A STUDY OF THE METRIC MEASURE SPACE OF PROBABILITY MEASURES VIA A PURELY ATOMIC SUPERPOSITION PRINCIPLE

ALESSANDRO PINZI

ABSTRACT. We study the continuity equation on the metric measure space $(\mathcal{P}_p(X), W_p, Q)$, when X is either the Euclidean space or a compact, oriented, and boundaryless Riemannian manifold, for some suitable reference measure $Q \in \mathcal{P}_p(\mathcal{P}_p(X))$, which by construction are concentrated over purely atomic measures. In fact, we consider the equation $\partial_t M_t + \operatorname{div}_{\mathcal{P}}(b_t M_t) = 0$, where $(M_t)_{t\in[0,T]} \subset \mathcal{P}(\mathcal{P}(X))$ and $b : [0,T] \times X \times \mathcal{P}(X) \to TX$, assuming that $M_t \ll Q$ for all $t \in [0,T]$, to then show when the purely atomic property is inherited by the liftings of the curve M_t given by the nested superposition principle.

On the Euclidean space, the main assumption is that the r-capacity of the diagonal $\Delta \subset \mathbb{R}^d \times \mathbb{R}^d$ is zero with respect to $\nu \otimes \nu$, where ν is the barycenter of the reference measure Q. We will give sufficient conditions to ensure it, and in particular, thanks to the Nash embedding theorem, this analysis will allow us to extent the main results from the Euclidean space to Riemannian manifolds.

Finally, we exploit this atomic superposition principle to show the lack of the Sobolev-to-Lipschitz property and the Poincaré inequality for functions in $W^{1,p}(\mathcal{P}_p(X), W_p, Q)$. Then, we complete the analysis showing that, however, the L^2-Wasserstein space endowed with suitable reference measure Q, satisfies a Bakry–Émery curvature condition.

Keywords: superposition principle, random measures, atomic measures, Wasserstein space, metric measure spaces.
2020 MSC: 60G57, 49Q22, 30L99, 35R06.

CONTENTS

Date: November 27, 2025.

1. Introduction

In recent years, the problem of endowing the space of propability measures with a "nice" reference measure attracted many researchers. It has been mostly studied to understand the closure of the associated Dirichlet form and then its corresponding diffusion process, as done in [Del22; Del24; DS25; Mar18; KR19; RS09; Stu25; RWW24; RW24]. In this paper, we study the continuity equation, and in general absolutely continuous curves of measures, on the metric measure space $(\mathcal{P}_p(X), W_p, Q)$, for suitable $p > 1$ and reference measure $Q \in \mathcal{P}(\mathcal{P}(X))$ concentrated over $\mathcal{P}_p(X)$, when X is either the Euclidean space $\mathbb{R}^d$ or a compact, boundaryless, and oriented Riemannian manifold. To our knowledge, this analysis is the first of his kind and makes it possible to consider a wide range of reference measures which include some of the ones built in the aforementioned articles.

The reference measures we consider are built in the following way. Define

$$\mathbf{T} := \left\{ \mathbf{a} = (a_i)_{i\in\mathbb{N}} \in [0,1]^{\mathbb{N}} \ : \ a_{i+1} \le a_i, \ \sum_{i=1}^{+\infty} a_i = 1 \right\}; \tag{1.1}$$

consider any probability measure $\pi \in \mathcal{P}(\mathbf{T})$. Now, given a Polish space X and an atomless probability measure $\nu \in \mathcal{P}(X)$; define the embedding function

$$\mathrm{em} : \mathbf{T} \times X^\infty \to \mathcal{P}(X), \quad \mathrm{em}(\mathbf{a}, \mathbf{x}) := \sum_{i=1}^{+\infty} a_i \delta_{x_i}, \tag{1.2}$$

to set $Q_{\pi,\nu} := (\mathrm{em}_\# \, \pi \otimes \nu^\infty)$. In particular, if $\nu \in \mathcal{P}_p(X)$, then $Q_{\pi,\nu}$ is concentrated over $\mathcal{P}_p(X)$. Then, we denote our class of reference measures as

$$\mathcal{Q}(X) := \{ Q_{\pi,\nu} \ : \ \pi \in \mathcal{P}(\mathbf{T}), \ \nu \in \mathcal{P}(X) \text{ atomless} \}. \tag{1.3}$$

It is important to stress the fact that any $Q \in \mathcal{Q}(X)$ is concentrated over the set of *purely atomic probability measures*, that can be defined as $\mathcal{P}^{\mathrm{pa}}(X) := \mathrm{em}(\mathbf{T} \times X^\infty)$.

We will mainly work with $X = \mathbb{R}^d$. Then, the main results will be transferred to the Riemannian setting exploiting the Nash isometric embedding theorem. So, we proceed by describing the main results and the strategy of the proofs in the Euclidean setting. The details about the Riemannian case can be found in Section 6.

Continuity equation and superposition principle. For a curve of (narrowly continuous) measures $(\mu_t)_{t\in[0,T]} \subset \mathcal{P}(\mathbb{R}^d)$ and a given Borel vector field $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, we say that the *continuity equation* $\partial_t \mu_t + \mathrm{div}(v_t \mu_t) = 0$ is satisfied if

$$\int_0^T \int_{\mathbb{R}^d} |v_t(x)| d\mu_t(x) dt < +\infty \quad \text{and} \quad \frac{d}{dt} \int_{\mathbb{R}^d} \phi(x) d\mu_t(x) = \int_{\mathbb{R}^d} v_t(x) \cdot \nabla \phi(x) d\mu_t(x), \tag{1.4}$$

in the sense of distributions in $(0,T)$, for all $\phi \in C^1_c(\mathbb{R}^d)$. A fundamental result for us is the Ambrosio's superposition principle [AGS08, Theorem 8.2.1]: for any solution of the continuity equation, there exists $\lambda \in \mathcal{P}(C([0,T],\mathbb{R}^d))$ which has marginals μ_t, that is

$$(\mathrm{e}_t)_\sharp \lambda = \mu_t \quad \forall t \in [0,T], \quad \text{where } \mathrm{e}_t(\gamma) := \gamma(t), \tag{1.5}$$

and that is concentrated over curves $\gamma \in C([0,T],\mathbb{R}^d)$ that satisfies

$$\gamma \in AC([0,T],\mathbb{R}^d), \quad \text{and} \quad \dot\gamma(t) = v_t(\gamma(t)) \quad \text{for a.e. } t \in [0,T]. \tag{1.6}$$

Returning to the main setting of the present paper, when we endow the space of probability measure with a measure $Q \in \mathcal{Q}(\mathbb{R}^d)$, it will be necessary to work with curves of probability measures (μ_t) that are purely atomic. In this setting, a natural question about the superposition principle can be raised.

Question 1.1. *If a (narrowly continuous) curve $(\mu_t)_{t\in[0,T]} \subset \mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)$ solves a continuity equation in the sense of* (1.4)*, does it exists a purely atomic superposition of curves $\lambda \in \mathcal{P}^{\mathrm{pa}}(C([0,T],\mathbb{R}^d))$ satisfying* (1.5) *and is concentrated over γ's as in* (1.6)*?*

The answer to this question is in general false: in Section 3, we build a W_∞-Lipschitz curve of measures in $\mathcal{P}^{\mathrm{pa}}(\mathbb{R})$, and in particular for which there exists a bounded vector field $v : [0,T]\times\mathbb{R} \to \mathbb{R}$ such that (1.4) is satisfied, but it admits a unique lifting $\lambda \in \mathcal{P}(C([0,T],\mathbb{R}))$ that is atomless.

This raises non-trivial problems in the study of evolutions in the metric measure space $(\mathcal{P}_p(\mathbb{R}^d), W_p, Q)$, for all $Q \in \mathcal{Q}(\mathbb{R}^d)$. A fundamental result to overcome this is to ask the continuity equation to be satisfied when tested against a more general class of functions. Notice that, a necessity condition for having a positive answer to the previous question is that the weights of the atoms are constant in time. In view of this, we enforce (1.4) asking that for all $\hat\phi = \hat\phi(x,r) \in C^1_c(\mathbb{R}^d \times (0,1])$

$$\frac{d}{dt}\int_{\mathbb{R}^d} \phi(x, \mu_t(\{x\}))d\mu_t(x) = \int_{\mathbb{R}^d} v_t(x)\cdot\nabla_x\phi(x,\mu_t(\{x\}))d\mu_t(x), \tag{1.7}$$

in the sense of distributions in $(0,T)$. Under this stronger assumption, we can give a positive answer to Question 1.1. For a more rigorous discussion, we refer to Lemma 4.6 and to Section 4.3.

Continuity equation for random measures and nested superposition principle. In [PS25], it has been introduced the continuity equation for general (narrowly continuous) curves of measures $(M_t)_{t\in[0,T]} \subset \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$. To recall it, we need to introduce what is a vector field in this setting and which is the natural class of test functions. The latter is given by the *cylinder functions*, see also [FSS23; Sod23].

Definition 1.2 (Cylinder functions)**.** *A functional $F : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ is called a C^1_c-cylinder function, and we write $F \in \mathrm{Cyl}^1_c(\mathcal{P}(\mathbb{R}^d))$, if there exists $k \in \mathbb{N}$, $\Phi = (\phi_1,\dots,\phi_k) \in C^1_c(\mathbb{R}^d;\mathbb{R}^k)$ and $\Psi \in C^1_b(\mathbb{R}^k)$ such that*

$$F(\mu) = \Psi\left(L_\Phi(\mu)\right), \quad L_\Phi(\mu) = \big(L_{\phi_1}(\mu),\dots,L_{\phi_k}(\mu)\big), \quad L_{\phi_i}(\mu) := \int_{\mathbb{R}^d}\phi_i(x)d\mu(x). \tag{1.8}$$

Its Wasserstein gradient is then defined as

$$\nabla_W F(x,\mu) := \sum_{i=1}^k \partial_i\Psi\left(L_\Phi(\mu)\right)\nabla\phi_i(x) \quad \forall x \in \mathbb{R}^d,\ \forall \mu \in \mathcal{P}(\mathbb{R}^d). \tag{1.9}$$

Then, it has been proved in [PS25, Section 4] that the natural choice for a vector field in this setting is a *non-local vector field*, that is a Borel map $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$. They are also called probability vector fields in literature, see [CSS23a; CSS23b]. Then, we say that the continuity equation $\partial_t M_t + \mathrm{div}_{\mathcal{P}}(b_t M_t) = 0$ is satisfied if

$$\begin{aligned} &\int_0^T \int_{\mathcal{P}} \int_{\mathbb{R}^d} |b_t(x,\mu)| d\mu(x) dM_t(\mu) dt < +\infty \quad \text{and} \\ &\forall F \in \mathrm{Cyl}^1_c(\mathcal{P}(\mathbb{R}^d)) \quad \frac{d}{dt} \int_{\mathcal{P}} F(\mu) dM_t(\mu) = \int_{\mathcal{P}} \int_{\mathbb{R}^d} \nabla_W F(x,\mu) \cdot b_t(x,\mu) d\mu(x) dM_t(\mu), \end{aligned} \tag{1.10}$$

in the sense of distributions in $(0,T)$. One of the main result of [PS25] is a version of the Ambrosio's superposition principle for this setting, which can be resumed as follows (for a more rigorous statement, see also Theorem 2.10). Given $(M_t)_{t\in[0,T]} \subset \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ a narrowly continuous curve that solves $\partial_t M_t + \mathrm{div}_{\mathcal{P}}(b_t M_t) = 0$, then it can be lifted to:

- $\Lambda \in \mathcal{P}(C([0,T], \mathcal{P}(\mathbb{R}^d)))$ such that Λa.e. $\boldsymbol{\mu}$ solves

$$\partial_t \mu_t + \mathrm{div}(b_t(\cdot,\mu_t)\mu_t) = 0. \tag{1.11}$$

- $\mathfrak{L} \in \mathcal{P}(\mathcal{P}(C([0,T],\mathbb{R}^d)))$ such that $\mathfrak{L}$-a.e. $\lambda \in \mathcal{P}(C([0,T],\mathbb{R}^d))$ is concentrated over absolutely continuous curves γ that are solutions of

$$\dot\gamma(t) = b(t, \gamma_t, (e_t)_\# \lambda) \quad \text{in } (0,T). \tag{1.12}$$

In this paper, we are interested in the evolution of curves $(M_t)_{t\in[0,T]} \subset \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ such that $M_t \ll Q$ for all $t \in [0,T]$, for some fixed $Q \in \mathcal{Q}(\mathbb{R}^d)$. Consequently, any lifting $\Lambda \in \mathcal{P}(C([0,T],\mathcal{P}(\mathbb{R}^d)))$ will be automatically concentrated on curves of purely atomic measures. Then, as for Question 1.1, it is natural to ask when there exists a lifting $\mathfrak{L} \in \mathcal{P}(\mathcal{P}(C([0,T],\mathbb{R}^d)))$ that satisfies the above properties and it is further concentrated over $\mathcal{P}^{\mathrm{pa}}(C([0,T],\mathbb{R}^d))$.

Again, this is a non-trivial question, and to positively answer to it we need further assumptions. In particular, assume that $Q = Q_{\pi,\nu}$, then we need to control that the diagonal in $\mathbb{R}^d \times \mathbb{R}^d$ is not too loaded by the measure $\nu \otimes \nu$, in a capacitary sense: to this aim, we introduce, for all $r \in [1,+\infty)$, the quantity

$$\mathrm{cap}_{r,\nu}(\Delta) := \inf\left\{ \int_{\mathbb{R}^{2d}} |h(x,y)|^r + |\nabla h(x,y)|^r d\nu\otimes\nu(x,y) : h \in C^1_b(\mathbb{R}^{2d}),\ h = 1 \text{ on } \Delta,\ h \le 1 \right\}.$$

Then, one of the main theorem of this paper, is the following, where we use the notations

$$\begin{aligned} &E : \mathcal{P}(C([0,T],\mathbb{R}^d)) \to C([0,T],\mathcal{P}(\mathbb{R}^d)), \quad E(\lambda) := ((\mathrm{e}_t)_\sharp \lambda)_{t\in[0,T]}, \\ &E_t : \mathcal{P}(C([0,T],\mathbb{R}^d)) \to \mathcal{P}(\mathbb{R}^d), \quad E_t(\lambda) := (\mathrm{e}_t)_\sharp \lambda, \\ &\mathfrak{e}_t : C([0,T],\mathcal{P}(\mathbb{R}^d)) \to \mathcal{P}(\mathbb{R}^d), \quad \mathfrak{e}_t(\boldsymbol{\mu}) := \mu_t. \end{aligned} \tag{1.13}$$

Theorem 1.3. *Let $\Lambda \in \mathcal{P}(C([0,T],\mathcal{P}(\mathbb{R}^d)))$ be concentrated over solutions of* (1.11) *and define $\boldsymbol{M} = (M_t)_{t\in[0,T]} \in C([0,T],\mathcal{P}(\mathcal{P}(\mathbb{R}^d)))$ as $M_t = (\mathfrak{e}_t)_\sharp \Lambda$. Assume:*

(1) $M_t = u_t Q$, with $u \in L^1_t(L^q(Q))$, with $q \in [1,+\infty]$;
(2) for some $p > 1$, it holds

$$\int \int_0^T \int_{\mathcal{P}(\mathbb{R}^d)} \int_{\mathbb{R}^d} |b_t(x,\mu)|^p d\mu(x) dM_t(\mu) dt < +\infty; \tag{1.14}$$

(3) $\mathrm{cap}_{r,\nu}(\Delta) = 0$, with $r \in [1,+\infty)$;
(4) $\frac{p'}{r} + \frac{1}{q} \le 1$, where p' is the conjugate exponent of p.

Then, there exists $\mathfrak{L} \in \mathcal{P}(\mathcal{P}(AC^p([0,T],\mathbb{R}^d)))$ *such that* $E_\sharp = \Lambda$*, it is concentrated over* $\Lambda \in \mathcal{P}^{\mathrm{pa}}(C[0,T],R^d)$ *that are, in turn, concentrated over solutions of* (1.12)*. In particular,* $(E_t)_\sharp \mathfrak{L} = M_t$ *for all* $t \in [0,T]$ *and* Λ *is concentrated over* $\mathcal{P}(C([0,T],\mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)))$*.*

The proof of the previous theorem relies on extending the class of functions that we can use to test the continuity equation $\partial_t M_t + \mathrm{div}_{\mathcal{P}}(b_t M_t) = 0$ and $\partial_t \mu_t + \mathrm{div}(b_t(\cdot,\mu_t)\mu_t) = 0$, for Λ-a.e. $(\mu_t)_{t\in[0,T]}$. Indeed, under the assumptions of Theorem 1.3, we can show that we can test them against the set of *generalized cylinder functions*, already introduced in [Del22; Del24].

Definition 1.4 (Generalized cylinder functions)**.** *Given* $\hat\phi \in C_c^1(\mathbb{R}^d \times (0,1])$*, define*

$$\hat L_{\hat\phi}(\mu) := \int_{\mathbb{R}^d} \phi(x,\mu[x])d\mu(x), \quad \mu[x] := \mu(\{x\}).$$

Then the set of cylinder functions $\mathrm{GC}_c^1(\mathcal{P}(\mathbb{R}^d))$*is made of functions* $\hat F : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ *is in* $\mathrm{GC}_c^1(\mathcal{P}(\mathbb{R}^d))$*, if*

$$\hat F(\mu) = \Psi\left(\hat L_{\hat\phi_1}(\mu),\dots,\hat L_{\hat\phi_k}(\mu)\right), \tag{1.15}$$

for some $k \in \mathbb{N}$*,* $\Psi \in C^1(\mathbb{R}^k)$*,* $\hat\phi_i \in C_c^1(X \times (0,1])$*. The Wasserstein differential of a generalized cylinder function is defined by*

$$\nabla_W \hat F(x,\mu) := \sum_{i=1}^k \partial_i \Psi(\hat L_{\hat{\boldsymbol\phi}}(\mu))\nabla_x \phi_i(x,\mu[x]). \tag{1.16}$$

In the case $\Psi \in C^\infty(\mathbb{R}^k)$ *and* $\hat\phi_1,\dots,\hat\phi_n \in C_c^\infty(\mathbb{R}^d \times (0,1])$*, we say that* $\hat F \in \mathrm{GC}_c^\infty(\mathcal{P}(\mathbb{R}^d))$*.*

The capacity assumption is needed to suitably approximate any generalized cylinder function with a sequence of cylinder functions, see Theorem 4.3, whose proof has been inspired by similar approximation results in [Del22; Del24]. This approximation will be the ingredient that will allow us to use them as test functions for the continuity equations, so that now we can rely on a positive answer to Question (1.1) for Λ-a.e. $(\mu_t)_{t\in[0,T]}$ to conclude the proof. The details are left for Section 4.2.

Metric counterpart and the Riemannian case. A metric version of the atomic nested superposition principle, Theorem 1.3, can be proved as well. We mean that, instead of working with a curve of measures $(M_t)_{t\in[0,T]} \subset \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ that solves some continuity equation, we assume that it is p-absolutely with respect to the p-Wasserstein on Wasserstein metric. The rigorous statement is given in Theorem 4.5. This result relies the strong link between curves of measures that solves a continuity equation and absolutely continuous curves, as shown in [AGS08; ST17; PS25].

In Section 5, the capacity assumption in Theorem 1.3 will be extensively studied. In particular, we will see how the exponent r is related to the geometry of the support of the measure ν. This analysis, will allow us to work with ν that are supported on an embedded compact Riemannian manifold without boundary, and in particular, the Nash isometric embedding theorem will give also an intrinsic version of Theorem 1.3, when the ambient space $\mathbb{R}^d$ is replaced with a compact, oriented, and boundaryless Riemannian manifold.

Application to the metric measure Wasserstein space. Assume now that X is either $\mathbb{R}^d$ or a compact, oriented, and boundaryless Riemannian manifold, and fix $Q = Q_{\pi,\nu} \in \mathcal{Q}(X)$, such that $\nu \in \mathcal{P}_p(X)$, for $p \in [1,+\infty)$. We use a metric-measure viewpoint to introduce the space of Sobolev functions over $\mathcal{P}_p(X)$, denoted $W^{1,p}(\mathcal{P}_p(X), W_p, Q)$. In particular, the definition through test plans fits particularly well in our setting, and it can be used to show that, if p is large enough, the Sobolev-to-Lipschitz property is not satisfied, that is: there exists $F \in W^{1,p}(\mathcal{P}_p(X), W_p, Q)$ with bounded gradient $|DF| \leq L$, for which it does not exist a L-Lipschitz representative. Our counterexample shows actually something stronger: in fact, we build a function $F \in W^{1,p}(\mathcal{P}_p(X), W_p, Q)$ that has null gradient Q-a.e., but that cannot admit a constant representative. In particular, this counterexample shows the lack of a Poincaré inequality as well: for all $N \in \mathbb{N}$ it holds

$$\int_{\mathcal{P}_p(X)} \left| F(\mu) - \int_{\mathcal{P}_p(X)} F dQ_{\pi,\nu} \right|^p dQ_{\pi,\nu}(\mu) > N \int_{\mathcal{P}_p(X)} |DF|^p(\mu) dQ_{\pi,\nu}. \tag{1.17}$$

The previous construction works under very general assumptions:

- if $X = \mathbb{R}^d$, with $d \geq 2$, and $\nu \ll \mathcal{L}^d$ with bounded density, then the previous counterexample applies for $p \geq \frac{d}{d-1}$, see Theorem 7.5 and Proposition 5.6;
- in the Riemannian case, say that $k \geq 2$ is the dimension of the manifold and that ν is absolutely continuous with respect to its volume measure, with bounded density. Then, the previous counterexample works for $p \geq \frac{k}{k-1}$, see Theorem 7.8.

In the previous assumptions, we can see that the case $p = 2$ is covered, implying that the metric measure space $(\mathcal{P}_2(X), W_2, Q)$ cannot be a RCD space, whose definition is recalled in Definition 7.3. However, thanks to [FSS23], it is always an infinitesimally Hilbertian space. Moreover, only for the Riemannian setting, in Section 7.2 we show that it also shares curvature property in the Bakry–Émery sense, if the base space (X, d, ν) does, see Theorem 7.13. A similar study was done in [DHS23], and in [DS21; DS22] for configuration spaces. The proof of the Bakry–Émery property relies on the characterization of the diffusion process associated with $(\mathcal{P}_2(X), W_2, Q)$ (and its heat semigroup) proved in [Del24], that has been proved only in the compact manifold case. If such a characterization holds in the Euclidean setting as well, then the Bakry–Émery property can be recovered in this setting as well. For more details on this discussion, see Section 7.

It is important to stress that the validity of the solely Bakry–Émery property is, in general, not strong enough. For example, under the ergodicity assumption of the reference measure Q, it can be shown that the Bakry–Émery condition gives the validity of a Poincaré inequality, see [BGL14; CL24]. In our case this cannot hold by absence of ergodicity, as already observed in [Del24, Theorem 1.4].

Plan of the paper. In **Section 2**, we give some preliminaries, in particular we fix the notation about the space of probability measures, together with its natural narrow topology and the Wasserstein distance. We also recall the atomic topology presented in [EK94], that will be useful throughout the paper. Then, we recall the standard and the nested superposition principle, and we define the notion of (p, ν)-capacity of the diagonal $\Delta \subset \mathbb{R}^d \times \mathbb{R}^d$.

In **Section 3**, we build the example of a W_∞-Lipschitz curve of purely atomic measures that does not admit a purely atomic lifting.

In **Section 4**, we move to the proof of our main results, the atomic nested superposition principles, in the Euclidean setting. In particular, we introduce the set of generalized cylinder functions and we show how the capacity assumption allows to use them to test the continuity.

Then, in **Section 5** we make an analysis on the capacity assumption (see also Appendix A).

In **Section 6**, we study the same problem when the underlying space is a boundaryless, oriented, and compact Riemannian manifold. The key tool is the Nash isometric embedding theorem, that allows to move back and forth all the objects between the manifold and the Euclidean space in which it is isometrically embedded.

Finally, in **Section 7**, we introduce general metric measure spaces, and in particular the definition of Sobolev space through test plans, and the synthetic notion of Ricci curvature bounded from below known as Bakry–Émery condition. We then prove the lackness of the Sobolev-to-Lipschitz property and the Poincaré inequality as a consequence of the atomic nested superposition principles (both in the Euclidean and Riemannian setting), and we conclude the analysis of the Wasserstein metric measure space showing the inheritance of the Bakry–Émery condition from the underlying space (in the Riemannian setting).

Acknowledgments. The author warmly thanks Lorenzo Dello Schiavo, Giuseppe Savaré and Giacomo Enrico Sodini for the fruitful discussions on this topic.

2. Preliminaries

Here is a list of the main notation used in the following:

$C_b(X)$, $(C_b(X;\mathbb{R}^n))$	continuous and bounded functions from X to $\mathbb{R}$ (resp. $\mathbb{R}^n$)
Δ, Δ_ε	resp., the diagonal and the ε-strip around the diagonal of $\mathbb{R}^d\times\mathbb{R}^d$
$\mathcal{B}(X)$	Borel sets of X
$\mathcal{P}(X)$	the space of Borel probability measures over a Polish space X
$\mathcal{P}^{\mathrm{pa}}(X)$	the space of purely atomic probability measures
$\mathcal{P}(B)$, for $B\in\mathcal{B}(x)$	probability measures in $\mathcal{P}(X)$ concentrated over B
$\mathcal{P}(\mathcal{P}(X))$	the space of random measures over a Polish space X
$g_\sharp$, $g_{\sharp\sharp}$	resp., the push forward and the double push-forward (see (2.3))
$\widetilde{M}$	see (2.4)
$W_{p,d}$	the p-Wasserstein distance built on the metric space (X,d)
$C([0,T],X)$	continuous curves in X
$AC^p([0,T],X)$	absolutely continuous curves in (X,d)
e_t, $\mathfrak{e}_t$, E_t	see §2.2.1

Let (X,τ) be a Polish space. We will denote by $\mathcal{B}(X)$ its Borel σ-algebra.

2.1. Probability measures, narrow topology and Wasserstein distance.

2.1.1. *Space of probability measures.* Let $\mathcal{P}(X)$ be the space of the Borel probability measures over X. We endow it with the narrow topology, i.e. the coarsest one for which

$$\mathcal{P}(X)\ni\mu\mapsto\int_X\phi\,d\mu \;\text{ is continuous for all } \phi\in C_b(X), \tag{2.1}$$

where $C_b(X)$ denotes the space of real-valued, bounded, and continuous functions.

If not otherwise specified, the space $\mathcal{P}(X)$ is always considered endowed with the narrow topology, denoted $\tau_{\mathtt{n}}$. In particular, we denote by $C([0,T],\mathcal{P}(X))$ the space of curves of probability measures that are narrowly continuous. For any $\mu_n,\mu\in\mathcal{P}(X)$, we write $\mu_n\to\mu$ if the convergence holds with respect to the narrow topology.

Given a measurable function $f: Z_1\to Z_2$, where $(Z_i,\mathcal{S}_i)$ are general measurable spaces, and a measure $\mu\in\mathcal{P}(Z_1)$, we denote with $f_\sharp\mu\in\mathcal{P}_+(Z_2)$ the push-forward measure, defined as

$$f_\sharp\mu(S):=\mu(f^{-1}(S))\quad\forall S\in\mathcal{S}_2.$$

The same construction works for the set of Borel non-negative measures $\mathcal{M}_+(X)$, that is a Polish space when endowed with the narrow topology.

We will often denote as $\mathcal{P}(B)$, for a given set $B \in \mathcal{B}(X)$, the probability measures $\mu \in \mathcal{P}(X)$ for which $\mu(B) = 1$.

2.1.2. *Space of (laws of) random probability measures.* The resulting space $\mathcal{P}(X)$ endowed with narrow topology is still a Polish space, so we can reiterate the construction and endow $\mathcal{P}(\mathcal{P}(X))$ with the so-called narrow over narrow topology. As above, with $C([0,T], \mathcal{P}(\mathcal{P}(X)))$ we denote the space of continuous curves of random probability measures.

Given $M \in \mathcal{P}(\mathcal{P}(X))$ we define the *barycenter measure* $\mathrm{bar}[M] \in \mathcal{P}(X)$ via

$$\int_X f(x) d\,\mathrm{bar}[M](x) = \int_{\mathcal{P}(X)} \int_X f(x) d\mu(x) dM(\mu) \quad \text{for all Borel function } f : X \to [0,1]. \tag{2.2}$$

Moreover, it will be useful the iterated push-forward of a random measure induced by a function $g : X \to Y$, where X and Y are two Polish spaces: given a random measure $M \in \mathcal{P}(\mathcal{P}(X))$, we denote

$$g_{\sharp\sharp} M := G_\sharp M, \quad G : \mathcal{P}(X) \to \mathcal{P}(Y),\ G(\mu) := g_\sharp \mu. \tag{2.3}$$

It is well defined due to [PS25, Appendix D], where the Borel measurability of G is shown. Moreover, for all $M \in \mathcal{P}(\mathcal{P}(X))$, we will denote by $\widetilde{M} \in \mathcal{P}(X \times \mathcal{P}(X))$ the measure defined through

$$\int H(x,\mu) d\widetilde{M}(x,\mu) = \int \int H(x,\mu) d\mu(x) dM(\mu) \text{ for all measurable } H : X \times \mathcal{P}(X) \to [0,1]. \tag{2.4}$$

2.1.3. *Wasserstein metric.* Fix $p \in [1, +\infty]$, and a complete and separable metric space (X, d). As usual, with $\mathcal{P}_p(X)$ we denote the space of probability measures μ for which

$$\|d(\cdot, x^*)\|_{L^p(\mu)} < +\infty \quad \text{for some (and then all) } x^* \in X. \tag{2.5}$$

Then, the p-Wasserstein distance between $\mu, \nu \in \mathcal{P}_p(X)$ is defined as

$$W_{p,d}(\mu,\nu) := \min \left\{ \|d(\cdot,\cdot)\|_{L^p(X\times X,\pi)} \ : \ \pi \in \Gamma(\mu,\nu) \right\},$$

$$\Gamma(\mu,\nu) := \Big\{ \pi \in \mathcal{P}(X \times X) \ : \ \pi(A \times X) = \mu(A), \quad \pi(X \times B) = \nu(B) \quad \forall A, B \in \mathcal{B}(X) \Big\}.$$

It is well-known [AGS08, Proposition 7.1.5] that $(\mathcal{P}_p(X), W_{p,d})$ is complete and separable, and for $p \in [1, +\infty)$ the convergence induced is

$$W_{p,d}(\mu_n, \mu) \to 0 \iff \mu_n \to \mu \quad \text{and} \quad \int_X d^p(x,x^*) d\mu_n(x) \to \int_X d^p(x,x^*) d\mu(x).$$

In particular, if the distance d is bounded, $\mathcal{P}_p(X) = \mathcal{P}(X)$ and $W_{p,d}$ metrizes the narrow topology. As for the narrow over narrow topology, we can reiterate the construction to build the *Wasserstein on Wasserstein* distance $\mathcal{W}_p := W_{p,W_p}$ over $\mathcal{P}_p(\mathcal{P}_p(X))$.

2.1.4. *Purely atomic measures and atomic topology.* Let (X, τ) be a Polish space. We introduce some notations used in [Del22; Del24]. To introduce the set of purely atomic probability measures, denote the infinite simplex $\mathbf{T} \subset [0,1]^{\mathbb{N}}$ as in (1.1), and its subset $\mathbf{T}_0$ as

$$\mathbf{T}_0 := \{ \mathbf{a} \in \mathbf{T} \ : \ a_{i+1} < a_i \ \forall i \in \mathbb{N} \}. \tag{2.6}$$

Moreover, denote by $X^\infty_{\neq} \subset X^{\mathbb{N}}$ the set of sequences $(x_n)_{n\in\mathbb{N}}$ such that $x_i \neq x_j$ for all $i \neq j$. Both $\mathbf{T}$ and $X^{\mathbb{N}}$ are endowed with the infinite product topology. We then consider the following spaces of purely atomic measures:

$$\begin{aligned}
\mathcal{P}^{\mathrm{pa}}(X) &:= \left\{ \mu \in \mathcal{P}(X) \ : \exists (z_i)_{i\in\mathbb{N}} \in X^\infty_{\neq},\ \exists (a_i)_{i\in\mathbb{N}} \in \mathbf{T},\ \text{s.t. } \mu = \sum_i a_i \delta_{z_i} \right\}, \\
\mathcal{P}^{\mathrm{pa}}_0(X) &:= \left\{ \mu \in \mathcal{P}(X) \ : \ \exists (z_i)_{i\in\mathbb{N}} \in X^\infty_{\neq},\ \exists (a_i)_{i\in\mathbb{N}} \in \mathbf{T},\ \text{s.t. } \mu = \sum_i a_i \delta_{z_i} \right\},
\end{aligned} \tag{2.7}$$

We will make use of an useful topology over the space of probability measures, called 'atomic topology'. We recall here the definition and its main properties, for which we refer to [EK94, §2]. Given $\mu \in \mathcal{P}(X)$, we denote

$$\mu^* := \sum_{x\in X} \mu(\{x\})^2 \delta_x \in \mathcal{M}_+(X). \tag{2.8}$$

Definition 2.1 (Atomic convergence)**.** *Let $\mu_n, \mu \in \mathcal{P}(X)$. We say that $\mu_n \xrightarrow{\tau_a} \mu$ if $\mu_n \to \mu$ narrowly and $\mu_n^*(X) \to \mu^*(X)$.*

Lemma 2.2. *Let d be any distance that induces τ and such that (X, d) is complete. Let $\Psi : [0, +\infty) \to [0, 1]$ be a continuous and non-increasing function. Then, define the distance $D_{p,d,\Psi}$ over $\mathcal{P}_p(X, d)$ as*

$$D_{p,d,\Psi}(\mu, \nu) := W_{p,d}(\mu, \nu) + \sup_{\varepsilon\in(0,1)} \left| \int_{X\times X} \Psi\left(\frac{d(x, y)}{\varepsilon}\right) d\big(\mu \otimes \mu - \nu \otimes \nu\big)(x, y) \right|. \tag{2.9}$$

Then $D_{p,d,\Psi}(\mu_n, \mu) \to 0$ if and only if $W_{p,d}(\mu_n, \mu) \to 0$ and $\mu_n^* \to \mu^*$. In particular, if d is bounded, $D_{p,d,\Psi}$ induces the atomic convergence and the induced topology $\tau_{\mathtt{a}}$, called atomic topology, is Polish. In particular, the Borel sets of $\mathcal{P}(X)$ built with respect to this topology, coincide with the ones built using the narrow topology.

