



**HAL**  
open science

# Recovering the homology of immersed manifolds

Raphaël Tinarrage

► **To cite this version:**

| Raphaël Tinarrage. Recovering the homology of immersed manifolds. 2019. hal-02396261v2

**HAL Id: hal-02396261**

**<https://hal.science/hal-02396261v2>**

Preprint submitted on 9 Dec 2019 (v2), last revised 2 Apr 2024 (v4)

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

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

# RECOVERING THE HOMOLOGY OF IMMERSED MANIFOLDS

Raphaël TINARRAGE  
Datashape, Inria Paris-Saclay

**Abstract.** Given a sample of an abstract manifold immersed in some Euclidean space, we describe a way to recover the singular homology of the original manifold. It consists in estimating its tangent bundle—seen as subset of another Euclidean space—in a measure theoretic point of view, and in applying measure-based filtrations for persistent homology. The construction we propose is consistent and stable, and does not involve the knowledge of the dimension of the manifold.

**Acknowledgements.** The author want to thank Frédéric Chazal, Marc Glisse and Théo Lacombe for fruitful discussions and corrections.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Statement of the problem . . . . .	2
1.2	Background on persistent homology . . . . .	4
1.3	Background on persistent homology for measures . . . . .	5
1.4	Notations and hypotheses . . . . .	6
<b>2</b>	<b>Reach of an immersed manifold</b>	<b>7</b>
2.1	Background on reach . . . . .	7
2.2	Geodesic bounds under curvature conditions . . . . .	8
2.3	Normal reach . . . . .	12
2.4	Probabilistic bounds under normal reach conditions . . . . .	15
<b>3</b>	<b>Tangent space estimation</b>	<b>21</b>
3.1	Local covariance matrix and $\check{\nu}$ . . . . .	22
3.2	Consistency of the estimation . . . . .	23
3.3	Stability of the estimation . . . . .	25
<b>4</b>	<b>Topological inference with <math>\check{\nu}</math></b>	<b>30</b>
4.1	An approximation theorem . . . . .	30
4.2	Homotopy type estimation with the DTM . . . . .	31
4.3	Persistent homology with DTM-filtrations . . . . .	32
<b>5</b>	<b>Conclusion</b>	<b>33</b>
<b>A</b>	<b>Supplementary material for Section 1</b>	<b>33</b>
<b>B</b>	<b>Supplementary material for Section 3</b>	<b>34</b>
	<b>References</b>	<b>40</b>

# 1 Introduction

## 1.1 Statement of the problem

Let  $\mathcal{M}_0$  be a compact  $\mathcal{C}^2$ -manifold of dimension  $d$ , and  $\mu_0$  a Radon probability measure on  $\mathcal{M}_0$ . Let  $E = \mathbb{R}^n$  be the Euclidean space and  $u : \mathcal{M} \rightarrow E$  be an immersion. We assume the following hypothesis: for every  $x_0, y_0 \in \mathcal{M}_0$  such that  $x_0 \neq y_0$  and  $u(x_0) = u(y_0)$ , the tangent spaces  $d_{x_0}u(T_{x_0}\mathcal{M}_0)$  and  $d_{y_0}u(T_{y_0}\mathcal{M}_0)$  are different. Define the image of the immersion  $\mathcal{M} = u(\mathcal{M}_0)$  and the pushforward measure  $\mu = u_*\mu_0$ . We suppose that we are observing the measure  $\mu$ , or a close measure  $\nu$ . Our goal is to infer the singular homology of  $\mathcal{M}_0$  (with coefficients in  $\mathbb{Z}_2$  for instance) from  $\nu$ .

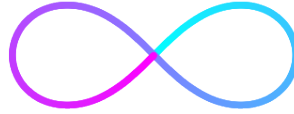


Figure 1: The abstract manifold  $\mathcal{M}_0$ , diffeomorphic to a circle

Figure 2: The immersion  $\mathcal{M} \subset \mathbb{R}^2$ , the Bernoulli's lemniscate

As shown in Figure 2, the immersion may self-intersect, hence the singular homology of  $\mathcal{M}_0$  and  $\mathcal{M}$  may differ. To get back to  $\mathcal{M}_0$ , we proceed as follows: let  $M(E)$  be the vector space of  $n \times n$  matrices, and  $\tilde{U} : \mathcal{M}_0 \rightarrow E \times M(E)$  the application  $x_0 \mapsto (u(x_0), \frac{1}{d+2}p_{T_x\mathcal{M}})$ , where  $p_{T_x\mathcal{M}}$  is the matrix representative of the orthogonal projection on  $T_x\mathcal{M} = d_{x_0}u(T_{x_0}\mathcal{M}_0) \subset E$ . Define  $\tilde{\mathcal{M}} = \tilde{U}(\mathcal{M}_0)$ . The set  $\tilde{\mathcal{M}}$  is a submanifold of  $E \times M(E)$ , diffeomorphic to  $\mathcal{M}_0$ .



Figure 3: The submanifold  $\tilde{\mathcal{M}} \subset \mathbb{R}^2 \times M(\mathbb{R}^2)$ , projected in a 3-dimensional subspace via PCA. Observe that it does not self-intersect

Suppose that one is able to estimate  $\tilde{\mathcal{M}}$  from  $\nu$ . Then one could consider the persistent homology of a filtration based on  $\tilde{\mathcal{M}}$ —say the Čech complex of  $\tilde{\mathcal{M}}$  in the ambient space  $E \times M(E)$  for example—and hope to read the singular homology of  $\mathcal{M}_0$  in the persistent barcode of  $\tilde{\mathcal{M}}$ .

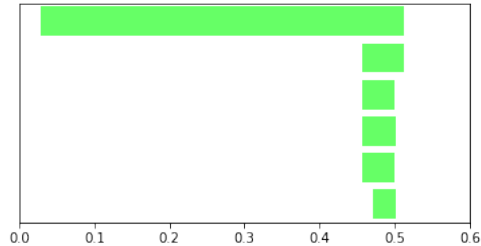
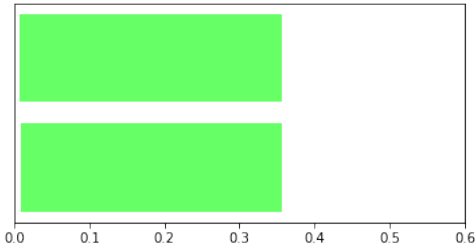


Figure 4: Persistence barcode of the 1-homology of the Čech filtration of  $\mathcal{M}$ . One reads the 1-homology of the lemniscate

Figure 5: Persistence barcode of the 1-homology of the Čech filtration of  $\tilde{\mathcal{M}}$ . One reads the 1-homology of a circle

Instead of estimating  $\check{\mathcal{M}}$ , we propose to estimate the measure  $\check{\mu}_0$ , defined as  $\check{\mu}_0 = \check{U}_* \mu_0$ . It is a measure on  $E \times M(E)$ , with support  $\check{\mathcal{M}}$ . Using measure-based filtrations—such as the DTM-filtrations—one can also hope to recover the singular homology of  $\mathcal{M}_0$ .

It is worth noting that  $\check{\mathcal{M}}$  can be naturally seen as a submanifold of  $E \times \mathcal{G}_d(E)$ , where  $\mathcal{G}_d(E)$  denotes the Grassmann manifold of  $d$ -dimensional linear subspaces of  $E$ . From this point of view,  $\check{\mu}_0$  can be seen as a measure on  $E \times \mathcal{G}_d(E)$ , i.e. a varifold. However, for computational reasons, we choose to work in  $M(E)$  instead of  $\mathcal{G}_d(E)$ .

Here is an alternative definition of  $\check{\mu}_0$ : for any  $\phi : E \times \mathcal{M}(E) \rightarrow \mathbb{R}$  with compact support,

$$\int \phi(x, A) d\check{\mu}_0(x, A) = \int \phi(u(x_0), \frac{1}{d+2} p_{T_x \mathcal{M}}) d\mu_0(x_0).$$

Getting back to the actual observed measure  $\nu$ , we propose to estimate  $\check{\mu}_0$  with the measure  $\check{\nu}$ , defined as follows: for any  $\phi : E \times \mathcal{M}(E) \rightarrow \mathbb{R}$  with compact support,

$$\int \phi(x, A) d\check{\nu}(x, A) = \int \phi(x, \bar{\Sigma}_\nu(x)) d\nu(x),$$

where  $\bar{\Sigma}_\nu(x)$ , the normalized local covariance matrix, is defined in Section 3. We prove that  $\bar{\Sigma}_\nu(x)$  can be used to estimate the tangent space  $\frac{1}{d+2} p_{T_x \mathcal{M}}$  of  $\mathcal{M}$  (Proposition 3.1), and that it is stable with respect to  $\nu$  (Equation 7). This estimation may be biased next to multiple points of  $\mathcal{M}$ , as shown in Figure 7. However, we prove a global estimation result, of the following form:  $\check{\mu}_0$  and  $\check{\nu}$  are close in the Wasserstein metric as long as  $\mu$  and  $\nu$  are (Theorem 4.1). As a consequence, the persistence diagrams of the DTM-filtrations based on  $\check{\mu}_0$  and  $\nu$  are close in bottleneck distance (Corollary 4.4).

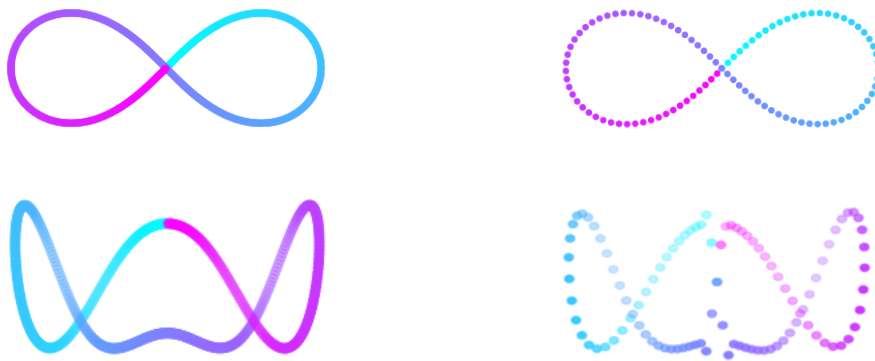


Figure 6: The sets  $\text{supp}(\mu) = \mathcal{M}$  and Figure 7: The sets  $\text{supp}(\nu)$  and  $\text{supp}(\check{\nu})$ , where  $\text{supp}(\check{\mu}_0) = \check{\mathcal{M}}$ , where  $\mu$  is the uniform measure on the lemniscate and  $\nu$  is the empirical measure on a sample of the lemniscate. Parameter  $r = 0.3$

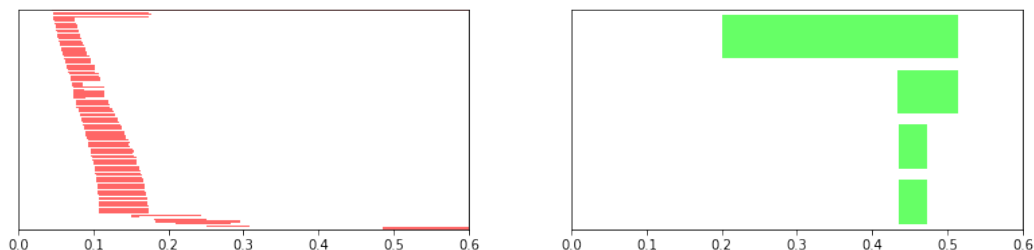


Figure 8: Persistence barcodes of the 0- and 1-homology of DTM-filtration of  $\check{\nu}$  with parameters  $r = 0.1$ ,  $\gamma = 3$  and  $m = 0.02$ . Observe that the 1-homology of the circle appears as a large feature of the barcode

## 1.2 Background on persistent homology

In the sequel, we consider interleavings of filtrations, interleavings of persistence modules and their associated pseudo-distances. Their definitions, restricted to the setting of the paper, are briefly recalled in this section. Let  $T = \mathbb{R}^+$  and  $E = \mathbb{R}^n$  endowed with the standard Euclidean norm.

**Filtrations of sets and simplicial complexes.** A family of subsets  $V = (V^t)_{t \in T}$  of  $E$  is a *filtration* if it is non-decreasing for the inclusion, i.e. for any  $s, t \in T$ , if  $s \leq t$  then  $V^s \subseteq V^t$ . Given  $\epsilon \geq 0$ , two filtrations  $V = (V^t)_{t \in T}$  and  $W = (W^t)_{t \in T}$  of  $E$  are  $\epsilon$ -*interleaved* if, for every  $t \in T$ ,  $V^t \subseteq W^{t+\epsilon}$  and  $W^t \subseteq V^{t+\epsilon}$ . The interleaving pseudo-distance between  $V$  and  $W$  is defined as the infimum of such  $\epsilon$ :

$$d_i(V, W) = \inf\{\epsilon, V \text{ and } W \text{ are } \epsilon\text{-interleaved}\}.$$

**Persistence modules.** Let  $k$  be a field. A *persistence module*  $\mathbb{V}$  over  $T = \mathbb{R}^+$  is a pair  $\mathbb{V} = ((\mathbb{V}^t)_{t \in T}, (v_s^t)_{s \leq t \in T})$  where  $(\mathbb{V}^t)_{t \in T}$  is a family of  $k$ -vector spaces, and  $(v_s^t : \mathbb{V}^s \rightarrow \mathbb{V}^t)_{s \leq t \in T}$  a family of linear maps such that:

- for every  $t \in T$ ,  $v_t^t : \mathbb{V}^t \rightarrow \mathbb{V}^t$  is the identity map,
- for every  $r, s, t \in T$  such that  $r \leq s \leq t$ ,  $v_s^t \circ v_r^s = v_r^t$ .

