ is closed, convex, and nonempty.

Assume by contradiction that $w \in \bigcap_{\sigma,\delta>0} K^{\sigma,\delta}_{m,D}(z) \setminus \mathrm{co} \bigcap_{\sigma,\delta>0} \overline{S^{\sigma,\delta}_{m,D}(z)}$. By Hahn-Banach separation theorem, there exist $\varepsilon > 0$ and $\bar{v} \in \mathbb{R}^d$ such that

$$\langle \bar{v}, w\rangle \geq \langle \bar{v}, \xi\rangle + \varepsilon, \text{ for all } \xi \in \mathrm{co} \bigcap_{\sigma,\delta>0} \overline{S^{\sigma,\delta}_{m,D}(z)},$$

in particular we have

$$\langle \bar{v}, w\rangle \geq \langle \bar{v}, \xi\rangle + \varepsilon, \text{ for all } \xi \in \bigcap_{\sigma,\delta>0} \overline{S^{\sigma,\delta}_{m,D}(z)}.$$

On the other hand, we have that

$$w \in \bigcap_{\sigma,\delta>0} \overline{\mathrm{co}}\, S^{\sigma,\delta}_{m,D}(z)$$

implies that for all $v \in \mathbb{R}^d$, $\sigma, \delta > 0$ we have

$$\langle v, w\rangle \leq \sup_{p \in \overline{\mathrm{co}}\, S^{\sigma,\delta}_{m,D}(z)} \langle v, p\rangle = \sup_{p \in S^{\sigma,\delta}_{m,D}(z)} \langle v, p\rangle,$$

so for every sequence $\sigma_i \to 0^+$ and $\delta_i \to 0$ we choose $\xi_i \in S^{\sigma_i,\delta_i}_{m,D}(z)$ such that

$$\sup_{p \in S^{\sigma_i,\delta_i}_{m,D}(z)} \langle v, p\rangle \leq \langle v, \xi_i\rangle + \frac{1}{i}.$$

Up to passing to a subsequence, we can assume that $\xi_i \to \bar{\xi}$. By construction, we have that $\bar{\xi} \in \overline{S^{\sigma,\delta}_{m,D}(z)}$ for all $\sigma, \delta > 0$, and

$$\langle v, w\rangle \leq \langle v, \bar{\xi}\rangle,$$

contradicting the fact that $\langle \bar{v}, w\rangle \geq \langle \bar{v}, \xi\rangle + \varepsilon$ for all $\xi \in \bigcap_{\sigma,\delta>0} \overline{S^{\sigma,\delta}_{m,D}(z)}$. □

Remark 3. In the case in which the maps $\mathbf{\Psi}_{\mathcal{K}} \ni \Psi_t : \mathbb{R}^d \to \mathbb{R}^d$ are continuous for all $t \in [0, T]$, then Theorem 3.8 holds also if instead of condition (H_1) we assume

(H_2) $\mu'(\overline{D}) = 1$ for all $\mu' \in \mathcal{K}$.

Indeed, in this case property (4) holds for all $x \in \overline{D}$ and not only for all $x \in D$, thanks to lower semicontinuity of g_σ. Furthermore, property (3) holds for μ-a.e. $z \in \tilde{E}_{m,D}$, where

$$\tilde{E}_{m,D} := \{z \in \overline{D} : \text{ there exists } \sigma_z, \delta_z > 0 \text{ such that } S^{\sigma_z,\delta_z}_{m,D}(z) \text{ is bounded}\}.$$

We notice also that if D is a dense subset of $\mathbb{R}^d$, condition (H_2) is trivially satisfied.

4. **Application to the composition of flows of vector fields.** As seen in the Introduction, in [17] the authors extended the definition of a Lie bracket of two C^1 vector fields to the case of two Lipschitz continuous vector fields X, Y, that is an assumption implying continuity of $\Psi_t(\cdot) := [\phi_t^X, \phi_t^Y](\cdot)$. In this case, the Lie bracket of the vector fields at every point turns out to be a set. Moreover, they provided in this framework an asymptotic formula for the flows and the generalization of other classical results holding for the Lie bracket of vector fields.

A natural question is to compare our construction with the one in [17] when the starting measure is reduced to a Dirac delta, in the spirit of the AGF formalism. The aim of this section is to perform such a comparison, showing that - roughly speaking - the density V of the measure theoretic bracket $V\mu$ is a L^p_μ-selection of the Rampazzo-Sussmann set-valued Lie bracket. In particular, when $\mu = \delta_q$, the two constructions are reduced to the same object.