Proof. The first part is a consequence of Lemma 2.1, Lemma 2.2 and Lemma 2.3 from [EK94]. The second part follows from [Sch73] and the fact that τ_a is stronger than τ_n. □

The following lemma shows how the atomic convergence actually forces convergence of the atoms and their weights. For the proof, we refer to [EK94, Lemma 2.5].

Lemma 2.3. *Assume that $\mu_n \to \mu$ narrowly. Let $\{a_i^{(n)}, x_i^{(n)}\}_{i\in\mathbb{N}}$ and $\{a_i, x_i\}_{i\in\mathbb{N}}$ be, respectively, the set of atoms of μ_n and μ, with locations $(x_i^{(n)})_{i\in\mathbb{N}}, (x_i)_{i\in\mathbb{N}} \in X^\infty_{\neq}$ and ordered weights $(a_i^{(n)})_{i\in\mathbb{N}}, (a_i)_{i\in\mathbb{N}} \in \mathbf{T}$. Then*

(1) $\mu_n \xrightarrow{\tau_a} \mu$ if and only if $a_i^{(n)} \to a_i$ for all $i \in \mathbb{N}$. Moreover, if $a_k > a_{k+1}$ for some $k \in \mathbb{N}$, then the set of locations $\{x_1^{(n)}, \dots, x_k^{(n)}\} \to \{x_1, \dots, x_k\}$;
(2) if $\mu \in \mathcal{P}^{\mathrm{pa}}(X)$, then $\mu_n \xrightarrow{\tau_a} \mu$ if and only if $\sum_i |a_i^{(n)} - a_i| \to 0$.

The next result can be found in [Del24]. Anyway, we give an original proof for completeness.

Lemma 2.4. *The sets $\mathcal{P}^{pa}(X)$ and $\mathcal{P}^{pa}_0(X)$ are Borel subsets of $\mathcal{P}(X)$.*

Proof. The set $\mathcal{P}^{\mathrm{pa}}(X)$ is closed in the atomic topology, and in particular it is a Borel subset of $\mathcal{P}(X)$. Regarding $\mathcal{P}^{\mathrm{pa}}_0(X)$, for any $j \in \mathbb{N}$, define $\mathcal{P}^{\mathrm{pa}}_{0,j}(X) := \{\mu \in \mathcal{P}^{\mathrm{pa}}(X) \ : \ a_j > a_{j+1}\}$ which is open because its complement $\{\mu \in \mathcal{P}^{\mathrm{pa}}(X) \ : \ a_j = a_{j+1}\}$ is closed thanks to Lemma 2.3. Then, simply notice that $\mathcal{P}^{\mathrm{pa}}_0(X) = \bigcap_{j\in\mathbb{N}} \mathcal{P}^{\mathrm{pa}}_{0,j}(X)$. □

A last simple but useful lemma is stated here.

Lemma 2.5. *Let X and Y be Polish space and $f : X \to Y$ a Borel map. If $\mu \in \mathcal{P}^{\mathrm{pa}}(X)$, then $f_\sharp \mu \in \mathcal{P}^{\mathrm{pa}}(Y)$.*

Proof. A simple calculation shows that, writing $\mu = \sum_i a_i \delta_{x_i}$, for all $G : Y \to [0,1]$ Borel function

$$\int_Y G(y) d(f_\sharp \mu)(y) = \sum_i a_i G(f(x_i)) = \int_Y G(y) d\Big(\sum_i a_i \delta_{f(x_i)} \Big),$$

giving that $f_\sharp \mu = \sum_i a_i \delta_{f(x_i)} \in \mathcal{P}^{\mathrm{pa}}(Y)$. □

2.2. **Curves of probability measures and probability over curves.**

2.2.1. *Probability measures over curves.* Let (X, τ) be a Polish space. The space $C([0,T], X)$ of continuous curves in X, endowed with the compact-open topology, is a Polish space. The function $\mathrm{e}_t : C([0,T], X) \to X$ corresponds to the evaluation at a given time $t \in [0,T]$.
The previous setting applies to the spaces $C([0,T], \mathcal{P}(X))$ and $\mathcal{P}(C([0,T], X))$, so that we can consider probability measures over them, that we will usually denote as $\Lambda \in \mathcal{P}(C([0,T], \mathcal{P}(X)))$ and $\mathfrak{L} \in \mathcal{P}(\mathcal{P}(C([0,T], X)))$.

Since these last two spaces will play a fundamental role for the rest of the paper, we fix some additional notation regarding them:

- the evaluation map from $C([0,T], \mathcal{P}(X))$ to $\mathcal{P}(X)$ at a fixed time $t \in [0,T]$, will be denoted as $\mathfrak{e}_t$;
- the push-forward of the evaluation map e_t will be denoted as $E_t := (\mathrm{e}_t)_\sharp$. In particular, $(E_t)_\sharp = (\mathrm{e}_t)_{\sharp\sharp}$

2.2.2. *Absolutely continuous curves.* Let (X, d) be a complete and separable metric space. We say that a continuous curve $\boldsymbol{x} : [0,T] \to X$ is absolutely continuous, and we write $\boldsymbol{x} \in AC([0,T], X)$, if it exists a function $g \in L^1(0,T)$ such that

$$d(\boldsymbol{x}(t), \boldsymbol{x}(s)) \leq \int_s^t g(r) dr \quad \text{whenever} \quad 0 \leq s \leq t \leq T. \tag{2.10}$$

If $g \in L^p(0,T)$, for some $p \in (1, +\infty]$, we say that $\boldsymbol{x} \in AC^p([0,T], X)$. The space $AC^p([0,T], X)$, for $p \in [1, +\infty]$ is a Borel subset of $C([0,T], X)$.

It is well known [AGS08] that absolutely continuous curves admit a *metric derivative*, that is,

$$\lim_{s \to t} \frac{d(\boldsymbol{x}(s), \boldsymbol{x}(t))}{|t-s|} =: |\dot{\boldsymbol{x}}|_d(t) \tag{2.11}$$

exists for $\mathcal{L}^1$-a.e. $t \in [0,T]$ and it provides the smallest g (in an almost everywhere pointwise sense) that fits in (2.10).

In particular, we can apply this definition to the Wasserstein space $(\mathcal{P}_p(X), W_p)$ and the Wasserstein on Wasserstein space $(\mathcal{P}_p(\mathcal{P}_p(X)), \mathcal{W}_p)$. In this paper we will deal with the spaces:

- $AC^p([0,T], \mathcal{P}_p(X))$ that is a Borel subset of $C([0,T], \mathcal{P}(X))$ for all $p \geq 1$;
- $AC^p([0,T], \mathcal{P}_p(\mathcal{P}_p(X)))$ that is a Borel subset of $C([0,T], \mathcal{P}(\mathcal{P}(X)))$ for all $p \geq 1$.

2.2.3. *Continuity equation for curve of measures.* Given a Borel measurable vector field $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and a curve of probability measures $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathbb{R}^d))$, we say that $\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0$ if

$$\begin{gathered}\int_0^T \int_{\mathbb{R}^d} |v_t(x)| d\mu_t(x) dt < +\infty \quad \text{and} \quad \text{for all } \xi \in C_c^1(0,T),\ \phi \in C_c^1(\mathbb{R}^d) \\ \int_0^T \xi'(t) \int_{\mathbb{R}^d} \phi(x) d\mu_t(x) dt = -\int_0^T \xi(t) \int_{\mathbb{R}^d} \nabla \phi(x) \cdot v(t,x) d\mu_t(x) dt.\end{gathered} \tag{2.12}$$

We recall some important properties of curves that are solution of a continuity equation: they characterize absolutely continuous curve and they can be represented as superposition of solution to the associated ordinary differential equation.

Theorem 2.6. *[AGS08, Theorem 8.3.1] Let $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathbb{R}^d))$ and assume that $\mu_0 \in \mathcal{P}_p(\mathbb{R}^d)$. If $p \in (1, +\infty)$, the following are equivalent:*

(1) $\boldsymbol{\mu} \in AC^p([0,T], \mathcal{P}_p(\mathbb{R}^d))$;
(2) there exists $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\int_0^T \int_{\mathbb{R}^d} |v(t,x)|^p d\mu_t(x) dt < +\infty \quad \text{and} \quad \partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0. \tag{2.13}$$

Theorem 2.7. *[AGS08, Theorem 8.2.1] Let $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathbb{R}^d))$ be satisfying* (2.12)*. Then, there exists a probability measure $\mathcal{P}(C([0,T], \mathbb{R}^d))$ such that $(\mathrm{e}_t)_\sharp \lambda = \mu_t$ for all $t \in [0,T]$ and λ is concentrated over curves $\gamma \in AC([0,T], \mathbb{R}^d)$ that satisfies*

$$\dot\gamma(t) = v_t(\gamma(t)) \quad \textit{for a.e. } t \in [0,T]. \tag{2.14}$$

Vice versa, if $\lambda \in \mathcal{P}(C([0,T], \mathbb{R}^d))$ is concentrated over absolutely continuous curves that satisfy (2.14) *and $\int \int_0^T |\dot\gamma(t)| dt d\lambda(\gamma) < +\infty$, then $\mu_t := (\mathrm{e}_t)_\sharp \lambda$ satisfies* (2.12)*.*

As a consequence of the previous two theorems, we recover the lifting result by S. Lisini [Lis07] in the Euclidean space.

Theorem 2.8. *Let $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in AC^p([0,T], \mathcal{P}_p(\mathbb{R}^d))$ with $p \in (1, +\infty)$. Then, there exists a lifting $\lambda \in \mathcal{P}(C([0,T], \mathbb{R}^d))$ such that $(\mathrm{e}_t)_\# \lambda = \mu_t$ for all $t \in [0,T]$ and λ is concentrated over $AC^p([0,T], \mathbb{R}^d)$. Moreover,*

$$\int \int_0^T |\dot\gamma(t)|^p dt d\lambda(\gamma) = \int_0^T |\dot{\boldsymbol{\mu}}|_{W_p}^p < +\infty. \tag{2.15}$$

On the other hand, for any $p \in [1, +\infty)$, given $\lambda \in \mathcal{P}(C_T(\mathbb{R}^d))$ concentrated over absolutely continuous curves, with $(e_0)_\# \lambda \in \mathcal{P}_p(\mathbb{R}^d)$, the curve $\boldsymbol{\mu} := ((e_t)_\# \lambda)_{t\in[0,T]}$ belongs to $AC_T^p(\mathcal{P}_p(\mathbb{R}^d))$.

For simplicity, we presented this structure in the space $\mathbb{R}^d$, but it can be recovered in more general metric spaces as well, as shown in [ST17].

2.2.4. *Continuity equation for random measures.* In [PS25], the authors showed that a similar structure holds at the level of random measures.

As already said in the introduction, here the role of the leading vector field is played by a non-local Borel measurable vector field $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$, while the class of 'smooth' functions to test the continuity equation is the one of cylinder functions, defined in 1.2. Now, let

$\boldsymbol{M} = (M_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathcal{P}(\mathbb{R}^d)))$ be a continuous curve of random probability measures: we say that $\partial_t M_t + \mathrm{div}_{\mathcal{P}}(b_t M_t) = 0$ if

$$\begin{aligned}&\int_0^T \int_{\mathcal{P}} \int_{\mathbb{R}^d} |b_t(x,\mu)| d\mu(x) dM_t(\mu) dt < +\infty \text{ and for all } \xi \in C_c^1(0,T),\ F \in \mathrm{Cyl}_c^1(\mathcal{P}(\mathbb{R}^d))\\ &\int_0^T \xi'(t) \int_{\mathcal{P}} F(\mu) dM_t(\mu) dt = -\int_0^T \xi(t) \int_{\mathcal{P}} \int_{\mathbb{R}^d} \nabla_W F(x,\mu) \cdot b_t(x,\mu) d\mu(x) dM_t(\mu) dt.\end{aligned} \tag{2.16}$$

Then, we recall some of the main results presented in [PS25] (see also [Pin25] for more general integrability conditions and for stochastic evolutions).

Theorem 2.9. *Let $\boldsymbol{M} = (M_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathcal{P}(\mathbb{R}^d)))$ and assume that $M_0 \in \mathcal{P}_p(\mathcal{P}_p(\mathbb{R}^d))$. If $p > 1$, the following are equivalent:*

(1) $\boldsymbol{M} \in AC^p([0,T], \mathcal{P}_p(\mathcal{P}_p(\mathbb{R}^d)))$;

(2) there exists $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ Borel measurable such that

$$\int_0^T \int_{\mathcal{P}} \int_{\mathbb{R}^d} |b_t(x,\mu)|^p d\mu(x) dM_t(\mu) dt < +\infty \quad \textit{and} \quad \partial_t M_t + \mathrm{div}_{\mathcal{P}}(b_t M_t) = 0. \tag{2.17}$$

Theorem 2.10. *Let $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ be Borel measurable and $\boldsymbol{M} = (M_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathcal{P}(\mathbb{R}^d)))$ be satisfying (2.16). Then, there exists $\Lambda \in \mathcal{P}(C([0,T], \mathcal{P}(\mathbb{R}^d)))$ and $\mathfrak{L} \in \mathcal{P}(\mathcal{P}(C([0,T], \mathbb{R}^d)))$ such that $(\mathfrak{e}_t)_\# \Lambda = M_t$ and $(E_t)_\# \mathfrak{L} = M_t$ for any $t \in [0,T]$, and:*

(1) Λ-a.e. $\boldsymbol{\mu}$ belongs to $AC([0,T], \mathcal{P}(\mathbb{R}^d))$ and solves

$$\partial_t \mu_t + \mathrm{div}(b_t(\cdot, \mu_t)\mu_t) = 0 \quad \textit{as in } (2.12).$$

(2) $\mathfrak{L}$-a.e. $\lambda \in \mathcal{P}(C([0,T], \mathbb{R}^d))$ is concentrated over absolutely continuous curves γ that are solutions of

$$\dot\gamma(t) = b(t, \gamma_t, (e_t)_\# \lambda) \quad \textit{in } (0,T).$$

Vice versa, if Λ satisfies (1) and $\int \int_0^T \int |b_t(x,\mu_t)| d\mu_t(x) dt d\Lambda(\boldsymbol{\mu}) < +\infty$, then $M_t := (\mathfrak{e}_t)_\sharp \Lambda$ satisfies (2.16). Similarly, if $\mathfrak{L}$ satisfies (2) and $\int \int \int_0^T |b_t(\gamma(t), (\mathrm{e}_t)_\sharp \lambda)| dt d\lambda(\gamma) d\mathfrak{L}(\gamma) < +\infty$, then $M_t := (E_t)_\sharp \mathfrak{L}$ satisfies (2.16).

Similarly to the classic continuity equation, a metric result can be recovered from the last two theorem, but we will not need it in the following.

Given a Borel non-local vector field $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$, we recall also the definition of the sets $\mathrm{CE}(b) \subset C([0,T], \mathcal{P}(\mathbb{R}^d))$ and $\mathrm{SPS}(b) \subset \mathcal{P}(C([0,T], \mathbb{R}^d))$:

$$\begin{aligned}\mathrm{CE}(b) := \Big\{ (\mu_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathbb{R}^d)) \ :\ &\int \int |b_t(x,\mu_t)| d\mu_t(x) dt < +\infty,\\ &\partial_t \mu_t + \mathrm{div}(b_t(\cdot,\mu_t)\mu_t) = 0 \Big\},\end{aligned} \tag{2.18}$$

$$\begin{aligned}\mathrm{SPS}(b) := \{ \lambda \in \mathcal{P}(C([0,T], \mathbb{R}^d)) \ :\ &\int \int |b_t(\gamma(t), (e_t)_\# \lambda)| d\lambda(\gamma) dt < +\infty,\\ &\lambda\big(\mathrm{AC}([0,T], \mathbb{R}^d)\big) = 1,\ \dot\gamma(t) = b_t(\gamma_t, (e_t)_\# \lambda)\ \mathcal{L}_T^1 \otimes \lambda\text{-a.e.} \}.\end{aligned} \tag{2.19}$$

As shown in [PS25, Section 5], they are Borel measurable. Moreover, in Theorem 2.10, (1) and (2) can be summarized, resp., as Λ concentrated on $\mathrm{CE}(b)$ and $\mathfrak{L}$ concentrated on $\mathrm{SPS}(b)$.

2.3. **The family $\mathcal{Q}$ of reference random atomic measures.** Let (X,τ) be a Polish space. We denote $\boldsymbol{X} := X^{\infty}_{\neq}$. Recall from (1.2) and (1.3), respectively, the definition of the map em and of the class of reference measures $\mathcal{Q}$. Then. we have the following results.

Lemma 2.11. *Endowing $\mathbf{T}\times\boldsymbol{X}$ with the product topology, the map* em *is continuous. Moreover, the image of* em *is $\mathcal{P}^{\mathrm{pa}}(X)$.*

Proof. It is clearly continuous considering the atomic topology over $\mathcal{P}(X)$, and thus Borel. □

Lemma 2.12. *Let $Q_{\pi,\nu}\in\mathcal{Q}$. The following hold:*

(1) $\mathrm{bar}[Q_{\pi,\nu}]=\nu$*;*
(2) let

$$c_1 := \int_{\mathbf{T}}\Big(\sum_{i\neq j} a_i a_j\Big)d\pi(\mathbf{a}),\quad c_2 := \int_{\mathbf{T}}\Big(\sum_{i=1}^{+\infty} a_i^2\Big)d\pi(\mathbf{a}).$$

Then for all $g : X\times X\to[0,1]$ Borel measurable it holds

$$\int_{\mathcal{P}(X)}\int_{X\times X} g(x,y)d\mu\otimes\mu(x,y)dQ_{\pi,\nu}(\mu) = c_1\int_{X\times X} g(x,y)d\nu\otimes\nu(x,y) + c_2\int_X g(x,x)d\nu(x). \tag{2.20}$$

In other words, given $C : \mathcal{P}(X)\to\mathcal{P}(X\times X)$ to be $C(\mu) := \mu\otimes\mu$, then $\mathrm{bar}[C_{\#}Q_{\pi,\nu}] = c_1\nu\otimes\nu + c_2(\mathrm{id},\mathrm{id})_{\#}\nu$.

Proof. For all $f : X\to[0,1]$ Borel measurable, we have

$$\begin{aligned}&\int_{\mathcal{P}(X)}\int_X f(x)d\mu(x)dQ_{\pi,\nu}(\mu) = \int_{\mathbf{T}}\left(\int_{\boldsymbol{X}}\sum_{i=1}^{+\infty} a_i f(x_i)d\nu^{\infty}(\mathbf{x})\right)d\pi(\mathbf{a})\\ &= \int_{\mathbf{T}}\left(\sum_{i=1}^{+\infty} a_i\int_{\boldsymbol{X}} f(x_i)d\nu^{\infty}(\mathbf{x})\right)d\pi(\mathbf{a}) = \int_{\mathbf{T}}\left(\sum_{i=1}^{+\infty} a_i\int_X f(x)d\nu(x)\right)d\pi(\mathbf{a}) = \int_X f(x)d\nu(x).\end{aligned}$$

Regarding the second part

$$\begin{aligned}&\int_{\mathcal{P}(X)}\int_{X^2} g(x,y)d\mu\otimes\mu(x,y)dQ_{\pi,\nu}(\mu) = \int_{\mathbf{T}}\int_{\boldsymbol{X}}\Big(\sum_{i,j\geq 1} a_i a_j g(x_i,x_j)\Big)d\nu^{\infty}(\mathbf{x})d\pi(\mathbf{a})\\ &= \int_{\mathbf{T}}\Big(\sum_{i\neq j} a_i a_j\int_{\boldsymbol{X}} g(x_i,x_j)d\nu^{\infty}(\mathbf{x}) + \sum_{i\geq 1} a_i^2\int_{\boldsymbol{X}} g(x_i,x_i)d\nu^{\infty}(\mathbf{x})\Big)d\pi(\mathbf{a})\\ &= \int_{\mathbf{T}}\Big(\sum_{i\neq j} a_i a_j\int_{X^2} g(x,y)d\nu\otimes\nu(x,y)\Big)d\pi(\mathbf{a}) + \int_{\mathbf{T}}\Big(\sum_{i\geq 1} a_i^2\int_X g(x,x)d\nu(x)\Big)d\pi(\mathbf{a})\\ &= c_1\int_{X^2} g(x,y)d\nu\otimes\nu(x,y) + c_2\int_X g(x,x)d\nu(x).\end{aligned}$$ □

We have many examples of measures $Q_{\pi,\nu}\in\mathcal{Q}$ coming from Bayesian non-parametric statistical theory. One is the Dirichlet-Ferguson measure $D_{\beta\nu}$ (see [Fer73; Del22]), where $\beta>0$: the measure $\pi\in\mathcal{P}(\mathbf{T})$ is determined by the stick-breaking procedure [Set94], and it is concentrated over $\mathbf{T}_0$, so that $D_{\beta\nu}$ is concentrated over $\mathcal{P}_0^{\mathrm{pa}}(X)$. Another example is the Poisson measure $P_{\lambda,\nu}$, with $\lambda>0$: a sample $\mu\sim P_\nu$ is built sampling $N\in\mathbb{N}$ according to a Poisson distribution with parameter λ and $\mathbf{x}\sim\nu^{\infty}$ independently, then $\mu = \frac{1}{N+1}\sum_{i=1}^{N+1}\delta_{x_i}$.

2.4. **Capacity of the diagonal with respect to $\nu \otimes \nu$.** The definition in this subsection are specific for $\mathbb{R}^d$. However, in Section 6 we will exploit it to bring on a similar study on compact Riemannian manifolds.

Given an atomless probability measure $\nu \in \mathcal{P}(\mathbb{R}^d)$, we will make use of the concept of *capacity* with respect to the product measure $\nu \otimes \nu$. In particular, given $r \geq 1$, we need the definition of (r,ν)-capacity of the diagonal set $\Delta := \{(x,x) \ : \ x \in \mathbb{R}^d\} \subset \mathbb{R}^d \times \mathbb{R}^d$, that we define as

$$\begin{aligned} \operatorname{cap}_{r,\nu}(\Delta) := \inf \bigg\{ \int_{\mathbb{R}^{2d}} |h(x,y)|^r + |\nabla_{x,y} h(x,y)|^r d\nu \otimes \nu(x,y) \ : \\ h \in C^1_b(\mathbb{R}^{2d}),\ h = 1 \text{ on } \Delta,\ h \leq 1 \bigg\}. \end{aligned} \tag{2.21}$$

A fundamental assumption for our main theorem will be that $\operatorname{cap}_{r,\nu}(\Delta) = 0$ for some $r \geq 1$. We will extensively study this assumption in Section 5, in which we will find sufficient conditions on ν and r for which that holds.

Before proceeding, we want to say that a more common definition for the capacity would be

$$\widetilde{\operatorname{cap}}_{r,\nu}(\Delta) := \inf \bigg\{ \int_{\mathbb{R}^{2d}} |h|^r + |\nabla_{x,y} h|^r d\nu \otimes \nu \ : h \in C^1_b(\mathbb{R}^{2d}),\ h \geq 1 \text{ on } \Delta \bigg\}. \tag{2.22}$$

In Appendix A, we show that the two definitions coincide, at least when $\nu \ll \mathcal{L}^d$.

3. Continuity equation for atomic measures

3.1. **An empirical superposition principle and the tangent space of purely atomic measures.** In this section, we will deal with narrowly continuous curves of atomic probability measures, i.e. $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in C([0,T],\mathcal{P}(\mathbb{R}^d))$ such that $\mu_t \in \mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)$ for all $t \in [0,T]$.

First, we recall a superposition principle for measures with a fixed finite number of atoms, shown in [CLOS22, Theorem C.1], that is an empirical version of Theorem 2.7. We denote

$$\mathcal{P}^N(\mathbb{R}^d) := \left\{ \mu \in \mathcal{P}(\mathbb{R}^d) \ : \ \exists x_1, \dots, x_n \in \mathbb{R}^d \text{ s.t. } \mu = \frac{1}{N}\sum_{i=1}^N \delta_{x_i} \right\} \quad \forall N \geq 1.$$

Theorem 3.1. *Let $(\mu_t)_{t\in[0,T]} \in C([0,T],\mathcal{P}(\mathbb{R}^d))$ be such that there exists $N \geq 1$ for which $\mu_t \in \mathcal{P}^N(\mathbb{R}^d)$ for all $t \in [0,T]$. Assume that there exists a Borel vector field $v : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and*

$$\int_0^T \int_{\mathbb{R}^d} |v(t,x)| d\mu_t(x) dt < +\infty \quad \textit{and} \quad \partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0. \tag{3.1}$$

Then there exists $\lambda \in \mathcal{P}^N(C([0,T],\mathbb{R}^d))$ such that $(\mathrm{e}_t)_\sharp \lambda = \mu_t$ for all $t \in [0,T]$ and such that

$$\lambda = \frac{1}{N}\sum_{i=1}^N \delta_{\gamma_i} \quad \textit{and} \quad \dot\gamma_i(t) = v(t,\gamma_i(t)) \textit{ for a.e. } t \in [0,T]. \tag{3.2}$$

Moreover, the vector field v for which (3.1) holds, is uniquely determined $\mu_t \otimes dt$-a.e.

Thanks to [AGS08, Proposition 8.4.5], the previous theorem can be used to show that for all $\mu \in \mathcal{P}^N(\mathbb{R}^d)$, the tangent space $\operatorname{Tan}_\mu \mathcal{P}_p(\mathbb{R}^d)$ is the whole $L^p(\mu;\mathbb{R}^d)$, for all $p > 1$. In the next proposition, we prove it for general $\mu \in \mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)$ directly from the definition of the tangent space.

Lemma 3.2. *For any $\mu \in \mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)$ and $p > 1$, we have*

$$\mathrm{CoTan}_\mu \mathcal{P}_p(\mathbb{R}^d) := \mathrm{Clos}_{L^{p'}(\mu;\mathbb{R}^d)} \left\{ \nabla f \ : \ f \in C_c^\infty(\mathbb{R}^d \right\} = L^{p'}(\mu;\mathbb{R}^d).$$

In particular, $\mathrm{Tan}_\mu \mathcal{P}_p(\mathbb{R}^d) := j_{p'}(\mathrm{CoTan}_\mu \mathcal{P}_p(\mathbb{R}^d)) = L^p(\mu;\mathbb{R}^d)$*, where* $j_{p'}(H) = |H|^{p'-2}H$ *for all* $H \in L^{p'}(\mu;\mathbb{R}^d)$

Proof. It suffices to show that if $H \in L^p(\mu;\mathbb{R}^d)$ satisfies

$$\int_{\mathbb{R}^d} H(x) \cdot \nabla f(x) d\mu(x) = 0 \quad \forall f \in C_c^\infty(\mathbb{R}^d), \tag{3.3}$$

then $H = 0$ μ-a.e. To fix the notation, let $\mu = \sum_{i\geq 1} a_i \delta_{x_i}$ with $(a_i)_{i\in\mathbb{N}} \in \mathbf{T}$ and $(x_i)_{i\in\mathbb{N}} \in (\mathbb{R}^d)^\infty_{\neq}$. Let $\rho \in C_c^\infty(\mathbb{R}^d)$ be a bump function, that is, such that $\operatorname{supp}\rho \subset B(0,2)$, $0 \leq \rho \leq 1$ and $\rho = 1$ on $B(0,1)$. Then, for all $i, m \in \mathbb{N}$ and $v \in \mathbb{R}^d$ let $f_{m,v,i}(x) := \rho(m(x-x_i))v\cdot(x-x_i) \in C_c^\infty(\mathbb{R}^d)$. Its gradient is given by $\nabla f_{m,v,i}(x) = \rho(m(x-x_i))v + mv\cdot(x-x_i)\nabla\rho(m(x-x_i))$, so that $\nabla f_{m,v,i}(x_i) = v$ and

$$|\nabla f_{m,v,i}(x_j)| \leq (|v|+m|v||x_j-x_i|\|\nabla\rho\|_\infty)\mathbb{1}_{B(0,\frac{2}{m})}(x_j-x_i) \leq |v|(1+2\|\nabla\rho\|_\infty)\mathbb{1}_{B(0,\frac{2}{m})}(x_j-x_i).$$

This shows that $\nabla f_{m,v,i}(x) \to v\mathbb{1}_{\{x_i\}}(x)$ as $m \to +\infty$ for μ-a.e. $x \in \mathbb{R}^d$, dominated by the right hand side in the previous computation, so that

$$0 = \int_{\mathbb{R}^d} H(x)\nabla f_{m,v,i}(x) d\mu(x) \longrightarrow H(x_i)\cdot v \quad \text{as } m \to +\infty$$

This shows that $H(x_i) = 0$ for all $i \in \mathbb{N}$ and concludes the proof. □

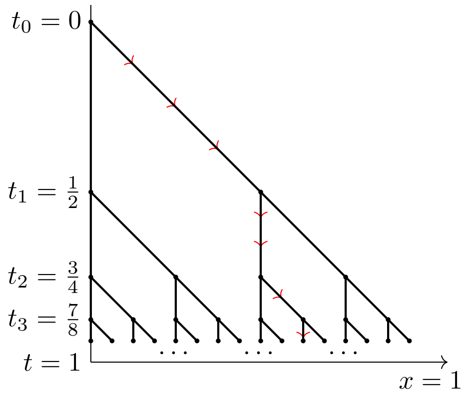
3.2. **Liftings of Lipschitz curves of purely atomic measures are not necessarily purely atomic measures of curves.** At this point, one is tempted to say that a version of Theorem 3.1 holds substituting $\mathcal{P}^N(\mathbb{R}^d)$ with $\mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)$. Unfortunately, this is not true, and in this section we build a counterexample for it. In fact, we build a curve of purely atomic probability measures that is Lipschitz with respect to W_∞, but for which a purely atomic lifting does not exist.

We proceed in the following way: consider the infinite product space $\Omega := \{0,1\}^{\mathbb{N}}$, endowed with the Borel σ-algebra generated by the product topology, and with the probability measure $\mathbb{P} := (\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1)^\infty$. Notice that $\mathbb{P}$ is an atomless probability measure. Consider the sequence of times $\{t_k\}_{k\geq 0}$ given by

$$t_0 = 0, \quad t_{k+1} = t_k + \frac{1}{2^{k+1}} \quad \forall k \geq 0$$

For any $\boldsymbol{\omega} = (\omega_0, \omega_1, \dots) \in \Omega$ let $\tilde{\gamma} \in C([0,1],\mathbb{R})$ be defined as

$$\tilde{\gamma}_{\boldsymbol{\omega}}(t) := \sum_{k\geq 0} \omega_k \big((t-t_k)_+ \wedge (t_{k+1}-t_k)\big). \tag{3.4}$$



The construction of these curves is summarized in the figure above, where the branches happen at the time t_k and the choice between left or right depends on the value of ω_k. In red, you can see the start of a curve corresponding to a sequence $\boldsymbol{\omega}$ with $\omega_1 = 1$, $\omega_2 = 0$, $\omega_3 = 1$ and $\omega_4 = 0$. It is clear that the map $\tilde{\iota} : \Omega \to C([0,1], \mathbb{R}^d)$ defined as $\tilde{\iota}(\boldsymbol{\omega}) = \tilde{\gamma}_{\boldsymbol{\omega}}$, is injective.

Then, define $\tilde{\lambda} := \tilde{\iota}_{\#}\mathbb{P} \in \mathcal{P}(C([0,1], \mathbb{R}))$ and $\tilde{\mu}_t := (\mathrm{e}_t)_\sharp \tilde{\lambda}$. Notice that $\tilde{\mu}_t \in \mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)$ for any $t \in [0,1)$: in fact, for any $t \in [0,1)$, there exists $n \in \mathbb{N}$ such that $t \in [t_n, t_{n+1}]$, and thanks to (3.4) we have

$$\tilde{\mu}_t = \frac{1}{2^{n+1}} \sum_{(\omega_0,\dots,\omega_n) \in \{0,1\}^{n+1}} \delta_{x_{\omega_0,\dots,\omega_n}(t)}, \quad \text{where } x_{\omega_0,\dots,\omega_n}(t) = \sum_{k=0}^{n-1} \frac{\omega_k}{2^k} + \omega_n(t - t_n).$$

Moreover, $(\mu_t)_{t\in[0,1]} \in \mathrm{Lip}\big([0,1], (\mathcal{P}([0,1]), W_\infty)\big)$: indeed, for any $t_n \le s < t \le t_{n+1}$, the transport plan

$$\pi_{s,t} := \frac{1}{2^{n+1}} \sum_{(\omega_0,\dots,\omega_n) \in \{0,1\}^{n+1}} \delta_{(x_{\omega_0,\dots,\omega_n}(t), x_{\omega_0,\dots,\omega_n}(s))}$$

is a coupling between μ_t and μ_s, and with a simple computation we can see that

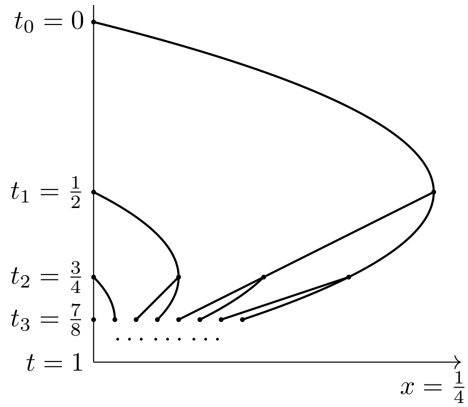
$$W_\infty(\mu_s, \mu_t) \le \|x - y\|_{L^\infty(\pi)} = (t - s).$$

Then, exploiting the triangular inequality and the continuity of the curve at the time $t = 1$, we achieve Lipschitzianity in the whole interval $[0,1]$.