Given  $\epsilon \geq 0$ , an  $\epsilon$ -*morphism* between two persistence modules  $\mathbb{V}$  and  $\mathbb{W}$  is a family of linear maps  $(\phi_t : \mathbb{V}^t \rightarrow \mathbb{W}^{t+\epsilon})_{t \in T}$  such that the following diagrams commute for every  $s \leq t \in T$ :

$$\begin{array}{ccc} \mathbb{V}^s & \xrightarrow{v_s^t} & \mathbb{V}^t \\ \downarrow \phi_s & & \downarrow \phi_t \\ \mathbb{W}^{s+\epsilon} & \xrightarrow{w_{s+\epsilon}^{t+\epsilon}} & \mathbb{W}^{t+\epsilon} \end{array}$$

An  $\epsilon$ -*interleaving* between two persistence modules  $\mathbb{V}$  and  $\mathbb{W}$  is a pair of  $\epsilon$ -morphisms  $(\phi_t : \mathbb{V}^t \rightarrow \mathbb{W}^{t+\epsilon})_{t \in T}$  and  $(\psi_t : \mathbb{W}^t \rightarrow \mathbb{V}^{t+\epsilon})_{t \in T}$  such that the following diagrams commute for every  $t \in T$ :

$$\begin{array}{ccc} \mathbb{V}^t & \xrightarrow{v_t^{t+2\epsilon}} & \mathbb{V}^{t+2\epsilon} \\ & \searrow \phi_t & \nearrow \psi_{t+\epsilon} \\ & \mathbb{W}^{t+\epsilon} & \end{array} \qquad \begin{array}{ccc} & \mathbb{V}^{t+\epsilon} & \\ \nearrow \psi_t & & \searrow \phi_{t+\epsilon} \\ \mathbb{W}^t & \xrightarrow{w_t^{t+2\epsilon}} & \mathbb{W}^{t+2\epsilon} \end{array}$$

The interleaving pseudo-distance between  $\mathbb{V}$  and  $\mathbb{W}$  is defined as

$$d_i(\mathbb{V}, \mathbb{W}) = \inf\{\epsilon \geq 0, \mathbb{V} \text{ and } \mathbb{W} \text{ are } \epsilon\text{-interleaved}\}.$$

A persistence module  $\mathbb{V}$  is said to be  $q$ -*tame* if for every  $s, t \in T$  such that  $s < t$ , the map  $v_s^t$  is of finite rank. The  $q$ -tameness of a persistence module ensures that we can define a notion of persistence diagram—see [CdSGO16]. Moreover, given two  $q$ -tame persistence modules  $\mathbb{V}, \mathbb{W}$  with persistence diagrams  $D(\mathbb{V}), D(\mathbb{W})$ , the so-called isometry theorem states that  $d_b(D(\mathbb{V}), D(\mathbb{W})) = d_i(\mathbb{V}, \mathbb{W})$  where  $d_b(\cdot, \cdot)$  denotes the bottleneck distance between diagrams ([CdSGO16, Theorem 4.11]).

**Relation between filtrations and persistence modules.** Applying the homology functor to a filtration gives rise to a persistence module where the linear maps between homology groups are induced by the inclusion maps between sets (or simplicial complexes). As a consequence, if two filtrations are  $\epsilon$ -interleaved then their associated homology persistence modules are also  $\epsilon$ -interleaved, the interleaving homomorphisms being induced by the interleaving inclusion maps. Moreover, if the modules are  $q$ -tame, then the bottleneck distance between their persistence diagrams is upperbounded by  $\epsilon$ .

### 1.3 Background on persistent homology for measures

In this subsection we define the distance to measure (DTM), based on [CCSM11], and the DTM-filtrations, based on [ACG<sup>+</sup>18]. Let  $T = \mathbb{R}^+$  and  $E = \mathbb{R}^n$  endowed with the standard Euclidean norm.

**Wasserstein distances.** Given two probability measures  $\mu$  and  $\nu$  over  $E$ , a transport plan between  $\mu$  and  $\nu$  is a probability measure  $\pi$  over  $E \times E$  whose marginals are  $\mu$  and  $\nu$ . Let  $p \geq 1$ . The  $p$ -Wasserstein distance between  $\mu$  and  $\nu$  is defined as

$$W_p(\mu, \nu) = \left( \inf_{\pi} \int_{E \times E} \|x - y\|^p d\pi(x, y) \right)^{\frac{1}{p}},$$

where the infimum is taken over all the transport plans  $\pi$ . If  $q \geq 1$  is such that  $p \leq q$ , then an application of Jensen's inequality shows that  $W_p(\mu, \nu) \leq W_q(\mu, \nu)$ .

**DTM.** Let  $\mu$  be a probability measure over  $E$ , and  $m \in [0, 1)$  a parameter. For every  $x \in E$ , let  $\delta_{\mu, m}$  be the function defined on  $E$  by  $\delta_{\mu, m}(x) = \inf\{r \geq 0, \mu(\overline{\mathcal{B}}(x, r)) > m\}$ . The DTM  $\mu$  of parameter  $m$  (and exponent 2) is the function  $d_{\mu, m} : E \rightarrow \mathbb{R}$  defined as:

$$d_{\mu, m}^2(x) = \frac{1}{m} \int_0^m \delta_{\mu, t}^2(x) dt.$$

When  $m$  is fixed, we write  $d_{\mu}$  instead of  $d_{\mu, m}$ . We cite two important properties of the DTM:

**Proposition 1.1** ([CCSM11], Corollary 3.7). *For every probability measure  $\mu$  and  $m \in [0, 1)$ ,  $d_{\mu}$  is 1-Lipschitz.*

**Theorem 1.2** ([CCSM11], Theorem 3.5). *Let  $\mu, \nu$  be two probability measures, and  $m \in (0, 1)$ . Then  $\|d_{\mu} - d_{\nu}\|_{\infty} \leq m^{-\frac{1}{2}} W_2(\mu, \nu)$ .*

The following theorem shows that the sublevel sets  $d_{\mu}^t$  of  $d_{\mu}$  can be used to estimate the homotopy type of  $\text{supp}(\mu)$ .

**Theorem 1.3** (Corollary 4.11 in [CCSM11]). *Let  $m \in (0, 1)$ ,  $\mu$  be any measure on  $E$ , and denote  $K = \text{supp}(\mu)$ . Suppose that  $\text{reach}(K) = \tau > 0$ , and that  $\mu$  satisfies the following hypothesis for  $r < m$ :  $\forall x \in K, \mu(\mathcal{B}(x, r)) \geq ar^d$ . Let  $\nu$  be another measure, and denote  $w = W_2(\mu, \nu)$ . Suppose that  $w \leq m^{\frac{1}{2}} \left( \frac{\tau}{9} - \left( \frac{m}{a} \right)^{\frac{1}{d}} \right)$ . Define  $\epsilon = \left( \frac{m}{a} \right)^{\frac{1}{d}} + m^{-\frac{1}{2}} w$  and choose  $t \in [4\epsilon, R - 4\epsilon]$ . Then  $d_{\mu, m}^t$  and  $K$  are homotopic equivalent.*

**DTM-filtrations.** We still consider a probability measure  $\mu$  on  $E$  and  $m \in [0, 1)$ . For every  $t \in T$ , consider the set

$$W^t[\mu] = \bigcup_{x \in \text{supp}(\mu)} \overline{\mathcal{B}}(x, (t - d_{\mu}(x))^+),$$

where  $\overline{\mathcal{B}}(x, r^+)$  denotes the closed ball of center  $x$  and of radius  $r$  if  $r \geq 0$ , or denotes the empty set if  $r < 0$ . The family  $W[\mu] = (W^t[\mu])_{t \geq 0}$  is a filtration of  $E$ . It is called the DTM-filtration with parameters  $(\mu, m, 1)$ . By applying the singular homology functor with coefficients in a field  $k$ , we obtain persistence module, denoted by  $\mathbb{W}[\mu]$ . If  $\text{supp}(\mu)$  is bounded, then  $\mathbb{W}[\mu]$  is  $q$ -tame.

We close this subsection with a stability result for the DTM-filtrations. If  $\mu$  is any measure, define the quantity  $c(\mu) = \sup_{x \in \text{supp}(\mu)} d_{\mu}(x)$ .

**Theorem 1.4** ([ACG<sup>+</sup>18], Theorem 4.5). *Consider two probability measures  $\mu, \nu$  on  $E$  with supports  $X$  and  $Y$ . Let  $\mu', \nu'$  be two probability measures with compact supports  $\Gamma$  and  $\Omega$  such that  $\Gamma \subseteq X$  and  $\Omega \subseteq Y$ . We have*

$$d_i(W[\mu], W[\nu]) \leq m^{-\frac{1}{2}} W_2(\mu, \mu') + m^{-\frac{1}{2}} W_2(\mu', \nu') + m^{-\frac{1}{2}} W_2(\nu', \nu) + c(\mu') + c(\nu').$$

The term  $c(\mu)$  is to be seen as a quantity controlling the regularity of  $\mu$ . In particular, if  $\mu$  is the uniform measure on a submanifold, it goes to 0 as  $m$  does, as shown by the following lemma.

**Lemma 1.5.** *Suppose that  $\mu$  satisfies the following for  $r < m$ :  $\forall x \in \text{supp}(\mu), \mu(\mathcal{B}(x, r)) \geq ar^d$ . Then  $c(\mu) \leq c_{1.5}m^{\frac{1}{d}}$  with  $c_{1.5} = a^{-(1+\frac{1}{d})}$ .*

We can restate Theorem 1.4 without mentioning the intermediate measures  $\mu'$  and  $\nu'$ . The proof is given in appendix.

**Corollary 1.6.** *Let  $\mu, \nu$  with  $W_2(\mu, \nu) = w \leq \frac{1}{4}$ . Suppose that  $\mu$  satisfies the following for  $r < m$ :  $\forall x \in \text{supp}(\mu), \mu(\mathcal{B}(x, r)) \geq ar^d$ . Then*

$$d_i(W[\mu], W[\nu]) \leq c_{1.6}\left(\frac{w}{m}\right)^{\frac{1}{2}} + 2c_{1.5}m^{\frac{1}{d}}$$

with  $c_{1.6} = (8D + 5)$  and  $D = \text{diam}(\text{supp}(\mu))$ .

## 1.4 Notations and hypotheses

**Notations.** Throughout the paper, we shall consider

- $n, d > 0$  integers
- If  $x, y \in \mathbb{R}$ ,  $x \wedge y$  is the minimum of  $x$  and  $y$
- $E = \mathbb{R}^n$  the Euclidean space,  $M(E)$  the vector space of  $n \times n$  matrices,  $\mathcal{G}_d(E)$  the Grassmannian
- For  $x, y \in E$ ,  $x \perp y$  denotes the orthogonality of  $x$  and  $y$
- If  $x, y \in E$ ,  $x \otimes y = x^t y \in M(E)$  is the outer product, and  $x^{\otimes 2} = x \otimes x$
- $\|\cdot\|$  the Euclidean norm on  $E$ ,  $\|\cdot\|_F$  the Frobenius norm on  $M(E)$ ,  $\|\cdot\|_\gamma$  the  $\gamma$ -norm on  $E \times M(E)$  (defined in Subsection 3.1)
- $W_p(\cdot, \cdot)$  the 1-Wasserstein distance between measures on  $E$ ,  $W_{p,\gamma}(\cdot, \cdot)$  the  $(p, \gamma)$ -Wasserstein distance between measures on  $E \times M(E)$  (defined in Subsection 3.1)
- $\mathcal{H}^d$  the  $d$ -dimensional Hausdorff measure on  $E$  or on a subspace  $T \subset E$
- If  $\mu$  is a measure of finite mass,  $|\mu|$  denotes its mass, and  $\bar{\mu} = \frac{1}{|\mu|}\mu$  is the associated probability measure
- If  $T$  is a subspace of  $E$ ,  $p_T$  denotes the orthogonal projection matrix on  $T$ .
- $\mathcal{B}(x, r)$  and  $\bar{\mathcal{B}}(x, r)$  the open and closed balls of  $E$ ,  $\partial\mathcal{B}(x, r)$  the sphere.  $V_d$  and  $S_{d-1}$  denote  $\mathcal{H}^d(\mathcal{B}(0, 1))$  and  $\mathcal{H}^{d-1}(\partial\mathcal{B}(0, 1))$  (note that  $S_{d-1} = dV_d$ )
- $\mathcal{M}_0$  is a Riemannian manifold, and  $\mathcal{B}_{\mathcal{M}_0}(x, r)$  and  $\bar{\mathcal{B}}_{\mathcal{M}_0}(x, r)$  denote the open and closed geodesics balls For  $x_0, y_0 \in \mathcal{M}_0$ ,  $d_{\mathcal{M}_0}(x_0, y_0)$  denotes the geodesic distance
- If  $T$  is a subspace of  $E$ ,  $\mathcal{B}_T(x, r)$  and  $\bar{\mathcal{B}}_T(x, r)$  denote the open and closed balls of  $T$  for the Euclidean distance
- if  $f$  is a map with values in  $\mathbb{R}$  and  $t \in \mathbb{R}$ ,  $f^t$  denotes the sublevel set  $f^t = f^{-1}((-\infty, t])$