We will take $\mathcal{K} = \mathscr{P}(\mathbb{R}^d)$ throughout the section, hence we will omit the condition (D_4) in Definition 3.1 since it follows from (D_1).

We recall the following definition from [17].

Definition 4.1 (Set-valued Lie brackets)**.** Let f, g be locally Lipschitz vector fields on $\mathbb{R}^d$. The (set-valued) Lie bracket of f and g at $x \in \mathbb{R}^d$ is

$$[f,g]_{\text{set}}(x) := \text{co}\Big\{ v \in \mathbb{R}^d : \text{ there exists a sequence } \{x_j\}_{j\in\mathbb{N}} \subseteq \text{dom}(Df) \cap \text{dom}(Dg), \\ \text{such that } x_j \to x \text{ and } v = \lim_{j\to\infty} [f,g](x_j) \Big\}$$

where $\text{dom}(Df)$ and $\text{dom}(Dg)$ denotes the set of differentiability points of f and g, respectively. Recalling Rademacher's Theorem, when f is Lipschitz continuous it is differentiable at a.e. $x \in \mathbb{R}^d$, thus $\text{dom}(Df) \cap \text{dom}(Dg)$ has full measure in $\mathbb{R}^d$.

According to Remark 3.6 in [17], the following equivalent definition can be given

$$[f,g]_{\text{set}}(x) = \{ Bf(x) - Ag(x) : (A,B) \in \partial(f \times g)(x) \},$$

where $f \times g$ is the map defined as $(f \times g)(x) = (f(x), g(x))$, and ∂ denotes the Clarke's generalized Jacobian, which for a Lipschitz continuous map $h : \mathbb{R}^k \to \mathbb{R}^m$ is defined as

$$\begin{aligned}\partial h(x) &:= \text{co}\left\{ L : \mathbb{R}^k \to \mathbb{R}^m : \begin{array}{l} \text{there exists } \{x_j\}_{j\in\mathbb{N}} \subseteq \text{dom}(Dh),\ x_j \to x, \\ \text{s.t. } L = \lim_{j\to\infty} Dh(x_j) \end{array} \right\}. \\ &= \text{co} \bigcap_{\delta>0} \overline{\{ Dh(y) : y \in \text{dom}(Dh) \cap B(x,\delta) \}}.\end{aligned}$$

Recall that in general $\partial(f \times g)(x) \subseteq \partial f(x) \times \partial g(x)$, and the inclusion may be strict.

We can recast the above definition by

$$[f,g]_{\text{set}}(x) = \text{co} \bigcap_{\delta>0} \overline{\{ Dg(y)f(y) - Df(y)g(y) : y \in \text{dom}(Df) \cap \text{dom}(Dg) \cap B(x,\delta) \}}.$$

Remark 4. Let v be a Lipschitz continuous vector field with Lipschitz constant $L > 0$. Fix a set of smooth mollifiers $\{s_\rho\}_{\rho>0}$ and set $v_\rho = v * s_\rho$. For any $\varepsilon > 0$

there exists $\rho > 0$ such that for all $0 \le t \le T$

$$\begin{aligned}|\phi_t^v(x)-\phi_t^{v_\rho}(y)| &\le |x-y| + \int_0^t |v(\phi_s^v(x)) - v_\rho(\phi_s^{v_\rho}(y))|\,ds \\ &\le |x-y| + \int_0^t |v(\phi_s^v(x)) - v(\phi_s^{v_\rho}(y))| + \int_0^t |v(\phi_s^{v_\rho}(y)) - v_\rho(\phi_s^{v_\rho}(y))|\,ds \\ &\le |x-y| + L\int_0^t |\phi_s^v(x) - \phi_s^{v_\rho}(y)| + \varepsilon T.\end{aligned}$$

By Gronwall's inequality,

$$|\phi_t^v(x) - \phi_t^{v_\rho}(y)| \le (|x-y| + \varepsilon T)\, e^{LT}$$

and so if $|x-y| \le C'\varepsilon$, there exists $C'' > 0$ such that $|\phi_t^v(x) - \phi_t^{v_\rho}(y)| \le C''\varepsilon$. The argument can be iterated for concatenation of flows of Lipschitz continuous vector fields.