At this point, the only thing we need to fix is the fact that $\tilde{\mu}_1$ is atomless, since $\tilde{\iota}$ is injective (it gives actually the uniform distribution of $[0,1]$). To do this, we consider:

(3.5)
$$\iota : \Omega \to \mathcal{P}(C([0,1], \mathbb{R})), \quad \iota(\boldsymbol{\omega}) = \gamma_{\boldsymbol{\omega}} \text{ where } \gamma_{\boldsymbol{\omega}}(t) = (1-t)\tilde{\gamma}_{\boldsymbol{\omega}}(t), \quad \lambda = \iota_{\#}\mathbb{P}, \quad \mu_t = (\mathrm{e}_t)_{\#}\lambda.$$

The distortion induced by the term $(1 - t)$ can be seen in the picture below (not on scale for the sake of visualization).



Now, it is not hard to see that:

(1) $\mu_t \in \mathcal{P}^{\mathrm{pa}}(\mathbb{R})$ for all $t \in [0,1]$ and $\mu_1 = \delta_0$;
(2) $(\mu_t)_{t\in[0,T]} \in \mathrm{Lip}\big([0,1], (\mathcal{P}(\mathbb{R}), W_\infty)\big)$;
(3) λ is atomless, since ι is injective as well.

We conclude thanks to the following lemma.

Lemma 3.3. *If $\lambda^1, \lambda^2 \in \mathcal{P}(C([0,1],\mathbb{R}))$ are such that $(\mathrm{e}_t)_\sharp\lambda^1 = (\mathrm{e}_t)_\sharp\lambda^2 = \mu_t$ for all $t \in [0,1]$, then $\lambda^1 = \lambda^2$. In particular, λ is the only admissible lifting of $(\mu_t)_{t\in[0,1]}$.*

Proof. Fix t_n for some $n \in \mathbb{N}$ and define, for $i = 1, 2$,

$$\Gamma_n := \big\{\gamma \in C([0,t_n],\mathbb{R}) \ : \ \gamma(t) \in \operatorname{supp}\mu_t \ \forall t \in [0,t_n]\big\}, \quad \lambda_n^i := (\mathrm{restr}_{[0,t_n]})_\sharp \lambda^i \in \mathcal{P}(C([0,t_n],\mathbb{R})),$$

where $\mathrm{restr}_{[0,t]} : C([0,T],\mathbb{R}) \to C([0,t],\mathbb{R})$ is the restriction map for all $t \in (0,T)$. Notice that λ_n is supported on Γ_n, that is a closed subset of $C([0,t_n],\mathbb{R})$. Moreover,

$$\operatorname{supp}\mu_{t_n} = \{(1-t_n)x_{a_0,\ldots,a_n}(t_n) \ : \ a_0,\ldots,a_{n-1} \in \{0,1\}\}$$

and the function $e_{t_n} : \Gamma_n \to \operatorname{supp}\mu_{t_n}$ is injective. Then, disintegrating both λ_n^1 and λ_n^2 with respect to e_{t_n} (see also [CSS25]), we see that

$$\mu_{t_n} = \frac{1}{2^n}\sum_{x\in\operatorname{supp}\mu_{t_n}} \delta_x \implies \lambda_n^1 = \lambda_n^2 = \frac{1}{2^n}\sum_{\gamma\in\Gamma_n}\delta_\gamma.$$

At this point, consider any $0 = s_0 < \cdots < s_k < s_{k+1} = 1$, and consider $t_n > s_k$, so that for any Borel set $A_0, \ldots A_{k+1}$ of $\mathbb{R}$, it holds

$$\lambda^i(\{\gamma \ : \ \gamma(s_j) \in A_j \ \forall j \le k+1\}) = \begin{cases} \lambda_n^i(\{\gamma \ : \ \gamma(s_j) \in A_j \ \forall j \le k\}) & \text{if } 0 \in A_{k+1} \\ 0 & \text{if } 0 \notin A_{k+1}, \end{cases}$$

and we conclude because the right hand side is independent on $i = 1, 2$, showing that $\lambda^1 = \lambda^2$. □

4. A superposition principle for laws of atomic random measures

From the result by S. Lisini in [Lis07], we know that the map $E = ((\mathrm{e}_t)_\sharp)_{t\in[0,T]}$ defined as

$$E : \mathcal{P}(C([0,T],\mathbb{R}^d)) \to C([0,T],\mathcal{P}(\mathbb{R}^d)), \quad E(\lambda) := \big((\mathrm{e}_t)_\sharp\lambda\big)_{t\in[0,T]},$$

is surjective seen as a map

$$E : \mathcal{P}(AC^p([0,T],\mathbb{R}^d)) \to AC^p([0,T],\mathcal{P}_p(\mathbb{R}^d)) \tag{4.1}$$

for all $p \in (1, +\infty)$. Similarly, thanks to the superposition result [AGS08, Theorem 8.2.1], E is surjective as a map

$$E : \mathrm{SPS}(b) \to \mathrm{CE}(b) \tag{4.2}$$

for all Borel $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$. In Section 3, we have shown that, if we restrict again the previous maps to purely atomic measures, we lose surjectivity: in fact, for $p > 1$

$$E : \mathcal{P}^{\mathrm{pa}}(C([0,T],\mathbb{R}^d)) \cap \mathcal{P}(AC^p([0,T],\mathbb{R}^d)) \to AC^p([0,T],\mathcal{P}_p(\mathbb{R}^d)) \cap C([0,T],\mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)) \tag{4.3}$$

is not onto, since the counterexample we built belongs to $AC^p([0,1],\mathcal{P}_p(\mathbb{R}^)) \cap C([0,1],\mathcal{P}^{\mathrm{pa}}(\mathbb{R}^))$ for all $p > 1$. Similarly, since for all $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in AC^p([0,T],\mathcal{P}_p(\mathbb{R}^d))$ there exists a vector field $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ such that $\boldsymbol{\mu} \in \mathrm{CE}(v)$, again the same counterexample shows that there exists a vector field v such that the following map is not onto:

$$E : \mathcal{P}^{\mathrm{pa}}(C([0,T],\mathbb{R}^d)) \cap \mathrm{SPS}(v) \to C([0,T],\mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)) \cap \mathrm{CE}(v). \tag{4.4}$$

Despite this lack of surjectivity, we will show that, for all $Q \in \mathcal{Q}$, given a non-local vector field $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ and a curve of random measures $(M_t)_{t\in[0,T]} \in C([0,T],\mathcal{P}(\mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)))$ that satisfies

$$\partial_t M_t + \mathrm{div}_{\mathcal{P}}(b_t M_t) = 0, \quad M_t \ll Q \ \text{ for all } t \in [0,T], \tag{4.5}$$

under the additional assumptions of Theorem 1.3, there exist:

(1) $\Lambda \in \mathcal{P}\big(C([0,T],\mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)) \cap \mathrm{CE}(b)\big)$ such that $(\mathfrak{e}_t)_\sharp \Lambda = M_t$ for all $t \in [0,T]$;
(2) $\mathfrak{L} \in \mathcal{P}\big(\mathcal{P}^{\mathrm{pa}}(C([0,T],\mathbb{R}^d)) \cap \mathrm{SPS}(b)\big)$ such that $(E_t)_\sharp \mathfrak{L} = M_t$ for all $t \in [0,T]$.

We will see that the key ingredient to find a lifting $\lambda \in \mathcal{P}^{\mathrm{pa}}(C([0,T],\mathbb{R}^d))$ for a curve $\boldsymbol{\mu} \in C([0,T],\mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)$ is to solve a continuity equation in duality with *generalized cylinder functions*, see Lemma 4.6.

4.1. **Approximation of generalized cylinder functions by cylinder functions.** For the whole section, we fix $\pi \in \mathcal{P}(\mathbf{T})$ and an atomless measure $\nu \in \mathcal{P}(\mathbb{R}^d)$, so that, for the sake of notation, we write $Q := Q_{\pi,\nu} \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$. Moreover, we denote $\mathbf{X} := (\mathbb{R}^d)^\infty_{\neq}$.

Recall from Definition 1.4, the notion of generalized cylinder functions. The aim of this section is to suitably approximate any generalized cylinder function with cylinder functions, with the aim of using them to test the continuity equation.

Notice that generalized cylinder functions are less regular than usual cylinder functions: they are still Borel measurable, since they are continuous with respect to the atomic topology, but not narrowly continuous. Moreover, they are 0 when evaluated in an atomless measure μ, see [Del24, Proposition 3.8] for more details. However, their Wasserstein differential is well-defined, since for all smooth vector field $\omega : \mathbb{R}^d \to \mathbb{R}^d$, given $\Psi^{\omega,t}$ its flow map at time t, it holds

$$\frac{d}{dt}\hat{F}(\Psi^{\omega,t}_{\#}\mu)|_{t=0} = \int_{\mathbb{R}^d} \nabla_W \hat{F}(x,\mu) \cdot \omega(x) d\mu(x).$$

Before proceeding, we need the following result, see also [Del22, Lemma 7.19].

Lemma 4.1. *Let $r \geq 1$. Given $\rho \in C^1(\mathbb{R})$, $f \in C^1_b(\mathbb{R}^d)$, $h \in C^1_b(\mathbb{R}^d \times \mathbb{R}^d)$, define*

$$F_{h,\rho,f}(\mu) := \int_{\mathbb{R}^d} f(x) \cdot \rho\left(\int_{\mathbb{R}^d} h(x,y) d\mu(y)\right) d\mu(x). \tag{4.6}$$

Then, there exists a sequence of cylinder functions $F_n \in \mathrm{Cyl}^1_c(\mathcal{P}(\mathbb{R}^d))$ such that

$$\int_{\mathcal{P}(\mathbb{R}^d)} |F_n(\mu) - F_{h,\rho,f}(\mu)|^r + \left(\int_{\mathbb{R}^d} |\nabla_W F_n - \nabla_W F_{h,\rho,f}|^r d\mu\right) dQ(\mu) \to 0, \tag{4.7}$$

where

$$\nabla_W F_{h,\rho,f}(x,\mu) := \rho\big(L_{h(x,\cdot)}(\mu)\big)\nabla f(x) + \rho'\big(L_{h(x,\cdot)}(\mu)\big) f(x)\int_{\mathbb{R}^d}\nabla_x h(x,y)d\mu(y) + \int_{\mathbb{R}^d}\rho'\big(L_{h(z,\cdot)}(\mu)\big)f(z)\nabla_y h(z,x)d\mu(z). \tag{4.8}$$

Proof. The definition of Wasserstein gradient is justified from the fact that for all smooth vector field $\omega:\mathbb{R}^d\to\mathbb{R}^d$, it holds

$$\begin{aligned}\frac{d}{dt}F_{h,\rho,f}(\Psi^{\omega,t}_{\#}\mu)|_{t=0} &= \int_{\mathbb{R}^d}\rho\left(\int_{\mathbb{R}^d}h(z,y)d\mu(y)\right)\nabla f(z)\cdot\omega(z)d\mu(z)+ \\ &+\int_{\mathbb{R}^d}f(z)\rho'\left(\int_{\mathbb{R}^d}h(z,y)d\mu(y)\right)\cdot\left(\int_{\mathbb{R}^d}\nabla_{z,y}h(z,y)\cdot(\omega(z),\omega(y))d\mu(y)\right)d\mu(z) \\ &= \int_{\mathbb{R}^d}\nabla_W F_{h,\rho,f}(x,\mu)\cdot\omega(x)d\mu(x).\end{aligned}$$

For the approximation result, consider $h_n\in\mathrm{Span}\big(g_1\otimes g_2\ :\ g_i\in C^1_b(\mathbb{R}^d)\big)\subset C^1_b(\mathbb{R}^d\times\mathbb{R}^d)$ converging to h on bounded sets w.r.t. $\|\cdot\|_{C^1}$ and such that $\|h_n\|_{C^1}\le 2\|h\|_{C^1}$. Consider also $\rho_n(a):=\sum_{j=1}^n c_j a^j$ be a sequence of polynomials converging to ρ uniformly on $[-2\|h\|_{C^1},2\|h\|_{C^1}]$ and define

$$F_n(\mu):=F_{h_n,\rho_n,f}(\mu).$$

Given the structure of the chosen functions h_n and ρ_n, the F_n's are cylinder functions, and it's not hard to see that they satisfy (4.7). □

We are going to use this result to approximate generalized cylinder functions with cylinder functions, under the additional assumption that the r-capacity of the diagonal $\Delta\subset\mathbb{R}^d\times\mathbb{R}^d$ is 0, see §2.4.

Proposition 4.2. *Let $r\in[1,+\infty)$ be such that $\mathrm{cap}_{r,\nu}(\Delta)=0$, where $\Delta:=\{(x,x)\in\mathbb{R}^d\times\mathbb{R}^d\ :\ x\in\mathbb{R}^d\}$. Then, for any $\hat F\in\mathrm{GC}^1_c(\mathcal{P}(\mathbb{R}^d))$, there exists a sequence $F_n\in\mathrm{Cyl}^1_c(\mathcal{P}(\mathbb{R}^d))$ such that*

$$\int_{\mathcal{P}(\mathbb{R}^d)}|F_n(\mu)-\hat F(\mu)|^r+\left(\int_{\mathbb{R}^d}|\nabla_W F_n(x,\mu)-\nabla_W\hat F(x,\mu)|^r d\mu(x)\right)dQ(\mu)\to 0. \tag{4.9}$$

Proof. It suffices to check the result for $\hat F=\hat L_{\hat\phi}$, for any $\hat\phi\in C^1_c(\mathbb{R}^d\times(0,1])$ of the form $\hat\phi(x,a)=f(x)\rho(a)$, with $f\in C^1_c(\mathbb{R}^d)$ and $\rho\in C^1_c((0,1])$. Moreover, thanks to Lemma 4.1, we can choose the approximating functions of the form (4.6).

Step 1: choice of the approximating sequence. Thanks to Lemma A.1, we can consider a sequence of functions $h_n\in C^1_b(\mathbb{R}^d\times\mathbb{R}^d)$ realizing the infimum in the definition of $\mathrm{cap}_{r,\nu}(\Delta)$ and such that $h_n\le 1$, $h_n(x,x)=1$ for all $x\in\mathbb{R}^d$ and

$$\int|h|^r+|\nabla h|^r d\nu\otimes\nu\to 0.$$

Then, define $F_n(\mu):=F_{h_n,\rho,f}(\mu)$.

Step 2: convergence of F_n to $\hat F$ in $L^r(Q)$.

$$\begin{aligned}\int_{\mathcal{P}(\mathbb{R}^d)}|F_n(\mu)-\hat F(\mu)|^r dQ(\mu) &= \int\left|\int_{\mathbb{R}^d}f(x)\left[\rho\left(\int_{\mathbb{R}^d}h_n(x,y)d\mu(y)\right)-\rho(\mu[x])\right]\right|^r dQ(\mu) \\ &\le\|f\|^r_\infty\int\int\left|\rho\left(\int_{\mathbb{R}^d}h_n(x,y)d\mu(y)\right)-\rho(\mu[x])\right|^r d\mu(x)dQ(\mu)\end{aligned}$$

$$
\begin{aligned}
&\leq \|f\|_\infty^r \|\rho'\|_\infty^r \int\int \left|\int_{\mathbb{R}^d} h_n(x,y)d\mu(y) - \mu[x]\right|^r d\mu(x)dQ(\mu) \\
&= \|f\|_\infty^r \|\rho'\|_\infty^r \int_{\mathbf{T}\times\boldsymbol{X}} \sum_{i=1}^\infty a_i \left|\sum_{j\neq i} a_j h_n(x_i,x_j)\right|^r d\pi\otimes\nu^\infty(\mathbf{a},\mathbf{x}) \\
&\leq \|f\|_\infty^r \|\rho'\|_\infty^r \int_{\mathbf{T}\times\boldsymbol{X}} \sum_{i\neq j} a_i a_j \left|h_n(x_i,x_j)\right|^r d\pi\otimes\nu^\infty(\mathbf{a},\mathbf{x}) \\
&\leq \|f\|_\infty^r \|\rho'\|_\infty^r \int_{\mathbb{R}^d\times\mathbb{R}^d} |h_n(x,y)|^r d\nu^2(x,y) \to 0.
\end{aligned}
$$

Step 3: convergence of the Wasserstein gradients.

$$
\begin{aligned}
&\int\int |\nabla_W F_n(x,\mu) - \nabla_W \hat{F}(x,\mu)|^r d\mu(x)dQ(\mu) = \int\int \Big|\rho\big(L_{h_n(x,\cdot)}(\mu)\big)\nabla f(x) - \rho(\mu[x])\nabla f(x) \\
&\quad + \rho'\big(L_{h_n(x,\cdot)}(\mu)\big)f(x)\int_{\mathbb{R}^d}\nabla_x h_n(x,z)d\mu(z) + \int_{\mathbb{R}^d}\rho'\big(L_{h_n(z,\cdot)}(\mu)\big)f(z)\nabla_y h_n(z,x)d\mu(z)\Big|^r d\mu(x)dQ(\mu) \\
&\leq 2^{r-1}\|\nabla f\|_\infty^r \int\int \Big|\rho\big(L_{h_n(x,\cdot)}(\mu)\big) - \rho(\mu[x])\Big|^r d\mu(x)dQ(\mu) \\
&\quad + 2^{r-1}\|\rho'\|_\infty^r\|f\|_\infty^r \int\int\int |\nabla_x h_n(x,z)|^r + |\nabla_y h_n(z,x)|^r d\mu(z)d\mu(x)dQ(\mu) \\
&= I_{1,n} + I_{2,n}.
\end{aligned}
$$

The first term $I_{1,n}$ goes to 0, exactly as in the previous step. For the second term, thanks to Lemma 2.12, we have

$$
\begin{aligned}
I_{2,n} &\leq C(r,\rho',f)\int\int\int |\nabla_{x,y}h_n(x,y)|^r d\mu\otimes\mu(x,y)dQ(\mu) \\
&= C(r,\rho',f)\left(c_1\int_{\mathbb{R}^d\times\mathbb{R}^d}|\nabla_{x,y}h_n(x,y)|^r d\nu\otimes\nu(x,y) + c_2\int_{\mathbb{R}^d}|\nabla_{x,y}h_n(x,x)|^r d\nu(x)\right) \\
&= C(r,\rho',f)c_1\int_{\mathbb{R}^d\times\mathbb{R}^d}|\nabla_{x,y}h_n(x,y)|^r d\nu\otimes\nu(x,y) \to 0,
\end{aligned}
$$

where the last equality follows from the fact that the gradient on diagonal terms vanishes. □

Now that we know that cylinder functions can approximate generalized cylinder functions in the sense of (4.9), we show that, under additional assumptions, any function in $\mathrm{GC}_c^1(\mathcal{P}(\mathbb{R}^d))$ can be used to test the continuity equation.

Theorem 4.3. *Let $b : [0,T]\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ be a Borel vector field. Let $\Lambda \in \mathcal{P}(C([0,T],\mathcal{P}(\mathbb{R}^d)))$ be concentrated on $\mathrm{CE}(b)$ and define $\boldsymbol{M} = (M_t)_{t\in[0,T]} \in C_T(\mathcal{P}(\mathcal{P}(\mathbb{R}^d)))$ as $M_t := (\mathfrak{e}_t)_\sharp\Lambda$. Assume that*

(1) $M_t = u_t Q$, with $u \in L^1_t(L^q(Q))$, with $q \in [1,+\infty]$. In other words

$$
\|u\|_{L^1_t(L^q(Q))} := \int_0^T \|u_t\|_{L^q(Q)}dt < +\infty; \tag{4.10}
$$

(2) $\|b\|^p_{L^p(\widetilde{M}_t\otimes dt)} = \int_0^T\int_{\mathcal{P}(\mathbb{R}^d)}\int_{\mathbb{R}^d}|b_t(x,\mu)|^p d\mu(x)dM_t(\mu)dt < +\infty$, with $p \in [1,+\infty]$;

(3) $\mathrm{cap}_{r,\nu}(\Delta) = 0$, with $r \in [1,+\infty)$;

(4) $\frac{p'}{r} + \frac{1}{q} \leq 1$.

Then, for Λ*-a.e.* $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathbb{R}^d))$ *and for all* $\hat{F} \in \mathrm{GC}^1_c(\mathcal{P}(\mathbb{R}^d))$*, it holds*

$$\frac{d}{dt}\hat{F}(\mu_t) = \int_{\mathbb{R}^d} \nabla_W \hat{F}(x,\mu_t)\cdot b(t,x,\mu_t)d\mu_t(x), \tag{4.11}$$

in the sense of distributions in $(0,T)$*. Moreover, it holds* $\partial_t M_t + \mathrm{div}_{\mathcal{P}}(b_t M_t) = 0$ *in the sense of (2.16), and in duality with generalized cylinder functions, that is for any* $\hat{F} \in \mathrm{GC}^1_c(\mathcal{P}(\mathbb{R}^d))$ *it holds*

$$\frac{d}{dt}\int_{\mathcal{P}(\mathbb{R}^d)} \hat{F}(\mu)dM_t(\mu) = \int_{\mathcal{P}(\mathbb{R}^d)}\int_{\mathbb{R}^d} \nabla_W \hat{F}(x,\mu)\cdot b(t,x,\mu)d\mu(x)dM_t(\mu). \tag{4.12}$$

Proof. Let $\hat{F} \in \mathrm{GC}^1_c(\mathcal{P}(\mathbb{R}^d))$ and consider $(F_n)_{n\in\mathbb{N}} \subset \mathrm{Cyl}^1_c(\mathcal{P}(\mathbb{R}^d))$ such that $F_n \longrightarrow \hat{F}$ in the sense of (4.9). For all $\xi \in C^1_c(0,T)$, let us first prove

$$0 = \int \left| \int_0^T \xi'(t)\hat{F}(\mu_t) + \xi(t)\left(\int_{\mathbb{R}^d} \nabla_W \hat{F}(x,\mu_t)\cdot b_t(x,\mu_t)d\mu_t(x)\right)dt\right| d\Lambda(\boldsymbol{\mu}). \tag{4.13}$$

Since (4.11) holds with F_n in place of $\hat{F}$ for Λ-a.e. $\boldsymbol{\mu}$, we have

$$\begin{aligned}
&\int \left| \int_0^T \xi'(t)\hat{F}(\mu_t) + \xi(t)\left(\int_{\mathbb{R}^d} \nabla_W \hat{F}(x,\mu_t)\cdot b_t(x,\mu_t)d\mu_t(x)\right)dt\right| d\Lambda(\boldsymbol{\mu})\\
&= \int \Bigg| \int_0^T \xi'(t)\hat{F}(\mu_t) - \xi'(t)F_n(\mu_t)dt\\
&\quad + \int_0^T \xi(t)\left(\int_{\mathbb{R}^d} \nabla_W \hat{F}(x,\mu_t)\cdot b_t(x,\mu_t) - \nabla_W F_n \cdot b_t(x,\mu_t)d\mu_t(x)\right)dt\Bigg| d\Lambda(\boldsymbol{\mu})\\
&\leq \int \left| \int_0^T \xi'(t)\hat{F}(\mu_t)dt - \int_0^T \xi'(t)F_n(\mu_t)dt\right| d\Lambda\big(\boldsymbol{\mu}\big)\\
&\quad + \int \left| \int_0^T \int_{\mathbb{R}^d} \xi(t)\nabla_W \hat{F}(x,\mu_t) - \nabla_W F_n(x,\mu_t)\cdot b_t(x,\mu_t)d\mu_t(x)dt\right| d\Lambda\big(\boldsymbol{\mu}\big)\\
&\leq \|\xi'\|_\infty \int_0^T \int_{\mathcal{P}(\mathbb{R}^d)} |\hat{F}(\mu) - F_n(\mu)|dM_t(\mu)dt\\
&\quad + \|\xi\|_\infty \int_0^T \int_{\mathcal{P}(\mathbb{R}^d)}\int_{\mathbb{R}^d} \left|\Big(\nabla_W \hat{F}(x,\mu) - \nabla_W F_n(x,\mu)\Big)\cdot b_t(x,\mu)\right| d\mu(x)dM_t(\mu)dt\\
&= I_{1,n} + I_{2,n}.
\end{aligned}$$

Both terms go to 0, in fact:

$$\begin{aligned}
I_{1,n} =& \|\xi'\|_\infty \int_0^T \int |\hat{F}(\mu) - F_n(\mu)| u_t(\mu)dQ(\mu)dt\\
\leq& \|\xi'\|_\infty \int_0^T \left(\int |\hat{F} - F_n|^r dQ\right)^{\frac{1}{r}} \left(\int_{\mathcal{P}} u_t^q dQ\right)^{\frac{1}{q}} dt \to 0;
\end{aligned}$$

$$I_{2,n} \leq \|\xi\|_\infty \|b\|_{L^p(\tilde{M}_t\otimes dt)} \left(\int_0^T \int\int |\nabla_W \hat{F} - \nabla_W F_n|^{p'} d\mu(x)dM_t(\mu)dt\right)^{\frac{1}{p'}}$$

$$\le \|\xi\|_\infty \|b\|_{L^p} \left(\int_0^T \left\{ \int \left[\int |\nabla_W \hat{F} - \nabla_W F_n|^{p'} d\mu(x) \right]^{q'} dQ(\mu) \right\}^{\frac{1}{q'}} \left\{ \int |u_t|^q dQ \right\}^{\frac{1}{q}} dt \right)^{\frac{1}{p'}}$$

$$= \|\xi\|_\infty \|b\|_{L^p} \|u\|^{1/p'}_{L^1_t(L^q(Q))} \left\{ \int \left[\int |\nabla_W \hat{F} - \nabla_W F_n|^{p'} d\mu(x) \right]^{q'} dQ(\mu) \right\}^{\frac{1}{p'q'}}$$

$$\le \|\xi\|_\infty \|b\|_{L^p} \|u\|^{1/p'}_{L^1_t(L^q(Q))} \left(\int \int |\nabla_W \hat{F} - \nabla_W F_n|^{p'q'} d\mu(x) dQ(\mu) \right)^{\frac{1}{p'q'}}$$

$$\le \|\xi\|_\infty \|b\|_{L^p} \|u\|^{1/p'}_{L^1_t(L^q(Q))} \left(\int \int |\nabla_W \hat{F} - \nabla_W F_n|^{r} d\mu(x) dQ(\mu) \right)^{1/r} \to 0$$

Exploiting (4.13) for a countable dense subset of $\{\xi_n\} \subset C^1_c(0,T)$, we have that (4.11) holds for the chosen $\hat{F}$, for all $\boldsymbol{\mu} \in \mathcal{N}_{\hat{F}} \subset C([0,T], \mathcal{P}(\mathbb{R}^d))$, with $\Lambda(\mathcal{N}_{\hat{F}}) = 0$. We need to find a set $\mathcal{N}$ that does not depend on the chosen generalized cylinder functions. To this aim, consider a countable dense subset $\{\hat{\phi}_n\}$ in $C^1_c(\mathbb{R}^d \times (0,1])$ and consider $\mathcal{N} := \bigcap_{n\in\mathbb{N}} \mathcal{N}_{\hat{L}_{\hat{\phi}_n}}$. By density, for all $\boldsymbol{\mu} \in \mathcal{N}$, (4.11) holds for $\hat{L}_{\hat{\phi}}$ for all $\hat{\phi} \in C^1_c(\mathbb{R}^d \times (0,1])$. Then, the same holds for any $\hat{F} \in \mathrm{GC}^1_c(\mathcal{P}(\mathbb{R}^d))$ simply by chain rule property. Finally, (4.12) follows from the previous computations. □

4.2. **Nested superposition principles for laws of purely atomic random measures.** The main goal of this subsection is to prove the following *atomic nested superposition principles*, that are, respectively, of a differential and a metric nature. The first theorem coincide with Theorem 1.3, and we restate it here for completeness.

Theorem 4.4. *Let $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ be Borel measurable, $\Lambda \in \mathcal{P}(C([0,T], \mathcal{P}(\mathbb{R}^d)))$ and $\boldsymbol{M} = (M_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathcal{P}(\mathbb{R}^d)))$ be satisfying the assumptions of Theorem 4.3. Then, there exists $\mathfrak{L} \in \mathcal{P}(\mathcal{P}(C([0,T], \mathbb{R}^d)))$ such that $E_\sharp \mathfrak{L} = \Lambda$ and $\mathfrak{L} \in \mathcal{P}\big(\mathcal{P}^{\mathrm{pa}}(C([0,T], \mathbb{R}^d)) \cap \mathrm{SPS}(b)\big)$. In particular, $(E_t)_\sharp \mathfrak{L} = M_t$ for all $t \in [0,T]$ and Λ is concentrated over $\mathrm{CE}(b) \cap \mathcal{P}(C([0,T], \mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)))$.*

Notice that, given $(M_t)_{t\in[0,T]}$ that satisfies the hypothesis of Theorem 4.3, then any $\Lambda \in \mathcal{P}(C([0,T], \mathcal{P}(\mathbb{R}^d)))$ given by Theorem (2.10) satisfies the hypothesis of the previous theorem.

Theorem 4.5. *Let $\Lambda \in \mathcal{P}(C([0,T], \mathcal{P}(\mathbb{R}^d)))$ be concentrated over $AC^p([0,T], \mathcal{P}_p(\mathbb{R}^d))$ and define $\boldsymbol{M} = (M_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathcal{P}(\mathbb{R}^d)))$ as $M_t = (\mathfrak{e}_t)_\sharp \Lambda$. Assume:*

1. *$M_t = u_t Q$, with $u \in L^1_t(L^q(Q))$, with $q \in [1, +\infty]$;*
2. *for some $p > 1$, it holds*

$$\int \int_0^T |\dot{\boldsymbol{\mu}}|^p_{W_p}(t) dt d\Lambda(\boldsymbol{\mu}) < +\infty; \tag{4.14}$$

3. *$\mathrm{cap}_{r,\nu}(\Delta) = 0$, with $r \in [1, +\infty)$;*
4. *$\frac{p'}{r} + \frac{1}{q} \le 1$.*

Then, there exists $\mathfrak{L} \in \mathcal{P}(\mathcal{P}^{\mathrm{pa}}(AC^p([0,T], \mathbb{R}^d)))$ such that $E_\sharp \mathfrak{L} = \Lambda$ and

$$\int \int \int_0^T |\dot{\gamma}|^p(t) dt d\lambda(\gamma) d\mathfrak{L}(\lambda) < +\infty \tag{4.15}$$

In particular, $(E_t)_\sharp \mathfrak{L} = M_t$ for all $t \in [0,T]$ and Λ is concentrated over $\mathcal{P}(C([0,T], \mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)))$.

Both the theorems are a consequence of the following lemma, whose proof is postponed to the next subsection.

Lemma 4.6. *Let $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathbb{R}^d))$ be satisfying:*

(H1) there exists $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ Borel such that $\int_0^T \int_{\mathbb{R}^d} |v_t(x)| d\mu_t(x) dt < +\infty$ and $\partial_t \mu_t + \operatorname{div}(v_t\mu_t) = 0$ in the classic sense and in duality with $\mathrm{GC}^1_c(\mathcal{P}(\mathbb{R}^d))$, that is

$$\frac{d}{dt}\hat{F}(\mu_t) = \int_{\mathbb{R}^d} \nabla_W \hat{F}(x, \mu_t) \cdot v_t(x) d\mu_t(x) \quad \forall \hat{F} \in \mathrm{GC}^1_c(\mathcal{P}(\mathbb{R}^d)); \tag{4.16}$$

(H2) there exists $I_{\boldsymbol{\mu}} \subset [0,T]$ of full Lebesgue measure such that $\mu_t \in \mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)$ for all $t \in I_{\boldsymbol{\mu}}$.