**Model.** We consider an abstract  $\mathcal{C}^2$ -manifold  $\mathcal{M}_0$  of dimension  $d$ , and an immersion  $u : \mathcal{M}_0 \rightarrow E$ . We denote  $\mathcal{M} = u(\mathcal{M}_0)$ . Moreover, we write  $T_{x_0}\mathcal{M}_0$  for the tangent space of  $\mathcal{M}_0$  at  $x_0$ , and  $T_x\mathcal{M}$  for  $d_{x_0}u(T_{x_0}\mathcal{M}_0)$ . Let  $\check{U}$  be the application

$$\begin{aligned} \check{U} : \mathcal{M}_0 &\longrightarrow E \times M(E) \\ x_0 &\longmapsto (x, p_{T_x\mathcal{M}}). \end{aligned}$$

We denote  $\check{\mathcal{M}} = \check{u}(\mathcal{M}_0)$ . We also consider a probability measure  $\mu_0$  on  $\mathcal{M}_0$ , and define  $\mu = u_*\mu_0$  and  $\check{\mu} = \check{U}_*\mu_0$ . These several sets and measures fit in the following commutative diagrams

$$\begin{array}{ccc} \mathcal{M}_0 & \xrightarrow{\check{u}} & \check{\mathcal{M}} \\ & \searrow u & \swarrow \text{proj} \\ & \mathcal{M} & \end{array} \qquad \begin{array}{ccc} \mu_0 & \xrightarrow{\check{u}_*} & \check{\mu} \\ & \searrow u_* & \swarrow \text{proj}_* \\ & \mu & \end{array}$$

Moreover, we endow  $\mathcal{M}_0$  with the Riemannian structure given by the immersion  $u$ . For every  $x_0 \in \mathcal{M}_0$ , the second fundamental form of  $\mathcal{M}_0$  at  $x_0$  is denoted  $II_{x_0} : (T_{x_0}\mathcal{M}_0)^2 \rightarrow (T_x\mathcal{M})^\perp$ , and the exponential map is denoted  $\exp_{x_0}^{\mathcal{M}_0} : T_{x_0}\mathcal{M}_0 \rightarrow \mathcal{M}_0$ . We also consider the application  $\exp_x^{\mathcal{M}} : T_x\mathcal{M} \rightarrow \mathcal{M}$ , defined as  $u \circ \exp_{x_0}^{\mathcal{M}_0} \circ (d_{x_0}u)^{-1}$ .

**Notation convention.** When considering a point  $x_0 \in \mathcal{M}_0$ , the notation  $x$  shall refer to  $u(x_0)$ , and the notation  $\check{x}$  to  $\check{u}(x_0)$ . Similarly, if  $\gamma_0 : I \rightarrow \mathcal{M}_0$  is a map, we shall use the notations  $\gamma = u \circ \gamma_0$  and  $\check{\gamma} = \check{u} \circ \gamma_0$ .

**Hypotheses.** In the sequel of the paper we shall refer to the following hypotheses.

*Hypothesis 1.* For every  $x_0, y_0 \in \mathcal{M}_0$  such that  $x_0 \neq y_0$  and  $x = y$ , we have  $T_{x_0}\mathcal{M} \neq T_{y_0}\mathcal{M}$ .

*Hypothesis 2.* The operator norm of the second fundamental form of  $\mathcal{M}_0$  at each point is bounded by  $\rho > 0$ .

*Hypothesis 3.*  $\mu_0$  admits a density  $f_0$  on  $\mathcal{M}_0$ . Moreover,  $f_0$  is  $L_0$ -Lipschitz (for the geodesic distance), and there exist  $f_{\min}, f_{\max} > 0$  such that  $\forall x_0 \in \mathcal{M}_0, f_{\min} \leq f_0(x_0) \leq f_{\max}$ .

Note that Hypothesis 1 ensure that  $\check{u}$  is injective, hence that the set  $\check{\mathcal{M}}$  is a submanifold of  $E \times M(E)$ . Hypothesis 2 implies the following key property: if  $\gamma_0 : I \rightarrow \mathcal{M}_0$  is an arc-length parametrized geodesic of class  $\mathcal{C}^2$ , then denoting  $\gamma = u \circ \gamma_0$ , we have  $\forall t \in I, \|\ddot{\gamma}(t)\| \leq \rho$  (see [NSW08, Section 6]).

In Subsection 2.3, we define an application  $\lambda_0 : \mathcal{M}_0 \rightarrow \mathbb{R}^+$ , called the normal reach. Remind that the notation  $\lambda_0^r$  refers to the sublevel set  $\lambda_0^{-1}([0, r])$ . We shall consider the following hypothesis:

*Hypothesis 4.* There exists  $c_4 \geq 0$  and  $r_0 > 0$  such that, for every  $r \in [0, r_0)$ ,  $\mu_0(\lambda_0^r) \leq c_4 r$ .

The author thinks that this hypothesis is a consequence of Hypotheses 1, 2 and 3, but has not been able to prove it yet.

## 2 Reach of an immersed manifold

In this section, we introduce a new notion of reach suitable for an immersed manifold.

### 2.1 Background on reach

Let us recall the definition of the reach of a subset  $A \subseteq E$ , as done in [Fed59, Definition 4.1]. Let  $x \in E \mapsto d(x, A) = \inf_{a \in A} \|x - a\|$  be the distance function to  $A$ . First, the medial axis of  $S$  is defined as

$$\text{med}(A) = \{x \in E, \exists a, b \in A \text{ s.t. } a \neq b \text{ and } \|x - a\| = \|x - b\| = d(x, A)\}.$$



**Definition 2.1.** Let  $a \in A$ . The *reach of  $A$  at  $a$*  (or *local feature size*) is defined as  $\text{reach}(A, a) = d(a, \text{med}(A))$ . The *reach of  $A$*  is defined as  $\text{reach}(A) = \inf_{a \in A} \text{reach}(A, a)$ .

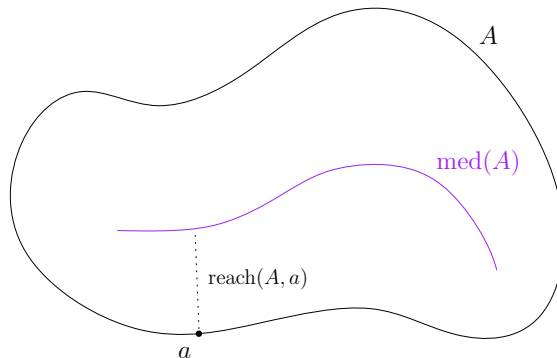


Figure 9: Medial axis and reach of a submanifold of  $\mathbb{R}^2$

Among the several properties of a set  $S$  with positive reach, a useful one is the approximation by tangent spaces. For a general set  $S$ , we define the tangent cone at  $x \in S$  as:

$$\text{Tan}(S, x) = \{0\} \cup \left\{ v \in E, \forall \epsilon > 0, \exists y \in S \text{ s.t. } y \neq x, \|y - x\| < \epsilon, \left\| \frac{v}{\|v\|} - \frac{y - x}{\|y - x\|} \right\| < \epsilon \right\}.$$

If  $S$  is a submanifold, we recover the usual notion of tangent space.

**Theorem 2.1** (Theorem 4.18(2) of [Fed59]). *A closed set  $S \subseteq E$  has positive reach  $\tau$  if and only if, for every  $x, y \in S$ ,*

$$d(y - x, \text{Tan}(S, x)) \leq \frac{1}{2\tau} \|y - x\|^2.$$

Using this property, it is shown in [ACLZ17] that if  $S = \mathcal{M}$  is a submanifold of positive reach  $\tau$ , one can estimate the tangent spaces of  $\mathcal{M}$  via its local covariance matrices. The quality of the estimation depends on  $\tau$ . However, in our case, the immersion  $u : \mathcal{M}_0 \rightarrow \mathcal{M}$  may be non-injective, and the set  $\mathcal{M}$  may be of reach 0. In Subsection 2.3, we define another notion of reach that will help us to estimate tangent spaces.

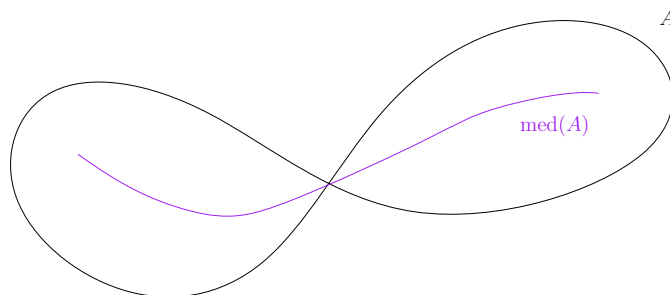


Figure 10: A subset of  $\mathbb{R}^2$  with zero reach

## 2.2 Geodesic bounds under curvature conditions

We now consider an immersion  $u : \mathcal{M}_0 \rightarrow \mathcal{M} \subset E$  which satisfies Hypothesis 2: the manifold  $\mathcal{M}_0$  is equipped with the Riemannian structure induced by  $u$ , and the operator norm of the second fundamental form of  $\mathcal{M}_0$  at each point is bounded by  $\rho$ . The aim of this subsection is to give technical results for the sequel of the paper.

First, we state Lemma 2.2, whose second point can be seen as an equivalent of Theorem 2.1 with respect to the geodesic distance on  $\mathcal{M}_0$ , where the quantity  $\frac{1}{\rho}$  plays the role of the reach  $\tau$ .

**Lemma 2.2.** *Let  $x_0 \in \mathcal{M}_0$  and  $\gamma_0 : I \rightarrow \mathcal{M}_0$  an arc-length parametrized geodesic starting from  $x_0$ . Let  $\gamma = u \circ \gamma_0$  and  $v = \dot{\gamma}(0)$ . For all  $t \in I$ , we have*

- $\|\gamma(t) - (x + tv)\| \leq \frac{\rho}{2}t^2$ .

As a consequence, for every  $y_0 \in \mathcal{M}_0$ , denoting  $\delta = d_{\mathcal{M}_0}(x_0, y_0)$ , we have

- $d(y - x, T_x\mathcal{M}) \leq \frac{\rho}{2}\delta^2$ ,
- $(1 - \frac{\rho}{2}\delta)\delta \leq \|x - y\|$ .

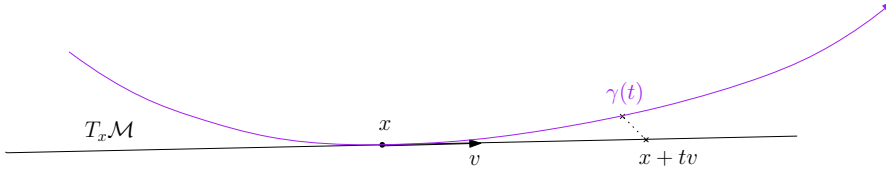


Figure 11: Deviation of a geodesic from its initial direction

*Proof.* Consider the application  $f : t \mapsto \|\gamma(t) - (x + tv)\|$ . Since  $\gamma$  is a geodesic, it is of class  $\mathcal{C}^2$ , and  $\sup_I \|\dot{\gamma}\| = \sup_I \|\dot{\gamma}_0\| \leq \rho$  (see [NSW08, Section 6]). Hence we can apply Taylor-Lagrange formula to get  $f(t) \leq \sup_I \|\ddot{\gamma}\| \frac{1}{2}t^2 \leq \frac{\rho}{2}t^2$ . Therefore, for all  $t \in I$ , we have  $\|\gamma(t) - (x + tv)\| \leq \frac{\rho}{2}t^2$ , and the first claim is proven.

Now, let  $\delta = d_{\mathcal{M}_0}(x_0, y_0)$ . By Hopf-Rinow Theorem ([dC92, Theorem 2.8 p146]), we can consider a length-minimizing geodesic  $\gamma_0$  from  $x_0$  to  $y_0$ . Using the last inequality for  $t = \delta$  yields

$$\|y - (x + \delta v)\| = \|\gamma(\delta) - (x + \delta v)\| \leq \frac{\rho}{2}\delta^2,$$

and we deduce that  $d(y - x, T_x\mathcal{M}) \leq \frac{\rho}{2}\delta^2$  since  $x + \delta v \in T_x\mathcal{M}$ .

To prove the last point, we apply the triangular inequality:

$$\|x - y\| \geq \|x - (x + \delta v)\| - \|(x + \delta v) - y\| \geq \delta - \frac{\rho}{2}\delta^2. \quad \square$$

*Remark 2.1.* The last point of Lemma 2.2 implies the following fact: for all  $x_0 \in \mathcal{M}_0$ , the map  $u$  is injective on  $\mathcal{B}_{\mathcal{M}_0}(x_0, \frac{2}{\rho})$ . Indeed, if  $x_0, y_0 \in \mathcal{M}_0$  are such that  $\delta = d_{\mathcal{M}_0}(x_0, y_0) < \frac{2}{\rho}$ , we get  $0 < (1 - \frac{\rho}{2}\delta)\delta \leq \|x - y\|$ , hence  $x \neq y$ .

We can also deduce the following: for every  $y_0 \in \mathcal{B}_{\mathcal{M}_0}(x_0, \frac{1}{\rho})$  such that  $y_0 \neq x_0$ , the vector  $y - x$  is not orthogonal to  $T_x\mathcal{M}$  nor  $T_y\mathcal{M}$ . To see this, notice that the inequality  $\delta < \frac{1}{\rho}$  and the second point of Lemma 2.2 yield

$$d(y - x, T_x\mathcal{M}) \leq \frac{\rho}{2}\delta^2 < \frac{1}{2}\delta.$$

Besides, the third point gives  $\delta < 2\|y - x\|$ , and we deduce that  $d(y - x, T_x\mathcal{M}) < \|y - x\|$ . Equivalently,  $y - x$  is not orthogonal to  $T_x\mathcal{M}$ . Similarly, one proves that  $y - x$  is not orthogonal to  $T_y\mathcal{M}$ .