Remark 5. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz continuous map. Then, if f is differentiable at $x \in \mathbb{R}^d$, we have $\nabla f_\rho(x) \to \nabla f(x)$, where $f_\rho(x) = (f * s_\rho)(x)$, and $\{s_\rho\}_{\rho>0}$ is any family of smooth mollifiers. It is enough to check the assertion for the directional derivatives of f, so let $v \in \mathbb{R}^d$, $\|v\| = 1$. Recalling that f_ρ converges uniformly to f on compact sets, we have

$$\begin{aligned}\{\partial_v f(x)\} &= \bigcap_{\sigma>0} \overline{\left\{\frac{f(x+tv)-f(x)}{t} \,:\, 0<t<\sigma\right\}} \\ &= \bigcap_{\sigma>0}\bigcap_{\rho>0} \overline{\left\{\frac{f_\tau(x+tv)-f_\tau(x)}{t} \,:\, 0<t<\sigma,\, 0<\tau<\rho\right\}} \\ &= \bigcap_{\rho>0}\bigcap_{\sigma>0} \overline{\left\{\frac{f_\tau(x+tv)-f_\tau(x)}{t} \,:\, 0<t<\sigma,\, 0<\tau<\rho\right\}} \\ &= \bigcap_{\rho>0} \overline{\{\partial_v f_\tau(x) \,:\, 0<\tau<\rho\}} = \left\{\lim_{\rho\to 0} \partial_v f_\rho(x)\right\}.\end{aligned}$$

We will show now a result stating the main connection between our construction and [17]. Indeed, we prove that in the same framework of [17], the two constructions agree.

Proposition 2. *Let now X, Y be locally Lipschitz continuous vector fields, set $\Psi_t(x) = [\phi_t^X, \phi_t^Y](x)$, then $\mathbf{\Psi} = \{\Psi_t(\cdot)\}_{t\in[0,T]}$ satisfies assumptions (D_1), (D_2), (D_3). For any $z \in \mathbb{R}^d$ and $V \in V_2^p(\delta_z, \mathbf{\Psi})$ we have*

$$V(z) \in [X,Y]_{\mathrm{set}}(z).$$

Proof. Let D be the set of differentiability points of X and Y, in particular it is dense in $\mathbb{R}^d$. Fix $z \in \mathbb{R}^d$. By Lemma 3.6, we can restrict ourselves to measures supported on a compact neighborhood of z, thus without loss of generality we can assume that X, Y are globally Lipschitz continuous.

Fix a smooth family of mollifiers $\{s_\rho\}_{\rho>0}$, and let $X^\rho = X * s_\rho$ and $Y^\rho = Y * s_\rho$. We set $\Psi_t^\rho(x) = [\phi_t^{X^\rho}, \phi_t^{Y^\rho}]$ and notice that $\mathbf{\Psi}^\rho$ converges uniformly to $\mathbf{\Psi}$ on every compact subset of $[0,T] \times \mathbb{R}^d$. Moreover, if $x \in D$ we have $\nabla X^\rho(x) \to \nabla X(x)$ as

$\rho \to 0^+$ by Remark 5. These two facts implies that

$$
\begin{aligned}
&\operatorname{co} \bigcap_{\sigma,\delta>0} \overline{S^{\sigma,\delta}_{m,D}(z)} \\
&= \operatorname{co} \bigcap_{\sigma,\delta>0} \bigcap_{\rho>0} \overline{\left\{ \frac{\Psi^\tau_t(x) - x}{t^2} : 0 < \tau < \rho,\, x \in B(z,\delta) \cap D,\, 0 < t < \sigma \right\}} \\
&= \operatorname{co} \bigcap_{\delta>0} \bigcap_{\rho>0} \bigcap_{\sigma>0} \overline{\left\{ \frac{\Psi^\tau_t(x) - x}{t^2} : 0 < \tau < \rho,\, x \in B(z,\delta) \cap D,\, 0 < t < \sigma \right\}} \\
&= \operatorname{co} \bigcap_{\delta>0} \bigcap_{\rho>0} \overline{\{[X^\tau, Y^\tau](x) : 0 < \tau < \rho,\, x \in B(z,\delta) \cap D\}} \\
&= \operatorname{co} \bigcap_{\delta>0} \overline{\{\nabla Y(x) \cdot X(x) - \nabla X(x) \cdot Y(x) : x \in B(z,\delta) \cap D\}} \\
&= [X,Y]_{\text{set}}(z).
\end{aligned}
$$

Hence we can conclude, thanks to Remark 3 and noticing that we have $\tilde{E}_{2,D} = \mathbb{R}^d$ by density of D in $\mathbb{R}^d$. □

Exploiting this representation formula, and the results of [17] (see in particular Theorem 5.3 for commutativity), the asymptotic result given by Corollary 1 can be refined as follows.