Then, there exists $\lambda \in \mathcal{P}^{\mathrm{pa}}(C([0,T], \mathbb{R}^d))$ such that $(\mathrm{e}_t)_\sharp \lambda = \mu_t$ for all $t \in [0,T]$ and that is concentrated over $\gamma \in AC([0,T], \mathbb{R}^d)$ solving $\dot{\gamma}(t) = v_t(\gamma(t))$ for a.e. $t \in [0,T]$. In particular $I_{\boldsymbol{\mu}} = [0,T]$, that is $\mu_t \in \mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)$ for all $t \in [0,T]$,

Proof of Theorem 4.4. Thanks to Theorem 4.3, Λ-a.e. $\boldsymbol{\mu}$ satisfies the hypothesis *(H1)* in Lemma 4.6, with $v(t,x) = b(t,x,\mu_t)$. On the other hand, the following computation shows that Λ-a.e. $\boldsymbol{\mu}$ satisfies hypothesis *(H2)* as well:

$$\begin{aligned}\int\int_0^T \mathbb{1}_{\{(t,\boldsymbol{\mu})\,:\,\mu_t\notin\mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)\}} dt d\Lambda(\boldsymbol{\mu}) &= \int_0^T \int \mathbb{1}_{\{(t,\boldsymbol{\mu})\,:\,\mu_t\notin\mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)\}} d\Lambda(\boldsymbol{\mu}) dt \\ = \int_0^T \int \mathbb{1}_{\{(t,\mu)\,:\,\mu\notin\mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)\}} dM_t(\mu) dt &= \int_0^T M_t\big(\mathcal{P}(\mathbb{R}^d) \setminus \mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)\big) dt = 0.\end{aligned}$$

Then, there exists $\lambda \in \mathcal{P}^{\mathrm{pa}}(C([0,T], \mathbb{R}^d)) \cap \mathrm{SPS}(b)$ such that $E(\lambda) = \boldsymbol{\mu}$. In particular Λ is concentrated over the set $E\big(\mathcal{P}^{\mathrm{pa}}(C([0,T], \mathbb{R}^d)) \cap \mathrm{SPS}(b)\big)$, that is a Souslin set (and thus Λ-measurable) because $\mathcal{P}^{\mathrm{pa}}(C([0,T], \mathbb{R}^d)) \cap \mathrm{SPS}(b)$ is Borel measurable, see [PS25, Proposition 5.8], and E is continuous. Then, using the measurable selection theorem [Bog07, Theorem 6.9.1], we can find a Souslin-Borel measurable right inverse $G : E(\mathcal{P}^{\mathrm{pa}}(C([0,T], \mathbb{R}^d)) \cap \mathrm{SPS}(b)) \to \mathcal{P}^{\mathrm{pa}}(C([0,T], \mathbb{R}^d)) \cap \mathrm{SPS}(b)$, i.e. such that $E(G(\boldsymbol{\mu})) = \boldsymbol{\mu}$ for all $\boldsymbol{\mu} \in E\big(\mathcal{P}^{\mathrm{pa}}(C([0,T], \mathbb{R}^d)) \cap \mathrm{SPS}(b)\big)$. Then, since Souslin sets are universally measurable, the measure $\mathfrak{L} := G_\sharp \Lambda$ is well-defined for any Λ as in Claim (1), and it clearly satisfies Claim (2). $\square$

Proof of Theorem 4.5. Similarly as before, Λ-a.e. $\boldsymbol{\mu}$ satisfies the hypothesis *(H2)* in Lemma 4.6. On the other hand, from [AGS08, Theorem 8.3.1], for Λ-a.e. $\boldsymbol{\mu}$ there exists $v_{\boldsymbol{\mu}} : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ Borel measurable and such that

$$\partial_t \mu_t + \operatorname{div}(v_{\boldsymbol{\mu},t}\mu_t) = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} |v_{\boldsymbol{\mu}}(t,x)|^p d\mu_t(x) = |\dot{\boldsymbol{\mu}}|^p_{W_p}(t) \quad \text{for a.e. } t \in [0,T]. \tag{4.17}$$

We want to show that the continuity equation $\partial_t \mu_t + \operatorname{div}(v_{\boldsymbol{\mu},t}\mu_t) = 0$ is satisfied in duality with generalized cylinder functions as well. Let $\hat{F} \in \mathrm{GC}^1_c(\mathcal{P}(\mathbb{R}^d))$ and consider $F_n \in \mathrm{Cyl}^1_c(\mathcal{P}(\mathbb{R}^d))$ such that $F_n \to F$ as in (4.9). By chain rule, we know that for all $\xi \in C^1_c(0,T)$ we have

$$\int_0^T \xi'(t) F_n(\mu_t) dt = -\int_0^T \xi(t) \int_{\mathbb{R}^d} \nabla_W F_n(x,\mu_t) \cdot v_{\boldsymbol{\mu},t}(x) d\mu_t(x) dt = 0.$$

Then, we are done if we prove that

$$\int_0^T \xi'(t) F_n(\mu_t) dt \to \int_0^T \xi'(t) \hat{F}(\mu_t) dt \tag{4.18}$$

and

$$\int_0^T \xi(t) \int_{\mathbb{R}^d} \nabla_W F_n(x,\mu_t)\cdot v_{\boldsymbol{\mu},t}(x)d\mu_t(x)dt \to \int_0^T \xi(t)\int_{\mathbb{R}^d} \nabla_W \hat{F}(x,\mu_t)\cdot v_{\boldsymbol{\mu},t}(x)d\mu_t(x)dt \tag{4.19}$$

for Λ-a.e. $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]}$, since, as for the proof of Theorem 4.3, it suffices to check it for a countable amount of $\xi \in C_c^1(0,T)$ and $\hat{F} \in \mathrm{GC}_c^1(\mathcal{P}(\mathbb{R}^d))$. Regarding (4.18), notice that

$$\int \left| \int_0^T \xi'(t)F_n(\mu_t)dt - \int_0^T \xi'(t)\hat{F}(\mu_t)dt\right| d\Lambda(\boldsymbol{\mu}) \le \|\xi'\|_\infty \int_0^T \int |F_n(\mu)-\hat{F}(\mu)|dM_t(\mu)dt \to 0,$$

as $I_{1,n}$ in the proof of Theorem (4.3), then (4.18) holds for Λ-a.e. $\boldsymbol{\mu}$. Regarding (4.19), we have

$$\begin{aligned}
&\left|\int_0^T \xi(t)\int_{\mathbb{R}^d} \nabla_W F_n(x,\mu_t)\cdot v_{\boldsymbol{\mu},t}(x)d\mu_t(x)dt - \int_0^T \xi(t)\int_{\mathbb{R}^d}\nabla_W\hat{F}(x,\mu_t)\cdot v_{\boldsymbol{\mu},t}(x)d\mu_t(x)dt\right| \\
&\le \|\xi\|_\infty \int_0^T \int_{\mathbb{R}^d} |v_{\boldsymbol{\mu},t}(x)||\nabla_W F_n(x,\mu_t) - \nabla_W \hat{F}(x,\mu_t)|d\mu_t(x)dt \\
&\le \|\xi\|_\infty \left(\int_0^T\int_{\mathbb{R}^d} |v_{\boldsymbol{\mu},t}|^p d\mu_t(x)dt\right)^{\frac{1}{p}} \left(\int_0^T\int_{\mathbb{R}^d} |\nabla_W F_n(x,\mu_t)-\nabla_W\hat{F}(x,\mu_t)|^{p'}d\mu_t(x)dt\right)^{\frac{1}{p'}} \\
&= \|\xi\|_\infty \left(\int_0^T |\dot{\boldsymbol{\mu}}|^p_{W_p}(t)dt\right)^{\frac{1}{p}} \left(\int_0^T\int_{\mathbb{R}^d} |\nabla_W F_n(x,\mu_t)-\nabla_W\hat{F}(x,\mu_t)|^{p'}d\mu_t(x)dt\right)^{\frac{1}{p'}},
\end{aligned}$$

and this last formula can be integrated with respect to Λ since it is measurable in the variable $\boldsymbol{\mu}$ (see [PS25, Appendix D]), so that

$$\begin{aligned}
&\int \left(\int_0^T |\dot{\boldsymbol{\mu}}|^p_{W_p}(t)dt\right)^{\frac{1}{p}} \left(\int_0^T\int_{\mathbb{R}^d} |\nabla_W F_n(x,\mu_t)-\nabla_W\hat{F}(x,\mu_t)|^{p'}d\mu_t(x)dt\right)^{\frac{1}{p'}} d\Lambda(\boldsymbol{\mu}) \\
&\le \left(\int\int_0^T |\dot{\boldsymbol{\mu}}|^p_{W_p}(t)dtd\Lambda(\boldsymbol{\mu})\right)^{\frac{1}{p}} \left(\int\int_0^T\int_{\mathbb{R}^d} |\nabla_W F_n(x,\mu_t)-\nabla_W\hat{F}(x,\mu_t)|^{p'}d\mu_t(x)dtd\Lambda(\boldsymbol{\mu})\right)^{\frac{1}{p'}}
\end{aligned}$$

and it goes to 0 following the same computations done for $I_{2,n}$ in the proof of Theorem 4.3. This shows that Λ-a.e. $\boldsymbol{\mu}$ satisfies *(H1)* as well. Then, applying Lemma 4.6, for Λ-a.e. $\boldsymbol{\mu}$ there exists $\lambda \in \mathcal{P}^{\mathrm{pa}}(C([0,T],\mathbb{R}^d))$ concentrated over absolutely continuous curves and such that $\dot\gamma(t) = v_{\boldsymbol{\mu},t}(\gamma(t))$ for λ-a.e. γ and a.e. $t\in[0,T]$. Moreover, thanks to (4.17), λ is supported over $AC^p([0,T],\mathbb{R}^d)$ and

$$\int\int_0^T |\dot\gamma|^p(t)dtd\lambda(\gamma) = \int_0^T |\dot{\boldsymbol{\mu}}|^p_{W_p}(t)dt.$$

In particular, Λ is concentrated over the set $E\big(\mathcal{P}^{\mathrm{pa}}(AC^p([0,T],\mathbb{R}^d))\cap\mathfrak{A}_{\min}\big)$, where

$$\mathfrak{A}_{\min} = \left\{\lambda\in\mathcal{P}(AC^p([0,T],\mathbb{R}^d)) \ : \ \int\int_0^T|\dot\gamma|^p(t)dtd\lambda(\gamma) = \int_0^T |\dot{\boldsymbol{\mu}}|^p_{W_p}(t)dt, \text{ with } \dot{\boldsymbol{\mu}} = E(\lambda)\right\}.$$

Such a set is Souslin, being the image of the intersection of two Borel subset (see [PS25, Appendix D]) through the continuous function E. Then, using again the measurable selection theorem [Bog07, Theorem 6.9.1], there exists a Souslin-Borel measurable map $G : E\big(\mathcal{P}^{\mathrm{pa}}(AC^p([0,T],\mathbb{R}^d))\big) \to \mathcal{P}^{\mathrm{pa}}(AC^p([0,T],\mathbb{R}^d))\cap\mathfrak{A}_{\min}$ such that $E(G(\boldsymbol{\mu})) = \boldsymbol{\mu}$ for Λ-a.e.

$\boldsymbol{\mu} \in E\big(\mathcal{P}^{\mathrm{pa}}(AC^p([0,T],\mathbb{R}^d)) \cap \mathfrak{A}_{\min}\big)$. Then, the measure $\mathfrak{L} := G_\sharp \Lambda$ is well-defined (because Souslin set are universally measurable) and clearly satisfies the requirements. □

4.3. **Proof of Lemma 4.6.** From the assumption *(H2)*, for all $t \in I_{\boldsymbol{\mu}}$ there exists $\mathbf{a}(t) = (a_i(t))_{i\in\mathbb{N}} \in \mathbf{T}$ and $\mathbf{x}(t) = (x_i(t))_{i\in\mathbb{N}} \in \mathbf{X}$ such that

$$\mu_t = \sum_{i\in\mathbb{N}} a_i(t)\delta_{x_i(t)} \quad \text{for all } t \in I_{\boldsymbol{\mu}}. \tag{4.20}$$

We need to find a measure $\lambda \in \mathcal{P}(C([0,T],\mathbb{R}^d))$ such that $(\mathrm{e}_t)_\sharp\lambda = \mu_t$ for all $t \in [0,T]$ and that has the form

$$\lambda = \sum_{i\in\mathbb{N}} a_i\delta_{\gamma_i}, \tag{4.21}$$

for some $\mathbf{a} \in \mathbf{T}$ and $(\gamma_i)_{i\in\mathbb{N}} \subset AC([0,T],\mathbb{R}^d)$ that satisfy $\dot\gamma_i(t) = b_t(\gamma_i(t),\mu_t)$. It is clear that, a posteriori, the weights $a_i(t)$ will not depend on t. The next lemma gets us closer to this property.

Lemma 4.7. *Let $\boldsymbol{\mu} \in C([0,T],\mathcal{P}(\mathbb{R}^d))$ be satisfying (H1). Then, we can use as test functions*

$$\hat L_{\phi,a}(\mu) := \int_{\mathbb{R}^d} \phi(x)\mathbb{1}_{\{a\}}(\mu[x])d\mu(x), \quad \text{with} \quad \nabla_W \hat L_{\phi,a}(x,\mu) := \nabla\phi(x)\mathbb{1}_{\{a\}}(\mu[x]) \tag{4.22}$$

for all $\phi \in C^1_b(\mathbb{R}^d)$ and $a \in (0,1]$, that is, for all $\xi \in C^1_c(0,T)$

$$\int_0^T \xi'(t)\hat L_{\phi,a}(\mu_t)dt = -\int_0^T \xi(t)\int_{\mathbb{R}^d} \nabla_W \hat L_{\phi,a}(x,\mu_t)d\mu_t(x)dt. \tag{4.23}$$

Moreover, there exists $J_{\boldsymbol{\mu}} \subset I_{\boldsymbol{\mu}}$ of full Lebesgue measure such that for all $a \in (0,1]$ there exists $C_a \in [0,1]$ such that

$$\int_{\mathbb{R}^d} \mathbb{1}_{\{a\}}(\mu_t[x])d\mu_t(x) = C_a \quad \text{for all } t \in J_{\boldsymbol{\mu}}. \tag{4.24}$$

Proof. Let $\phi \in C^1_c(\mathbb{R}^d)$ and $a \in (0,1]$, the function $\mathbb{1}_{\{a\}}$ can be approximated pointwise by $\rho_n \in C^1_c((0,1])$. The assumptions allow us to use the generalized cylinder function $\hat L_{\phi,\rho_n}(\mu) := \int_{\mathbb{R}^d} \phi(x)\rho_n(\mu\{x\})d\mu(x)$ as a test function, to obtain that $(\mu_t)_{t\in[0,T]} \in C_T(\mathcal{P}(\mathbb{R}^d))$ satisfies

$$\int_0^T \xi'(t)\int_{\mathbb{R}^d} \phi(x)\rho_n(\mu_t[x])d\mu_t(x)dt = -\int_0^T \xi(t)\int_{\mathbb{R}^d} \rho_n(\mu_t[x])\nabla\phi(x)\cdot b(t,x,\mu_t)d\mu_t(x)dt,$$

for all $\xi \in C^1_c(0,T)$. Using the fact that $\rho_n \to \mathbb{1}_{\{a\}}$ pointwise and dominated convergence theorem, it is not hard to obtain

$$\int_0^T \xi'(t)\int_{\mathbb{R}^d} \phi(x)\mathbb{1}_{\{a\}}(\mu_t[x])d\mu_t(x)dt = -\int_0^T \xi(t)\int_{\mathbb{R}^d} \mathbb{1}_{\{a\}}(\mu_t[x])\nabla\phi(x)\cdot b(t,x,\mu_t)d\mu_t(x)dt.$$

The case in which $\phi \in C^1_b(\mathbb{R}^d)$ can be obtained through a standard cut-off argument, considering the approximating sequence $\phi_R(x) := \phi(x)f(x/R)$ as $R \to +\infty$, where $f \in C^1_c(\mathbb{R}^d)$ and $f \equiv 1$ in $B(0,1)$.

To prove (4.24), consider $\{\rho^{(n)}\}_{n\in\mathbb{N}} \subset C^1_c([0,1])$ a dense and countable subset of $C_0([0,1])$, that are continuous functions $\rho : [0,1] \to \mathbb{R}$ such that $\rho(0) = 0$. We know that, for all $n \in \mathbb{N}$,

$t \mapsto \int_{\mathbb{R}^d} \rho^{(n)}(\mu_t[x])d\mu_t(x)$ is weakly differentiable with null derivative, and so we can find a set $J_{\boldsymbol{\mu}} \subset I_{\boldsymbol{\mu}}$ where those maps are constant, say $C^{(n)}$. Now, for all $\rho \in C_0([0,1])$, the map

$$J_{\boldsymbol{\mu}} \ni t \mapsto \int_{\mathbb{R}^d} \rho(\mu_t[x]) \ \text{ is constant.}$$

Indeed, let $\rho^{(n_k)}$ be a sequence that converges uniformly to ρ, which implies that for all $t \in J_{\boldsymbol{\mu}}$

$$\lim_{k\to+\infty} C^{(n_k)} = \lim_{k\to+\infty} \int_{\mathbb{R}^d} \rho^{(n_k)}(\mu_t[x])d\mu_t(x) = \int_{\mathbb{R}^d} \rho(\mu_t[x])d\mu_t(x),$$

and in particular the right hand side does not depend on $t \in J_{\boldsymbol{\mu}}$, and we denote it by C_ρ. Similarly, for all $a \in (0,1]$, consider $\rho_n \in C_0([0,1])$ converging pointwise to $\mathbb{1}_{\{a\}}$ as above, and then, by dominated convergence theorem, we have that for all $t \in J_{\boldsymbol{\mu}}$

$$C_{\rho_n} = \int_{\mathbb{R}^d} \rho_n(\mu_t[x])d\mu_t(x) \to \int_{\mathbb{R}^d} \mathbb{1}_{\{a\}}(\mu_t[x])d\mu_t(x),$$

and in particular it does not depend on $t \in J_{\boldsymbol{\mu}}$. $\square$

Now, fix $\bar t \in J_{\boldsymbol{\mu}}$, and define

$$S_a := \{i \in \mathbb{N} \ : \ a_i(\bar t) = a\}, \quad a \in \mathcal{A} := \{a \in (0,1] \ : \ a = a_i(\bar t) \text{ for some } i \in \mathbb{N}\}.$$

Consider $N_a = \sharp S_a$ for all $a \in \mathcal{A}$. Thanks to Lemma 4.7 and $\bar t \in J_{\boldsymbol{\mu}}$, for all $a \in \mathcal{A}$

$$\int_{\mathbb{R}^d} \mathbb{1}_{\{a\}}(\mu_t[x])d\mu_t(x) = aN_a \quad \text{for all } t \in J_{\boldsymbol{\mu}}.$$

This argument leads to the equality $a_i(t) = a_i(\bar t)$ for all $t \in J_{\boldsymbol{\mu}}$, since for all $t \in J_{\boldsymbol{\mu}}$, the sequence $\{a_i(t)\}_{i\in\mathbb{N}}$ is non-increasing in i and there exist $\sharp S_a$ points in $\mathbb{R}^d$ to which μ_t is assigning mass equal to a. In particular, thanks to Lemma 2.3, the map

$$J_{\boldsymbol{\mu}} \ni t \mapsto \mu_t \text{ is continuous in the atomic topology.} \tag{4.25}$$

Thanks to [DGT04] and Lemma 2.3, we could also rearrange the positions $\{x_i(t)\}_{i\in\mathbb{N}}$ with permutations of S_a, to have that $J_{\boldsymbol{\mu}} \ni t \mapsto x_i(t)$ is continuous for all $i \in \mathbb{N}$. However, we do not need that to pursue our strategy and it will be a consequence of Lemma 4.8.

Now, let us exploit again Lemma 4.7 using $\hat L_{\phi,a}$ as test function, with general $\phi \in C^1_b(\mathbb{R}^d)$ and $a \in \mathcal{A}$. This gives the following: for all $\xi \in C^1_c(0,T)$, $\phi \in C^1_b(\mathbb{R}^d)$ and $a \in \mathcal{A}$

$$a\int_0^T \xi'(t) \sum_{i\in S_a} \phi(x_i(t))dt = -a\int_0^T \xi(t) \sum_{i\in S_a} \nabla\phi(x_i(t)) \cdot b_t(x_i(t),\mu_t)dt. \tag{4.26}$$

This can be rewritten more compactly as

$$\partial_t \nu_t^{(a)} + \operatorname{div}(v_t^{\boldsymbol{\mu}} \nu_t^{(a)}) = 0, \quad \nu_t^{(a)} := \frac{1}{N_a}\sum_{i\in S_a} \delta_{x_i(t)} \in \mathcal{P}^{N_a}(\mathbb{R}^d), \quad v_t^{\boldsymbol{\mu}}(x) := b_t(x,\mu_t). \tag{4.27}$$

The measure $\nu_t^{(a)}$ is well defined for all $t \in J_{\boldsymbol{\mu}}$ and there it is narrowly continuous because of (4.25) and Lemma 2.3. Thanks to [AGS08, Lemma 8.1.2], it exists a unique narrowly continuous representative $\tilde\nu_t^{(a)}$ on $[0,T]$, that coincides with $\nu_t^{(a)}$ in $J_{\boldsymbol{\mu}}$.

Lemma 4.8. *For all $t \in [0,T]$, $\tilde\nu_t^{(a)} \in \mathcal{P}^{N_a}(\mathbb{R}^d)$.*

Proof. Take $t \in [0,T] \setminus J_{\boldsymbol{\mu}}$ and consider $t_n \in J_{\boldsymbol{\mu}}$ such that $t_n \to t$ as $n \to +\infty$. By tightness of $\{\nu_{t_n}^{(a)}\}_{n\in\mathbb{N}}$, there exists a compact set $K \subset \mathbb{R}^d$ such that $x_i(t_n) \in K$ for all $i, n \in \mathbb{N}$. Then, there exists a subsequence n_k and such that for all $i \in \mathbb{N}$ $x_i(t_{n_k}) \to x_i$ as $k \to +\infty$, for some $x_i \in \mathbb{R}^d$. Then, it holds

$$\tilde{\nu}_t^{(a)} = \frac{1}{N_a} \sum_{i \in S_a} \delta_{x_i}. \qquad \square$$

At this point, notice that

$$[0,T] \ni t \mapsto \sum_{a \in \mathcal{A}} a N_a \tilde{\nu}_t^{(a)} \in \mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d) \text{ is narrowly continuous.}$$

Then, it must coincide with μ_t for all $t \in [0,T]$ by uniqueness of the continuous representative, since it is narrowly continuous and it coincides with μ_t for all $t \in J_{\boldsymbol{\mu}}$. In particular, $\mu_t \in \mathcal{P}^{\mathrm{pa}}(\mathbb{R}^d)$ for all $t \in [0,T]$.
Moreover, for all $a \in \mathcal{A}$, we can apply Theorem 3.1 to the curve $(\tilde{\nu}_t^{(a)})_{t\in[0,T]}$, to obtain a measure

$$\lambda^{(a)} = \frac{1}{N_a} \sum_{i \in S_a} \delta_{\gamma_i} \in \mathcal{P}^{N_a}(C([0,T], \mathbb{R}^d))$$

for some $\gamma_i \in AC([0,T], \mathbb{R}^d)$ satisfying $\dot{\gamma}_i(t) = v_t^{\boldsymbol{\mu}}(\gamma_i(t))$, and such that $(e_t)_\sharp \lambda^{(a)} = \tilde{\nu}_t^{(a)}$ for all $t \in [0,T]$. Then, the following lemma concludes the proof.

Lemma 4.9. *The measure*

$$\lambda := \sum_{a \in \mathcal{A}} a N_a \lambda^{(a)} = \sum_{i \in \mathbb{N}} a_i(\bar{t}) \delta_{\gamma_i} \in \mathcal{P}^{\mathrm{pa}}(C([0,T], \mathbb{R}^d)) \tag{4.28}$$

belongs to $\mathrm{SPS}(b)$ *and satisfies* $(\mathrm{e}_t)_\# \lambda = \mu_t$ *for all* $t \in [0,T]$. *In particular,* $\boldsymbol{\mu}$ *belongs to the image of* $\mathcal{P}^{\mathrm{pa}}(C([0,T], \mathbb{R}^d)) \cap \mathrm{SPS}(b)$ *through the function* E.

Proof. For all $t \in [0,T]$ and $g : \mathbb{R}^d \to [0,1]$ Borel measurable, it holds

$$\int f(\gamma(t)) d\lambda(\gamma) = \sum_{a \in \mathcal{A}} a N_a \int f(\gamma(t)) d\lambda^{(a)}(\gamma) = \sum_{a \in \mathcal{A}} \int f(x) d\tilde{\nu}_t^{(a)}(x) = \int f(x) d\mu_t(x),$$

so that

$$\int \int_0^T |b_t(\gamma(t), (\mathrm{e}_t)_\sharp \lambda)| dt d\lambda(\gamma) = \int_0^T \int |b_t(x, \mu_t)| d\mu_t(x) dt < +\infty.$$

Moreover, each γ_i is absolutely continuous and solves $\dot{\gamma}_i(t) = b_t(\gamma_i(t), \mu_t)$, showing that $\lambda \in \mathrm{SPS}(b)$. $\square$

Afterwards, we can assess that $I_{\boldsymbol{\mu}} = [0,T]$, but we cannot say that $J_{\boldsymbol{\mu}} = [0,T]$. Indeed, $J_{\boldsymbol{\mu}}$ is characterized by (4.25) and it may happen that for some $a \in \mathcal{A}$ and $i, j \in S_a$, the curves γ_i and γ_j cross at a time $t \in [0,T] \setminus J_{\boldsymbol{\mu}}$, raising a singularity for $(\mu_t)_{t\in[0,T]}$ with respect to the atomic topology.

5. On the capacity of the diagonal

We will analyze closer the hypothesis (3) in Theorem 4.3. In particular, we will see sufficient conditions on ν that implies it, and how it is connected to geometric properties of the measure. Moreover, we show what we can deduce about the evolution of random measures under stronger assumptions.

5.1. **Measure of an ε-strip around the diagonal.** In the following, for all ε we denote by $\Delta_\varepsilon \subset \mathbb{R}^d \times \mathbb{R}^d$ the ε-strip around the diagonal Δ, that is

$$\Delta_\varepsilon := \{(x,y) \in \mathbb{R}^d \times \mathbb{R}^d \ : \ |x-y| < \varepsilon\}. \tag{5.1}$$

Proposition 5.1. *Let $\nu \in \mathcal{P}(\mathbb{R}^d)$ be an atomless measure. Assume that, for some $\tilde{r} \in [1,+\infty)$, it holds*

$$\nu \otimes \nu(\Delta_\varepsilon) \leq C\varepsilon^{\tilde{r}}, \tag{5.2}$$

for a constant $C > 0$ independent of $\varepsilon > 0$. Then,

$$\operatorname{cap}_{r,\nu}(\Delta) = 0 \quad \textit{for all } r < \tilde{r}. \tag{5.3}$$

Proof. Let $\rho \in C^1_c(\mathbb{R}^{2d})$ be a standard mollifier, and $\rho_\varepsilon(\cdot) := \frac{1}{\varepsilon^{2d}}\rho(\cdot/\varepsilon)$ for all $\varepsilon > 0$. Then, define

$$h_\varepsilon(x,y) := \mathbb{1}_{\Delta_{2\varepsilon}} * \rho_{\varepsilon/2}(x).$$

It is immediate to verify that $0 \leq h_\varepsilon \leq 1$, $h_\varepsilon(x,y) = 1$ for all $(x,y) \in \Delta_\varepsilon$ and $h_\varepsilon(x,y) = 0$ for all $(x,y) \notin \Delta_{3\varepsilon}$, which implies that $\nabla h_\varepsilon(x,y) = 0$ if $(x,y) \in \Delta_\varepsilon \cup \Delta^c_{3\varepsilon}$. On the other hand, if $(x,y) \in \Delta_{3\varepsilon} \setminus \Delta_\varepsilon$, it holds

$$\begin{aligned}
|\nabla h_\varepsilon(x,y)| &= \left| \int_{\mathbb{R}^{2d}} \mathbb{1}_{\Delta_{2\varepsilon}}(x-z, y-w) \nabla \rho_{\varepsilon/2}(z,w) dz dw \right| \\
&\leq \frac{2^{2d+1}}{\varepsilon^{2d+1}} \int_{\{|(z,w)| \leq \varepsilon/2\}} \mathbb{1}_{\Delta_{2\varepsilon}}(x-z, y-w) \left| \nabla \rho \left(\frac{z}{\varepsilon}, \frac{w}{\varepsilon} \right) \right| dz dw \\
&\leq \frac{2^{2d+1} \|\nabla \rho\|_\infty}{\varepsilon^{2d+1}} \mathcal{L}^{2d}\big(\{(z,w) \ : \ |(z,w)| \leq \varepsilon/2\} \big) \leq C(d) \varepsilon^{-1}.
\end{aligned}$$

Then, for all $r < \tilde{r}$, it holds

$$\begin{aligned}
&\int_{\mathbb{R}^{2d}} |h_\varepsilon(x,y)|^r + |\nabla h_\varepsilon(x,y)|^r d\nu \otimes \nu(x,y) \leq \nu \otimes \nu(\Delta_{3\varepsilon}) + \int_{\Delta_{3\varepsilon}} |\nabla h_\varepsilon(x,y)|^r d\nu \otimes \nu(x,y) \\
&\leq C\varepsilon^{\tilde{r}} + \frac{C(d)^r}{\varepsilon^r} \nu \otimes \nu(\Delta_{3,\varepsilon}) \leq \left(C + 3^{\tilde{r}} C^{\tilde{r}} C(d)^r \right) \left(\varepsilon^r + \varepsilon^{\tilde{r}-r} \right) \to 0.
\end{aligned}$$

□

Some corollaries immediately follow from the previous Proposition.

Corollary 5.2. *Let $\nu \in \mathcal{P}(\mathbb{R}^d)$ be an atomless measure and $\tilde{r} \in [1,+\infty)$, be such that*

$$\sup_{x \in \mathbb{R}^d} \nu\big(B(x,\varepsilon)\big) \leq C\varepsilon^{\tilde{r}}, \tag{5.4}$$

for some $C > 0$. Then $\nu \otimes \nu(\Delta_\varepsilon) \leq C\varepsilon^{\tilde{r}}$, and in particular $\operatorname{cap}_{r,\nu}(\Delta) = 0$ for all $r < \tilde{r}$.

Corollary 5.3. *The thesis of Theorem 4.4 and Theorem 4.5 holds if replace assumptions (3) and (4), respectively, by*

- *(3') $\nu \otimes \nu(\Delta_\varepsilon) \leq C\varepsilon^{\tilde{r}}$ for all $\varepsilon > 0$;*
- *(4') $\frac{p'}{\tilde{r}} + \frac{1}{q} < 1$.*

Corollary 5.4. *Let $E \subset \mathbb{R}^d$ and $s > 0$ such that the Hausdorff measure $\mathcal{H}^s(E) \in (0,+\infty)$. Assume that $\nu \in \mathcal{P}(\mathbb{R}^d)$ is such that $\nu \ll \mathcal{H}^s|_E$, with a density $f \in L^\infty(\mathcal{H}^s|_E)$. Then (5.4) and (5.2) are satisfied with $\tilde{r} = s$.*

Proof. It is well known that $\mathcal{H}^s(E \cap B(x,\varepsilon)) \leq \varepsilon^s$, which immediately gives

$$\nu(B(x,\varepsilon)) \leq \|f\|_{L^\infty(\mathcal{H}^s|_E)} \varepsilon^s.$$

□

We can exploit Proposition 5.2 to find examples of ν and $r \geq 1$ that satisfy $\mathrm{cap}_{r,\nu} = 0$. In fact, we see how $\tilde{r}$ for which (5.2) is valid, is related to the dimension of the measure ν.

Corollary 5.5. *Let $\nu \in \mathcal{P}(\mathbb{R}^d)$ and assume that $\nu \ll \mathcal{L}^d$, with density $f \in L^p(\mathbb{R}^d, \mathcal{L}^d)$. Then:*

(i) if $p \geq 2$, then (5.2) holds with $\tilde{r} = d$;
(ii) if $p \in (1,2)$, then (5.2) holds with $\tilde{r} = \frac{2d}{p'}$.

Proof. Let $\tilde{f}(x) := f(-x)$ for all $x \in \mathbb{R}^d$. For all $\varepsilon > 0$ it holds

$$\begin{aligned}\nu \otimes \nu(\Delta_\varepsilon) &= \int_{\mathbb{R}^d} \nu(B(x,\varepsilon)) d\nu(x) = \int_{\mathbb{R}^d} f(x) \left(\int_{B(x,\varepsilon)} f(y) dy \right) dx \\ &= \int_{\mathbb{R}^d} f(x) \left(\int_{B(0,\varepsilon)} f(x-y) dy \right) dx = \int_{B(0,\varepsilon)} \left(\int_{\mathbb{R}^d} f(x-y) f(x) dx \right) dy \\ &= \int_{B(0,\varepsilon)} \tilde{f} * f(y) dy.\end{aligned}$$

Then, by Young inequality and the fact that $f, \tilde{f} \in L^1(\mathcal{L}^d) \cap L^p(\mathcal{L}^d)$, it holds:

(i) if $p \geq 2$, then $p' \in [1,p]$ and $\tilde{f} \in L^{p'}$, which gives $\tilde{f} * f \in L^\infty$ and $\nu \otimes \nu(\Delta_\varepsilon) \leq \omega_d \|\tilde{f} * f\|_{L^\infty} \varepsilon^d$;
(ii) if $p \in (1,2)$, then $\tilde{f} * f \in L^q(\mathcal{L}^d)$ with $q > 1$ satisfying $\frac{1}{q} + 1 = \frac{2}{p}$, that is $q = \frac{p}{2-p}$. Then, by Holder inequality, we have

$$\nu \otimes \nu(\Delta_\varepsilon) \leq \omega_d^{1/q'} \|\tilde{f} * f\|_{L^q} \varepsilon^{\frac{d}{q'}},$$

and we can verify that $\frac{d}{q'} = \frac{2d}{p'}$.

□

A different technique shows that, in case $\nu \ll \mathcal{L}^d$ has a bounded density, then $\mathrm{cap}_{d,\nu}(\Delta) = 0$. Before stating the result, let us introduce the 'logarithmic-type' test functions that will make the job.

Consider $\eta \in C^\infty(\mathbb{R})$ non-increasing and such that $\eta(x) = 1$ for all $x \leq 0$, $\eta(x) = 0$ for all $x \geq 1$ and $|\eta'(x)| \leq 2$ for all $x \in (0,1)$. Then, for all $0 < \varepsilon < R < +\infty$ define

$$f_{\varepsilon,R}(r) := \eta\left(\frac{\log r - \log \varepsilon}{\log R - \log \varepsilon} \right) \in C^\infty((0,+\infty)), \quad h_{\varepsilon,R}(x,y) := f_{\varepsilon,R}(|x-y|). \tag{5.5}$$

The functions $h_{\varepsilon,R}$ are well-defined and smooth in $\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta$. However, since $\eta \equiv 1$ on negative numbers, we have that $h_{\varepsilon,R} \equiv 1$ in $\Delta_\varepsilon \setminus \Delta$, and in particular it can be smoothly extended in the whole $\mathbb{R}^d \times \mathbb{R}^d$ simply putting $h_{\varepsilon,R}(x,x) = 1$ for all $x \in \mathbb{R}^d$.