Consider two points  $x_0, y_0 \in \mathcal{M}_0$ . We wish to compare their geodesic distance  $d_{\mathcal{M}_0}(x_0, y_0)$  and their Euclidean distance  $\|y - x\|$ . A first inequality is true in general:

$$\|y - x\| \leq d_{\mathcal{M}_0}(x_0, y_0).$$

Moreover, if they are close enough in geodesic distance—say  $d_{\mathcal{M}_0}(x_0, y_0) \leq \frac{1}{\rho}$  for instance—then Lemma 2.2 third point yields

$$d_{\mathcal{M}_0}(x_0, y_0) \leq 2\|x - y\|.$$

However, without any assumption on  $d_{\mathcal{M}_0}(x_0, y_0)$ , such an inequality does not hold in general. Figure 12 represents a pair of points which are close in Euclidean distance, but far away with respect to the geodesic distance. In the next subsection, we prove such an inequality, but imposing a constraint on  $\|y - x\|$  instead of  $d_{\mathcal{M}_0}(x_0, y_0)$  (Lemma 2.5).

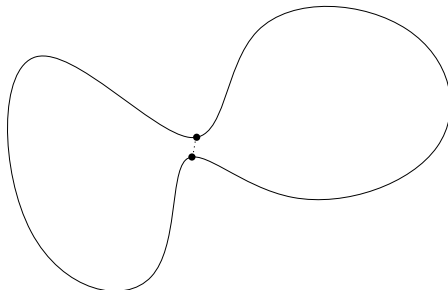


Figure 12: Pair of points for which the geodesic distance is large compared to the Euclidean distance

We close this subsection with the following technical lemma.

**Lemma 2.3.** *Let  $x_0, y_0 \in \mathcal{M}_0$  and  $\gamma_0 : I \mapsto \mathcal{M}_0$  an arc-length parametrized geodesic with  $\gamma_0(0) = y_0$ . Define  $v = \dot{\gamma}(0)$ . Define  $l = \|y - x\|$ , and let  $r$  be such that  $l \leq r < \frac{1}{\rho}$ . Consider the application  $\phi : t \in I \mapsto \|\gamma(t) - x\|^2$ .*

- *If  $\langle v, y - x \rangle \geq 0$ , then  $\phi > \phi(0)$  on  $(0, T_1)$ , where  $T_1 = \frac{2}{\rho} \sqrt{1 - \rho l}$*
- *If  $\langle v, y - x \rangle = 0$ , then  $\phi$  is increasing on  $[0, T_2]$  where  $T_2 = \frac{\sqrt{2}}{\rho} \sqrt{2 - \sqrt{3 + \rho^2 l^2}}$ .*

Let  $b$  be the first value of  $t$  such that  $\|\gamma(t) - x\| = r$ .

- *For all  $t \in [0, b]$ , we have  $\ddot{\phi}(t) \geq 2(1 - \rho r)$ .*
- *If  $\langle v, y - x \rangle \leq 0$ , then  $b \geq (1 + \rho r)^{-\frac{1}{2}} \sqrt{r^2 - l^2}$ .*
- *If  $\langle v, y - x \rangle \geq 0$ , then  $b \leq (\frac{1 - \rho r}{2})^{-\frac{1}{2}} \sqrt{r^2 - l^2}$ . Note that if  $r < \frac{1}{2\rho}$ , then  $b < 2r < \frac{1}{\rho}$ .*

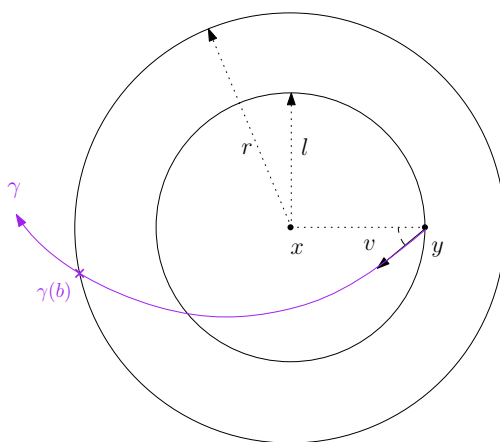


Figure 13: Illustration of Lemma 2.3 fourth point

*Proof. Point (1):* We use the triangle inequality, the Pythagorean Theorem and Lemma 2.2 to get

$$\begin{aligned}\|\gamma(t) - x\| &\geq \|(y + tv) - x\| - \|\gamma(t) - (y + tv)\| \\ &\geq \sqrt{\|tv\|^2 + \|y - x\|^2 + 2\langle tv, y - x \rangle} - \frac{\rho}{2}t^2 \\ &\geq \sqrt{t^2 + l^2} - \frac{\rho}{2}t^2.\end{aligned}$$

Now, a computation shows that the function  $t \mapsto \sqrt{t^2 + l^2} - \frac{\rho}{2}t^2$  is greater than  $l$  on  $(0, T_1)$ . Hence for  $t \in (0, T_1)$ , we have  $\phi(t) = \|\gamma(t) - x\|^2 > l^2 = \phi(0)$ .

*Point (2):* Observe that  $\dot{\phi}(t) = 2\langle \dot{\gamma}(t), \gamma(t) - x \rangle$ , and that

$$\ddot{\phi}(t) = 2\langle \ddot{\gamma}(t), \gamma(t) - x \rangle + 2\langle \dot{\gamma}(t), \dot{\gamma}(t) \rangle.$$

By Cauchy-Schwarz inequality,  $\langle \ddot{\gamma}(t), \gamma(t) - x \rangle \geq -\|\ddot{\gamma}(t)\|\|\gamma(t) - x\|$ . Note that  $\langle \dot{\gamma}(t), \dot{\gamma}(t) \rangle = 1$  and  $\|\ddot{\gamma}(t)\| \leq \rho$ . Hence we get

$$\ddot{\phi}(t) \geq 2(1 - \rho\|\gamma(t) - x\|). \quad (1)$$

Now, since  $\langle v, y - x \rangle = 0$ , we have

$$\begin{aligned}\|\gamma(t) - x\| &\leq \|(y + tv) - x\| + \|\gamma(t) - (y + tv)\| \\ &\leq \sqrt{\|tv\|^2 + \|y - x\|^2} + \frac{\rho}{2}t^2 \\ &= \sqrt{t^2 + l^2} + \frac{\rho}{2}t^2.\end{aligned}$$

A computation shows that the function  $t \mapsto \sqrt{t^2 + l^2} + \frac{\rho}{2}t^2$  is lower than  $\frac{1}{\rho}$  on  $(0, T_2)$ . Hence for  $t \in (0, T_2)$ , we have  $\ddot{\phi}(t) \geq 0$ . And since  $\dot{\phi}(0) = 0$ , we have that  $\phi$  is increasing.

*Point (3):* For all  $t \in (0, b)$ , it holds that  $\|\gamma(t) - x\| \leq r$ , hence Equation 1 gives  $\ddot{\phi}(t) \geq 2(1 - \rho r)$ .

*Point (4):* Assume that  $\langle v, y - x \rangle \leq 0$ . We still have the inequality

$$\|\gamma(t) - x\| \leq \sqrt{t^2 + l^2} + \frac{\rho}{2}t^2. \quad (2)$$

Consider  $t^*$ , the first non-negative root of  $\sqrt{t^2 + l^2} + \frac{\rho}{2}t^2 = r$ . According to Equation 2,  $b \geq t^*$ . Now, a computation gives

$$t^* = \frac{\sqrt{2}}{\rho} \sqrt{1 + \rho r - \sqrt{(1 + \rho r)^2 - \rho^2(r^2 - l^2)}}.$$

Using the inequality  $\sqrt{B} - \sqrt{A} = \frac{1}{\sqrt{B} + \sqrt{A}}(B - A) \geq \frac{1}{2\sqrt{B}}(B - A)$ , where  $A < B$ , we get

$$1 + \rho r - \sqrt{(1 + \rho r)^2 - \rho^2(r^2 - l^2)} \geq \frac{1}{2(1 + \rho r)}\rho^2(r^2 - l^2),$$

and we conclude that  $t^* \geq \frac{1}{\sqrt{1 + \rho r}}\sqrt{r^2 - l^2}$ .

*Point (5):* Assume that  $\langle v, y - x \rangle \geq 0$ . In the same vein as Point 4, we have  $\|\gamma(t) - x\| \geq \sqrt{t^2 + l^2} - \frac{\rho}{2}t^2$ , and we deduce  $b \leq t^*$ , where  $t^*$  is the first positive root of  $\sqrt{t^2 + l^2} - \frac{\rho}{2}t^2 = r$ . Solving this equation leads to

$$t^* = \frac{\sqrt{2}}{\rho} \sqrt{1 - \rho r - \sqrt{(1 - \rho r)^2 - \rho^2(r^2 - l^2)}}.$$

We use the inequality  $\sqrt{B} - \sqrt{A} = \frac{1}{\sqrt{A} + \sqrt{B}}(B - A) \leq \frac{1}{\sqrt{B}}(B - A)$ , where  $A < B$ , to get

$$1 - \rho r - \sqrt{(1 - \rho r)^2 - \rho^2(r^2 - l^2)} \leq \frac{1}{1 - \rho r}\rho^2(r^2 - l^2)$$

and we conclude that  $t^* \leq \frac{\sqrt{2}}{\sqrt{1 - \rho r}}\sqrt{r^2 - l^2}$ .  $\square$

### 2.3 Normal reach

We still consider an immersion  $u : \mathcal{M}_0 \rightarrow \mathcal{M} \subset E$  which satisfies Hypothesis 2.

**Definition 2.2.** For every  $x_0 \in \mathcal{M}_0$ , we set  $\Lambda(x_0) = \{y_0 \in \mathcal{M}_0, y_0 \neq x_0, x - y \perp T_y \mathcal{M}\}$ . The normal reach of  $\mathcal{M}_0$  at  $x_0$  is defined as:

$$\lambda_0(x_0) = \inf_{y_0 \in \Lambda(x_0)} \|x - y\|.$$

Note that  $\Lambda(x_0)$  is closed and hence this infimum is attained. Indeed, we can write  $\Lambda(x_0) = L \setminus \{x_0\}$ , with  $L = \{y_0 \in \mathcal{M}_0, x - y \perp T_y \mathcal{M}\}$ .  $L$  is a closed set since it is the preimage of  $\{0\}$  by the continuous function  $y_0 \mapsto \|p_{T_y \mathcal{M}}(x - y)\|$ . Furthermore,  $\{x_0\}$  is an isolated point of  $\Lambda(x_0)$ , since Remark 2.1 says that, for every  $y_0$  in the geodesic ball  $\mathcal{B}_{\mathcal{M}_0}(x_0, \frac{1}{\rho})$  such that  $y_0 \neq x_0$ , the vector  $x - y$  is not orthogonal to  $T_y \mathcal{M}$ , and hence  $y_0 \notin L$ .

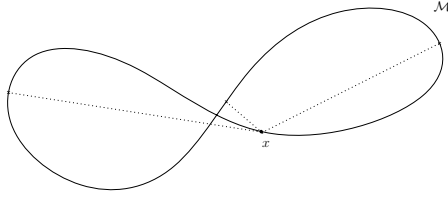


Figure 14: The set  $\Lambda(x_0)$  from Definition 2.2

Observe that if a point  $x \in \mathcal{M}$  has several preimages by  $u$ , then for all  $x_0 \in u^{-1}(\{x\})$ , we have  $\lambda_0(x_0) = 0$ . Hence we can define a map  $\lambda : \mathcal{M} \rightarrow \mathbb{R}$  as

$$\lambda(x) = \begin{cases} \lambda_0(u^{-1}(x)) & \text{if } x \text{ has only one preimage,} \\ 0 & \text{else.} \end{cases}$$

It satisfies the relation  $\lambda_0 = \lambda \circ u$ .

Here is a key property of the normal reach:

**Lemma 2.4.** Let  $x_0 \in \mathcal{M}_0$ . Let  $r > 0$  such that  $r < \lambda(x)$ . Then  $u^{-1}(\mathcal{M} \cap \overline{\mathcal{B}}(x, r))$  is connected.

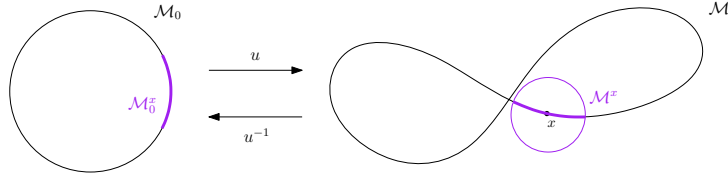


Figure 15: The set  $u^{-1}(\mathcal{M} \cap \overline{\mathcal{B}}(x, r))$ , with  $r < \lambda(x)$ , is connected

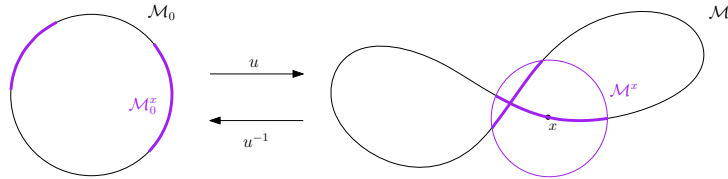


Figure 16: The set  $u^{-1}(\mathcal{M} \cap \overline{\mathcal{B}}(x, r))$ , with  $r \geq \lambda(x)$ , may not be connected

*Proof.* Denote  $\mathcal{M}^x = \overline{\mathcal{B}}(x, r) \cap \mathcal{M}$  and  $\mathcal{M}_0^x = u^{-1}(\mathcal{M}^x)$ . Let us prove that  $\mathcal{M}_0^x$  is connected. Suppose that it is not the case. Let  $C \subset \mathcal{M}_0^x$  be a connected component which does not contain  $x_0$ . Since  $C$  is compact, we can consider a minimizer  $y_0$  of  $\{\|x - y\|, y_0 \in C\}$ . Let us show that  $x - y \perp T_y \mathcal{M}$ , which will lead to a contradiction.