Corollary 2. *Let $T > 0$, $m \in \mathbb{N}$, $m \geq 1$, $p \geq 1$, and let X, Y be locally Lipschitz continuous vector fields. Set $\Psi_t(x) = [\phi^X_t, \phi^Y_t](x)$, $\boldsymbol{\Psi}_t = \{\Psi_t(\cdot)\}_{t\in[0,T]}$. Then, if $V^p_2(\mu, \boldsymbol{\Psi}) = \{0\}$ for all $\mu \in \mathscr{P}_p(\mathbb{R}^d)$ we have $(\phi^X_t \circ \phi^Y_t)\sharp\mu = (\phi^Y_t \circ \phi^X_t)\sharp\mu$ for all $\mu \in \mathscr{P}_p(\mathbb{R}^d)$, $t \in [0,T]$.*

Apparently, the construction of Proposition 2 can be extended to any formal bracket by using Theorem 2.6. However, it has been pointed out in [17] that the step between the definition of the *single* set-valued bracket, and the definition of *higher order* bracket is quite nontrivial. Indeed, we can give just a partial answer to this issue.

Definition 4.2. Let $k \in \mathbb{N} \setminus \{0,1\}$, and $X_1, \ldots, X_k$ be vector fields of class $C^{k-2,1}(\mathbb{R}^d)$. Let $\mathscr{S} := \{\phi^{X_i}_t : i = 1, \ldots, k\}$ and consider a formal bracket $B(\phi^{X_1}_t, \ldots, \phi^{X_k}_t)$ of order k w.r.t. $\mathscr{S}$. Let $D \subseteq \mathbb{R}^d$. We define for any $z \in \mathbb{R}^d$

$$
\begin{aligned}
&B_{\text{set}}(X_1, \ldots, X_k)(z) \\
&= \operatorname{co} \bigcap_{\delta>0} \bigcap_{\rho>0} \overline{\{B(X^\tau_1, \ldots, X^\tau_k)(x) : x \in B(z,\delta) \cap D,\, 0 < \tau < \rho\}}.
\end{aligned} \tag{5}
$$

The motivation for such a definition is the following.

Remark 6. Set $\Psi_t(x) = B(\phi^{X_1}_t, \ldots, \phi^{X_k}_t)(x)$ and let D be the set of differentiability points for all the vector fields involved and for their derivatives up to the order appearing in the bracket B. In particular, D is dense in $\mathbb{R}^d$. By Theorem 3.8, for all $z \in \mathbb{R}^d$ we have

$$V(z) \in \operatorname{co} \bigcap_{\sigma,\delta>0} \overline{S^{\sigma,\delta}_{k,D}(z)},$$

for all $V \in V^p_k(\delta_z, \boldsymbol{\Psi})$. Thus it make sense to define

$$B_{\text{set}}(X_1, \ldots, X_k)(z) = \operatorname{co} \bigcap_{\sigma,\delta>0} \overline{S^{\sigma,\delta}_{k,D}(z)},$$

indeed, equality follows by the very same argument of Proposition 2.

When z is a differentiability point for all the vector fields involved and for their derivatives up to the order appearing in the bracket B, we can refine (5), in the spirit of Proposition 2, i.e., we set D as the set of common differentiability points for all the vector fields and their derivatives, and we have for all $z \in D$

$$B_{\text{set}}(X_1, \dots, X_k)(z) = \text{co} \bigcap_{\delta>0} \overline{\{B(X_1, \dots, X_k)(x) : x \in B(z, \delta) \cap D\}}. \tag{6}$$

However, in general, the definition given in (6) is *not consistent with the asymptotic formula when $z \notin D$*, in the following sense: to have

$$\text{co} \bigcap_{\delta>0} \overline{\{B(X_1, \dots, X_k)(x) : x \in B(y, \delta) \cap D\}} = 0$$

for all y in a neighborhood of z, in general *does not imply* that $\lim\limits_{t \to 0} \dfrac{\Psi_t(z) - z}{t^m} = 0$, as showed with a counterexample in Section 7.1 and Section 7.2 of [17], where the possibility to extend the construction of [17] to higher order brackets respecting the asymptotic formulas is extensively studied.

On the other hand, (5) is coherent with the asymptotic formula at all $z \in \mathbb{R}^d$, by construction, but lacks of a simpler representation.