Now, η takes values in $[0,1]$, and its properties give $f_{\varepsilon,R}(r) = 1$ for all $r \leq \varepsilon$ and $f_{\varepsilon,R}(r) = 0$ for all $r \geq R$. Then, for all $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$

$$0 \leq h_{\varepsilon,R}(x,y) \leq 1, \quad |\nabla h_{\varepsilon,R}(x,y)| = \eta'\left(\frac{\log \frac{r}{\varepsilon}}{\log \frac{R}{\varepsilon}} \right) \frac{|\nabla_{x,y}|x-y||}{|x-y| \log \frac{R}{\varepsilon}} \leq \frac{2}{|x-y| \log \frac{R}{\varepsilon}} \mathbb{1}_{(\varepsilon,R)}(|x-y|). \tag{5.6}$$

Proposition 5.6. *Let $\nu \in \mathcal{P}(\mathbb{R}^d)$ and assume that $\nu \ll \mathcal{L}^d$, with density $f \in L^\infty(\mathbb{R}^d, \mathcal{L}^d)$. Then, if $d \geq 2$, $\mathrm{cap}_{d,\nu}(\Delta) = 0$.*

Proof. For all $0 < \varepsilon < R < +\infty$, the function $h_{\varepsilon,R}$ is a competitor for the minimum problem defining $\mathrm{cap}_{d,\nu}(\Delta)$. We need to prove that

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |h_{\varepsilon,R}(x,y)|^d + |\nabla h_{\varepsilon,R}(x,y)|^d d\nu\otimes\nu(x,y) \to 0,$$

carefully choosing a way for letting ε and R go to 0. The first term goes to 0, as $h_{\varepsilon,R}(x,y) = 0$ outside Δ_R and is bounded by 1, so that

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |h_{\varepsilon,R}(x,y)|^d d\nu\otimes\nu(x,y) \le \nu\otimes\nu(\Delta_R) \to 0, \quad \text{as } R\to 0.$$

On the other hand, thanks to (5.6)

$$\begin{aligned}\int_{\mathbb{R}^d\times\mathbb{R}^d}|h_{\varepsilon,R}(x,y)|^d d\nu\otimes\nu(x,y) \le & \frac{2^d}{\left(\log\frac{R}{\varepsilon}\right)^d}\int_{\mathbb{R}^d}\int_{\{x\in\mathbb{R}^d:|x-y|\in(\varepsilon,R)\}}\frac{1}{|x-y|^d}f(x)dxd\nu(y)\\ \le & \frac{2^d\|f\|_\infty}{\left(\log\frac{R}{\varepsilon}\right)^d}\int_{\{z\in\mathbb{R}^d:|z|\in(\varepsilon,R)\}}\frac{1}{|z|^d}dz = \frac{d\omega_d 2^d\|f\|_\infty}{\left(\log\frac{R}{\varepsilon}\right)^d}\int_\varepsilon^R\frac{1}{r}dr = \frac{d\omega_d 2^d\|f\|_\infty}{\left(\log\frac{R}{\varepsilon}\right)^{d-1}},\end{aligned}$$

where ω_d is the volume of the unit ball in $\mathbb{R}^d$. Choosing $R = \sqrt{\varepsilon}$ and allowing $\varepsilon \to 0$, the proof is concluded. $\square$

5.2. **Absolutely continuous curves of random measures with respect to a new metric.** In this subsection, we will show that we can obtain further information on the evolution of random measures $(M_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathcal{P}(\mathbb{R}^d)))$ that satisfy a continuity equation, whenever $M_t \ll Q_{\pi,\nu}$ and we have control of small strips around the diagonal as in (5.2). This information is encoded by a natural distance over random measures, similar to (2.9).

Definition 5.7. *Let $\Psi : [0,+\infty) \to [0,1]$ be a continuous and non-increasing function with compact support, satisfying $\Psi(0) = 1$. Then, for all $M, N \in \mathcal{P}_p(\mathcal{P}_p(\mathbb{R}^d))$ we define*

$$\mathcal{D}_{p,\Psi}(M,N) := \mathcal{W}_p(M,N) + \sup_{\varepsilon\in(0,1)}\left|\int F_{\varepsilon,\Psi}(\mu)dM(\mu) - \int F_{\varepsilon,\Psi}(\mu)dN(\mu)\right| \tag{5.7}$$

where

$$F_{\varepsilon,\Psi}(\mu) := \int_{\mathbb{R}^{2d}}\Psi\left(\frac{|x-y|}{\varepsilon}\right)d\mu\otimes\mu(x,y). \tag{5.8}$$

Lemma 5.8. *The metric space $(\mathcal{P}_p(\mathcal{P}_p(\mathbb{R}^d)), \mathcal{D}_{p,\Psi})$ is complete and separable.*

Proof. It is immediate to see that $\mathcal{D}_{p,\Psi}$ is a distance. To show separability, we compare it with the distance $W_{p,D_{p,|\cdot|,\Psi}}$, i.e. the p-Wasserstein distance defined over the distance $D_{p,|\cdot|,\Psi}$, which induces a separable topology. In fact, let $M, N \in \mathcal{P}_p(\mathcal{P}_p(\mathbb{R}^d))$ and consider $\Pi \in \Gamma(M,N)$ be optimal for $W_{p,D_{p,|\cdot|,\Psi}}$, so that

$$\begin{aligned}\mathcal{D}^p_{p,\Psi}(M,N) \le & 2^p\left(\mathcal{W}^p_p(M,N) + \sup_{\varepsilon\in(0,1)}\left|\int F_{\varepsilon,\Psi}(\mu)dM(\mu) - \int F_{\varepsilon,\Psi}(\nu)dN(\nu)\right|^p\right)\\ \le & 2^p\left(\int W^p_p(\mu,\nu)d\Pi(\mu,\nu) + \sup_{\varepsilon\in(0,1)}\int|F_{\varepsilon,\Psi}(\mu)-F_{\varepsilon,\Psi}(\nu)|^p d\Pi(\mu,\nu)\right)\\ \le & 2^p\left(\int W^p_p(\mu,\nu)d\Pi(\mu,\nu) + \int\sup_{\varepsilon\in(0,1)}|F_{\varepsilon,\Psi}(\mu)-F_{\varepsilon,\Psi}(\nu)|^p d\Pi(\mu,\nu)\right)\end{aligned}$$

$$\leq 2^p \left(\int D^p_{p,|\cdot|,\Psi} d\Pi(\mu,\nu) \right) = 2^p W^p_{p,D_{p,|\cdot|,\Psi}}(M,N).$$

In particular, if $\{M_n\}_{n\in\mathbb{N}} \subset \mathcal{P}_p(\mathcal{P}_p(\mathbb{R}^d))$ is dense with respect to $W_{p,D_{p,|\cdot|,\Psi}}$, then it is dense for $\mathcal{D}_{p,\Psi}$ as well.

Regarding completeness, let $\{M_n\}_{n\in\mathbb{N}} \subset \mathcal{P}_p(\mathcal{P}_p(\mathbb{R}^d))$ be a Cauchy sequence with respect to $\mathcal{D}_{p,\Psi}$. Then, it is also a Cauchy sequence with respect to $\mathcal{W}_p$, so that there exists $M \in \mathcal{P}_p(\mathcal{P}_p(\mathbb{R}^d))$ for which $\mathcal{W}_p(M_n, M) \to 0$. To show that $\mathcal{D}_{p,\Psi}(M_n, M) \to 0$ we need to prove that

$$\alpha_n := \sup_{\varepsilon\in(0,1)} \left| \int F_{\varepsilon,\Psi}(\mu) dM_n(\mu) - \int F_{\varepsilon,\Psi}(\nu) dM(\nu) \right| \to 0.$$

First, notice that the function F_Ψ is narrowly continuous, so that for all $\varepsilon \in (0,1)$, $\int F_{\varepsilon,\Psi} dM_n \to \int F_{\varepsilon,\Psi} dM$. Then, for all $n \in \mathbb{N}$ consider $\varepsilon_n \in (0,1)$ such that

$$\alpha_n \leq \left| \int F_{\varepsilon_n,\Psi}(\mu) dM_n(\mu) - \int F_{\varepsilon_n,\Psi}(\nu) dM(\nu) \right| + \frac{1}{n}.$$

Then we conclude thanks to the following inequalities:

$$\begin{aligned}
\limsup_{n\to+\infty} \alpha_n \leq& \limsup_{n\to+\infty} \left| \int F_{\varepsilon_n,\Psi}(\mu) dM_N(\mu) - \int F_{\varepsilon_n,\Psi}(\nu) dN(\nu) \right| \\
=& \limsup_{n\to+\infty} \lim_{m\to+\infty} \left| \int F_{\varepsilon_n,\Psi}(\mu) dM_n(\mu) - \int F_{\varepsilon_n,\Psi}(\nu) dM_m(\nu) \right| \\
\leq& \limsup_{n\to+\infty} \limsup_{m\to+\infty} \mathcal{D}_{p,\Psi}(M_n, M_m) = 0.
\end{aligned}$$ □

Notice that the topology $\tau_{p,\Psi}$ induced by $\mathcal{D}_{p,\Psi}$ is between the narrow over narrow topology τ_{nn} and the narrow over atomic topology τ_{na}. We know $\tau_{nn} \subsetneq \tau_{na}$, but it is not clear whether $\tau_{p,\Psi}$ coincides with one of them. In the following, we will again fix a reference measure $Q = Q_{\pi,\nu} \in \mathcal{Q}$.

Theorem 5.9. *Let $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ be a Borel vector field and $\boldsymbol{M} = (M_t)_{t\in[0,T]} \in C_T(\mathcal{P}(\mathcal{P}(\mathbb{R}^d)))$ be satisfying $\frac{d}{dt} M_t + \mathrm{div}_{\mathcal{P}}(b_t M_t) = 0$ and $M_0 \in \mathcal{P}_p(\mathcal{P}_p(\mathbb{R}^d))$. Assume that*

(1) $M_t = u_t Q$, with $u \in L^1_t(L^q(Q))$, with $q \in [1,+\infty]$. In other words

$$\|u\|_{L^1_t(L^q(Q))} := \int_0^T \|u_t\|_{L^q(Q)} dt < +\infty; \tag{5.9}$$

(2) $\|b\|^p_{L^p(\tilde{M}_t\otimes dt)} = \int_0^T \int_{\mathcal{P}(\mathbb{R}^d)} \int_{\mathbb{R}^d} |b_t(x,\mu)|^p d\mu(x) dM_t(\mu) dt < +\infty$, with $p \in [1,+\infty]$;

(3) $\nu \otimes \nu(\Delta_\varepsilon) \leq C\varepsilon^{\tilde{r}}$ for some $\tilde{r} \geq 1$;

(4) $\frac{p'}{\tilde{r}} + \frac{1}{q} \leq 1$.

Further assuming that $\Psi \in C^1_c([0,\infty),[0,1])$ and $\Psi'(0) = 0$, we have that $\boldsymbol{M}$ is absolutely continuous with respect to $\mathcal{D}_{p,\Psi}$. Moreover, if the inequality in (4) is strict, then $\boldsymbol{M}$ is ℓ-absolutely continuous with respect to $\mathcal{D}_{p,\Psi}$, where $\ell = \frac{p(\tilde{r}-q')}{\tilde{r}}$.

Proof. Thanks to [PS25, Proposition 4.10], the curve $\boldsymbol{M}$ is p-absolutely continuous with respect to $\mathcal{W}_p$. Noticing that $\ell \in [1,p)$, we will be done if we show that there exists $g \in L^\ell(0,T)$ such that for all $0 \leq t_1 < t_2 \leq T$ it holds

$$\sup_{\varepsilon\in(0,1)} \left| \int F_\Psi(\mu) dM_{t_1}(\mu) - \int F_\Psi(\mu) dM_{t_2}(\mu) \right| \leq \int_{t_1}^{t_2} g(s) ds. \tag{5.10}$$

To this aim, notice that $F_{\varepsilon,\Psi}$ is an admissible test function for all $\varepsilon \in (0,1)$ because of Lemma 4.1, for which it holds

$$\frac{d}{dt}\int F_{\varepsilon,\Psi}(\mu)dM_t(\mu) = \int\int_{\mathbb{R}^d} \nabla_W F_{\varepsilon,\Psi}(x,\mu)\cdot b_t(x,\mu)d\mu(x)dM_t(\mu)$$

in the sense of distribution, where

$$\nabla_W F_{\varepsilon,\Psi}(x,\mu) = \frac{2}{\varepsilon}\int_{\mathbb{R}^d} \Psi'\left(\frac{|x-y|}{\varepsilon}\right)\frac{x-y}{|x-y|}d\mu(y),$$

that is well-defined since $\Psi'(0)=0$. Then, consider

$$\tilde{p} := \frac{r}{r-q'}, \quad \text{so that} \quad \frac{\tilde{p}'}{\tilde{r}} + \frac{1}{q} = 1.$$

Notice that $1 \le \tilde{p} \le p$ and that the equality holds if and only if (4) holds with equality. Moreover, let $R>0$ be such that $\operatorname{supp}\Psi \subset [0,R]$. Then, for all $\varepsilon \in (0,1)$, similar computations as in the proof of Theorem 4.3 give

$$\begin{aligned}
&\left|\int F_{\varepsilon,\Psi}(\mu)dM_{t_1}(\mu) - \int F_{\varepsilon,\Psi}(\mu)dM_{t_2}(\mu)\right| \le \int_{t_1}^{t_2}\int\int_{\mathbb{R}^d} |b_t(x,\mu)||\nabla_W F_{\varepsilon,\Psi}(x,\mu)|d\mu(x)dM_t(\mu)dt\\
&\le \int_{t_1}^{t_2} \|b_t(\cdot,\cdot)\|_{L^{\tilde{p}}(\widetilde{M}_t)}\left(\int\int |\nabla_W F(x,\mu)|^{\tilde{p}'}d\mu(x)u_t(\mu)dQ(\mu)\right)^{\frac{1}{\tilde{p}'}}dt\\
&\le \int_{t_1}^{t_2} \|b_t(\cdot,\cdot)\|_{L^{\tilde{p}}(\widetilde{M}_t)}\|u_t\|_{L^q(Q)}^{\frac{1}{\tilde{p}'}}\left(\int\int |\nabla_W F(x,\mu)|^{\tilde{p}'q'}d\mu(x)dQ(\mu)\right)^{\frac{1}{\tilde{p}'q'}}dt\\
&\le 2\int_{t_1}^{t_2} \|b_t(\cdot,\cdot)\|_{L^{\tilde{p}}(\widetilde{M}_t)}\|u_t\|_{L^q(Q)}^{\frac{1}{\tilde{p}'}}\left(\int\int_{\mathbb{R}^{2d}} \frac{1}{\varepsilon^{\tilde{p}'q'}}\left|\Psi\left(\frac{|x-y|}{\varepsilon}\right)\right|d\mu\otimes\mu(x,y)dQ(\mu)\right)^{\frac{1}{\tilde{p}'q'}}dt\\
&\le 2\int_{t_1}^{t_2} \|b_t(\cdot,\cdot)\|_{L^{\tilde{p}}(\widetilde{M}_t)}\|u_t\|_{L^q(Q)}^{\frac{1}{\tilde{p}'}}\left(\frac{1}{\varepsilon^{\tilde{p}'q'}}\nu\otimes\nu(\Delta_{R\varepsilon})\right)^{\frac{1}{\tilde{p}'q'}}dt\\
&\le 2CR^{\frac{\tilde{r}}{\tilde{p}'q'}}\|u\|_{L^1_t(L^q(Q))}\int_{t_1}^{t_2}\int\int |b_t(\cdot,\cdot)|^{\tilde{p}}d\mu(x)dM_t(\mu)dt,
\end{aligned}$$

where in the last inequality we used Holder inequality and the fact that $\nu\otimes\nu(\Delta_{R\varepsilon}) \le CR^{\tilde{r}}\varepsilon^{\tilde{r}}$. At this point, simply notice that $\ell = \frac{p}{\tilde{p}}$, so that

$$\int_0^T\left(\int\int |b_t(\cdot,\cdot)|^{\tilde{p}}d\mu(x)dM_t(\mu)\right)^{\ell}dt \le \int_0^T\int\int |b_t(\cdot,\cdot)|^p d\mu(x)dM_t(\mu)dt < +\infty.$$

□

6. Nested superposition principles on Riemannian manifolds

In this subsection, we establish a version of Theorems 2.10, 4.4 and 4.5 in a Riemannian setting. For the rest of the section, if not differently specified, $(\mathbb{S},\mathsf{g})$ will be a compact boundaryless and oriented Riemannian manifold of dimension $k\ge 2$, with induced distance and volume measure denoted as $d_{\mathbb{S}}$ and $\mathrm{vol}_{\mathbb{S}}$. Moreover, we will use the same construction of reference random measures showed in §2.3: fix $\pi \in \mathbf{T}$ and $\nu \in \mathcal{P}(\mathbb{S})$ an atomless probability measure, and denote $Q = Q_{\pi,\nu} \in \mathcal{P}(\mathcal{P}(\mathbb{S}))$.

The general idea is to use the Nash embedding theorem for transporting all the notions over $\mathbb{S}$ in $\mathbb{R}^d$ and then use the superposition principles in the Euclidean setting. To this end, here we collect the notation and the main properties of the Nash embedding theorem [Nas54]: there exists a smooth map (the Nash isometry)

$$\jmath : \mathbb{S} \to \mathbb{R}^d \tag{6.1}$$

satisfying the following properties.

(1) It is a local isometry, and in particular for all $y \in \mathbb{S}$ and $v, w \in T_x\mathbb{S}$,

$$\mathsf{g}_{\mathrm{Eucl}}\big(d_y\jmath(v), d_y\jmath(w)\big) = \mathsf{g}_y(v, w), \tag{6.2}$$

where $\mathsf{g}_{\mathrm{Eucl}}$ is the standard Euclidean Riemannian metric on $\mathbb{R}^d$ and $d_y\jmath : T_y\mathbb{S} \to T_{\jmath(x)}S \subset \mathbb{R}^d$ is the differential operator of the map $\jmath$ in the point $y \in \mathbb{S}$.

(2) Its image $S := \jmath(\mathbb{S}) \subset \mathbb{R}^d$ is compact. Moreover, it inherits the differential and Riemannian structure from $\mathbb{S}$ and $\jmath^{-1} : S \to \mathbb{S}$ is well-defined and smooth;

(3) The intrinsic distance of S as a subset of $\mathbb{R}^d$ is defined as

$$d_S(x_0, x_1) := \inf\{\mathrm{length}(\gamma) \ : \ \gamma \in C([0,1], S),\ \gamma(0) = x_0,\ \gamma(1) = x_1\}, \tag{6.3}$$

where the length is evaluated looking at γ as a curve in $\mathbb{R}^d$. The space (S, d_S) is Polish and isometric to $(\mathbb{S}, d_{\mathbb{S}})$ through $\jmath$. Moreover, there exists a constant $C \geq 1$ such that

$$|x - y| \leq d_S(x, y) \leq C|x - y| \quad \forall x, y \in S; \tag{6.4}$$

(4) The volume measure $\mathrm{vol}_{\mathbb{S}}$ is transported to $\mathrm{vol}_S = \mathcal{H}^k|_S$, that is

$$\jmath_\sharp \mathrm{vol}_{\mathbb{S}} = \mathcal{H}^k|_S. \tag{6.5}$$

In particular, if $\nu \ll \mathrm{vol}_{\mathbb{S}}$, then $\tilde{\nu} := \jmath_\sharp\nu$ is absolutely continuous w.r.t. $\mathcal{H}^k|_S$, with density given by $\frac{d\nu}{d\,\mathrm{vol}_{\mathbb{S}}} \circ \jmath^{-1}$;

(5) For all $p \in [1, +\infty]$, the spaces $(\mathcal{P}_p(\mathbb{S}), W_{p,d_{\mathbb{S}}})$ and $(\mathcal{P}_p(S), W_{p,d_S})$ are isometric through the push forward function $\boldsymbol{\jmath} := \jmath_\sharp$, and $\boldsymbol{\jmath}^{-1} = \jmath_\sharp^{-1}$. Similarly, for the spaces of random measures $((\mathcal{P}_p(\mathcal{P}_p(\mathbb{S})), \mathcal{W}_{p,d_{\mathbb{S}}}))$ and $(\mathcal{P}_p(\mathcal{P}_p(S)), \mathcal{W}_{p,d_S})$ through $\boldsymbol{\jmath}_\sharp = \jmath_{\sharp\sharp}$. Moreover $\mathcal{P}_p(S)$ and $\mathcal{P}_p(\mathcal{P}_p(S))$ are, respectively, embedded in $\mathcal{P}_p(\mathbb{R}^d)$ and $\mathcal{P}_p(\mathcal{P}_p(\mathbb{R}^d))$, and (6.4) implies that

$$\begin{aligned} W_p(\mu, \nu) \leq W_{p,d_S}(\mu, \nu) \leq C W_p(\mu, \nu) \quad &\forall \mu, \nu \in \mathcal{P}_p(S), \\ \mathcal{W}_p(M, N) \leq \mathcal{W}_{p,d_S}(M, N) \leq C\mathcal{W}_p(M, N) \quad &\forall M, N \in \mathcal{P}_p(\mathcal{P}_p(S)); \end{aligned} \tag{6.6}$$

(6) The reference measure $Q_{\pi,\nu}$ is transported as

$$\jmath_{\sharp\sharp} Q_{\pi,\nu} = Q_{\pi, \jmath_\sharp\nu}. \tag{6.7}$$

6.1. **Continuity equation and nested superposition principle for random measures on compact Riemannian manifolds.** Let $b : [0,T] \times \mathbb{S} \times \mathcal{P}(\mathbb{S}) \to T\mathbb{S}$ be a Borel vector field (with the natural topology over the tangent bundle $T\mathbb{S}$), i.e. such that $b(t, y, \nu) \in T_y\mathbb{S}$ for any $(t, y, \nu) \in [0,T] \times \mathbb{S} \times \mathcal{P}(\mathbb{S})$.

Definition 6.1. *Let $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathbb{S}))$. We say that the continuity equation $\partial_t\mu_t + \mathrm{div}_{\mathbb{S}}(b_t(\cdot, \mu_t)\mu_t) = 0$ if*

$$\int_0^T \int_{\mathbb{S}} \|b(t, y, \mu_t)\|_{\mathsf{g}_y} d\mu_t(y) dt < +\infty \tag{6.8}$$

and for all $\xi \in C_c^1(0,T)$ and $\phi \in C^1(\mathbb{S})$

$$\int_0^T \xi'(t) \int_{\mathbb{S}} \phi(y) d\mu_t(y) dt = -\int_0^T \xi(t) \int_{\mathbb{S}} d_y\phi(b(t, y, \mu_t)) d\mu_t(y) dt. \tag{6.9}$$

As for the Euclidean case, we need to introduce the cylinder functions.

Definition 6.2 (Cylinder functions and Wasserstein differential)**.** *A functional $F : \mathcal{P}(\mathbb{S}) \to \mathbb{R}$ is called a cylinder function if there exists $k \in \mathbb{N}$, $\Phi = (\phi_1, \dots, \phi_k) \in C^1(\mathbb{S}; \mathbb{R}^k)$ and $\Psi \in C^1(\mathbb{R}^k)$ such that*

$$F(\mu) = \Psi\left(L_\Phi(\mu)\right), \quad L_\Phi(\mu) = \big(L_{\phi_1}(\mu), \dots, L_{\phi_k}(\mu)\big), \quad L_{\phi_i}(\mu) := \int_{\mathbb{S}} \phi_i(y) d\mu(y). \tag{6.10}$$

Its Wasserstein differential is then defined as

$$D_W F(y, \mu) := \sum_{i=1}^k \partial_i \Psi\left(L_\Phi(\mu)\right) d_y \phi_i \quad \forall y \in \mathbb{S},\ \forall \mu \in \mathcal{P}(\mathbb{S}). \tag{6.11}$$

The collection of all the cylinder functions its called $\mathrm{Cyl}^1(\mathcal{P}(\mathbb{S}))$*. If $\Psi \in C^\infty(\mathbb{R}^k)$ and $\phi_1, \dots, \phi_k \in C^\infty(\mathbb{S})$, we write that $F \in \mathrm{Cyl}^\infty(\mathcal{P}(\mathbb{S}))$.*

Similarly, we may define the sets of generalized cylinder functions $\mathrm{GC}^1_c(\mathcal{P}(\mathbb{S}))$ and $\mathrm{GC}^\infty_c(\mathcal{P}(\mathbb{S}))$: we say that $\hat{F} \in \mathrm{GC}^1_c(\mathcal{P}(\mathbb{S}))$ (resp. $\mathrm{GC}^\infty_c(\mathcal{P}(\mathbb{S}))$) is there exist $k \geq 1$, $\Psi \in C^1(\mathbb{R}^k)$ (resp. $C^\infty(\mathbb{R}^k)$) and $\hat{\Phi} = (\hat{\phi}_1, \dots, \hat{\phi}_k) \in C^1_c(\mathbb{S} \times (0,1]; \mathbb{R}^k)$ (resp. $C^\infty_c(\mathbb{S} \times (0,1]; \mathbb{R}^k)$) such that

$$\hat{F}(\mu) = \Psi\left(\hat{L}_{\hat{\Phi}}(\mu)\right), \quad \hat{L}_{\hat{\Phi}}(\mu) = \big(\hat{L}_{\hat{\phi}_1}(\mu), \dots, \hat{L}_{\hat{\phi}_k}(\mu)\big), \quad \hat{L}_{\hat{\phi}_i}(\mu) := \int_{\mathbb{S}} \hat{\phi}_i(y, \mu[y]) d\mu(y). \tag{6.12}$$

Their Wasserstein differential is then defined as

$$D_W \hat{F}(y, \mu) := \sum_{i=1}^k \partial_i \Psi\left(\hat{L}_{\hat{\Phi}}(\mu)\right) d_y \hat{\phi}_i(\cdot, \mu[y]) \quad \forall y \in \mathbb{S},\ \forall \mu \in \mathcal{P}(\mathbb{S}). \tag{6.13}$$

Definition 6.3. *Let $\boldsymbol{M} = (M_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathcal{P}(\mathbb{S})))$. We say that the continuity equation $\partial_t M_t + \mathrm{div}_{\mathcal{P}(\mathbb{S})}(b_t M_t) = 0$ holds, if*

$$\int_0^T \int_{\mathcal{P}(\mathbb{S})} \int_{\mathbb{S}} \|b(t, y, \mu)\|_{\mathsf{g}_y} d\mu(y) dM_t(\mu) dt < +\infty \tag{6.14}$$

and for all $\xi \in C^1_c(0,T)$ and $F \in \mathrm{Cyl}^1(\mathcal{P}(\mathbb{S}))$

$$\begin{aligned} \int_0^T \xi'(t) \int_{\mathcal{P}(\mathbb{S})} F(\mu) dM_t(\mu) dt &= -\int_0^T \xi(t) \int_{\mathcal{P}(\mathbb{S})} \int_{\mathbb{S}} D_W F(y, \mu)\big(b(t, y, \mu)\big) d\mu(y) dM_t(\mu) dt \\ &= -\int_0^T \xi(t) \int_{\mathcal{P}(\mathbb{S})} \sum_{i=1}^k \Psi(L_\Phi(\mu)) \int_{\mathbb{S}} d_y \phi_i\big(b(t, y, \mu)\big) d\mu(y) dM_t(\mu) dt. \end{aligned} \tag{6.15}$$

Given a Borel vector field $b : [0,T] \times \mathbb{S} \times \mathcal{P}(\mathbb{S}) \to T\mathbb{S}$, define the Borel vector field $\tilde{b} : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ as

$$\tilde{b}(t, x, \tilde{\mu}) = \begin{cases} d_{\jmath^{-1}(x)} \jmath\big(b(t, \jmath^{-1}(x), \boldsymbol{\jmath}^{-1}(\tilde{\mu}))\big) & \text{if } (x, \tilde{\mu}) \in S \times \mathcal{P}(S) \\ 0 & \text{otherwise.} \end{cases} \tag{6.16}$$

Moreover, given a curve of random measures on $\mathbb{S}$, $\boldsymbol{M} = (M_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathcal{P}(\mathbb{S})))$, define

$$\tilde{M}_t := \boldsymbol{\jmath}^{-1}_\sharp M_t = \jmath^{-1}_{\sharp\sharp} M_t \in \mathcal{P}(\mathcal{P}(S)). \tag{6.17}$$

Proposition 6.4. *Let $b : [0,T] \times \mathbb{S} \times \mathcal{P}(\mathbb{S}) \to T\mathbb{S}$ be a Borel vecotr field and $\boldsymbol{M} = (M_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathcal{P}(\mathbb{S})))$ be a solution of the continuity equation $\partial_t M_t + \mathrm{div}_{\mathcal{P}(\mathbb{S})}(b_t M_t) = 0$ in the sense of Def. 6.3. Then, the curve $(\tilde{M}_t)_{t\in[0,T]} \subset C([0,T], \mathcal{P}(\mathcal{P}(\mathbb{R}^d)))$ solves the continuity equation $\partial_t \tilde{M}_t + \mathrm{div}_{\mathcal{P}}(\tilde{b}_t \tilde{M}_t) = 0$ in the sense of (2.16).*

Proof. The local isometry property of $\jmath$ implies that, for all $\mu \in \mathcal{P}(\mathbb{S})$, for μ-a.e. y and for a.e. $t \in [0,T]$, it holds $\|b_t(y,\mu)\|_{\mathsf{g}_y} = |\tilde{b}_t(\jmath(y), \jmath_\sharp\mu)|$, which gives

$$\int_0^T \int_{\mathcal{P}(\mathbb{R}^d)} \int_{\mathbb{R}^d} |\tilde{b}(t,x,\tilde{\mu})| d\tilde{\mu}(x) d\tilde{M}_t(\tilde{\mu}) dt = \int_0^T \int_{\mathcal{P}(\mathbb{S})} \int_{\mathbb{S}} \|b(t,y,\mu)\| d\mu(y) dM_t(\mu) dt < +\infty.$$

Let $\xi \in C_c^1(0,T)$ and $\tilde{F} = \Psi \circ L_{\tilde{\Phi}} \in \mathrm{Cyl}_c^1(\mathcal{P}(\mathbb{R}^d))$, with $\tilde{\Phi} = (\tilde{\phi}_1, \dots, \tilde{\phi}_k) \in C_c^1(\mathbb{R}^d, \mathbb{R}^k)$ and $\Psi \in C_c^1(\mathbb{R}^k)$, then, exploiting the properties of $\jmath$ and $\boldsymbol{\jmath}$, we have

$$\begin{aligned}
\int_0^T \psi'(t) \int_{\mathcal{P}(\mathbb{R}^d)} \tilde{F}(\tilde{\mu}) d\tilde{M}_t(\tilde{\mu}) dt &= \int_0^T \xi'(t) \int_{\mathcal{P}(S)} \tilde{F}(\tilde{\mu}) d\tilde{M}_t(\tilde{\mu}) dt = \int_0^T \xi'(t) \int_{\mathcal{P}(\mathbb{S})} \tilde{F}(\jmath_\sharp \mu) dM_t(\mu) dt \\
&= \int_0^T \xi'(t) \int_{\mathcal{P}(\mathbb{S})} \Psi\left(\int_{\mathbb{S}} \tilde{\phi}_1 \circ \jmath d\mu, \dots, \int_{\mathbb{S}} \tilde{\phi}_k \circ \jmath d\mu \right) dM_t(\mu) dt \\
&= -\int_0^T \xi(t) \int_{\mathcal{P}(\mathbb{S})} \sum_{i=1}^k \left(\partial_i \Psi(L_{\tilde{\Phi}\circ\jmath}(\mu)) \int_{\mathbb{S}} d_y(\tilde{\phi}_i \circ \jmath)\big(b(t,y,\mu)\big) d\mu(y) \right) dM_t(\mu) dt \\
&= -\int_0^T \xi(t) \int_{\mathcal{P}(\mathbb{S})} \sum_{i=1}^k \left(\partial_i \Psi(L_{\tilde{\Phi}\circ\jmath}(\mu)) \int_{\mathbb{S}} d_{\jmath(y)}\tilde{\phi}_i \circ d_y\jmath\big(b(t,y,\mu)\big) d\mu(y) \right) dM_t(\mu) dt \\
&= -\int_0^T \xi(t) \int_{\mathcal{P}(S)} \sum_{i=1}^k \left(\partial_i \Psi(L_{\tilde{\Phi}}(\tilde{\mu})) \int_{\mathbb{S}} d_{\jmath(y)}\tilde{\phi}_i \circ d_y\jmath\big(b(t,y,\jmath_\sharp^{-1}\tilde{\mu})\big) d\jmath_\sharp^{-1}\tilde{\mu}(y) \right) d\tilde{M}_t(\mu) dt \\
&= -\int_0^T \xi(t) \int_{\mathcal{P}(S)} \sum_{i=1}^k \left(\partial_i \Psi(L_{\tilde{\Phi}}(\tilde{\mu})) \int_S d_x\tilde{\phi}_i \circ d_{\jmath^{-1}(x)}\jmath\big(b(t,\jmath^{-1}(x),\jmath_\sharp^{-1}\tilde{\mu})\big) d\mu(x) \right) d\tilde{M}_t(\tilde{\mu}) dt \\
&= -\int_0^T \xi(t) \int_{\mathcal{P}(S)} \sum_{i=1}^k \left(\partial_i \Psi(L_{\tilde{\Phi}}(\tilde{\mu})) \int_S d_x\tilde{\phi}_i\big(\tilde{b}(t,x,\tilde{\mu})\big) d\tilde{\mu}(x) \right) d\tilde{M}_t(\tilde{\mu}) dt \\
&= -\int_0^T \xi(t) \int_{\mathcal{P}(S)} \sum_{i=1}^k \left(\partial_i \Psi(L_{\tilde{\Phi}}(\tilde{\mu})) \int_S \nabla\tilde{\phi}_i \cdot \tilde{b}(t,x,\tilde{\mu}) \; d\tilde{\mu}(x) \right) d\tilde{M}_t(\tilde{\mu}) dt \\
&= -\int_0^T \xi(t) \int_{\mathcal{P}(S)} \int_S \nabla_W \tilde{F}(x,\tilde{\mu}) \cdot \tilde{b}(t,x,\tilde{\mu}) d\tilde{\mu}(x) d\tilde{M}_t(\tilde{\mu}) dt.
\end{aligned}$$