Two cases may occur:  $y \in \mathcal{B}(x, r)$  or  $y \in \partial \mathcal{B}(x, r)$ . If  $y \in \mathcal{B}(x, r)$ , then there exists a neighborhood  $V_0 \subseteq \mathcal{M}_0$  of  $y$  such that  $V_0 \subseteq \mathcal{M}_0^x$ . Hence  $y$  satisfies  $x - y \perp T_y \mathcal{M}$ , otherwise it would not be a local minimizer. Now, suppose that  $y \in \partial \mathcal{B}(x, r)$ . Since  $y_0$  is a minimizer, there exists a neighborhood  $V_0 \subseteq C$  of  $y_0$  such that  $V \cap \mathcal{B}(x, r) = \emptyset$ . We deduce the existence of a neighborhood  $V'_0 \subseteq \mathcal{M}_0$  of  $y_0$  such that  $V' \cap \mathcal{B}(x, r) = \emptyset$ . For instance, take a ball  $\mathcal{B} = \mathcal{B}_{\mathcal{M}_0}(y_0, s)$  such that  $\mathcal{B} \cap C \subseteq V_0$ , and define  $V'_0 = \mathcal{B}$ . We deduce that  $y - x \perp T_y \mathcal{M}$ .

To conclude, the properties  $x - y \perp T_y \mathcal{M}$  and  $x_0 \neq y_0$  imply that  $\|x - y\| \geq \lambda(x)$ , which contradicts  $r < \lambda(x)$ .  $\square$

The following lemma is an equivalent of [NSW08, Proposition 6.3] for the normal reach. It allows to compare the geodesic and Euclidean distance by only imposing a condition on the last one.

**Lemma 2.5.** *Let  $x_0, y_0 \in \mathcal{M}_0$ . Denote  $r = \|x - y\|$  and  $\delta = d_{\mathcal{M}_0}(x_0, y_0)$ . Suppose that  $\|x - y\| < \frac{1}{2\rho} \wedge \lambda(x)$ . Then*

$$\delta \leq c_{2.5}(\rho r) \quad \text{where} \quad c_{2.5}(t) = \frac{1}{t}(1 - \sqrt{1 - 2t}).$$

*In other words, the following inclusion holds:  $u^{-1}(\overline{\mathcal{B}}(x, r)) \subseteq \overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, c_{2.5}(\rho r)r)$ .*

A computation shows that, for  $t < \frac{1}{2}$ , we have the inequalities  $1 \leq c_{2.5}(t) \leq 1 + 2t < 2$ .

*Proof.* Denote  $\mathcal{M}^x = \overline{\mathcal{B}}(x, r) \cap \mathcal{M}$ ,  $\mathcal{M}_0^x = u^{-1}(\mathcal{M}^x)$  and  $\delta = d_{\mathcal{M}_0}(x, y)$ .

*Step 1:* Let us prove that  $\mathcal{M}_0^x \cap \partial \mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon) = \emptyset$ , with  $\delta_{\min} = c_{2.5}(\rho r)r$ , where  $c_{2.5}(t) = \frac{1}{t}(1 - \sqrt{1 - 2t})$  and  $\epsilon$  is small enough. Choose  $y_0 \in \partial \mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon)$ . According to Lemma 2.2, we have

$$\|x - y\| \geq (1 - \frac{\rho}{2}(\delta_{\min} + \epsilon))(\delta_{\min} + \epsilon). \quad (3)$$

Consider the polynomial  $\phi : t \mapsto (1 - \frac{\rho}{2}t)t - r$ . Its discriminant is  $1 - 2\rho r > 0$ , and we deduce that  $\phi(t)$  is positive if and only if  $t \in (\frac{1}{\rho}(1 - \sqrt{1 - 2\rho r}), \frac{1}{\rho}(1 + \sqrt{1 - 2\rho r}))$ . Observe that the first value  $\frac{1}{\rho}(1 - \sqrt{1 - 2\rho r})$  is equal to  $c_{2.5}(\rho r)r = \delta_{\min}$ . Hence  $\phi(\delta_{\min} + \epsilon) > 0$  for  $0 < \epsilon < \frac{2}{\rho}\sqrt{1 - 2\rho r}$ , and Inequality 3 gives  $\|x - y\| > r$ .

In other words,  $y \notin \overline{\mathcal{B}}(x, r)$ . This being true for every  $y_0 \in \partial \mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon)$ , we have  $\mathcal{M}_0^x \cap \partial \mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon) = \emptyset$ .

*Step 2:* Let us deduce that  $\mathcal{M}_0^x \subseteq \mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min})$ . By contradiction, if a point  $z_0 \in \mathcal{M}_0$  with  $\|z - x\| > \delta_{\min}$  were to be in  $\mathcal{M}_0^x$ , it would be in the connected component of  $x_0$  in  $\mathcal{M}_0^x$ , since it is connected by Lemma 2.4. But since  $\mathcal{M}_0$  is a manifold, this would imply the existence of a continuous path from  $x_0$  to  $z_0$  in  $\mathcal{M}_0^x$ . But such a path would go through a sphere  $\partial \mathcal{B}_{\mathcal{M}_0}(x_0, \delta_{\min} + \epsilon)$ , which contradicts Step 1.  $\square$

The following proposition connects the normal reach to the usual notion of reach.

**Proposition 2.6.** *Suppose that  $u : \mathcal{M}_0 \rightarrow \mathcal{M} \subset E$  is an embedding. Let  $\tau > 0$  be the reach of  $\mathcal{M}$ . We have*

$$\tau = \frac{1}{\rho_*} \wedge \frac{1}{2}\lambda_*.$$

where  $\rho_*$  is the supremum of the operator norms of the second fundamental forms of  $\mathcal{M}_0$ , and  $\lambda_* = \inf_{x \in \mathcal{M}} \lambda(x)$ .

*Proof.* We first prove that  $\tau \geq \frac{1}{\rho_*} \wedge \frac{1}{2}\lambda_*$ . According to [AKC<sup>+</sup>19, Theorem 3.4], two cases may occur: the reach is either caused by a bottleneck or by curvature. In the first case, there exists  $x, y \in \mathcal{M}$  and  $z \in \text{med}(\mathcal{M})$  with  $\|x - y\| = 2\tau$  and  $\|x - z\| = \|y - z\| = \tau$ . We deduce that  $z - y \perp T_y \mathcal{M}$ , and that  $x - y \perp T_y \mathcal{M}$ . Hence by definition of  $\lambda(x)$ ,

$$\lambda(x) \leq \|x - y\| = 2\|x - z\| \leq 2\tau.$$

In the second case, there exists  $x \in \mathcal{M}$  and an arc-length parametrized geodesic  $\gamma : I \rightarrow \mathcal{M}$  such that  $\gamma(0) = x$  and  $\|\dot{\gamma}(0)\| = \frac{1}{\tau}$ . But  $\|\ddot{\gamma}(0)\| \leq \rho_*$ , hence  $\frac{1}{\tau} \leq \rho_*$ .

This disjunction shows that  $\tau \geq \frac{1}{\rho} \wedge \frac{1}{2}\lambda_{\min}$ .

We now prove that  $\tau \leq \frac{1}{\rho_*} \wedge \frac{1}{2}\lambda_*$ . The inequality  $\tau \leq \frac{1}{\rho_*}$  appears in [NSW08, Proposition 6.1]. To prove  $\tau \leq \frac{1}{2}\lambda_*$ , consider any  $x_0 \in \mathcal{M}_0$ . Let  $y_0 \in \Lambda(x_0)$  such that  $\|x - y\|$  is minimal. Using Theorem 2.1 and the property  $x - y \perp T_y \mathcal{M}$ , we immediately have

$$\tau \leq \frac{\|x - y\|^2}{2d(y - x, T_y \mathcal{M})} = \frac{\|x - y\|}{2} = \frac{\lambda(x)}{2}. \quad \square$$

In the case where  $u$  is not an embedding,  $\mathcal{M}$  may have reach 0. However, as shown by the following theorem, the normal reach gives a scale at which  $\mathcal{M}$  still behaves well. Note that we will not make use of this result in the sequel of the paper.

**Theorem 2.7.** *Let  $x \in \mathcal{M}_0$  and  $r < \frac{1}{4\rho} \wedge \lambda(x)$ . Then  $\overline{\mathcal{B}}(x, r) \cap \mathcal{M}$  is a set of reach at least  $\frac{1-2\rho r}{\rho}$ . In particular, it is greater than  $\frac{1}{2\rho}$ .*

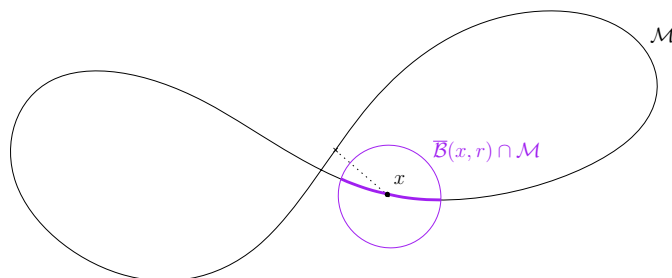


Figure 17: The set  $\overline{\mathcal{B}}(x, r) \cap \mathcal{M}$  has positive reach

*Proof.* Denote  $\mathcal{M}^x = \overline{\mathcal{B}}(x, r) \cap \mathcal{M}$  and  $\mathcal{M}_0^x = u^{-1}(\mathcal{M}^x)$ .

*Step 1:* Let us prove that for every  $y_0, z_0 \in \mathcal{M}_0^x$ ,

$$d(z - y, T_y \mathcal{M}) \leq \frac{\rho}{2(1 - 2\rho r)} \|z - y\|^2.$$

Let  $y_0, z_0 \in \mathcal{M}_0^x$ , and  $\delta = d_{\mathcal{M}_0}(y_0, z_0)$ . Lemma 2.2 Point 3 gives  $\delta \leq \frac{1}{1 - \frac{\rho}{2}\delta} \|y - z\|$ . Moreover,  $\delta \leq d_{\mathcal{M}_0}(y_0, x_0) + d_{\mathcal{M}_0}(x_0, z_0) \leq 2c_{2.5}(\rho r)r$ . Hence,

$$\frac{1}{1 - \frac{\rho}{2}\delta} \leq \frac{1}{1 - c_{2.5}(\rho r)\rho r} = \frac{1}{\sqrt{1 - 2\rho r}},$$

and we deduce that

$$\delta \leq \frac{1}{\sqrt{1 - 2\rho r}} \|y - z\|. \quad (4)$$

Besides, Lemma 2.2 Point 2 gives  $d(z - y, T_y \mathcal{M}) \leq \frac{\rho}{2}\delta^2$ , and combining these two inequalities yields  $d(z - y, T_y \mathcal{M}) \leq \frac{\rho}{2(1 - 2\rho r)} \|z - y\|^2$ .

*Step 2:* Let us prove that

$$d(z - y, \text{Tan}(\mathcal{M}^x, y)) \leq \frac{\rho}{2(1 - 2\rho r)} \|z - y\|^2,$$

where  $\text{Tan}(\mathcal{M}^x, y)$  is the tangent cone at  $y$  of the closed set  $\mathcal{M}^x$ .

If  $y \in \mathcal{B}(x, r)$ , then  $\text{Tan}(\mathcal{M}^x, y) = T_y \mathcal{M}$ , and the inequality follows from Step 1. Otherwise, suppose that  $y \in \partial \mathcal{B}(x, r)$  and that  $z \neq y$ . Let  $\delta = d_{\mathcal{M}_0}(y_0, z_0)$ . According to Equation 4, the inequality  $\|y - z\| \leq 2r$  and the assumption  $r < \frac{1}{4\rho}$ , we have  $\delta < \frac{1}{\rho}$ . Consider a length-minimizing geodesic  $\gamma_0 : [0, \delta] \rightarrow \mathcal{M}_0$  from  $y_0$  to  $z_0$ , and denote  $v = \dot{\gamma}(0)$ . Let us show that  $v \in \text{Tan}(\mathcal{M}^x, y)$ , and we will conclude with Step 1.

Since  $\mathcal{M}^x = \overline{\mathcal{B}}(x, r) \cap \mathcal{M}$ ,  $v \in \text{Tan}(\mathcal{M}^x, y)$  is implied by  $\langle v, y - x \rangle < 0$ . Suppose by contradiction that  $\langle v, y - x \rangle \geq 0$ . Hence, according to Lemma 2.3 Point 1, with  $l = r < \frac{1}{2\rho}$ , we have  $T_1 = \frac{2}{\rho} \sqrt{1 - \rho l} > \frac{\sqrt{2}}{\rho} > \delta$ , and

$$\|z - x\| = \|\gamma(\delta) - x\| > \|\gamma(0) - x\| = \|y - x\| = r.$$

We deduce the contradiction  $z \notin \overline{\mathcal{B}}(x, r)$ .