The problem for the pointwise set-valued bracket has been partially treated in [10], and will be concluded in [11], by using different techniques w.r.t. this paper. We just point out here that a useful tool to study the cluster points of $B(X_1^\tau, \dots, X_k^\tau)(x)$ as $\tau \to 0$ is provided by the following result, which is a simplified version of Theorem 9.67 in [19].

Proposition 3. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally Lipschitz function, and let $\{s_\rho\}_{\rho>0}$ be a sequence of smooth mollifiers. Set $f_\rho = f * s_\rho$. Then*

$$\text{co} \bigcap_{\delta>0} \bigcap_{\rho>0} \overline{\{\nabla f_\tau(x') : x' \in B(x, \delta),\, 0 < \tau < \rho\}} = \partial_C f(x).$$

5. **An example.**

In this section we provide an example illustrating our approach.

In the example below, we first consider the case in which the measure μ is blind w.r.t. the singularity set H of the vector fields, i.e. the singularities of the vector fields are contained in a μ-negligible closed set. In this case, roughly speaking, we can neglect them and perform the computations exactly as in the classical case. In the same setting, we then analyze the behaviour of the system on the singular set H. To this aim, we will set $D = \mathbb{R}^d \setminus H$.

Example 5.1. In $\mathbb{R}^2$, set $H := \{(x, y) \in \mathbb{R}^2 : xy = 0\}$ and consider two Borel vector fields safifying for $(x, y) \in D$

$$X(x, y) := \sqrt{\frac{3}{5}} \cdot \frac{x}{y^{2/3}} \cdot (1, 1), \qquad Y(x, y) := X(y, x).$$

Since in the open set D these vector fields are smooth, we can set $\Psi_t(x, y) = [\phi_t^X, \phi_t^Y](x, y)$ for $(x, y) \in D$ and t small enough, thus for all $(x, y) \in D$ we have

$$\lim_{\substack{(u,w)\to(x,y)\\ t\to 0^+}} \frac{\Psi_t(u, v) - (u, v)}{t^2} = [X, Y](x, y) = \frac{x - y}{x^{2/3} y^{2/3}}(1, 1).$$

According to the representation formula, we have that if $V_2^p(\mu,\mathbf{\Psi})\neq\emptyset$, we must have

$$V(x,y)=\frac{x-y}{x^{2/3}y^{2/3}}(1,1),\ \text{for }\mu\text{-a.e. }(x,y)\in D\text{ and all }V\in V_2^p(\mu,\mathbf{\Psi}).$$

Thus if the map $(x,y)\mapsto\dfrac{x-y}{x^{2/3}y^{2/3}}(1,1)\in L^p_\mu(\mathbb{R}^d)$ and $\mu(D)=1$, we obtain that $V_2^p(\mu,\mathbf{\Psi})$ is reduced to the singleton $(x,y)\mapsto\dfrac{x-y}{x^{2/3}y^{2/3}}(1,1)$. For istance, this holds for $1\le p<3/2$ and any $\mu\ll\mathscr{L}$ with compact support.

Fix $x_0\neq 0$. For every $\delta,\sigma>0$ the set $\overline{S^{\sigma,\delta}_{2,D}(x_0,0)}$ is unbounded, since

$$\overline{S^{\sigma,\delta}_{2,D}(x_0,0)}\supseteq\bigcap_{\sigma'>0}\overline{S^{\sigma',\delta}_{2,D}(x_0,0)}=\overline{\left\{\frac{x-y}{x^{2/3}y^{2/3}}(1,1):\,(x,y)\in B((x_0,0),\delta)\cap D\right\}}.$$

According to the representation formula, we have that if $V^p_{2,D}(\mu,\mathbf{\Psi})\neq\emptyset$, we must have for μ-a.e. $(x_0,0)\in\mathbb{R}^2$

$$V(x_0,0)\in\bigcap_{\sigma,\delta>0}\operatorname{co}\overline{S^{\sigma,\delta}_{2,D}(x_0,0)},$$

but this set is empty. Thus if $\mu(\{(x_0,0):\,x_0>0\})>0$ we have that $V^p_{2,D}(\mu,\mathbf{\Psi})=\emptyset$. However, it is easy to show that for $1<m<2$ we have

$$\bigcap_{\sigma,\delta>0}\operatorname{co}\overline{S^{\sigma,\delta}_{m,D}(x_0,0)}=\{\lambda(1,1):\,\lambda\ge 0\}.$$

We can reason in a similar way on all the points of $H\setminus\{(0,0)\}$.