□

Before proceeding, let us give some preliminary results: define the maps,

$$J : C([0,T], \mathcal{P}(\mathbb{S})) \to C([0,T], \mathcal{P}(S)), \quad (\mu_t)_{t\in[0,T]} \mapsto (\jmath_\sharp \mu_t)_{t\in[0,T]},$$

$$\ell : C([0,T], \mathbb{S}) \to C([0,T], S), \quad \ell(\boldsymbol{y}) = (\jmath(y_t))_{t\in[0,T]}, \tag{6.18}$$

$$\mathfrak{J} : \mathcal{P}(C([0,T], \mathbb{S})) \to \mathcal{P}(C([0,T], S)), \quad \mathfrak{J}\lambda = \ell_\sharp \lambda.$$

and their inverse

$$J^{-1} : C([0,T], \mathcal{P}(S)) \to C([0,T], \mathcal{P}(\mathbb{S})), \quad (\tilde{\mu}_t)_{t\in[0,T]} \mapsto (\jmath_\sharp^{-1} \tilde{\mu}_t)_{t\in[0,T]},$$

$$\ell^{-1} : C([0,T],S) \to C([0,T],\mathbb{S}), \quad \ell^{-1}(\boldsymbol{x}) = (\jmath^{-1}(x_t))_{t\in[0,T]}, \tag{6.19}$$
$$\mathfrak{J}^{-1} : \mathcal{P}(C([0,T],S)) \to \mathcal{P}(C([0,T],\mathbb{S})), \quad \mathfrak{J}^{-1}\tilde{\lambda} = \ell^{-1}_{\sharp}\tilde{\lambda}.$$

Proposition 6.5. *The following hold:*

(1) $\boldsymbol{x} = (x_t)_{t\in[0,T]} \in AC^p([0,T],S)$ *if and only if* $\boldsymbol{y} := \ell^{-1}(\boldsymbol{x}) \in AC^p([0,T],\mathbb{S})$*. Moreover*

$$|\dot{x}|_{d_S}(t) = |\dot{y}|_{d_{\mathbb{S}}}(t) \quad \textit{for a.e. } t \in [0,T]; \tag{6.20}$$

(2) $\tilde{\boldsymbol{\mu}} = (\tilde{\mu}_t)_{t\in[0,T]} \in AC^p([0,T],\mathcal{P}_p(S))$ *if and only if* $\boldsymbol{\mu} := J^{-1}(\tilde{\boldsymbol{\mu}}) \in AC^p([0,T],\mathcal{P}_p(\mathbb{S}))$*. Moreover*

$$|\dot{\tilde{\mu}}|_{W_{p,d_S}}(t) = |\dot{\mu}|_{W_{p,d_{\mathbb{S}}}}(t) \quad \textit{for a.e. } t \in [0,T]; \tag{6.21}$$

(3) $\tilde{\lambda} \in \mathcal{P}(C([0,T],S))$ *is concentrated over* $\boldsymbol{x} = (x_t)_{t\in[0,T]} \in AC([0,T],S)$ *that solve*

$$\dot{x}(t) = \tilde{b}(t, x_t, (e_t)_\sharp\tilde{\lambda}), \tag{6.22}$$

if and only if $\lambda := \mathfrak{J}^{-1}\tilde{\lambda} \in \mathcal{P}(C([0,T],\mathbb{S}))$ *is concentrated over* $\boldsymbol{y} \in AC([0,T],\mathbb{S})$ *that solve*

$$\dot{y}(t) = b(t, y_t, (e_t)_\sharp\lambda); \tag{6.23}$$

(4) $\tilde{\boldsymbol{\mu}} = (\tilde{\mu}_t)_{t\in[0,T]} \in C([0,T],\mathcal{P}(S)) \subset C([0,T],\mathcal{P}(\mathbb{R}^d))$ *is a solution of*

$$\partial_t\tilde{\mu}_t + \mathrm{div}(\tilde{b}_t(\cdot,\tilde{\mu}_t)\mu_t) = 0 \tag{6.24}$$

if and only if $\boldsymbol{\mu} := J^{-1}(\tilde{\boldsymbol{\mu}})$ *is a solution of* $\partial_t\mu_t + \mathrm{div}_{\mathbb{S}}(b_t(\cdot,\mu_t)\mu_t) = 0$*.*

Proof. (1) and *(2)* follows directly from the metric properties of the Nash isometry, i.e. $\jmath$ is an isometry between the metric spaces $(\mathbb{S}, d_{\mathbb{S}})$ and (S, d_s), and $\jmath_\sharp$ is an isometry between $(\mathcal{P}(\mathbb{S}), W_{p,d_{\mathbb{S}}})$ and $(\mathcal{P}(S), W_{p,d_S})$. Regarding *(3)*, since $\mathfrak{J}^{-1} = \ell^{-1}_\sharp$, for the direct implication it suffices to show that whenever $\boldsymbol{x} \in AC([0,T],S)$ solves (6.22), then $\boldsymbol{y} := \ell^{-1}(\boldsymbol{x})$ solves (6.23). This follows from noticing that for all $f \in C^1(\mathbb{S})$

$$\begin{aligned}\frac{d}{dt}f(y_t) =& \frac{d}{dt}f\circ\jmath^{-1}(x_t) = d_{\jmath^{-1}(x_t)}f \circ d_{x_t}\jmath^{-1}\big(\tilde{b}(t,x_t,(e_t)_\sharp\tilde{\lambda})\big)\\ =& d_{\jmath^{-1}(x_t)}f \circ d_{x_t}\jmath^{-1} \circ d_{\jmath^{-1}(x_t)}\jmath\big(b(t,\jmath^{-1}(x_t),(\jmath^{-1}\circ e_t)_\sharp\tilde{\lambda})\big)\\ =& d_{y_t}f\big(b(t,y_t,(e_t\circ\ell^{-1})_\sharp\tilde{\lambda})\big) = d_{y_t}f\big(b(t,y_t,(e_t)_\sharp\lambda)\big),\end{aligned}$$

where we used the trivial equality $e_t \circ \ell^{-1} = \jmath^{-1} \circ e_t$. We conclude thanks to the fact that for a curve $\boldsymbol{y} : [0,T] \to \mathbb{S}$, being absolutely continuous is equivalent to ask that $f \circ \boldsymbol{y} \in AC(0,T)$ for all $f \in C^1(\mathbb{S})$. Moreover, its derivative is characterized as follows: for a.e. $t \in (0,T)$

$$\dot{\boldsymbol{y}} = v_t \in T_{y_t}\mathbb{S} \quad \text{for a.e. } t \iff \frac{d}{dt}f\circ y(t) = d_{y_t}f\big(v_t\big) \quad \text{for a.e. } t.$$

The converse implication in *(3)* follows from a similar computation. Regarding *(4)*, the local isometry property of $\jmath$ implies

$$\int_0^T\int_{\mathbb{S}} \|b(t,y,\mu_t)\|_{T_y\mathbb{S}}d\mu_t(y)dt = \int_0^T\int_S |\tilde{b}(t,x,\tilde{\mu}_t)|d\tilde{\mu}_t(x)dt,$$

so that the integrability assumptions for both the continuity equations are satisfied. Now, let $\xi \in C^1_c(0,T)$ and $\phi \in C^1(\mathbb{S})$, then

$$\int_0^T \xi'(t)\int_{\mathbb{S}}\phi(y)d\mu_t(y)dt + \int_0^T\xi(t)\int_{\mathbb{S}}d_y\phi\big(b(t,y,\mu_t)\big)d\mu_t(y)dt$$

$$
\begin{aligned}
&= \int_0^T \xi'(t) \int_{\mathbb{S}} \phi(y) d\jmath_\sharp^{-1} \tilde{\mu}_t(y) dt + \int_0^T \xi(t) \int_{\mathbb{S}} d_y \phi\big(b(t, y, \jmath_\sharp^{-1} \tilde{\mu}_t)\big) d\jmath_\sharp^{-1} \tilde{\mu}_t(y) dt \\
&= \int_0^T \xi'(t) \int_S \phi(\jmath^{-1}(x)) d\tilde{\mu}_t(x) dt + \int_0^T \xi(t) \int_S d_{\jmath^{-1}(x)} \phi\big(b(t, \jmath^{-1}(x), \jmath_\sharp^{-1} \tilde{\mu}_t)\big) d\tilde{\mu}_t(x) dt \\
&= \int_0^T \xi'(t) \int_S \phi(\jmath^{-1}(x)) d\tilde{\mu}_t(x) dt + \int_0^T \xi(t) \int_S d_{\jmath^{-1}(x)} \phi\Big(d_x \jmath^{-1} \circ d_{\jmath^{-1}(x)} \jmath\big(b(t, \jmath^{-1}(x), \jmath_\sharp^{-1} \tilde{\mu}_t)\big)\Big) d\tilde{\mu}_t(x) dt \\
&= \int_0^T \xi'(t) \int_S \phi(\jmath^{-1}(x)) d\tilde{\mu}_t(x) dt + \int_0^T \xi(t) \int_S d_x (\phi \circ \jmath^{-1}) \big(\tilde{b}(t, x, \tilde{\mu}_t)\big) d\tilde{\mu}_t(x) dt.
\end{aligned}
$$

Now, for the direct implication, since $\phi \circ \jmath^{-1} \in C^1(S)$, the last row is equal to 0 whatever its extension to the whole $\mathbb{R}^d$ is. For the converse implication, choosing $\phi = \tilde{\phi} \circ \jmath$ with $\tilde{\phi} \in C_c^1(\mathbb{R}^d)$, again the previous computation allows to conclude because $\phi \in C^1(\mathbb{S})$. □

Theorem 6.6. *Let $\boldsymbol{M} = (M_t)_{t \in [0,T]} \in C([0,T], \mathcal{P}(\mathcal{P}(\mathbb{S})))$ be a solution of the continuity equation $\partial_t M_t + \mathrm{div}_{\mathcal{P}(\mathbb{S})}(b_t M_t) = 0$ in the sense of Definition 6.3. Then, there exists $\Lambda \in \mathcal{P}(C([0,T], \mathcal{P}(\mathbb{S})))$ and $\mathfrak{L} \in \mathcal{P}(\mathcal{P}(C([0,T], \mathbb{S})))$ such that:*

(1) $(\mathfrak{e}_t)_\sharp \Lambda = M_t$ for any $t \in [0,T]$, with $\mathfrak{e}_t(\boldsymbol{\mu}) = \mu_t$ and Λ-a.e. $\boldsymbol{\nu} = (\mu_t)_{t \in [0,T]} \in C([0,T], \mathcal{P}(\mathbb{S}))$ solves $\partial_t \mu_t + \mathrm{div}_{\mathbb{S}}(b_t(\cdot, \mu_t)\mu_t) = 0$, in the sense of Def. 6.1;

(2) $(E_t)_\sharp \mathfrak{L} = M_t$, where $E_t = (e_t)_\sharp$ and $e_t(\boldsymbol{y}) = y_t$ for any $\boldsymbol{y} = (y_t)_{t \in [0,T]} \in C([0,T], \mathbb{S})$. Moreover, for $\mathfrak{L}$-a.e. $\lambda \in \mathcal{P}(C([0,T], \mathbb{S}))$, it holds that λ is concentrated over $(y_t)_{t \in [0,T]} \in AC([0,T], \mathbb{S})$ that are solutions of $\dot{\boldsymbol{y}}(t) = b(t, y_t, (e_t)_\sharp \lambda)$.

Proof. Define $(\tilde{M}_t)_{t \in [0,T]} \in C([0,T], \mathcal{P}(\mathcal{P}(\mathbb{R}^d)))$ and $\tilde{b} : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ as in, respectively, (6.17) and (6.16). Thanks to Proposition 6.4, the continuity equation $\partial_t \tilde{M}_t + \mathrm{div}(\tilde{b}_t \tilde{M}_t) = 0$ is satisfied.

Thanks to Theorem 2.10, there exists $\tilde{\Lambda} \in \mathcal{P}(C_T(\mathcal{P}(\mathbb{R}^d)))$ such that $(\mathfrak{e}_t)_\sharp \tilde{\Lambda} = \tilde{M}_t$ and $\tilde{\Lambda}$-a.e. $\tilde{\boldsymbol{\mu}} = (\tilde{\mu}_t)_{t \in [0,T]} \in C([0,T], \mathcal{P}(\mathbb{R}^d))$ solves

$$\partial_t \tilde{\mu}_t + \mathrm{div}(\tilde{b}_t(\cdot, \tilde{\mu}_t)\tilde{\mu}_t) = 0.$$

In particular, note that $\tilde{\Lambda} \in \mathcal{P}(C([0,T], \mathcal{P}(S)))$, since $S \subset \mathbb{R}^d$ is closed. Then, we set $\Lambda := J_\sharp^{-1} \tilde{\Lambda}$. We now show that Λ satisfies the properties in 1. For all $F : \mathcal{P}(\mathbb{S}) \to [0, +\infty]$ Borel, it holds

$$
\begin{aligned}
\int_{C([0,T], \mathcal{P}(\mathbb{S}))} F(\mathfrak{e}_t \boldsymbol{\mu}) d\Lambda(\boldsymbol{\mu}) &= \int_{C([0,T], \mathcal{P}(\mathbb{S}))} F(\mathfrak{e}_t \boldsymbol{\mu}) dJ_\sharp^{-1} \tilde{\Lambda}(\boldsymbol{\mu}) = \int_{C([0,T], \mathcal{P}(S))} F(\mathfrak{e}_t(J\tilde{\boldsymbol{\mu}})) d\tilde{\Lambda}(\tilde{\boldsymbol{\mu}}) \\
&= \int_{C([0,T], \mathcal{P}(S))} F(\jmath_\sharp^{-1} \tilde{\mu}_t) d\tilde{\Lambda}(\tilde{\boldsymbol{\mu}}) = \int_{\mathcal{P}(S)} F(\jmath_\sharp^{-1} \tilde{\mu}) d\tilde{M}_t(\tilde{\mu}) = \int_{\mathcal{P}(\mathbb{S})} F(\mu) dM_t(\mu).
\end{aligned}
$$

The fact that it is concentrated over solutions of $\partial_t \mu_t + \mathrm{div}_{\mathbb{S}}(b_t(\cdot, \mu_t)\mu_t) = 0$, is a consequence of Proposition 6.5, *(4)*.

Again by Theorem 2.10, there exists a random measure $\tilde{\mathfrak{L}} \in \mathcal{P}(\mathcal{P}(C([0,T], \mathbb{R}^d)))$ satisfying $(E_t)_\sharp \tilde{\mathfrak{L}} = M_t$ and $\tilde{\mathfrak{L}}$-a.e. $\tilde{\lambda} \in \mathcal{P}(C([0,T], \mathbb{R}^d))$ is concentrated over curves $\boldsymbol{x} = (x_t)_{t \in [0,T]} \in AC([0,T], \mathbb{R}^d)$ that are solutions of $\dot{x}(t) = \tilde{b}(t, x_t, (e_t)_\sharp \tilde{\lambda})$. Notice that $\tilde{\mathfrak{L}} \in \mathcal{P}(\mathcal{P}(C([0,T], S)))$, since $S \subset \mathbb{R}^d$ is closed. Then, set $\mathfrak{L} := \mathfrak{J}_\sharp^{-1} \tilde{\mathfrak{L}} = \ell_{\sharp\sharp}^{-1} \tilde{\mathfrak{L}}$, for which we check that satisfies the properties in 2. For all $F : \mathcal{P}(\mathbb{S}) \to [0,1]$ Borel, it holds

$$\int_{\mathcal{P}(C([0,T], \mathbb{S}))} F((e_t)_\sharp \eta) d\mathfrak{L}(\eta) = \int_{\mathcal{P}(C([0,T], \mathbb{S}))} F((e_t)_\sharp \lambda) d\mathfrak{J}_\sharp^{-1} \tilde{\mathfrak{L}}(\lambda)$$

$$= \int_{\mathcal{P}(C([0,T],S))} F((e_t \circ \ell^{-1})_\sharp \tilde{\lambda}) d\tilde{\mathfrak{L}}(\tilde{\lambda}) = \int_{\mathcal{P}(C([0,T],S))} F((\jmath^{-1} \circ e_t)_\sharp \tilde{\lambda}) d\tilde{\mathfrak{L}}(\tilde{\lambda})$$
$$= \int_{\mathcal{P}(C([0,T],S))} F(\jmath_\sharp^{-1}(E_t\tilde{\lambda})) d\tilde{\mathfrak{L}}(\tilde{\lambda}) = \int_{\mathcal{P}(S)} F(\jmath_\sharp^{-1}\tilde{\mu}) d\tilde{M}_t(\tilde{\mu}) = \int_{\mathcal{P}(\mathbb{S})} F(\mu) dM_t(\mu),$$

where we used $e_t \circ \ell^{-1} = \jmath^{-1} \circ e_t$. The other property follows from Proposition 6.5, *(3)*. □

Remark 6.7. *By construction, the following property holds as well:* $\Lambda = E_\sharp \mathfrak{L}$, *where* $E : \mathcal{P}(C([0,T],\mathbb{S})) \to C([0,T],\mathcal{P}(\mathbb{S}))$ *is defined as* $E(\lambda) = ((e_t)_\sharp \lambda)_{t\in[0,T]}$, *for any* $\lambda \in \mathcal{P}(C([0,T],\mathbb{S}))$. *Moreover, there exists a (Souslin-Borel) measurable map* $G : C([0,T],\mathcal{P}(\mathbb{S})) \to \mathcal{P}(C([0,T],\mathbb{S}))$ *such that* $\mathfrak{L} = G_\sharp \Lambda$, *and for* Λ*-a.e.* $\boldsymbol{\mu}$ *it holds* $E(G(\boldsymbol{\mu})) = \boldsymbol{\mu}$. *Indeed, such a property holds in* $\mathbb{R}^d$, *as shown in [PS25], and it is inherited thanks to the properties of the Nash isometry.*

6.2. **Atomic nested superposition principle for compact Riemannian manifolds.** In this subsection, we exploit again the Nash isometry theorem to prove a version of Theorem 4.4 and 4.5 in the ambient space $\mathbb{S}$.

The next proposition shows how the capacity assumption is related to the dimension of the manifold.

Proposition 6.8. *Let* $(\mathbb{S}, g_\mathbb{S})$ *be a compact Riemannian manifold of dimension* $k \geq 2$, *with* $\mathrm{vol}_\mathbb{S}$ *its natural volume measure. Then, for all* $f \in L^\infty(\mathrm{vol}_\mathbb{S})$, *given* $\nu = f\,\mathrm{vol}_\mathbb{S}$, *it holds* $\mathrm{cap}_{k,\jmath_\sharp\nu}(\Delta) = 0$. *In particular,* $\mathrm{cap}_{r,\jmath_\sharp\nu}(\Delta) = 0$ *for all* $r \leq k$.

Proof. The proof relies again on the logarithmic cutoff functions introduced in (5.5). As in Proposition 5.6, writing $\tilde{\nu} := \jmath_\sharp \nu$ we show that

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |h_{\varepsilon,R}(x_1,x_2)|^k + |\nabla h_{\varepsilon,R}(x_1,x_2)|^k d\tilde{\nu}\otimes\tilde{\nu}(x_1,x_2) \to 0,$$

when choosing $R = \sqrt{\varepsilon}$ and letting $\varepsilon \to 0$. The first term is again trivial, since

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |h_{\varepsilon,R}(x_1,x_2)|^k d\tilde{\nu}\otimes\tilde{\nu}(x_1,x_2) = \tilde{\nu}\otimes\tilde{\nu}(\Delta_R) \to \tilde{\nu}\otimes\tilde{\nu}(\Delta) = 0,$$

as $R \to 0$, since $\tilde{\nu}$ is atomless. The second term is a bit more delicate, and we will use many properties of the Nash embedding $\jmath$, in particular (6.4), (6.5) and the fact that $\{x \in S : |x_1 - x_2| \in (\varepsilon, R)\} \subset \{x_1 \in S : d_S(x_1,x_2) \in (\varepsilon, CR)\}$, as well as the estimate (5.6):

$$\begin{aligned}
\int_{\mathbb{R}^d\times\mathbb{R}^d} |\nabla h_{\varepsilon,R}|^k d\tilde{\nu}\otimes\tilde{\nu} &\leq \frac{2^k}{\left(\log\frac{R}{\varepsilon}\right)^k} \int_S \int_{\{x_1\in S:|x_1-x_2|\in(\varepsilon,R)\}} \frac{1}{|x_1-x_2|^k} f\circ\jmath^{-1}(x_1) d\mathcal{H}^k(x_1) d\tilde{\nu}(x_2) \\
&\leq \frac{2^k\|f\|_\infty}{\left(\log\frac{R}{\varepsilon}\right)^k} \sup_{x_2\in S} \int_{\{x_1\in S:|x_1-x_2|\in(\varepsilon,R)\}} \frac{1}{|x_1-x_2|^k} d\mathcal{H}^k(x_1) \\
&\leq \frac{2^k\|f\|_\infty}{C^k\left(\log\frac{R}{\varepsilon}\right)^k} \sup_{x_2\in S} \int_{\{x_1\in S:d_S(x_1,x_2)\in(\varepsilon,CR)\}} \frac{1}{d_S^k(x_1,x_2)} d\mathcal{H}^k(x_1) \\
&= \frac{2^k\|f\|_\infty}{C^k\left(\log\frac{R}{\varepsilon}\right)^k} \sup_{y_2\in\mathbb{S}} \int_{\{y_1\in\mathbb{S}:d_\mathbb{S}(y_1,y_2)\in(\varepsilon,CR)\}} \frac{1}{d_\mathbb{S}^k(y_1,y_2)} d\,\mathrm{vol}_\mathbb{S}(y_1) \\
&= \frac{2^k\|f\|_\infty}{C^k\left(\log\frac{R}{\varepsilon}\right)^k} \sup_{y_2\in\mathbb{S}} \int_\varepsilon^{CR} \frac{1}{r^k} \mathcal{H}^{k-1}(\{y_1\in\mathbb{S} : d_\mathbb{S}(y_1,y_2) = r\}) dr.
\end{aligned}$$

Notice that, in the last row, the measure $\mathcal{H}^{k-1}$ is the $(k-1)$-dimensional Hausdorff measure built on the manifold $\mathbb{S}$. Now, thanks to [Gra74, Equation (11)], for all $y_2 \in \mathbb{S}$ there exists a constant $\tilde{C} > 0$ such that, for $r > 0$ small enough, it holds $\mathcal{H}^{k-1}(\{y_1 \in \mathbb{S} : d_{\mathbb{S}}(y_1, y_2) = r\}) \leq \alpha_k r^{k-1} + \tilde{C} r^{k+1}$, where α_k is the $(k-1)$-dimensional measure of a unit sphere in $\mathbb{R}^k$. Then

$$\begin{aligned}\int_{\mathbb{R}^d \times \mathbb{R}^d} |\nabla h_{\varepsilon,R}|^k d\tilde{\nu} \otimes \tilde{\nu} \leq & \frac{2^k \|f\|_\infty}{C^k \left(\log \frac{R}{\varepsilon}\right)^k} \int_\varepsilon^{CR} \frac{\alpha_k}{r} + \tilde{C} r dr \\ = & \frac{2^k \|f\|_\infty}{C^k} \left(\frac{\alpha_k}{\left(\log \frac{R}{\varepsilon}\right)^{k-1}} + \frac{\log C + C^2 R^2 - \varepsilon^2}{\left(\log \frac{R}{\varepsilon}\right)^k} \right),\end{aligned}$$

and we conclude choosing $R = \sqrt{\varepsilon}$ and letting $\varepsilon \to 0$. □

Note that the last part of the statement about the r-capacity with $r < k$, would easily follow from Corollary 5.4 as well.

We are now ready to state and prove the atomic nested superposition principle, in both the differential and metric versions, when the ambient space is a Riemannian manifold.

Theorem 6.9. *Let* $(\mathbb{S}, g_{\mathbb{S}})$ *be a compact, boundaryless and oriented Riemannian manifold. Assume* $\pi \in \mathcal{P}(\mathbf{T})$ *and* $\nu \in \mathcal{P}(\mathbb{S})$ *with bounded density with respect to* $\mathrm{vol}_{\mathbb{S}}$.

Let $\Lambda \in \mathcal{P}(C([0,T], \mathcal{P}(\mathbb{S})))$ *be concentrated over solutions of* $\partial_t \mu_t + \mathrm{div}_{\mathbb{S}}(b_t(\cdot, \mu_t)\mu_t) = 0$ *and define* $\boldsymbol{M} = (M_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathcal{P}(\mathbb{S})))$ *as* $M_t = (\mathfrak{e}_t)_\sharp \Lambda$. *Assume:*

(1) $M_t = u_t Q_{\pi,\nu}$, *with* $u \in L^1_t(L^q(Q_{\pi,\nu}))$, *with* $q \in [1, +\infty]$;
(2) for some $p > 1$, *it holds*

$$\int_0^T \int_{\mathcal{P}(\mathbb{S})} \int_{\mathbb{S}} \|b_t(y,\mu)\|^p_{T_y\mathbb{S}} d\mu(y) dM_t(\mu) dt < +\infty; \tag{6.25}$$

(3) $\frac{p'}{k} + \frac{1}{q} \leq 1$.

Then, there exists $\mathfrak{L} \in \mathcal{P}(\mathcal{P}^{\mathrm{pa}}(AC^p([0,T], \mathbb{S})))$ *such that* $E_\sharp \mathfrak{L} = \Lambda$ *and is concentrated over* $\lambda \in \mathcal{P}(C([0,T], \mathbb{S}))$ *that are, in turn, concentrated over absolutely continuous curves* $\boldsymbol{y} = (y_t)_{t\in[0,T]}$ *that are solutions of* $\dot{\boldsymbol{y}}(t) = b_t(y_t, (e_t)_\sharp \lambda)$.
In particular, $(E_t)_\sharp \mathfrak{L} = M_t$ *for all* $t \in [0,T]$ *and* Λ *is concentrated over* $C([0,T], \mathcal{P}^{\mathrm{pa}}(\mathbb{S}))$.

Proof. Define $\tilde{\Lambda} := J_\sharp \Lambda \in \mathcal{P}(C([0,T], \mathcal{P}(S))) \subset \mathcal{P}(C([0,T], \mathcal{P}(\mathbb{R}^d)))$. Thanks to Proposition 6.5, *(4)*, $\tilde{\Lambda}$ is concentrated on $\mathrm{CE}(\tilde{b})$. Next, we show that it satisfies the hypothesis of Theorem 4.4.

(1) The equality $\mathfrak{e}_t \circ J = \jmath_\sharp \circ \mathfrak{e}_t$ for all $t \in [0,T]$, gives that $(\mathfrak{e}_t)_\sharp \tilde{\Lambda} = \tilde{M}_t := \jmath_{\sharp\sharp} M_t$. Moreover, recalling (6.7), we have that

$$\tilde{M}_t = \tilde{u}_t Q_{\pi, \jmath_\sharp \nu}, \quad \tilde{u}_t \circ \jmath_\sharp^{-1},$$

which satisfy $\tilde{u} \in L^1_t(L^q(Q_{\pi,\jmath_\sharp\nu}))$ from the following computation:

$$\int_0^T \|u_t\|_{L^q(Q_{\pi,\nu})} dt = \int_0^T \|u_t \circ \boldsymbol{\jmath}^{-1} \circ \boldsymbol{\jmath}\|_{L^q(Q_{\pi,\nu})} dt = \int_0^T \|u_t \circ \boldsymbol{\jmath}^{-1}\|_{L^q(Q_{\pi,\boldsymbol{\jmath}(\nu)})} dt.$$

(2) Again, the local isometry property of $\jmath$ implies $\|b_t(y,\mu)\|_{\mathsf{g}} = |\tilde{b}_t(\jmath(y), \jmath_\sharp \mu)|$, and in particular $\|b_t(\cdot,\mu)\|_{L^p(\mu)} = \|\tilde{b}_t(\cdot, \jmath_\sharp \mu)\|_{L^p(\jmath_\sharp \mu)}$. Then, since $\tilde{M}_t = \jmath_{\sharp\sharp} M_t$, we finally have

$$\int_0^T \int_{\mathcal{P}(\mathbb{S})} \int_{\mathbb{S}} \|b_t(y,\mu)\|^p_{\mathsf{g}} d\mu(y) dM_t(\mu) dt = \int_0^T \int_{\mathcal{P}(\mathbb{R}^d)} \int_{\mathbb{R}^d} |\tilde{b}_t(x, \tilde{\mu})|^p d\tilde{\mu}(x) d\tilde{M}_t dt;$$

(3) Proposition 6.8 gives that $\mathrm{cap}_{k,\jmath_\sharp\nu}(\Delta) = 0$.
(4) The balance between the exponents works for any $r \leq k$.

Then, there exists $\tilde{\mathfrak{L}} \in \mathcal{P}(\mathcal{P}^{\mathrm{pa}}(C([0,T],\mathbb{R}^d)))$ that is concentrated over $\mathrm{SPS}(\tilde{b})$ and satisfies $E_\sharp\tilde{\mathfrak{L}} = \tilde{\Lambda}$, which implies that $\tilde{\mathfrak{L}} \in \mathcal{P}(\mathcal{P}(C([0,T],S)))$. Then, defining $\mathfrak{L} := \mathfrak{J}_\sharp^{-1}\tilde{\mathfrak{L}} \in \mathcal{P}(\mathcal{P}(C([0,T],\mathbb{S})))$, we have: it is concentrated over $\lambda \in \mathcal{P}^{\mathrm{pa}}(C([0,T],\mathbb{S}))$ thanks to Lemma 2.5; it is concentrated over λ that, in turn, are concentrated over $\boldsymbol{y} = (y_t)_{t\in[0,T]} \in AC([0,T],\mathbb{S})$ that are solutions of $\dot{y}(t) = b_t(y_t,(e_t)_\sharp\lambda)$, thanks to Proposition 6.5, *(3)*; the trivial equality $E \circ \mathfrak{J}^{-1} = J^{-1} \circ E$ gives that $E_\sharp\mathfrak{L} = (E \circ \mathfrak{J}^{-1})_\sharp\tilde{\mathfrak{L}} = (J^{-1} \circ E)_\sharp\tilde{\mathfrak{L}} = J_\sharp^{-1}\tilde{\Lambda} = \Lambda$. □

Theorem 6.10. *Let $(\mathbb{S}, g_{\mathbb{S}})$ be a compact, boundaryless and oriented Riemannian manifold. Assume $\pi \in \mathcal{P}(\mathbf{T})$ and $\nu \in \mathcal{P}(\mathbb{S})$ with bounded density with respect to $\mathrm{vol}_{\mathbb{S}}$.*

Let $\Lambda \in \mathcal{P}(C([0,T],\mathcal{P}(\mathbb{S})))$ be concentrated over $AC^p([0,T],\mathcal{P}_p(\mathbb{S}))$ and define $\boldsymbol{M} = (M_t)_{t\in[0,T]} \in C([0,T],\mathcal{P}(\mathcal{P}(\mathbb{S})))$ as $M_t = (\mathfrak{e}_t)_\sharp\Lambda$. Assume:

(1) $M_t = u_tQ_{\pi,\nu}$, with $u \in L^1_t(L^q(Q_{\pi,\nu}))$, with $q \in [1,+\infty]$;
(2) for some $p > 1$, it holds

$$\int\int_0^T |\dot{\boldsymbol{\mu}}|^p_{W_{p,d_{\mathbb{S}}}}(t)dtd\Lambda(\boldsymbol{\mu}) < +\infty; \tag{6.26}$$

(3) $\frac{p'}{k} + \frac{1}{q} \leq 1$.

Then, there exists $\mathfrak{L} \in \mathcal{P}(\mathcal{P}^{\mathrm{pa}}(AC^p([0,T],\mathbb{S})))$ such that $E_\sharp\mathfrak{L} = \Lambda$ and

$$\int\int\int_0^T |\dot{\boldsymbol{y}}|^p_{d_{\mathbb{S}}}(t)dtd\lambda(\boldsymbol{y})d\mathfrak{L}(\lambda) < +\infty \tag{6.27}$$

In particular, $(E_t)_\sharp\mathfrak{L} = M_t$ for all $t \in [0,T]$ and Λ is concentrated over $C([0,T],\mathcal{P}^{\mathrm{pa}}(\mathbb{S}))$.