To conclude the proof, it follows from Theorem 2.1 that  $\mathcal{M}^x$  has reach at least  $\frac{1-2\rho r}{\rho}$ .  $\square$

## 2.4 Probabilistic bounds under normal reach conditions

We now consider  $\mathcal{M}_0$  and  $\mu_0$  which satisfy the hypotheses 1, 2 and 3. The aim of this subsection is to provide a quantitative control of  $\mu = u_* \mu_0$  (Propositions 2.11 and 2.12). To do so, Lemmas 2.8 and 2.9 bound differential quantities related to the exponential map of  $\mathcal{M}_0$ , denoted  $\exp_{x_0}^{\mathcal{M}_0} : T_{x_0} \mathcal{M}_0 \rightarrow \mathcal{M}_0$ . We shall then consider the pull-back of  $\mu$  on the tangents spaces  $T_x \mathcal{M}$ , where it is simpler to compute integrals (Lemma 2.10).

An application of the coarea formula shows that  $\mu$  admits  $f$  as a density against  $\mathcal{H}_{\mathcal{M}}^d$ , the  $d$ -dimensional Hausdorff measure restricted to  $\mathcal{M}$ , where  $f(x) = \sum_{x_0 \in u^{-1}(\{x\})} f_0(x_0)$ . In particular, if  $x$  has only one preimage by  $u$ —i.e. if  $\lambda(x) > 0$ —then  $f(x) = f_0 \circ u^{-1}(x)$ .

*Remark 2.2.* Recall that the density  $f_0$  is  $L_0$ -Lipschitz with respect to the geodesic distance: for all  $x_0, y_0 \in \mathcal{M}_0$ ,

$$|f_0(x_0) - f_0(y_0)| \leq L_0 d_{\mathcal{M}_0}(x_0, y_0).$$

We can deduce the following: for all  $x_0, y_0 \in \mathcal{M}_0$  such that  $\|x - y\| < \frac{1}{2\rho} \wedge \lambda(x)$ , we have

$$|f(x) - f(y)| \leq L \|x - y\|$$

with  $L = 2L_0$ . To prove this, assume that  $y$  has only one preimage by  $u$ . Hence we can write  $|f(x) - f(y)| = |f_0 \circ u^{-1}(x) - f_0 \circ u^{-1}(y)| \leq L_0 d_{\mathcal{M}_0}(u^{-1}(x), u^{-1}(y)) \leq 2L_0 \|x - y\|$ , where we used Lemma 2.5 on the last inequality. Now we prove that  $\|x - y\| < \frac{1}{2\rho} \wedge \lambda(x)$  implies that  $y$  has only one preimage. Let  $r = \|x - y\|$ , and suppose by contradiction that  $y_0, y_1$  are two distinct preimages. According to Remark 2.1,  $d_{\mathcal{M}_0}(y_0, y_1) \geq \frac{2}{\rho}$ . But Lemma 2.5 says that  $u^{-1}(\mathcal{B}(x, r)) \subseteq \mathcal{B}_{\mathcal{M}_0}(x_0, 2r) \subseteq \mathcal{B}_{\mathcal{M}_0}(x_0, \frac{1}{\rho})$ , which contradicts  $d_{\mathcal{M}_0}(y_0, y_1) \geq \frac{2}{\rho}$ .

**Lemma 2.8** ([Aam18], Proposition C.1). *Let  $x_0 \in \mathcal{M}_0$ . The exponential map  $\exp_{x_0}^{\mathcal{M}_0}$  is injective on  $\mathcal{B}_{T_{x_0}}(0, \frac{\pi}{\rho})$  (and is a diffeomorphism onto its image). Let  $r \leq \frac{\pi}{2\sqrt{2}\rho}$ . For all  $v \in \mathcal{B}_{T_{x_0} \mathcal{M}_0}(x_0, r)$ ,*

*the Jacobian  $J_v = \sqrt{\det((d_v \exp_{x_0}^{\mathcal{M}_0})^t (d_v \exp_{x_0}^{\mathcal{M}_0}))}$  of  $\exp_{x_0}^{\mathcal{M}_0}$  is bounded by*

$$\left(1 - \frac{(r\rho)^2}{6}\right)^d \leq J_v \leq (1 + (r\rho)^2)^d.$$

We obtain the following lemma as an immediate consequence.



**Lemma 2.9.** Let  $x_0 \in \mathcal{M}_0$ . Define the map  $\exp_x^{\mathcal{M}} = u \circ \exp_{x_0}^{\mathcal{M}_0} \circ (d_{x_0}u)^{-1} : T_x\mathcal{M} \rightarrow \mathcal{M}$ . Moreover, define  $\overline{\exp}_x^{\mathcal{M}}$  to be the map  $\exp_x^{\mathcal{M}}$  restricted to  $\overline{\mathcal{B}}_{T_x\mathcal{M}}(0, \frac{\pi}{\rho})$ . It is injective.

Let  $r < \frac{1}{2\rho} \wedge \lambda(x)$ . Denote  $\overline{\mathcal{B}} = \overline{\mathcal{B}}(x, r)$  and  $\overline{\mathcal{B}}^T = (\overline{\exp}_x^{\mathcal{M}})^{-1}(\overline{\mathcal{B}})$ . These maps fit in the following commutative diagrams:

$$\begin{array}{ccc} \mathcal{M}_0 & \xrightarrow{u} & \mathcal{M} \\ \exp_{x_0}^{\mathcal{M}_0} \uparrow & & \uparrow \exp_x^{\mathcal{M}} \\ T_{x_0}\mathcal{M}_0 & \xrightarrow{d_{x_0}u} & T_x\mathcal{M} \end{array} \qquad \begin{array}{ccc} \overline{\mathcal{B}} \cap \mathcal{M} & & \\ \uparrow \overline{\exp}_x^{\mathcal{M}} & & \\ \overline{\mathcal{B}}^T \subset T_x\mathcal{M} & & \end{array}$$

We have  $\overline{\mathcal{B}}_{T_x\mathcal{M}}(0, r) \subseteq \overline{\mathcal{B}}^T \subseteq \overline{\mathcal{B}}_{T_x\mathcal{M}}(0, c_{2.5}(\rho)r)$ . For all  $v \in \overline{\mathcal{B}}^T$ , the Jacobian of  $\overline{\exp}_x^{\mathcal{M}}$ ,  $J_v = \sqrt{\det((d_v\overline{\exp}_x^{\mathcal{M}})^t(d_v\overline{\exp}_x^{\mathcal{M}}))}$ , is bounded by

$$\left(1 - \frac{(r\rho)^2}{6}\right)^d \leq J_v \leq (1 + (r\rho)^2)^d,$$

and these terms are bounded by  $J_{\min} = (\frac{23}{24})^d$  and  $J_{\max} = (\frac{5}{4})^d$ .

**Lemma 2.10.** Let  $x_0 \in \mathcal{M}_0$  and  $r < \frac{1}{2\rho} \wedge \lambda(x)$ . Consider the map  $\overline{\exp}_x^{\mathcal{M}}$  defined in Lemma 2.9. Consider  $\mu_x$ , the measure  $\mu$  restricted to  $\overline{\mathcal{B}}(x, r)$ , and define  $\nu_x = (\overline{\exp}_x^{\mathcal{M}})_*^{-1}\mu_x$ . The measure  $\nu_x$  admits a density  $g$  against the  $d$ -dimensional Hausdorff measure  $\mathcal{H}^d$  on  $T_x\mathcal{M} \subset E$ , where

$$g(v) = f(\overline{\exp}_x^{\mathcal{M}}(v)) \cdot J_v \cdot 1_{\overline{\mathcal{B}}^T}(v).$$

For all  $v \in \overline{\mathcal{B}}^T$ , the map  $g$  satisfies  $|g(v) - g(0)| \leq c_{2.10}r$ , where  $c_{2.10} = (4L_0J_{\max} + \frac{d}{2}\rho f_{\max})r$ .

*Proof.* The expression of  $g$  comes from the area formula ([Fed14, Theorem 3.2.5]). To prove the inequality, observe that we can decompose

$$\begin{aligned} g(v) - g(0) &= f(\exp_x^{\mathcal{M}}(v))J_v - f(\exp_x^{\mathcal{M}}(0))J_0 \\ &= [f(\exp_x^{\mathcal{M}}(v)) - f(\exp_x^{\mathcal{M}}(0))]J_v + (J_v - J_0)f(\exp_x^{\mathcal{M}}(0)) \end{aligned}$$

On the one hand, using Remark 2.2, we get

$$\begin{aligned} |f(\overline{\exp}_x^{\mathcal{M}}(v)) - f(\overline{\exp}_x^{\mathcal{M}}(0))| &\leq L\|\overline{\exp}_x^{\mathcal{M}}(v) - \overline{\exp}_x^{\mathcal{M}}(0)\| \\ &= L\|u \circ \exp_{x_0}^{\mathcal{M}_0}(v) - u \circ \exp_{x_0}^{\mathcal{M}_0}(0)\| \\ &\leq L \cdot d_{\mathcal{M}_0}(\exp_{x_0}^{\mathcal{M}_0}(v), x_0) = L\|v\|. \end{aligned}$$

On the other hand,  $J_0 = 1$  and  $(1 - \frac{(r\rho)^2}{6})^d \leq J_v \leq (1 + (r\rho)^2)^d$  yield  $|J_v - J_0| \leq d(\rho r)^2 \leq \frac{d}{2}\rho r$ . We eventually obtain

$$g(v) - g(0) \leq L\|v\|J_{\max} + f_{\max}\frac{d}{2}\rho r \leq (2LJ_{\max} + f_{\max}\frac{d}{2}\rho)r. \quad \square$$

*Remark 2.3.* In the same vein as Lemma 2.10, we can define  $\overline{\exp}_{x_0}^{\mathcal{M}_0}$  to be the map  $\exp_{x_0}^{\mathcal{M}_0}$  restricted to  $\overline{\mathcal{B}}_{T_{x_0}\mathcal{M}_0}(0, \frac{\pi}{\rho})$ . For any  $x_0 \in \mathcal{M}_0$ , let  $\mu_0^{x_0}$  be the measure  $\mu_0$  restricted to  $\overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, \frac{1}{\rho})$ , and define the measure  $\nu_0 = (\overline{\exp}_{x_0}^{\mathcal{M}_0})^{-1}\mu_0^{x_0}$ . Then  $\nu_0$  admits a density  $g_0$  over the  $d$ -dimensional Hausdorff measure  $\mathcal{H}^d$  on  $T_{x_0}\mathcal{M}_0$ , where

$$g_0(v) = f_0(\overline{\exp}_{x_0}^{\mathcal{M}_0}(v)) \cdot J_v \cdot 1_{\overline{\mathcal{B}}_{T_{x_0}\mathcal{M}_0}(0, \frac{1}{\rho})}(v).$$

Using the density  $g$ , we can derive explicit bounds on  $\mu$ .

**Proposition 2.11.** Let  $x_0 \in \mathcal{M}_0$ ,  $r \leq \frac{1}{2\rho} \wedge \lambda(x)$  and  $s \in [0, r]$ . We have

- $\mu(\overline{\mathcal{B}}(x, r)) \geq ar^d$
- $|\frac{\mu(\overline{\mathcal{B}}(x, r))}{V_d r^d} - f(x)| \leq c_{2.11}r$
- $\mu(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s)) \leq br^{d-1}(r - s)$

with  $a = f_{\min} J_{\min} V_d$ ,  $c_{2.11} = c_{2.10} + f_{\max} J_{\max} d 2^d \rho$  and  $b = d 2^d f_{\max} J_{\max} V_d$ .

*Proof.* Consider the map  $\overline{\text{exp}}_x^{\mathcal{M}}$  and the measure  $\nu_x$  as defined in Lemma 2.10. In the following, we write  $T = T_x \mathcal{M}$ , and  $\overline{\mathcal{B}}^T = (\overline{\text{exp}}_x^{\mathcal{M}})^{-1}(\overline{\mathcal{B}}(x, r))$ .

*Point (1):* We have  $\mu(\overline{\mathcal{B}}(x, r)) = \nu_x(\overline{\mathcal{B}}^T)$ . Writing down the density yields

$$\begin{aligned} \nu_x(\overline{\mathcal{B}}^T) &= \int_{\overline{\mathcal{B}}^T} g(v) d\mathcal{H}^d(v) \geq \int_{\overline{\mathcal{B}}^T} f_{\min} J_{\min} d\mathcal{H}^d(v) \\ &= f_{\min} J_{\min} \mathcal{H}^d(\overline{\mathcal{B}}^T). \end{aligned}$$

And since  $\overline{\mathcal{B}}^T \supset \overline{\mathcal{B}}_T(0, r)$ , we get  $\nu_x(\overline{\mathcal{B}}^T) \geq f_{\min} J_{\min} V_d r^d$ .

*Point (2):* Observe that  $\int_{\overline{\mathcal{B}}_T(0, r)} f(x) d\mathcal{H}^d(v) = f(x) V_d r^d$ . Hence

$$\begin{aligned} &|\mu(\overline{\mathcal{B}}(x, r)) - f(x) V_d r^d| \\ &= \left| \int_{\overline{\mathcal{B}}^T} g(v) d\mathcal{H}^d(v) - \int_{\overline{\mathcal{B}}_T(0, r)} f(x) d\mathcal{H}^d(v) \right| \\ &\leq \underbrace{\left| \int_{\overline{\mathcal{B}}_T(0, r)} (f(x) - g(v)) d\mathcal{H}^d(v) \right|}_{(1)} + \underbrace{\left| \int_{\overline{\mathcal{B}}^T \setminus \overline{\mathcal{B}}_T(0, r)} g(v) d\mathcal{H}^d(v) \right|}_{(2)} \end{aligned}$$

To bound Term (1), notice that  $g(0) = f(\text{exp}_x^{\mathcal{M}}(0)) J_0 = f(x)$ . Hence we can write:

$$\left| \int_{\overline{\mathcal{B}}_T(0, r)} (f(x) - g(v)) d\mathcal{H}^d(v) \right| \leq \int_{\overline{\mathcal{B}}_T(0, r)} |g(0) - g(v)| d\mathcal{H}^d(v).$$

Lemma 2.10 gives  $|g(v) - g(0)| \leq c_{2.10}r$ , and we obtain  $\left| \int_{\overline{\mathcal{B}}_T(0, r)} (f(x) - g(v)) d\mathcal{H}^d(v) \right| \leq c_{2.10}r V_d r^d$ .