Concerning the origin, we notice that

$$\bigcap_{\sigma,\delta>0}\operatorname{co}\overline{S^{\sigma,\delta}_{2,D}(0,0)}=\mathbb{R}^2,$$

thus in the case that $\mu(H\setminus\{(0,0)\})=0$, we are able to define again $V(\cdot)\in V^p_{2,D}(\mu,\mathbf{\Psi})$ provided that $(x,y)\mapsto\dfrac{x-y}{x^{2/3}y^{2/3}}(1,1)\in L^p_\mu(\mathbb{R}^d\setminus\{(0,0)\})$ (we can simply set $V(0,0)=0$).

Acknowledgments. We would like to thank prof. Franco Rampazzo (University of Padova, Italy) for the useful discussion on the subject.

REFERENCES

[1] L. Ambrosio, The flow associated to weakly differentiable vector fields: recent results and open problems, *Nonlinear Conservation Laws and Applications*, IMA Vol. Math. Appl., Springer, New York, **153** (2011), 181–193.

[2] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2^{nd} edition, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008.

[3] G. Buttazzo, *Semicontinuity, Relaxation and Integral Representation in the Calculus of Variations*, Pitman Research Notes in Mathematics Series, **207**, Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1989.

[4] G. Cavagnari, Regularity results for a time-optimal control problem in the space of probability measures, *Mathematical Control and Related Fields*, **7** (2017), 213–233.

[5] G. Cavagnari and A. Marigonda, Time-optimal control problem in the space of probability measures, *Lecture Notes in Comput. Sci.*, **9374** (2015), 109–116.

[6] G. Cavagnari, A. Marigonda, K. T. Nguyen and F. S. Priuli, Generalized control systems in the space of probability measures, *Set-Valued and Variational Analysis* (2018). Published online.

[7] G. Cavagnari, A. Marigonda and G. Orlandi, Hamilton-Jacobi-Bellman equation for a time-optimal control problem in the space of probability measures, in *System Modeling and Optimization. CSMO 2015. IFIP Advances in Information and Communication Technology* (eds. L. Bociu, JA. Désidéri, A. Habbal), Springer, Cham, **494** (2016), 200–208.

[8] G. Cavagnari, A. Marigonda and B. Piccoli, Averaged time-optimal control problem in the space of positive Borel measures, *ESAIM: COCV* (2018). Published online.

[9] G. Cavagnari, A. Marigonda and B. Piccoli, Optimal syncronization problem for a multi-agent system, *Networks and Heterogeneous Media*, **12** (2017), 277–295.

[10] E. Feleqi and F. Rampazzo, Integral representations for bracket-generating multi-flows, *Discrete Contin. Dyn. Syst.*, **35** (2015), 4345–4366.

[11] E. Feleqi and F. Rampazzo, Iterated Lie brackets for nonsmooth vector fields, *Nonlinear Differ. Equ. Appl.*, **24** (2017), Art. 61, 43 pp.

[12] V. Jurdjevic, *Geometric Control Theory*, Cambridge Studies in Advanced Mathematics, **52**, Cambridge University Press, Cambridge, 1997.

[13] A. Marigonda and S. Rigo, Controllability of some nonlinear systems with drift via generalized curvature properties, *SIAM J. Control Optim.*, **53** (2015), 434–474.

[14] A. Marigonda and Thuy T. T. Le, Small-time local attainability for a class of control systems with state constraints, *ESAIM: Control, Optimization and Calc. of Var.*, **23** (2017), 1003–1021.

[15] M. Mauhart and P. W. Michor, Commutators of flows and fields, *Arch. Math. (Brno)*, **28** (1992), 229–236.

[16] F. Rampazzo, Frobenius-type theorems for Lipschitz distributions, *J. Differential Equations*, **243** (2007), 270–300.

[17] F. Rampazzo and H. J. Sussmann, Commutators of flow maps of nonsmooth vector fields, *J. Differential Equations*, **232** (2007), 134–175.

[18] F. Rampazzo and H.J. Sussmann, Set-valued differentials and a nonsmooth version of Chow's theorem, in *Proc. of the 40th IEEE Conf. on Decision and Control*, Orlando, FL, December 2001, **3** (2001), IEEE Publications, New York, 2613–2618.

[19] R.T. Rockafellar and R.J.-B. Wets, *Variational Analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], **317**, Springer-Verlag, Berlin, 1998.

E-mail address: giulia.cavagnari@rutgers.edu
E-mail address: antonio.marigonda@univr.it