Proof. The proof follows the same line of the one of Theorem 6.9. Define $\tilde{\Lambda} := J_\sharp\Lambda \in \mathcal{P}(C([0,T],\mathcal{P}(S)))$. Thanks to Proposition 6.5, *(2)*, $\tilde{\Lambda}$ is concentrated over $AC^p([0,T],S)$, which coincides with $AC^p([0,T],\mathbb{R}^d) \cap C([0,T],S)$ because $|x-y| \leq d_S(x,y) \leq C|x,y|$ for all $x,y \in S$. Let us verify that $\tilde{\Lambda}$ satisfies the assumptions of Theorem 4.5:

(1) As in the proof of Theorem 6.9, $(\mathfrak{e}_t)_\sharp\tilde{\Lambda} = \tilde{M}_t := \jmath_{\sharp\sharp}M_t$, and $\tilde{M}_t = \tilde{u}_tQ_{\pi,\jmath_\sharp\nu}$, with $\tilde{u} \in L^1_t(L^q(Q_{\pi,\jmath_\sharp\nu}))$.
(2) Thanks to (6.21) and $W_p(\tilde{\mu}_1,\tilde{\mu}_2) \leq CW_{p,d_S}(\tilde{\mu}_1,\tilde{\mu}_2)$ for all $\tilde{\mu}_1,\tilde{\mu}_2 \in \mathcal{P}(S)$, we have

$$\int\int_0^T |\dot{\tilde{\boldsymbol{\mu}}}|^p_{W_p}(t)dtd\tilde{\Lambda}(\tilde{\boldsymbol{\mu}}) \leq C\int\int_0^T |\dot{\tilde{\boldsymbol{\mu}}}|^p_{W_{p,d_S}}(t)dtd\tilde{\Lambda}(\tilde{\boldsymbol{\mu}}) = C\int\int_0^T |\dot{\boldsymbol{\mu}}|^p_{W_{p,d_{\mathbb{S}}}}(t)dtd\Lambda(\boldsymbol{\mu}) < +\infty.$$

(3) and (4) They both easily follow as in the proof of Theorem 6.9.

Then, there exists $\tilde{\mathfrak{L}} \in \mathcal{P}(\mathcal{P}^{\mathrm{pa}}(AC^p([0,T],\mathbb{R}^d)))$ satisfying $E_\sharp\tilde{\mathfrak{L}} = \tilde{\Lambda}$, and in particular $\tilde{\mathfrak{L}} \in \mathcal{P}(\mathcal{P}(C([0,T],S)))$, so that

$$\int\int\int_0^T |\dot{x}|^p(t)dtd\tilde{\lambda}(\boldsymbol{x})d\tilde{\mathfrak{L}}(\tilde{\lambda}) \leq C\int\int\int_0^T |\dot{x}|^p_{d_S}(t)dtd\tilde{\lambda}(\boldsymbol{x})d\tilde{\mathfrak{L}}(\tilde{\lambda}) < +\infty.$$

Define $\mathfrak{L} := \mathfrak{J}_\sharp^{-1}\tilde{\mathfrak{L}}$. Again, we have: $E_\sharp\mathfrak{L} = (E \circ \mathfrak{J}^{-1})_\sharp\tilde{\mathfrak{L}} = (J^{-1} \circ E)_\sharp\tilde{\mathfrak{L}} = J_\sharp^{-1}\tilde{\Lambda} = \Lambda$; $\mathfrak{L} \in \mathcal{P}(\mathcal{P}^{\mathrm{pa}}(C([0,T],\mathbb{S})))$ thanks to Lemma 2.5; thanks to Proposition 6.20, *(1)*, $\mathfrak{L} \in \mathcal{P}(\mathcal{P}(AC^p([0,T],\mathbb{S})))$ and (6.27) holds as well. □

7. A metric measure viewpoint on the Wasserstein space

Let us recall some general notions about Sobolev spaces on metric measure spaces, for which we mainly follow [GP20; AILP24]. In particular, over all the equivalent definitions of Sobolev spaces, we will give the one that exploits *test plans*, that will be particularly useful for our setting.
Let $(X, d, \mathtt{m})$ be a metric probability measure space, that is (X, d) is a complete and separable metric space and $\mathtt{m} \in \mathcal{P}(X)$.

Definition 7.1 (p'-test plan)**.** *Let $p \in [1, +\infty)$. A p'-test plan is a probability measure over curves $\Pi \in \mathcal{P}(C([0,1], X))$ satisfying:*

(1) there exists $C \geq 0$ (compression constant) such that $(\mathrm{e}_t)_\sharp \Pi \leq C\mathtt{m}$ for all $t \in [0,1]$;
(2) it has finite p'-energy, that is

$$\int \int_0^1 |\dot{\boldsymbol{x}}|_d^{p'}(t) dt d\Pi(\boldsymbol{x}) < +\infty. \tag{7.1}$$

Definition 7.2 (p-Sobolev space)**.** *Let $p \in [1, +\infty)$ and $f \in L^p(\mathtt{m})$. We say that $f \in W^{1,p}(X, d, \mathtt{m})$ if there exists a non-negative function $G \in L^p(\mathtt{m})$ such that for all p'-test plan Π, it holds $f \circ \boldsymbol{x} \in W^{1,1}(0,1)$ for Π-a.e. $\boldsymbol{x}$ and*

$$|(f \circ \boldsymbol{x})'(t)| \leq G(\boldsymbol{x}_t)|\dot{\boldsymbol{x}}|_d(t), \quad \text{for } \Pi \otimes \mathcal{L}^1\text{-a.e. } (\boldsymbol{x}, t) \in C([0,1], X) \times [0,1]. \tag{7.2}$$

Any G as above is called a p-weak upper gradient, and in the class of weak upper gradients we can consider the $\mathtt{m}$-a.e. pointwise minimum, denoted by $|Df|$. Then, $W^{1,p}(X, d, \mathtt{m})$ is a Banach space when endowed with the norm

$$\|f\|_{W^{1,p}(X,d,\mathtt{m})} := \left(\|f\|^p_{L^p(\mathtt{m})} + \||Df|\|^p_{L^p(\mathtt{m})} \right)^{1/p}. \tag{7.3}$$

A different, but equivalent, approach for defining the Sobolev spaces is by relaxing Lipschitz functions [AILP24, Section 5.1], giving in particular that globally L-Lipschitz functions f are in $W^{1,p}(X, d, \mathtt{m})$, with $|Df| \leq L$ $\mathtt{m}$-a.e. The vice versa is called *Sobolev-to-Lipschitz* property and does not always hold:

$$\begin{gathered} W^{1,p}(X, d, \mathtt{m}) \text{ satisfies the } p\text{-Sobolev-to-Lipschitz property} \\ \Longleftrightarrow \\ \text{for all } f \in W^{1,p}(X, d, \mathtt{m}), \ \ |Df| \leq L \implies \exists \tilde{f} : X \to \mathbb{R} \ L\text{-Lip. with } f = \tilde{f} \ \mathtt{m}\text{-a.e.} \end{gathered} \tag{Sob-Lip}$$

Another important property regards the Hilbertianity of the 2-Sobolev space. It is not always true, as in the usual Sobolev theory, that $W^{1,2}(X, d, \mathtt{m})$ is an Hilbert space. We thus recall the following definition due to N. Gigli [Gig13]:

$$(X, d, \mathtt{m}) \text{ is } \textbf{infinitesimally Hilbertian} \iff W^{1,2}(X, d, \mathtt{m}) \text{ is a Hilbert space.} \tag{inf-H}$$

From now on, we assume that $(X, d, \mathtt{m})$ is infinitesimally Hilbertian, which implies that the associated Cheeger energy

$$\mathrm{Ch}(f) := \begin{cases} \int_X |Df|^2 d\mathtt{m}, & f \in W^{1,2}(X, d, \mathtt{m}) \\ +\infty, & f \in L^2(\mathtt{m}) \setminus W^{1,2}(X, d, \mathtt{m}) \end{cases}$$

is quadratic. Moreover, it is lower semicontinuous with respect to the L^2-convergence. In particular, we can build its gradient flow in the Lebesgue space $L^2(\mathtt{m})$, obtaining a contractive linear semigroup $\mathtt{H}_t : L^2(\mathtt{m}) \to L^2(\mathtt{m})$, $t \geq 0$, usually called *heat semigroup* associated to $(X, d, \mathtt{m})$, performing the *minimizing movement scheme*: for all $f \in L^2(\mathtt{m})$ and $\tau > 0$, define $f^{(\tau)}_{(0)} := f$ and

$$f^{(\tau)}_{(n+1)} \in \operatorname{argmin}\left(\mathrm{Ch}(\cdot) + \frac{\|\cdot - f^{(\tau)}_{(n)}\|^2}{2\tau}\right), \quad f^{(\tau)}_t := f^{(\tau)}_{(n)}, \text{ with } t \in [n\tau, (n+1)\tau). \tag{7.4}$$

Then, $\mathtt{H}_t f = \lim_{\tau\to 0} f^{(\tau)}_t$. Its main properties are listed in [GP20, Theorem 5.1.12]. Here we list the ones that are of our interest:

(1) for all $t \geq 0$, $\mathtt{H}_t$ is well-defined as an operator from $L^p(\mathtt{m})$ to $L^p(\mathtt{m})$, as shown in [GP20, Proposition 5.2.15];
(2) for all $t > 0$ and $f \in L^2(\mathtt{m})$, $\mathtt{H}_t f \in W^{1,2}(X, d, \mathtt{m})$;
(3) for all $p \in [1, +\infty)$ and $t > 0$, $\|\mathtt{H}_t f\|_{L^p(\mathtt{m})} \leq \|f\|_{L^p(\mathtt{m})}$ for all $L^2(\mathtt{m}) \cap L^p(\mathtt{m})$. In particular, it can be extended to a linear and continuous operator $\mathtt{H}_t : L^p(\mathtt{m}) \to L^p(\mathtt{m})$;
(4) it has the Markov property, that is $\mathtt{H}_t \mathbb{1} = \mathbb{1}$ and $\mathtt{H}_t f \geq 0$ $\mathtt{m}$-a.e., whenever $f \geq 0$ $\mathtt{m}$-a.e. Both properties can be directly verified from the minimizing movement scheme (7.4).

The last property we introduce about general metric measure spaces is the *Bakry–Émery curvature condition*, for some $K \in \mathbb{R}$:

$$|D(\mathtt{H}_t f)|^2 \leq e^{-2Kt}\mathtt{H}_t(|Df|^2) \quad \mathtt{m}\text{-a.e.}, \qquad \forall f \in W^{1,2}(X, d, \mathtt{m}),\ \forall t \geq 0. \tag{BE(K,∞)}$$

The three conditions above provides one of the equivalent definitions of $\mathrm{RCD}(K, \infty)$ spaces (see for example [Sav14; AGS15]).

Definition 7.3. *Let $K \in \mathbb{R}$. A metric measure space $(X, d, \mathtt{m})$ is said to be* $\mathrm{RCD}(K, \infty)$ *if it satisfies* (inf-H), (Sob-Lip) *with $p = 2$ and* (BE(K, ∞)).

In the next section, we will see how our atomic nested superposition principles can be exploited to show that the L^2-Wasserstein space, endowed with suitable reference measures of the form $Q_{\pi,\nu}$, does not satisfy (Sob-Lip), and therefore it cannot be a RCD space.

Before applying this general theory to the Wasserstein space, let us state a useful lemma.

Lemma 7.4. *Let $(X, d, \mathtt{m})$ be infinitesimally Hilbertian, and $K \in \mathbb{R}$. Let $\mathcal{A} \subset W^{1,2}(X, d, \mathtt{m})$ be dense in energy, that is, for all $f \in W^{1,2}(X, d, \mathtt{m})$ there exist $f_n \in \mathcal{A}$ such that*

$$f_n \xrightarrow{L^2(\mathtt{m})} f, \quad |Df_n| \xrightarrow{L^2(\mathtt{m})} |Df|. \tag{7.5}$$

Then, it suffices to verify (BE(K, ∞)) *for $f \in \mathcal{A}$.*

Proof. First, we show that for all Borel function $\varphi : X \to [0, +\infty)$ such that there exist $0 < c < C < +\infty$ with $c \leq \varphi \leq C$, it holds

$$\mathrm{Ch}_\varphi : L^2(\mathtt{m}) \to [0, +\infty], \quad \mathrm{Ch}_\varphi(f) := \begin{cases} \int_X \varphi |Df|^2 d\mathtt{m}, & f \in W^{1,2}(X, d, \mathtt{m}) \\ +\infty, & f \in L^2(\mathtt{m}) \setminus W^{1,2}(X, d, \mathtt{m}) \end{cases} \tag{7.6}$$

is lower semicontinuous with respect to the L^2-convergence. Take $f_n \to f$ in $L^2(\mathtt{m})$. If $\liminf \mathrm{Ch}_\varphi(f_n) = +\infty$, there is nothing to prove, thus we assume it is finite and, up to consider to a subsequence, we assume that the limit exists. Moreover, define the finite positive measure $\mathtt{m}_\varphi := \varphi\mathtt{m}$, which is equivalent to $\mathtt{m}$, since $c\mathtt{m} \leq \mathtt{m}_\varphi \leq C\mathtt{m}$. Now, since the limit of $\mathrm{Ch}_\varphi(f_n)$ is finite, there exists a subsequence f_{n_k} and $G \in L^2(\mathtt{m})$ such that $|Df_{n_k}| \rightharpoonup G$ weakly in $L^2(\mathtt{m}_\varphi)$.

Moreover, the same limit holds weakly in $L^2(\mathtt{m})$, implying that $G \geq |Df|$, see for example [Gig18, Proposition 2.1.13]. In particular

$$\lim_{n\to+\infty} \mathrm{Ch}_\varphi(f_n) = \lim_{n\to+\infty} \||Df_n|\|^2_{L^2(\mathtt{m}_\varphi)} \geq \|G\|^2_{L^2(\mathtt{m}_\varphi)} = \int_X \varphi G^2 \mathtt{m} \geq \int_X \varphi |Df|^2 d\mathtt{m}.$$

Now, we are done if we prove that

$$\int_X \varphi |D(\mathtt{H}_t f)|^2 d\mathtt{m} \leq e^{-2Kt} \int_X \varphi\ \mathtt{H}_t(|Df|^2) d\mathtt{m},$$

for all non-negative and bounded functions φ. Consider $f_n \in \mathcal{A}$ such that $f_n \to f$ as in (7.5). Then, $|Df_n|^2 \to |Df|^2$ in $L^1(\mathtt{m})$, and by the L^1-contraction property of $\mathtt{H}_t$, also $\mathtt{H}_t(|Df_n|^2) \to \mathtt{H}_t(|Df|^2)$ in $L^1(\mathtt{m})$. Then, for all $\varepsilon > 0$, we have

$$\begin{aligned}\int_X \varphi |D(\mathtt{H}_t f)|^2 d\mathtt{m} &\leq \int_X (\varphi+\varepsilon) |D(\mathtt{H}_t f)|^2 d\mathtt{m} \liminf_{n\to+\infty} \int_X (\varphi+\varepsilon)|D(\mathtt{H}_t f_n)|^2 d\mathtt{m} \\ &\leq \liminf_{n\to+\infty} e^{-2Kt} \int_X (\varphi+\varepsilon)\ \mathtt{H}_t(|Df_n|^2) d\mathtt{m} = e^{-2Kt}\int_X \varphi\ \mathtt{H}_t(|Df|^2)d\mathtt{m} + \varepsilon \int_X \mathtt{H}_t(|Df|^2) d\mathtt{m},\end{aligned}$$

and we conclude by arbitrarinees of $\varepsilon > 0$, since $\int_X \mathtt{H}_t(|Df|^2) d\mathtt{m} < +\infty$. □

7.1. **Lack of Sobolev-to-Lipschitz and Poincaré inequality.**

7.1.1. *Euclidean case.* Let $\pi \in \mathbf{T}$ and $\nu \in \mathcal{P}_p(\mathbb{R}^d)$ be an atomless measure. Here we show that, for suitable $p > 1$, the p-Sobolev-to-Lipschitz and the Poincaré inequality properties fail in the metric measure space $(\mathcal{P}_p(\mathbb{R}^d), W_p, Q_{\pi,\nu})$. In the next theorem, we build a function that will be a counterexample to both properties.

Theorem 7.5. *Let $p \in (1,+\infty)$ and $r \in (1,+\infty)$ be such that $p \geq \frac{r}{r-1}$ and* $\mathrm{cap}_{r,\nu}(\Delta) = 0$*. Assume that $\pi \neq \delta_{\mathbf{a}}$ for all $\mathbf{a} \in \mathbf{T}$.*
Then, there exists a function $F \in W^{1,p}(\mathcal{P}_p(\mathbb{R}^d), W_p, Q_{\pi,\nu})$ for which $|DF| \equiv 0$, but it does not admit a constant representative.

Proof. For all $a \in (0,1]$, consider the function $F_a(\mu) := \int_{\mathbb{R}^d} \mathbb{1}_{[a,1]}(\mu[x]) d\mu(x)$, for all $\mu \in \mathcal{P}_p(\mathbb{R}^d)$. We prove that $F_a \in W^{1,p}\big(\mathcal{P}_p(\mathbb{R}^d), W_p, Q_{\pi,\nu}\big)$ and $|DF_a| \equiv 0$. To show that, consider a p'-test plan $\Lambda \in \mathcal{P}(C([0,1], \mathcal{P}_p(\mathbb{R}^d)))$ and notice that it falls in the hypothesis of Theorem 4.5, so that there exists $\mathfrak{L} \in \mathcal{P}(\mathcal{P}^{\mathrm{pa}}(AC^{p'}([0,1],\mathbb{R}^d)))$ such that $E_\sharp \mathfrak{L} = \Lambda$. In particular, (7.2) with $G \equiv 0$ is implied by the fact that for $\mathfrak{L}$-a.e. $\lambda \in \mathcal{P}^{\mathrm{pa}}(AC^{p'}([0,1],\mathbb{R}^d))$, there exist $\mathbf{a} \in \mathbf{T}$ and $\{\gamma_i\} \in AC^{p'}([0,1],\mathbb{R}^d)$ such that

$$\lambda = \sum_{i=1}^{+\infty} a_i \delta_{\gamma_i} \quad \text{and} \quad F_a((\mathrm{e}_t)_\sharp \lambda) = \sum_{i=1}^{+\infty} \mathbb{1}_{[a,1]}(a_i),$$

which is constant in time, and in particular is in $W^{1,1}(0,1)$, thus $G \equiv 0$ is a p-weak upper gradient, which must coincide with $|DF_a|$ by minimality.
Now, if (Sob-Lip) holds, then for all $a \in (0,1]$ there exists a constant $C_a \geq 0$ such that $F_a(\mu) = C_a$ for $Q_{\pi,\nu}$-a.e. $\mu \in \mathcal{P}_p(\mathbb{R}^d)$. We show that the hypothesis that π is not a Dirac measure prevents this. Indeed, define

$$\underline{i} := \min\left\{ j+1 \ :\ \nexists (s_1,\dots,s_j) \in [0,1]^j \text{ s.t. } \pi(\{s_1\}\times\cdots\times\{s_j\}\times[0,1]^\infty) = 1 \right\}.$$

Then, there exists $1 \geq a_1 \geq \cdots \geq a_{\underline{i}} \geq \underline{a} > 0$ such that

$$A^+ := \{\mathbf{a} \in \mathbf{T} \ :\ a_{\underline{i}+1} \geq \underline{a}\}, \quad 0 < \pi(A^+) < 1, \quad \pi(\{a_1\}\times\cdots\times\{a_{\underline{i}}\}\times[0,1]^\infty) = 1.$$

Define then $A^- := \mathbf{T} \setminus A^+$, so that

$$Q_{\pi,\nu}\big(\mathrm{em}(A^+ \times \mathbf{X})\big) > 0, \quad Q_{\pi,\nu}\big(\mathrm{em}(A^- \times \mathbf{X})\big) > 0.$$

At this point, consider $\mu^+ = \sum_j a_j^+ \delta_{x_j^+} \in A^+ \times \mathbf{X}$ and $\mu^- = \sum_j a_j^- \delta_{x_j^-} \in A^- \times \mathbf{X}$, for some $\mathbf{a}^\pm \in \mathbf{T}$ and $\mathbf{x}^\pm \in \mathbf{X}$. Notice that $a_j^- < \underline{a}$ for all $j \geq \underline{i}+1$. Then

$$F_{\underline{a}}(\mu^+) = \left(\sum_{j=1}^{\underline{i}} \mathbb{1}_{[\underline{a},1]}(a_j)\right) + \mathbb{1}_{[\underline{a},1]}(a^+_{\underline{i}+1}) + \left(\sum_{j \geq \underline{i}+2} \mathbb{1}_{[\underline{a},1]}(a_j^+)\right) \geq \underline{i}+1$$

$$F_{\underline{a}}(\mu^-) = \left(\sum_{j=1}^{\underline{i}} \mathbb{1}_{[\underline{a},1]}(a_j)\right) + \mathbb{1}_{[\underline{a},1]}(a^-_{\underline{i}+1}) + \left(\sum_{j \geq \underline{i}+2} \mathbb{1}_{[\underline{a},1]}(a_j^-)\right) = \underline{i}.$$

In particular $F_{\underline{a}}$ cannot admit a constant representative that coincides with it $Q_{\pi,\nu}$-a.e.. $\square$

Corollary 7.6. *Let $p \in (1,+\infty)$ and $r \in (1,+\infty)$ be such that $p \geq \frac{r}{r-1}$ and $\mathrm{cap}_{r,\nu}(\Delta) = 0$. Assume that $\pi \neq \delta_{\mathbf{a}}$ for all $\mathbf{a} \in \mathbf{T}$. Then, $W^{1,p}\big(\mathcal{P}_p(\mathbb{R}^d), W_p, Q_{\pi,\nu}\big)$ does not satisfy* (Sob-Lip).

Proof. If (Sob-Lip), then the function F from Theorem 7.5 would admit a representative $\tilde{F}$ satisfying $|\tilde{F}(\mu) - \tilde{F}(\nu)| \leq \varepsilon W_p(\mu,\nu)$ for all $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$. Letting $\varepsilon \to 0$, gives that $\tilde{F}$ is constant, giving a contradiction. $\square$

Corollary 7.7. *Let $p \in (1,+\infty)$ and $r \in (1,+\infty)$ be such that $p \geq \frac{r}{r-1}$ and $\mathrm{cap}_{r,\nu}(\Delta) = 0$. Assume that $\pi \neq \delta_{\mathbf{a}}$ for all $\mathbf{a} \in \mathbf{T}$. Then, there exists a function $F \in W^{1,p}(\mathcal{P}_p(\mathbb{R}^d), W_p, Q_{\pi,\nu})$ such that*

$$\int_{\mathcal{P}_p(\mathbb{R}^d)} \left| F(\mu) - \int_{\mathcal{P}_p(\mathbb{R}^d)} F dQ_{\pi,\nu} \right|^p dQ_{\pi,\nu}(\mu) > N \int_{\mathcal{P}_p(\mathbb{R}^d)} |DF|^p(\mu) dQ_{\pi,\nu} \quad \text{for all } N \in \mathbb{N}. \tag{7.7}$$

In particular, the metric measure space $(\mathcal{P}_p(\mathbb{R}^d), W_p, Q_{\pi,\nu})$ does not support a Poincaré inequality.

Proof. The function F from Theorem 7.5 is not constant, implying that the left hand side in (7.7) is strictly positve. On the other hand, $|DF| \equiv 0$ implies that the right hand side in (7.7) is always null. $\square$

7.1.2. *Riemannian case.*

Theorem 7.8. *Let $(\mathbb{S}, g_{\mathbb{S}})$ be a boundaryless compact Riemannian manifold of dimension $k \geq 2$, and $\nu \in \mathcal{P}(\mathbb{S})$ be a probability measure that has a bounded density with respect to $\mathrm{vol}_{\mathbb{S}}$. Assume that $\pi \neq \delta_{\mathbf{a}}$ for all $\mathbf{a} \in \mathbf{T}$.*
Then, there exists a function $F \in W^{1,p}(\mathcal{P}_p(\mathbb{S}), W_p, Q_{\pi,\nu})$ for all $p \in [\frac{k}{k-1}, +\infty)$, for which $|DF| \equiv 0$ and that does not admit a constant representative.

Proof. The strategy is the same as for the Euclidean case, with the difference that we will make use of the atomic nested superposition principles in Section 6.2. For all $a \in (0,1]$, consider the function $F_a(\mu) := \int_{\mathbb{S}} \mathbb{1}_{[a,1]}(\mu[y]) d\mu(y)$, for all $\mu \in \mathcal{P}(\mathbb{S})$. Consider a p'-test plan for the metric measure space $(\mathcal{P}(\mathbb{S}), W_p, Q_{\pi,\nu})$, that is $\Lambda \in \mathcal{P}(C([0,1], \mathcal{P}(\mathbb{S})))$, which falls in the hypothesis of Theorem 6.10, with $q = +\infty$. Then, there exists $\mathfrak{L} \in \mathcal{P}(\mathcal{P}^{\mathrm{pa}}(AC^{p'}([0,1],\mathbb{S})))$ such that $E_\sharp \mathfrak{L} = \Lambda$. In particular, for $\mathfrak{L}$-a.e. $\lambda \in \mathcal{P}^{\mathrm{pa}}(AC^{p'}([0,1],\mathbb{S}))$ there exist $\boldsymbol{a} \in \mathbf{T}$ and $\boldsymbol{y}_i \in AC^{p'}([0,1],\mathbb{S})$ such that

$$\lambda = \sum_{i=1}^{+\infty} a_i \delta_{\boldsymbol{y}_i} \quad \text{and} \quad F_a((e_t)_\sharp \lambda) = \sum_{i=1}^{+\infty} \mathbb{1}_{[a,1]}(a_i).$$

Since it does not depend on $t \in [0, T]$, this shows that $G \equiv 0$ is a p-upper gradient for the function F_a. Then, the proof can be concluded as in Theorem 7.5. $\square$

As for the Euclidean case, the previous theorem provides a counterexample for the p-Sobolev-to-Lipschitz and the Poincaé inequality.

Corollary 7.9. *Let $(\mathbb{S}, g_{\mathbb{S}})$ be a boundaryless compact Riemannian manifold of dimension $k \geq 2$, and $\nu \in \mathcal{P}(\mathbb{S})$ be a probability measure that has a bounded density with respect to* $\mathrm{vol}_{\mathbb{S}}$. *Assume that $\pi \neq \delta_{\mathbf{a}}$ for all $\mathbf{a} \in \mathbf{T}$.*
Then, the space $W^{1,p}(\mathcal{P}(\mathbb{S}), W_p, Q_{\pi,\nu})$ does not satisfy (Sob-Lip) *for all $p \in [\frac{k}{k-1}, +\infty)$.*

Corollary 7.10. *Let $(\mathbb{S}, g_{\mathbb{S}})$ be a boundaryless compact Riemannian manifold of dimension $k \geq 2$, and $\nu \in \mathcal{P}(\mathbb{S})$ be a probability measure that has a bounded density with respect to* $\mathrm{vol}_{\mathbb{S}}$. *Assume that $\pi \neq \delta_{\mathbf{a}}$ for all $\mathbf{a} \in \mathbf{T}$.*
Then, there exists a function $F \in W^{1,p}(\mathcal{P}_p(\mathbb{S}), W_p, Q_{\pi,\nu})$ such that for all $p \in [\frac{k}{k-1}, +\infty)$

$$\int_{\mathcal{P}_p(\mathbb{S})} \left| F(\mu) - \int_{\mathcal{P}_p(\mathbb{S})} F dQ_{\pi,\nu} \right|^p dQ_{\pi,\nu}(\mu) > N \int_{\mathcal{P}_p(\mathbb{S})} |DF|^p(\mu) dQ_{\pi,\nu}, \quad \textit{for all } N \in \mathbb{N}. \tag{7.8}$$

In particular, the metric measure space $(\mathcal{P}_p(\mathbb{S}), W_p, Q_{\pi,\nu})$ does not support a Poincaré inequality.

7.2. **Bakry–Émery curvature condition.** In this subsection, we establish a Bakry–Émery curvature dimension condition for the space $(\mathcal{P}(\mathbb{S}), W_2, Q_{\pi,\nu})$, under the following assumption.

Assumption 7.11. $(\mathbb{S}, g)$ *is a compact, oriented and boundaryless Riemannian manifold of dimensione $k \geq 2$, with associated volume measure* $\mathrm{vol}_{\mathtt{g}}$. *Moreover, $\nu = \rho \,\mathrm{vol}_{\mathtt{g}} \in \mathcal{P}(\mathbb{S})$ with $\rho = e^{-V}$ with $V \in C^\infty(\mathbb{S})$.*

The necessity of working with a compact Riemannian manifold comes from the characterization of the diffusion process, and consequently of its associated semigroup $(\mathtt{H}_t^{\pi,\nu})_{t \geq 0}$, induced by the Cheeger energy in $(\mathcal{P}(\mathbb{S}), W_2, Q_{\pi,\nu})$, that is one of the main result of [Del24, Theorem 3.24]. Thus, that hypothesis can be relaxed as soon as such a characterization is recovered: it is still open for non-compact underlying spaces (see [Del24, Example 3.40]).

Next, we introduce the main notation that we need to describe the aforementioned characterization.

7.2.1. *Rescaled and product Riemannian metric.* We will use the following notation:

(1) $\nabla^{\mathtt{g}}$ the gradient operator of $(\mathbb{S}, \mathtt{g})$, which is characterized by requiring that for all $f \in C^1(\mathbb{S})$, $y \in \mathbb{S}$ and $v \in T_y\mathbb{S}$, $d_y f(v) = \mathtt{g}_y(\nabla^g f(y), v)$;
(2) for all $f \in C^1(\mathbb{S})$ and $y \in \mathbb{S}$, denote $\Gamma^{\mathtt{g}} f(y) := \mathtt{g}_y(\nabla^{\mathtt{g}} f(y), \nabla^{\mathtt{g}} f(y))$. It is actually well defined for all $f \in W^{1,2}(\mathbb{S}, \mathtt{g})$ and it is called the *carré du champ* associated to its Dirichlet form;
(3) for all $y \in \mathbb{S}$, we can uniquely associate to the measure ν a diffusion process $(Y_t^{y,\nu})_{t \geq 0}$ with $Y_0^{y,\nu} = y$. Formally, it is a solution of the stochastic differential equation $dY_t^{y,\nu} = -\nabla V dt + \sqrt{2} dW_t^y$, where $(W_t^y)_{t \geq 0}$ is the Wiener process of the manifold exiting from y;
(4) the associated linear and contractive heat semigroup is denoted, for all $t \geq 0$, $\mathtt{h}_t^{\mathtt{g},\nu} : L^p(\nu) \to L^p(\nu)$, for $p \in [1, +\infty)$;
(5) the Ricci tensor of $(\mathbb{S}, \mathtt{g})$ is denoted as $\mathrm{Ric}^{\mathtt{g}}$. We say that it is bounded from below by $K \in \mathbb{R}$ if $\mathrm{Ric}^{\mathtt{g}} \geq K \mathtt{g}$;

(6) we say that the triple $(\mathbb{S}, d_{\mathtt{g}}, \nu)$ satisfies the *Bakry–Émery curvature condition* $\mathrm{BE}(K,\infty)$, with $K \in \mathbb{R}$, if

$$\Gamma^{\mathtt{g}}(\mathtt{h}_t^{\mathtt{g},\nu} f) \le e^{-2Kt}\mathtt{h}_t^{\mathtt{g},\nu}(\Gamma^{\mathtt{g}} f), \quad \forall f \in C^1(\mathbb{S}),\ \forall t \ge 0. \tag{7.9}$$

Recall that, in this smooth Riemannian setting, if $\nu = \mathrm{vol}_{\mathtt{g}}$, $\mathrm{BE}(K,\infty)$ is satisfied if and only if $\mathrm{Ric}^{\mathtt{g}} \ge K\mathtt{g}$.

For all $a > 0$, we define the rescaled metric $\mathtt{g}^a := a\mathtt{g}$, for which the objects introduced above rescale as (we omit the term $\mathtt{g}$ for the sake of notation)

$$\begin{gathered}\nabla^a = \frac{1}{a}\nabla, \quad \Gamma^a = \frac{1}{a}\Gamma, \quad \mathrm{vol}_a = a^{\frac{k}{2}}\,\mathrm{vol},\\ \mathtt{h}_t^a = \mathtt{h}_{t/a}, \quad Y_t^{y,\nu,a} = Y_{t/a}^{y,\nu}, \quad \mathrm{Ric}^a = \mathrm{Ric}\,.\end{gathered} \tag{7.10}$$

In particular, if $(\mathbb{S},\mathtt{g})$ has Ricci curvature bounded from below by K, then $(\mathbb{S},\mathtt{g}^a)$ has Ricci curvature bounded from below by K/a.

We proceed looking at the curvature properties of the product between rescaled spaces. Fix $\mathbf{a} = (a_1, a_2, \dots) \in \mathbf{T}_0$ and, for all $n \ge 1$, on the manifold $\mathbb{S}^n$ we consider the Riemannian metric $\mathtt{g}^{n,\mathbf{a}} := \mathtt{g}^{a_1} \oplus \cdots \oplus \mathtt{g}^{a_n}$, which raises the objects

$$\begin{gathered}\nabla^{n,\mathbf{a}} := \nabla^{\mathtt{g}^{n,\mathbf{a}}} = \nabla^{a_1} \oplus \cdots \oplus \nabla^{a_n}, \quad \Gamma^{n,\mathbf{a}} := \Gamma^{\mathtt{g}^{n,\mathbf{a}}} = \Gamma^{a_1} \oplus \cdots \oplus \Gamma^{a_n},\\ \mathrm{vol}_{n,\mathbf{a}} := \mathrm{vol}_{\mathtt{g}^{n,\mathbf{a}}} = \mathrm{vol}_{a_1} \otimes \cdots \otimes \mathrm{vol}_{a_n}, \quad Y_t^{\mathrm{y},\nu,n,\mathbf{a}} := \big(Y_{t/a_1}^{y_1,\nu}, \dots, Y_{t/a_n}^{y_n,\nu}\big),\\ \mathtt{H}^{n,\mathbf{a}} := \mathtt{h}^{\mathtt{g}^{n,\mathbf{a}}} = \big(\mathtt{h}_{t/a_1}, \dots, \mathtt{h}_{t/a_n}\big), \quad \mathrm{Ric}_{n,\mathbf{a}} := \mathrm{Ric}_{\mathtt{g}^{n,\mathbf{a}}} = \mathrm{Ric} \oplus \cdots \oplus \mathrm{Ric}\,.\end{gathered} \tag{7.11}$$

where $\mathrm{y} = (y_1, y_2, \dots) \in \mathbb{S}^\infty$. Notice that the diffusion process $Y_t^{\mathrm{y},\nu,n,\mathbf{a}}$ is well-posed (existence and uniqueness for any starting vector y) thanks to [Del24, Corollary 2.25].