On the other hand, we bound Term (2) thanks to the inclusion  $\overline{\mathcal{B}}^T \subseteq \overline{\mathcal{B}}_T(0, c_{2.5}(\rho r)r)$ :

$$\begin{aligned} \int_{\overline{\mathcal{B}}^T \setminus \overline{\mathcal{B}}_T(0, r)} g(v) d\mathcal{H}^d(v) &\leq \int_{\overline{\mathcal{B}}_T(0, c_{2.5}(\rho r)r) \setminus \overline{\mathcal{B}}_T(0, r)} g(v) d\mathcal{H}^d(v) \\ &\leq f_{\max} J_{\max} (\mathcal{H}^d(\overline{\mathcal{B}}_T(0, c_{2.5}(\rho r)r) \setminus \overline{\mathcal{B}}_T(0, r))) \\ &= f_{\max} J_{\max} V_d (c_{2.5}(\rho r)^d - 1) r^d \end{aligned}$$

Now we use  $c_{2.5}(\rho r) \leq 1 + 2\rho r \leq 2$  and the inequality  $A^d - 1 \leq d(A - 1)A^{d-1}$ , where  $A \geq 1$ , to get

$$\begin{aligned} (c_{2.5}(\rho r)^d - 1) &\leq d \cdot (c_{2.5}(\rho r) - 1) \cdot c_{2.5}(\rho r)^{d-1} \\ &\leq d \cdot 2\rho r \cdot 2^{d-1}. \end{aligned}$$

We deduce that  $\left| \int_{\overline{\mathcal{B}}^T \setminus \overline{\mathcal{B}}_T(x, r)} g(v) d\mathcal{H}^d(v) \right| \leq f_{\max} J_{\max} V_d r^d d \cdot \rho r 2^d$ . To conclude, we obtained

$$|\mu(\overline{\mathcal{B}}(x, r)) - f(x) V_d r^d| \leq r(c_{2.10} + f_{\max} J_{\max} d \rho 2^d) V_d r^d.$$

*Point (3):* Let us write

$$\begin{aligned} \mu(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s)) &= \nu_x((\overline{\text{exp}}_x^{\mathcal{M}})^{-1}(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s))) \\ &= \int_{(\overline{\text{exp}}_x^{\mathcal{M}})^{-1}(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s))} g(v) d\mathcal{H}^d(v). \end{aligned}$$

In spherical coordinates, this integral reads

$$\int_{(\overline{\exp_x \mathcal{M}})^{-1}(\overline{\mathcal{B}(x,r)} \setminus \overline{\mathcal{B}(x,s)})} g(v) d\mathcal{H}^d(v) = \int_{v \in \partial \mathcal{B}_T(0,1)} \int_{t=a(v)}^{b(v)} g(tv) t^{d-1} dt dv,$$

where  $a$  and  $b$  are defined as follows: for every  $v \in \partial \mathcal{B}_T(0,1) \subset T_x \mathcal{M}$ , let  $\gamma_0$  be a arc-length parametrized geodesic with  $\gamma(0) = x$  and  $\dot{\gamma}(0) = v$ , and set  $a(v)$  and  $b(v)$  to be the first positive values such that  $\|\gamma(a(v)) - x\| = s$  and  $\|\gamma(b(v)) - x\| = r$ .

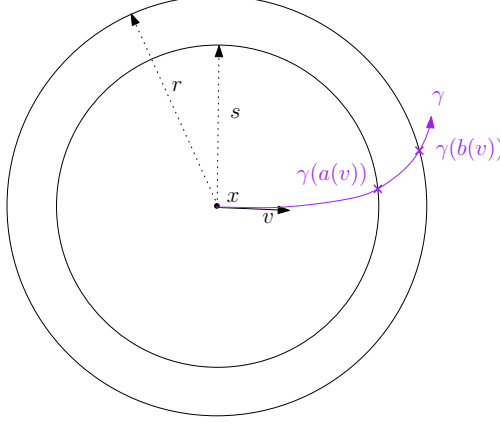


Figure 18: Illustration of  $a(v)$  and  $b(v)$

Let us show that  $b(v) - a(v) \leq \frac{1}{1-\rho r}(r - s)$ . Consider the application  $\phi : t \mapsto \|\gamma(t) - x\|^2$ . According to Lemma 2.3 Point 3 with  $l = 0$ , we have  $\ddot{\phi}(t) \geq 2(1 - \rho r)$  for  $t \in [0, b(v)]$ . It follows that  $\dot{\phi}(t) \geq 2(1 - \rho r)t$ , and that

$$\begin{aligned} \phi(b(v)) - \phi(a(v)) &= \int_{a(v)}^{b(v)} \dot{\phi}(t) dt \geq \int_{a(v)}^{b(v)} 2(1 - \rho r)t dt \\ &= (1 - \rho r)(b(v)^2 - a(v)^2). \end{aligned}$$

Since  $r^2 - s^2 = \phi(b(v)) - \phi(a(v))$ , we deduce that

$$r^2 - s^2 \geq (1 - \rho r)(b(v)^2 - a(v)^2). \quad (5)$$

Writing  $r^2 - s^2 = (r + s)(r - s)$  and  $b(v)^2 - a(v)^2 = (b(v) + a(v))(b(v) - a(v))$  leads to

$$(r - s) \frac{1}{1 - \rho r} \frac{r + s}{b(v) + a(v)} \geq b(v) - a(v).$$

But  $b(v) + a(v) \geq r + s$ , hence  $(r - s) \frac{1}{1 - \rho r} \geq b(v) - a(v)$ , as wanted.

Now, notice that we have  $b(v) \leq 2r$ . Indeed,  $b < \frac{1}{\rho}$  by Lemma 2.3 Point 5, and we conclude with Lemma 2.2 Point 2. We then write

$$\begin{aligned} \int_{t=a(v)}^{b(v)} g(tv) t^{d-1} dt &\leq \int_{t=a(v)}^{b(v)} f_{\max} J_{\max}(2r)^{d-1} dt \\ &\leq \frac{1}{1 - \rho r} (r - s) f_{\max} J_{\max}(2r)^{d-1}. \end{aligned}$$

and we deduce

$$\begin{aligned} \int_{v \in \partial \mathcal{B}(0,1)} \int_{t=a(v)}^{b(v)} g(tv) t^{d-1} dt dv &\leq \frac{1}{1 - \rho r} (r - s) f_{\max} J_{\max}(2r)^{d-1} \int_{v \in \partial \mathcal{B}(0,1)} dv \\ &= \frac{1}{1 - \rho r} (r - s) f_{\max} J_{\max}(2r)^{d-1} \cdot dV_d. \end{aligned}$$

We obtained

$$\mu(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s)) = \frac{2^{d-1} d V_d f_{\max} J_{\max}}{1 - \rho r} (r - s) r^{d-1},$$

and we conclude with  $r \leq \frac{1}{2\rho}$ :  $\mu(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s)) = 2^d d V_d f_{\max} J_{\max} (r - s) r^{d-1}$ .  $\square$

The following proposition gives probability bounds without normal reach conditions.

**Proposition 2.12.** *Let  $x_0 \in \mathcal{M}_0$ ,  $r \leq \frac{1}{2\rho}$  and  $s \in [0, r]$ . We have*

- $\mu(\overline{\mathcal{B}}(x, r)) \geq ar^d$
- $\mu(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s)) \leq b' r^{d-\frac{1}{2}} (r - s)^{\frac{1}{2}}$

with  $a = f_{\min} J_{\min} V_d$  and  $b' = \frac{f_{\max} J_{\max}}{f_{\min} J_{\min}} \left( \frac{\rho}{\sqrt{4-\sqrt{13}}} \right)^d d 2^{2d} \sqrt{3}$ .

*Proof.* Let  $\mathcal{M}^x = \mathcal{M} \cap \overline{\mathcal{B}}(x, r)$  and  $\mathcal{M}_0^x = u^{-1}(\mathcal{M}^x)$ . Lemma 2.5 does not apply: it is not true that  $\mathcal{M}_0^x \subseteq \overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, c_{2.5}(\rho)r)$ . However, we can decompose  $\mathcal{M}_0^x$  in connected components  $C_0^i, i \in I$ .

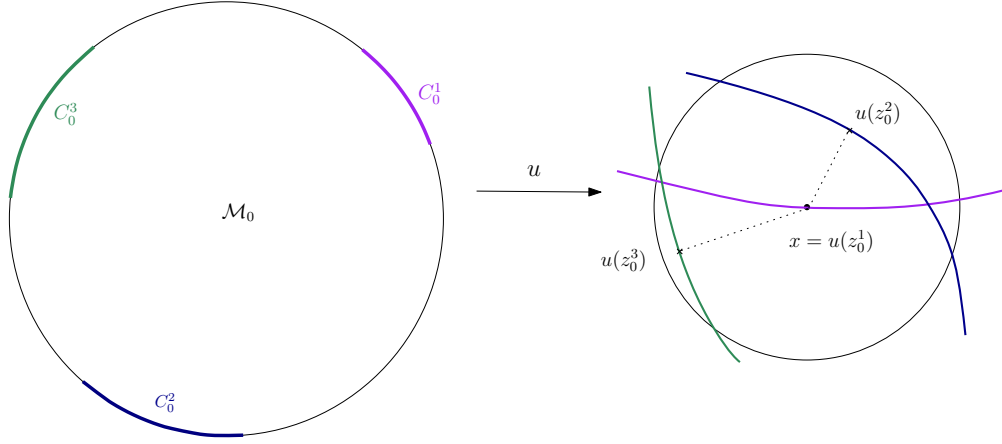


Figure 19: The connected components  $C_0^i$

For every  $i \in I$ , let  $z_0^i$  be a minimizer of  $z_0 \mapsto \|z - x\|$  on  $C_0^i$ . We have  $x - z^i \perp T_{z^i} \mathcal{M}$ , hence according to Lemma 2.3 Point 5,  $C_0^i \subseteq \overline{\mathcal{B}}_{\mathcal{M}_0}(z_0^i, \frac{1}{\rho})$ . For all  $i \in I$ , consider  $\mu_0^i$ , the measure  $\mu_0$  restricted to  $C_0^i$ , and define  $\nu_0^i = (\overline{\exp}_{z_0^i}^{\mathcal{M}_0})_*^{-1} \mu_0^i$ , as in Remark 2.3. The measure  $\nu_0^i$  admits  $g_0^i$  as a density over the  $d$ -dimensional Hausdorff measure on  $T_{z_0^i} \mathcal{M}_0$ , where

$$g_0^i(v) = f_0(\overline{\exp}_{z_0^i}^{\mathcal{M}_0}(v)) \cdot J_v \cdot 1_{(\overline{\exp}_{z_0^i}^{\mathcal{M}_0})^{-1}(C_0^i)}(v).$$

*Point (1):* We can write

$$\mu(\overline{\mathcal{B}}(x, r)) = \mu_0(u^{-1}(\overline{\mathcal{B}}(x, r))) = \sum_{i \in I} \mu_0(C_0^i).$$

Let  $i_* \in I$  be the index of the connected component of  $\mathcal{M}_0^x$  which contains  $x_0$ . We have  $C_0^{i_*} \supset \overline{\mathcal{B}}_{\mathcal{M}_0}(x_0, r)$ , and we deduce that

$$\begin{aligned} \mu_0(C_0^{i_*}) &\geq \int_{(\overline{\exp}_{z_0}^{\mathcal{M}_0})^{-1}(C_0^{i_*})} g_0^{i_*} d\mathcal{H}^d \\ &\geq f_{\min} J_{\min} \mathcal{H}^d((\overline{\exp}_{z_0}^{\mathcal{M}_0})^{-1}(C_0^{i_*})) = f_{\min} J_{\min} V_d r^d. \end{aligned}$$

Therefore,  $\mu(\overline{\mathcal{B}}(x, r)) \geq f_{\min} J_{\min} V_d r^d$ .

*Point (2):* We now prove the second point.

*Step 1:* Let us show that the cardinal of  $I$  is lower than  $\frac{1}{f_{\min} J_{\min} V_d} (\frac{2\rho}{\alpha})^d$ , with  $\alpha = \sqrt{4 - \sqrt{13}}$ . We shall prove that for every  $i, j \in I$  such that  $i \neq j$ ,  $d_{\mathcal{M}_0}(z_0^i, z_0^j) \geq \frac{\alpha}{\rho}$ .