Lemma 7.12. *Assume that $\mathrm{Ric}^{\mathtt{g}} \ge K\mathtt{g}$. Then, the Riemannian manifold $(\mathbb{S}^{n,\mathbf{a}}, \mathtt{g}^{n,\mathbf{a}})$ has Ricci curvature bounded from below by*

$$K^{n,\mathbf{a}} := \begin{cases} \frac{K}{a_1} & \text{if } K \ge 0;\\ \frac{K}{a_n} & \text{if } K < 0.\end{cases}$$

Moreover, if $\nu = e^{-V}\,\mathrm{vol}^{\mathtt{g}} \in \mathcal{P}(\mathbb{S})$ is as in Assumption 7.11 and V is λ-convex, then the metric measure space $(\mathbb{S}^n, \mathtt{g}^{n,\mathbf{a}}, \nu^n)$ satisfies $\mathrm{BE}(K^{n,\mathbf{a},\lambda},\infty)$, with

$$K^{n,\mathbf{a},\lambda} := \begin{cases} \frac{K+\lambda}{a_1} & \text{if } K+\lambda \ge 0;\\ \frac{K+\lambda}{a_n} & \text{if } K+\lambda < 0.\end{cases}$$

Proof. For all $y_1, \dots, y_n \in \mathbb{S}$ and $v_i \in T_{y_i}\mathbb{S}$, it holds

$$\mathrm{Ric}(v_1,v_1) + \cdots + \mathrm{Ric}(v_n,v_n) \ge \frac{K}{a_1}\mathtt{g}^{a_1}(v_1,v_1) + \cdots + \frac{K}{a_n}\mathtt{g}^{a_n}(v_n,v_n), \tag{7.12}$$

which concludes the first statement, thanks to (7.11). Regarding the second statement, notice that

$$\nu^n = e^{-V\oplus\cdots\oplus V}\,\mathrm{vol}_{\mathtt{g}}^n = e^{-\left(V\oplus\cdots\oplus V + \frac{k}{2}\sum_{i=1}^n \log(a_i)\right)}\,\mathrm{vol}_{n,\mathbf{a}}\,.$$

On the other hand, recall that V is λ-convex if and only if the Hessian tensor $\mathrm{Hess}_{\mathtt{g}} V \ge \lambda\mathtt{g}$. The Hessian is invariant with respect to rescaling, so V λ-convex implies that for all $i \le n$, $\mathrm{Hess}_{\mathtt{g}^{a_i}} V = \mathrm{Hess}_{\mathtt{g}} V \ge \frac{\lambda}{a_i}\mathtt{g}^{a_i}$. Putting everything together, we conclude thanks to the first part of the statement and [Stu06, Theorem 4.9], since the $\mathrm{CD}(K,\infty)$ condition introduced there implies the $\mathrm{BE}(K,\infty)$ for infinitesimally Hilbertian metric measure spaces, see [Sav14]. □

7.2.2. *Characterization for the heat semigroup of* $(\mathcal{P}(\mathbb{S}), W_2, Q_{\pi,\nu})$. Under Assumption 7.11, further assume that $\pi \in \mathcal{P}(\mathbf{T}_0)$, and define $Q_{\pi,\nu}$ as usual (see (2.3)). Thanks to [Del24, Theorem 3.39], the quadratic pre-Dirichlet form $\big(\widehat{\mathcal{E}}, \mathrm{GC}_c^\infty(\mathcal{P}(\mathbb{S}))\big)$ defined as

$$\widehat{\mathcal{E}}(\hat{F}) := \int_{\mathcal{P}(\mathbb{S})} \int_{\mathbb{S}} |\nabla_W \hat{F}(y,\mu)|^2_{\mathsf{g}} d\mu(y) dQ_{\pi,\nu}(\mu) \quad \forall \hat{F} \in \mathrm{GC}_c^\infty, \tag{7.13}$$

is closable. Moreover, by [Del24, Proposition 3.31] and [FSS23, Theorem 6.2], it coincides with the Cheeger energy, in the sense that $\mathcal{D}(\widehat{\mathcal{E}}) = W^{1,2}(\mathcal{P}(\mathbb{S}), W_2, Q_{\pi,\nu})$ and $\widehat{\mathcal{E}}(F) = \mathrm{Ch}(F)$ for all $F \in W^{1,2}(\mathcal{P}(\mathbb{S}), W_2, Q_{\pi,\nu})$. In particular, for all $\widehat{F} \in \mathrm{GC}_c^1(\mathcal{P}(\mathbb{S}))$, it holds

$$\int_{\mathbb{S}} |\nabla_W \hat{F}(y,\mu)|^2_{\mathsf{g}} d\mu(y) = |D\widehat{F}|^2(\mu) \quad \text{ for } Q_{\pi,\nu}\text{-a.e. } \mu. \tag{7.14}$$

Note that the equivalence with the Cheeger energy can also be recovered applying [FSS23, Theorem 6.2], Proposition 6.8 and Proposition 4.2, after using the Nash isometry to embed $\mathbb{S}$ in some $\mathbb{R}^d$.

The equivalence between the Cheeger energy and the Dirichlet form $\widehat{\mathcal{E}}$ allows us to exploit [Del24, Theorems 3.15 & 3.24], which in particular gives the following characterization of the associated diffusion process: again by [Del24, Corollary 2.25], for all $\mathbf{a} = (a_1, a_2, \dots) \in \mathbf{T}_0$ and $\mathrm{y} = (y_1, y_2, \dots) \in \mathbb{S}^\infty_{\neq}$, we can find a countable number of independent processes $(Y_t^{y_i,\nu})_{t\geq 0}$, all defined on a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, for all $\mu \in \mathcal{P}^{\mathrm{pa}}(\mathbb{S})$, $\mu = \mathrm{em}(\mathbf{a}, \mathrm{y})$, the diffusion process associated to $(\mathcal{P}(\mathbb{S}), W_2, Q_{\pi,\nu})$ exiting from μ is given by

$$\mu_t := \mathrm{em}\left(\mathbf{a}, \mathrm{Y}_t^{\mathbf{a},\mathrm{y},\nu}\right) = \sum_{i=1}^{+\infty} a_i \delta_{Y^{y_i,\nu}_{t/a_i}}, \quad \mathrm{Y}_t^{\mathbf{a},\mathrm{y},\nu} := \left(Y^{y_i,\nu}_{t/a_i}\right)_{i\geq 1}. \tag{7.15}$$

In other words, the masses are kept fixed, while the positions move according to the underlying geometry, and the lighter the particles are (smaller weight a_i) the faster they move. We can use this to gain information on the heat semigroup $(H_t^{\pi,\nu})_{t\geq 0}$: for our purposes, it suffices to see how it acts on generalized cylinder functions. First, we introduce the subsets of generalized cylinder function given $a \in (0,1)$

$$\mathrm{GC}_a^\infty(\mathcal{P}(\mathbb{S})) := \left\{ \Psi \circ \hat{L}_{\hat{\Phi}} \ : \ k \geq 1, \ \Psi \in C^\infty(\mathbb{R}^k), \ \hat{\Phi} \in C_c^\infty(\mathbb{S} \times (a,1]; \mathbb{R}^k) \right\}.$$

Notice that $\mathrm{GC}_c^\infty(\mathcal{P}(\mathbb{S})) = \cup_{n\geq 2} \mathrm{GC}_{1/n}^\infty(\mathcal{P}(\mathbb{S}))$. Let $n \geq 2$, $\hat{F} = \Psi(\hat{L}_{\hat{\phi}_1}, \dots, \hat{L}_{\hat{\phi}_k}) \in \mathrm{GC}_{1/n}^\infty(\mathcal{P}(\mathbb{S}))$, and $\mu = \mathrm{em}(\mathbf{a}, \mathrm{y})$, with $\mathbf{a} \in \mathbf{T}_0$ and $\mathrm{y} \in \mathbb{S}^\infty_{\neq}$. Then define

$$\hat{F}^{n,\mathbf{a}}(y_1, \dots, y_n) = \hat{F}\big(\mathrm{em}\left(\mathbf{a}, \mathrm{y}\right)\big),$$

which is a function that truly depends just on the first n-coordinates of y (possibly even less coordinates may contribute) whatever $\mathbf{a} \in \mathbf{T}_0$, and it is smooth on $\mathbb{S}^n$, so that

$$\begin{aligned} \mathtt{H}_t^{\pi,\nu} \hat{F}(\mu) =& \mathbb{E}\left[\hat{F}\big(\mathrm{em}\left(\mathbf{a}, \mathrm{Y}_t^{\mathbf{a},\mathrm{y},\nu}\right)\big)\right] = \mathbb{E}\left[\hat{F}^{n,\mathbf{a}}\big(Y^{y_1,\nu}_{t/a_1}, \dots, Y^{y_n,\nu}_{t/a_n}\big)\right] \\ =& \mathtt{H}_t^{n,\mathbf{a},\nu} \hat{F}^{n,\mathbf{a}}(y_1, \dots, y_n). \end{aligned} \tag{7.16}$$

Moreover,

$$
\begin{aligned}
|D\hat{F}|^2(\mu) &= \int_{\mathbb{S}} |\nabla_W \hat{F}(y,\mu)|^2_{\mathtt{g}} d\mu(y) = \sum_{i=1}^n a_i |\nabla_W \hat{F}(y_i,\mu)|^2_{\mathtt{g}} \\
&= \sum_{i=1}^n a_i \left| \frac{1}{a_i} \nabla^{\mathtt{g}}_{y_i} \hat{F}^{n,\mathbf{a}}(y_1,\dots,y_n) \right|^2_{\mathtt{g}} = \sum_{i=1}^n \frac{1}{a_i} \Gamma^{\mathtt{g}}_{y_i} \hat{F}^{n,\mathbf{a}}(y_1,\dots,y_n) \\
&= \sum_{i=1}^n \Gamma^{a_i}_{y_i} \hat{F}^{n,\mathbf{a}}(y_1,\dots,y_n) = \Gamma^{n,\mathbf{a}} \hat{F}^{n,\mathbf{a}}(y_1,\dots,y_n).
\end{aligned}
\tag{7.17}
$$

More generally, the intertwining property shown in [Del24, Theorem 3.15], gives that

$$
|D(\mathtt{H}^{\pi,\nu}_t \hat{F})|^2(\mu) = \Gamma^{n,\mathbf{a}}\big(\mathtt{H}^{n,\mathbf{a},\nu}_t \hat{F}^{n,\mathbf{a}}\big)(y_1,\dots,y_n) \tag{7.18}
$$

We can finally state and prove the Bakry–Émery condition for the metric measure space $(\mathcal{P}(\mathbb{S}), W_2, Q_{\pi,\nu})$.

Theorem 7.13. *Let $(\mathbb{S},\mathtt{g})$ be a compact, oriented and boundaryless Riemannian manifold of dimension $k \geq 2$ and with Ricci curvature bounded from below by $K \in \mathbb{R}$. Let $\pi \in \mathcal{P}(\mathbf{T}_0)$ and $\nu = e^{-V}\,\mathrm{vol}_{\mathtt{g}} \in \mathcal{P}(\mathbb{S})$ as in Assumption 7.11.*
Assume that V is λ-convex, with $K+\lambda \geq 0$. Then $(\mathcal{P}(\mathbb{S}), W_2, Q_{\pi,\nu})$ satisfies $\mathrm{BE}(K+\lambda,\infty)$.

Proof. Thanks to Lemma 7.4, it suffices to prove the Bakry–Émery condition for all $\hat{F} \in \mathrm{GC}^\infty_c(\mathcal{P}(\mathbb{S}))$. Indeed, the algebra of classic cylinder functions $\mathrm{Cyl}^\infty(\mathcal{P}(\mathbb{S}))$ is dense in energy thanks to [FSS23, Section 6.2] (see also [Sod23, Corollary 4.20]), and for all cylinder function $F(\cdot) = \Psi(L_{\phi_1}(\cdot),\dots,L_{\phi_n}(\cdot))$, it is approximated in the norm of $W^{1,2}(\mathcal{P}(\mathbb{S}), W_2, Q_{\pi,\nu})$ by the generalized cylinder functions $\hat{F}_n \in \mathrm{GC}^\infty_c(\mathcal{P}(\mathbb{S}))$ defined as

$$
\hat{F}_n(\mu) := \Psi\left(\hat{L}_{\hat{\phi}_1}(\mu),\dots,\hat{L}_{\hat{\phi}_n}(\mu)\right), \quad \hat{\phi}_i(y,r) := \phi_i(y)\eta_n(r),
$$

where $\eta_n \in C^\infty_c((0,1])$ is such that $\eta_n \equiv 1$ in $[1/n,1]$. Then, fix $n \geq 2$ and $\hat{F} \in \mathrm{GC}^\infty_{1/n}(\mathcal{P}(\mathbb{S}))$. Thanks to (7.18), applying Lemma 7.12, we have

$$
|D(\mathtt{H}^{\pi,\nu}_t \hat{F})|^2(\mu) = \Gamma^{n,\mathbf{a}}\big(\mathtt{H}^{n,\mathbf{a},\nu}_t \hat{F}^{n,\mathbf{a}}\big)(y_1,\dots,y_n) \leq e^{-2\frac{(K+\lambda)}{a_1}t} \mathtt{H}^{n,\mathbf{a},\nu}\big(\Gamma^{n,\mathbf{a}}\hat{F}^{n,\mathbf{a}}\big)(y_1,\dots,y_n).
$$

Now, thanks to (7.17), $\Gamma^{n,\mathbf{a}}\hat{F}^{n,\mathbf{a}}(y_1,\dots,y_n) = |D\hat{F}|^2(\mu)$, and the latter is again a generalized cylinder function in $\mathrm{GC}^\infty_c(\mathcal{P}(\mathbb{S}))$ satisfying $\Gamma^{n,\mathbf{a}}\hat{F}^{n,\mathbf{a}}(y_1,\dots,y_n) = |D\hat{F}|^2\big(\mathrm{em}(\mathbf{a},\mathrm{y})\big)$, so that (7.16) gives

$$
\mathtt{H}^{n,\mathbf{a},\nu}\big(\Gamma^{n,\mathbf{a}}\hat{F}^{n,\mathbf{a}}\big)(y_1,\dots,y_n) = \mathtt{H}^{\pi,\nu}|D\hat{F}|^2(\mu).
$$

We conclude observing that $\frac{(K+\lambda)}{a_1} \geq K+\lambda$ for all $\mathbf{a} = (a_1,a_2,\dots) \in \mathbf{T}_0$. □

Appendix A. On the definition of capacity

Lemma A.1. *Assume that $\nu \ll \mathcal{L}^d$. Then for all $p \geq 1$*

$$
\mathrm{cap}_{p,\nu}(\Delta) = \widetilde{\mathrm{cap}}_{p,\nu}(\Delta).
$$

Proof. Clearly, $\mathrm{cap}_{p,\nu}(\Delta) \leq \widetilde{\mathrm{cap}}_{p,\nu}(\Delta)$, so we focus on the other inequality. Let us call $\mathcal{C}(h) := \int |h^p| + |\nabla h|^p d\nu\otimes\nu$. We are going to prove that for any $h \in C^1_b(\mathbb{R}^{2d})$ satisfying $h \geq 1$ on Δ and for any $\delta > 0$, there exists a function $g \in C^1_b(\mathbb{R}^{2d})$ such that $g \leq 1$ everywhere, $g = 1$ on Δ and

$$
\mathcal{C}(g) \leq \mathcal{C}(h) + \delta. \tag{A.1}
$$

Step 1: define $\tilde{h}(x,y) := h(x,y) \wedge 1$. Notice that $\tilde{h}$ is a Lipschitz function, and satisfies $\tilde{h} \leq 1$, $\tilde{h} = 1$ on Δ and $\mathcal{C}(\tilde{h}) \leq \mathcal{C}(h)$, thanks to the locality of the gradient for Lipschitz functions, i.e.

$$\nabla \tilde{h}(x,y) = \nabla h(x,y) \mathbb{1}_{\{h \leq 1\}}(x,y) \quad \mathcal{L}^{2d}\text{-a.e.,} \tag{A.2}$$

and the fact that $\nu \ll \mathcal{L}^d$.

Step 2: we claim that there exists a Lipschitz function $\overline{h}$ such that $\overline{h} \leq 1$, $\mathcal{C}(\overline{h}) \leq \mathcal{C}(\tilde{h}) + \frac{\delta}{2}$ and $\overline{h} = 1$ on $\Delta_\varepsilon := Delta + B(0,\varepsilon)$, for some $\varepsilon > 0$, i.e. there exists a 'strip' of length ε around the diagonal Δ where the function $\overline{h}$ is constantly 1.
To this aim, define π_Δ and $\pi_{\Delta^\perp}$, respectively, the orthogonal projection over Δ and $\Delta^\perp$, i.e.

$$\pi_\Delta(x,y) = \left(\frac{x+y}{2}, \frac{x+y}{2}\right), \quad \pi_{\Delta^\perp}(x,y) = \left(\frac{x-y}{2}, -\frac{x-y}{2}\right),$$
$$(x,y) = \pi_\Delta(x,y) + \pi_{\Delta^\perp}(x,y), \quad \|(x,y)\|^2 = \|\pi_\Delta(x,y)\|^2 + \|\pi_{\Delta^\perp}(x,y)\|^2.$$

For any $\varepsilon > 0$, define $p_\varepsilon : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ as follows

$$p_\varepsilon(x,y) := \begin{cases} \pi_\Delta(x,y) & \text{if } (x,y) \in \Delta_\varepsilon \\ (x,y) - \varepsilon \frac{\pi_{\Delta^\perp}(x,y)}{\|\pi_{\Delta^\perp}(x,y)\|} & \text{otherwise.} \end{cases} \tag{A.3}$$

In the next step, we will prove that p_ε is 1-Lipschiz for any $\varepsilon > 0$. For now, assume that is true and we continue defining $h_\varepsilon(x,y) := \tilde{h}(p_\varepsilon(x,y))$. We conclude the proof of this step if

$$\limsup_{\varepsilon \searrow 0} \mathcal{C}(h_\varepsilon) \leq \mathcal{C}(\tilde{h}).$$

To this aim, notice that $h_\varepsilon \to \tilde{h}$ pointwise as $\varepsilon \searrow 0$, so that

$$\int |h_\varepsilon|^p d\nu \otimes \nu \to \int |\tilde{h}|^p d\nu \otimes \nu$$

by dominated convergence theorem. Regarding the gradients, notice that $|\nabla \tilde{h}(x,y)| \leq |\nabla h(x,y)| \mathbb{1}_{\{h \leq 1\}}(x,y)$ almost everywhere, and the right hand side is an upper semicontinuous function. Then, thanks to the straightforward fact that $(p_\varepsilon)_\#(\nu \otimes \nu) \to \nu \otimes \nu$ in the narrow convergence and

$$|\nabla h_\varepsilon(x,y)| \leq |\nabla \tilde{h}(p_\varepsilon(x,y))| = |\nabla h(p_\varepsilon(x,y))| \mathbb{1}_{\{h \leq 1\}}(p_\varepsilon(x,y)) \quad \mathcal{L}^{2d}\text{-a.e.,}$$

we have that

$$\limsup_{\varepsilon \searrow 0} \int |\nabla h_\varepsilon|^p(x,y) d\nu \otimes \nu(x,y) \leq \limsup \int |\nabla h(x,y)|^p \mathbb{1}_{\{h \leq 1\}}(x,y) d\big((p_\varepsilon)_\#(\nu \otimes \nu)\big)(x,y)$$
$$\leq \int \nabla |h(x,y)|^p \mathbb{1}_{\{h \leq 1\}}(x,y) d\nu \otimes \nu(x,y) = \int |\nabla \tilde{h}(x,y)|^p d\nu \otimes \nu(x,y).$$

Step 3: proof that p_ε is 1-Lipschitz for all $\varepsilon > 0$. We need to verify thay $\|p_\varepsilon(x,y) - p_\varepsilon(z,w)\| \leq \|(x,y) - (z,w)\|$ for all $(x,y), (z,w) \in \mathbb{R}^d \times \mathbb{R}^d$, where $\|\cdot\|$ is the Euclidean norm of $\mathbb{R}^{2d}$. There are three cases:

- $(x,y), (z,w) \in \Delta_\varepsilon$, then it is straightforward because any orthogonal projection is 1-Lipschitz;
- $(x,y) \in \Delta_\varepsilon$ and $(z,w) \notin \Delta_\varepsilon$, then notice that $\|\pi_{\Delta^\perp}(x,y)\| < \varepsilon$ and $\|\pi_{\Delta^\perp}(z,w)\| \geq \varepsilon$. Then

$$\begin{aligned}\|p_\varepsilon(x,y) - p_\varepsilon(z,w)\|^2 =& \|\pi_\Delta(x,y) - \pi_\Delta(z,w) - \left(1 - \frac{\varepsilon}{\|\pi_{\Delta^\perp}(z,w)\|}\right) \pi_{\Delta^\perp}(z,w)\|^2 \\ =& \|\pi_\Delta(x,y) - \pi_\Delta(z,w)\|^2 + (\|\pi_{\Delta^\perp}(z,w)\| - \varepsilon)^2\end{aligned}$$

$$
\begin{aligned}
(\star) \leq & \|\pi_\Delta(x,y) - \pi_\Delta(z,w)\|^2 + \|\pi_{\Delta^\perp}(x,y) - \pi_{\Delta^\perp}(z,w)\|^2 \\
= & \|(x,y)-(z,w)\|^2,
\end{aligned}
$$

where in $(\star)$ we used that $\|\pi_{\Delta^\perp}(z,w)\| - \varepsilon \leq \|\pi_{\Delta^\perp}(x,y)\| - \|\pi_{\Delta^\perp}(z,w)\| \leq \|\pi_{\Delta^\perp}(x,y) - \pi_{\Delta^\perp}(z,w)\|$;

- $(x,y),(z,w) \notin \Delta_\varepsilon$. Then

$$
\begin{aligned}
\|p_\varepsilon(x,y) - p_\varepsilon(z,w)\|^2 = & \|\pi_\Delta(x,y) - \pi_\Delta(z,w)\|^2 \\
& + \left\| \left(1 - \frac{\varepsilon}{\|\pi_{\Delta^\perp}(x,y)\|}\right)\pi_{\Delta^\perp}(x,y) - \left(1 - \frac{\varepsilon}{\|\pi_{\Delta^\perp}(z,w)\|}\right)\pi_{\Delta^\perp}(z,w)\right\|^2 .
\end{aligned}
\tag{A.4}
$$

For simplicity, call $v_1 = \pi_{\Delta^\perp}(x,y)$ and $v_2 = \pi_{\Delta^\perp}(z,w)$, together with $\alpha_i := 1 - \varepsilon/\|v_i\|$, for $i = 1,2$. Then

$$
\|\alpha_1 v_1 - \alpha_2 v_2\| = \|(\alpha_1 - \alpha_2)v_1 + \alpha_2(v_1 - v_2)\| \leq |\alpha_1 - \alpha_2|\|v_1\| + \alpha_2\|v_1 - v_2\|
$$

and

$$
|\alpha_1 - \alpha_2|\|v_1\| = \varepsilon \left| \frac{1}{\|v_2\|} - \frac{1}{\|v_1\|} \right| \|v_1\| \leq \varepsilon \frac{\|v_1 - v_2\|}{\|v_2\|},
$$

so that

$$
\|\alpha_1 v_1 - \alpha_2 v_2\| \leq \frac{\varepsilon}{\|v_2\|}\|v_1 - v_2\| + \|v_1 - v_2\| - \frac{\varepsilon}{\|v_2\|}\|v_1 - v_2\| = \|v_1 - v_2\|.
$$

Substituting in (A.4) we conclude that p_ε is 1-Lipschitz.

Step 4: convolution and conclusion. Consider $\overline{h}$ as in step 2. Then, for any $0 < \alpha < \varepsilon$, define the functions $g_\alpha := \overline{h} * \rho_\alpha$, where $\rho_\alpha(x,y) := \frac{1}{\alpha^{2d}}\rho\left(\frac{(x,y)}{\alpha}\right)$ and ρ is a mollifier with support in $B(0,1)$. Notice that each $g_\alpha \in C_b^1(\mathbb{R}^{2d})$ satisfies $h \leq 1$ and $h = 1$ on Δ. We conclude observing that, up to selecting a suitable subsequence $\alpha_n \to 0$, $|g_\alpha|^p + |\nabla g_\alpha|^p \to |\overline{h}|^p + |\nabla \overline{h}|^p$ pointwise almost everywhere, and then by dominated convergence theorem we have that

$$
\lim_{\alpha \to 0} \mathcal{C}(g_\alpha) \to \mathcal{C}(\overline{h}).
$$

□

References

[AGS08] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Second. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008, pp. x+334. ISBN: 978-3-7643-8721-1.

[AGS15] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. "Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds". In: *Ann. Probab.* 43.1 (2015), pp. 339–404. ISSN: 0091-1798. DOI: 10.1214/14-AOP907.

[AILP24] Luigi Ambrosio, Toni Ikonen, Danka Lučić, and Enrico Pasqualetto. *Metric Sobolev spaces I: equivalence of definitions.* cvgmt preprint. 2024.

[BGL14] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators.* Vol. 348. Grundlehren der mathematischen Wissenschaften, Springer, 2014. DOI: doi:10.1007/978-3-319-00227-9.

[Bog07] V. I. Bogachev. *Measure theory. Vol. I, II.* Springer-Verlag, Berlin, 2007, Vol. I: xviii+500 pp., Vol. II: xiv+575. ISBN: 978-3-540-34513-8. DOI: 10.1007/978-3-540-34514-5.

[CL24] Djalil Chafaıˆ and Joseph Lehec. "Logarithmic sobolev inequalities essentials". In: *https://djalil.chafai.net/docs/M2/chafai-lehec-m2-lsie-lecture-notes.pdf* (2024). Lecture notes.

[CLOS22] Giulia Cavagnari, Stefano Lisini, Carlo Orrieri, and Giuseppe Savaré. "Lagrangian, Eulerian and Kantorovich formulations of multi-agent optimal control problems: equivalence and gamma-convergence". In: *Journal of Differential Equations* 322 (2022), pp. 268–364.

[CSS23a] Giulia Cavagnari, Giuseppe Savaré, and Giacomo Enrico Sodini. "A Lagrangian approach to totally dissipative evolutions in Wasserstein spaces". In: (2023). DOI: `10.48550/arXiv.2305.05211`. arXiv: `2305.05211 [math.FA]`.

[CSS23b] Giulia Cavagnari, Giuseppe Savaré, and Giacomo Enrico Sodini. "Dissipative probability vector fields and generation of evolution semigroups in Wasserstein spaces". In: *Probability Theory and Related Fields* 185.3 (2023), pp. 1087–1182.

[CSS25] Giulia Cavagnari, Giuseppe Savaré, and Giacomo Enrico Sodini. "Stochastic Euler Schemes and Dissipative Evolutions in the Space of Probability Measures". In: *arxiv preprint, https://arxiv.org/abs/2505.20801* (2025).

[Del22] Lorenzo Dello Schiavo. "The Dirichlet–Ferguson diffusion on the space of probability measures over a closed Riemannian manifold". In: *The Annals of Probability* 50.2 (2022), pp. 591–648.

[Del24] Lorenzo Dello Schiavo. "Massive particle systems, Wasserstein Brownian motion and the Dean-Kawasaki equation". In: *arxiv preprint, https://arxiv.org/abs/2411.14936* (2024).

[DGT04] Camillo De Lellis, Carlo Romano Grisanti, and Paolo Tilli. "Regular selections for multiple-valued functions". In: *Mat. Pura Appl. IV* (2004), pp. 79–95. DOI: `https://doi.org/10.1007/s10231-003-0081-5`.

[DHS23] Lorenzo Dello Schiavo, Ronan Herry, and Kohei Suzuki. "Wasserstein geometry and Ricci curvature bounds for Poisson spaces". In: *arXiv preprint arXiv:2303.00398* (2023).

[DS21] Lorenzo Dello Schiavo and Kohei Suzuki. "Configuration spaces over singular spaces–I. Dirichlet-Form and Metric Measure Geometry". In: *arXiv preprint arXiv:2109.03192* (2021).

[DS22] Lorenzo Dello Schiavo and Kohei Suzuki. "Configuration Spaces over Singular Spaces–II. Curvature". In: *arXiv preprint arXiv:2205.01379* (2022).

[DS25] Lorenzo Dello Schiavo and Giacomo Enrico Sodini. "The Hellinger-Kantorovich metric measure geometry on spaces of measures". In: *arXiv preprint arXiv:2503.07802* (2025).

[EK94] SN Ethier and Thomas G Kurtz. "Convergence to Fleming-Viot processes in the weak atomic topology". In: *Stochastic Processes and their Applications* 54.1 (1994), pp. 1–27.

[Fer73] Thomas S Ferguson. "A Bayesian analysis of some nonparametric problems". In: *The Annals of Statistics* (1973), pp. 209–230.

[FSS23] Massimo Fornasier, Giuseppe Savaré, and Giacomo Enrico Sodini. "Density of subalgebras of Lipschitz functions in metric Sobolev spaces and applications to Wasserstein Sobolev spaces". In: *Journal of Functional Analysis* 285.11 (2023), p. 110153.

[Gig13] Nicola Gigli. "The splitting theorem in non-smooth context". In: *arXiv preprint arXiv:1302.5555* (2013).

[Gig18] Nicola Gigli. *Nonsmooth differential geometry–An approach tailored for spaces with Ricci curvature bounded from below*. Vol. 251. 1196. American Mathematical Society, 2018.

[GP20] Nicola Gigli and Enrico Pasqualetto. *Lectures on nonsmooth differential geometry*. Vol. 2. Springer, 2020. ISBN: 978-3-030-38612-2. DOI: `https://doi.org/10.1007/978-3-030-38613-9`.

[Gra74] Alfred Gray. "The volume of a small geodesic ball of a Riemannian manifold". In: *Michigan Math. J.* (1974), pp. 329–344. DOI: `10.1307/mmj/1029001150`.

[KR19] Vitalii Konarovskyi and Max-K von Renesse. "Modified massive Arratia flow and Wasserstein diffusion". In: *Communications on pure and applied mathematics* 72.4 (2019), pp. 764–800.

[Lis07] Stefano Lisini. "Characterization of absolutely continuous curves in Wasserstein spaces". In: *Calculus of variations and partial differential equations* 28.1 (2007), pp. 85–120.

[Mar18] Victor Marx. "A new approach for the construction of a Wasserstein diffusion". In: (2018).

[Nas54] John Nash. "C^1 isometric imbeddings". In: *Annals of mathematics* 60.3 (1954), pp. 383–396.

[Pin25] Alessandro Pinzi. "First order equation on random measures as superposition of weak solutions to the McKean-Vlasov equation". In: *arXiv, arxiv.org/abs/2510.07542* (2025).

[PS25] Alessandro Pinzi and Giuseppe Savarè. "Nested superposition principle for random measures and the geometry of the Wasserstein on Wasserstein space". In: *arXiv* (2025).

[RS09] Max-K von Renesse and Karl-Theodor Sturm. "Entropic measure and Wasserstein diffusion". In: *The Annals of Probability* 37.3 (2009), pp. 1114–1191. DOI: `10.1214/08-AOP430`.

[RW24] Panpan Ren and Feng-Yu Wang. "Ornstein- Uhlenbeck type processes on Wasserstein spaces". In: *Stochastic Processes and their Applications* 172 (2024), p. 104339.

[RWW24] Panpan Ren, Feng-Yu Wang, and Simon Wittmann. "Diffusion Processes on p-Wasserstein Space over Banach Space". In: *arXiv preprint arXiv:2402.15130* (2024).

[Sav14] Giuseppe Savaré. "Self-improvement of the Bakry-Émery condition and Wasserstein contraction of the heat flow in $\mathrm{RCD}(K,\infty)$ metric measure spaces". In: *Discrete Contin. Dyn. Syst.* 34.4 (2014), pp. 1641–1661. ISSN: 1078-0947. DOI: `10.3934/dcds.2014.34.1641`.

[Sch73] L Schwartz. "Radon measures on arbitrary topological spaces and cylindrical measures". In: *Oxford Univ.* (1973).

[Set94] Jayaram Sethuraman. "A constructive definition of Dirichlet priors". In: *Statistica sinica* (1994), pp. 639–650.

[Sod23] Giacomo Enrico Sodini. "The general class of Wasserstein Sobolev spaces: density of cylinder functions, reflexivity, uniform convexity and Clarkson's inequalities". In: *Calculus of Variations and Partial Differential Equations* 62.7 (2023), p. 212.

[ST17] Eugene Stepanov and Dario Trevisan. "Three superposition principles: currents, continuity equations and curves of measures". In: *J. Funct. Anal.* 272.3 (2017), pp. 1044–1103. ISSN: 0022-1236. DOI: `10.1016/j.jfa.2016.10.025`.

[Stu06] Karl-Theodor Sturm. "On the geometry of metric measure spaces I". In: *Acta Math.* 196.1 (2006), pp. 65–131. DOI: `https://doi.org/10.1007/s11511-006-0002-8`.

[Stu25] Karl-Theodor Sturm. “Wasserstein Diffusion on Multidimensional Spaces”. In: *Annals of Probability, to appear* (2025+).

(Alessandro Pinzi) Università Commerciale Luigi Bocconi, Dipartimento di Scienze delle Decisioni, via Roentgen 1, 20136 Milano, Italy
e-mail address: alessandro.pinzi@phd.unibocconi.it
Orcid: https://orcid.org/0009-0007-9146-5434