Let  $\gamma_0$  be a geodesic from  $z_0^i$  to  $z_0^j$ , with  $\gamma(0) = z^i$ ,  $\gamma(T) = z^j$ , and  $\dot{\gamma}_0(0) = v_0$ . Consider the application  $\phi : t \mapsto \|\gamma(t) - x\|^2$ . Since  $C_0^i$  and  $C_0^j$  are disjoint connected components, there must be a  $t^* < T$  such that  $\|\gamma(t^*) - x_0\| > r$ . Moreover, according to Lemma 2.3 Point 2,  $\phi$  is increasing on  $[0, T_2]$  where  $T_2 = \frac{\sqrt{2}}{\rho} \sqrt{2 - \sqrt{3 + \rho^2 l^2}}$ . Since  $\phi(T) \leq r$ , we deduce that  $T$  is greater than  $T_2$ . Note that the assumption  $r \leq \frac{1}{2\rho}$  yields  $T_2 \geq \frac{\alpha}{\rho}$ .

This implies that the geodesic balls  $\mathcal{B}_{\mathcal{M}_0}(z_0^i, \frac{\alpha}{2\rho})$  are disjoint. Therefore,

$$1 \geq \mu_0\left(\bigcup_i \mathcal{B}_{\mathcal{M}_0}(z_0^i, \frac{\alpha}{2\rho})\right) \geq |I| f_{\min} J_{\min} V_d \left(\frac{\alpha}{2\rho}\right)^d,$$

and we deduce  $|I| \leq \frac{1}{f_{\min} J_{\min} V_d} (\frac{2\rho}{\alpha})^d$ .

*Step 2:* Let  $i \in I$ , and define  $D_0^i = C_0^i \cap u^{-1}(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s))$ . Let us show that

$$\mu_0(D_0^i) \leq f_{\max} J_{\max} 2^{d-1} \sqrt{6} d V_d \cdot r^{d-1} \sqrt{r^2 - s^2}.$$

Let us distinguish two cases:  $l_i \geq s$  or  $l_i < s$ .

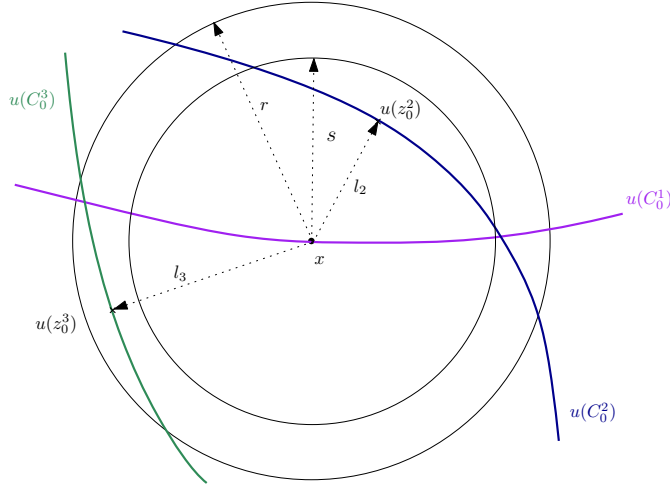


Figure 20: Illustration of the cases  $l_i \geq s$  and  $l_i < s$

First, assume that  $l_i < s$ . Let  $\gamma$  be a geodesic starting from  $z_0^i$ , denote  $v = \dot{\gamma}(0)$  and consider the application  $\phi : t \mapsto \|\gamma(t) - x\|^2$ . Let  $a(v), b(v)$  be the first values of  $t \geq 0$  such that  $\|\gamma(t) - x\| = s$  and  $\|\gamma(t) - x\| = r$ . As in the proof of Proposition 2.11 Point 3, we still have Equation 5:

$$r^2 - s^2 \geq (1 - \rho r)(b(v)^2 - a(v)^2),$$

from which we deduce  $b(v) - a(v) \leq \frac{1}{1 - \rho r} \frac{1}{b(v) + a(v)} (r^2 - s^2)$ . According to Lemma 2.3 Point 4,  $b(v) + a(v) \geq b(v) \geq (1 + \rho r)^{-\frac{1}{2}} \sqrt{r^2 - l_i^2} \geq (1 + \rho r)^{-\frac{1}{2}} \sqrt{r^2 - s^2}$ , and we obtain

$$b(v) - a(v) \leq \frac{(1 + \rho r)^{\frac{1}{2}}}{1 - \rho r} \sqrt{r^2 - s^2}. \quad (6)$$

Now, we write

$$\mu_0(D_0^i) = \nu_0^i((\exp_{z_0^i}^{\mathcal{M}_0})^{-1}(D_0^i)).$$

In spherical coordinates, this measure reads

$$\int_{(\overline{\text{exp}}_{z_0^i} \mathcal{M}_0)^{-1}(D_0^i)} g_0^i(y) d\mathcal{H}^d(y) = \int_{v \in \partial \mathcal{B}(0,1)} \int_{t=a(v)}^{b(v)} g_0^i(tv) t^{d-1} dt dv.$$

We can now conclude as in the proof of Proposition 2.11 Point 3. We still have  $b(v) \leq 2r$ , and we write

$$\int_{t=a(v)}^{b(v)} g_0^i(tv) t^{d-1} dt \leq \int_{t=a(v)}^{b(v)} f_{\max} J_{\max}(2r)^{d-1} dt.$$

Using Equation 6, we obtain

$$\int_{t=a(v)}^{b(v)} f_{\max} J_{\max}(2r)^{d-1} dt \leq \frac{(1+\rho r)^{\frac{1}{2}}}{1-\rho r} \sqrt{r^2 - s^2} f_{\max} J_{\max}(2r)^{d-1}.$$

Therefore,

$$\int_{v \in \partial \mathcal{B}(0,1)} \int_{t=a(v)}^{b(v)} g_0^i(tv) t^{d-1} dt dv \leq \frac{(1+\rho r)^{\frac{1}{2}}}{1-\rho r} \sqrt{r^2 - s^2} f_{\max} J_{\max}(2r)^{d-1} dV_d.$$

The assumption  $r < \frac{1}{2\rho}$  yields  $\frac{(1+\rho r)^{\frac{1}{2}}}{1-\rho r} < \sqrt{6}$ , and we finally obtain

$$\mu_0(D_0^i) \leq f_{\max} J_{\max} 2^{d-1} \sqrt{6} dV_d \cdot r^{d-1} \sqrt{r^2 - s^2}.$$

Now, assume that  $l_i \geq s$ . This case is similar to the first one. One has

$$\mu_0(D_0^i) \leq \int_{(\overline{\text{exp}}_{z_0^i} \mathcal{M}_0)^{-1}(D_0^i)} g_0^i(y) d\mathcal{H}^d(y) = \int_{v \in \partial \mathcal{B}(0,1)} \int_{t=0}^{b(v)} g_0^i(tv) t^{d-1} dt dv.$$

and Lemma 2.3 Point 5 gives  $b(v) \leq (\frac{1-\rho r}{2})^{-\frac{1}{2}} \sqrt{r^2 - l^2} \leq (\frac{1-\rho r}{2})^{-\frac{1}{2}} \sqrt{r^2 - s^2}$ . Note that  $(\frac{1-\rho r}{2})^{-\frac{1}{2}}$  is not greater than 2 when  $r < \frac{1}{2\rho}$ . One deduces that

$$\mu_0(D_0^i) \leq f_{\max} J_{\max} 2^{d-1} 2 dV_d \cdot r^{d-1} \sqrt{r^2 - s^2}.$$

*Step 3:* We conclude: Since  $u^{-1}(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s)) = \bigcup_i D_0^i$ , Step 1 and 2 yield

$$\begin{aligned} \mu(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s)) &= \sum_{i \in I} \mu_0(D_0^i) \leq |I| f_{\max} J_{\max} 2^{d-1} \sqrt{6} dV_d \cdot r^{d-1} \sqrt{r^2 - s^2} \\ &\leq \frac{1}{f_{\min} J_{\min} V_d} \left(\frac{2\rho}{\alpha}\right)^d f_{\max} J_{\max} 2^{d-1} \sqrt{6} dV_d \cdot r^{d-1} \sqrt{r^2 - s^2}. \end{aligned}$$

Finally, the inequality  $\sqrt{r^2 - s^2} \leq \sqrt{2r} \sqrt{r - s}$  yields

$$\mu(\overline{\mathcal{B}}(x, r) \setminus \overline{\mathcal{B}}(x, s)) \leq \frac{f_{\max} J_{\max}}{f_{\min} J_{\min}} \left(\frac{\rho}{\alpha}\right)^d d 2^{2d} \sqrt{3} r^{d-\frac{1}{2}} \sqrt{r - s}. \quad \square$$

### 3 Tangent space estimation

In this section, we show that one can estimate the tangent spaces of  $\mathcal{M}$  based on a sample of it, via the computation of local covariance matrices. We study the consistency of this estimation in Subsection 3.2, which is based on the results of the last section. In Subsection 3.3 we prove that this estimation is stable, based on lighter hypotheses.

### 3.1 Local covariance matrix and $\check{\nu}$

**Definition 3.1.** Let  $\nu$  be any probability measure on  $E$ . Let  $r > 0$  and  $x \in \text{supp}(\nu)$ . The *local covariance matrix of  $\nu$  around  $x$  at scale  $r$*  is the following matrix:

$$\Sigma_\nu(x) = \int_{\overline{\mathcal{B}}(x,r)} (x-y)^{\otimes 2} \frac{d\nu(y)}{\nu(\overline{\mathcal{B}}(x,r))}$$

We also define the *normalized local covariance matrix* as  $\overline{\Sigma}_\nu(x) = \frac{1}{r^2} \Sigma_\nu(x)$ .

Note that  $\Sigma_\nu(x)$  and  $\overline{\Sigma}_\nu(x)$  depend on  $r$ , which is not made explicit in the notation. The normalization factor  $\frac{1}{r^2}$  of the normalized local covariance matrix is justified by Proposition 3.1. Moreover, we introduce the following notations: for every  $r > 0$  and  $x \in \text{supp}(\nu)$ ,

- $\nu_x$  is the restriction of  $\nu$  to the ball  $\overline{\mathcal{B}}(x,r)$ ,
- $\overline{\nu}_x = \frac{1}{\nu(\overline{\mathcal{B}}(x,r))} \nu_x$  is the corresponding probability measure.

Thus the local covariance matrix can be written as  $\Sigma_\nu(x) = \int (x-y)^{\otimes 2} d\overline{\nu}_x(y)$ .

The collection of probability measures  $\{\overline{\nu}_x\}_{x \in \text{supp}(\nu)}$  is called in [MSW19, Section 3.3] the local truncation of  $\nu$  at scale  $r$ . The application  $x \mapsto \Sigma_\nu(x)$  is called in [MMM18, Section 2.2] the multiscale covariance tensor field of  $\nu$  associated to the truncation kernel.

Remind that the aim of this paper is to estimate the measure  $\check{\mu}_0$ , defined on  $E \times M(E)$  as  $\check{\mu}_0 = \check{U}_* \mu_0$ . In other words,  $\check{\mu}_0 = (u_* \mu_0)(x_0) \otimes \delta_{\frac{1}{d+2} p_{T_x \mathcal{M}}}$ . Here is another alternative definition of  $\check{\mu}_0$ : for any  $\phi : E \times \mathcal{M}(E) \rightarrow \mathbb{R}$  with compact support,

$$\int \phi(x, A) d\check{\mu}_0(x, A) = \int \phi(u(x_0), \frac{1}{d+2} p_{T_x \mathcal{M}}) d\mu_0(x_0).$$

To do so, we consider the following construction.

**Definition 3.2.** if  $\nu$  is any measure on  $E$ , we denote by  $\check{\nu}$  the measure on  $E \times \mathcal{M}(E)$  defined by  $\check{\nu} = \nu(x) \otimes \delta_{\overline{\Sigma}_\nu(x)}$ . In other words, for every  $\phi : E \times \mathcal{M}(E) \rightarrow \mathbb{R}$  with compact support, we have

$$\int \phi(x, A) d\check{\nu}(x, A) = \int \phi(x, \overline{\Sigma}_\nu(x)) d\nu(x).$$

In accordance with the covariance matrices, the measure  $\check{\nu}$  depends on the parameter  $r$  which is not made explicit in the notation. In order to compare these measures, we consider a Wasserstein-type distance on the space  $E \times M(E)$ . Let  $\|\cdot\|_\gamma$  be the Euclidean norm  $\|(x, A)\|_\gamma = (\|x\|^2 + \gamma^2 \|A\|_{\mathbb{F}}^2)^{\frac{1}{2}}$  on  $E \times M(E)$ , where  $\gamma > 0$  is a fixed constant. Let  $p \geq 1$ . We denote by  $W_{p,\gamma}(\cdot, \cdot)$  the  $p$ -Wasserstein distance with respect to this metric. By definition, if  $\alpha, \beta$  are probability measures on  $E \times M(E)$ , then  $W_{p,\gamma}(\alpha, \beta)$  can be written as

$$W_{p,\gamma}(\alpha, \beta) = \inf_{\pi} \left( \int_{(E \times \mathcal{M}(E))^2} \|(x, A) - (y, B)\|_\gamma^p d\pi((x, A), (y, B)) \right)^{\frac{1}{p}},$$

where the infimum is taken over all measures  $\pi$  on  $(E \times \mathcal{M}(E))^2$  with marginals  $\alpha$  and  $\beta$ .

The sequel of this section consists in showing that

- if  $\mu_0$  is a measure satisfying the Hypotheses 1, 2 and 3, then  $W_{p,\gamma}(\check{\mu}_0, \check{\mu})$  is small (Proposition 3.4),
- in addition, if  $\nu$  is a measure on  $E$  such that  $W_1(\mu, \nu)$  is small, then so is  $W_{p,\gamma}(\check{\mu}, \check{\nu})$  (Proposition 3.5).

### 3.2 Consistency of the estimation

In this subsection, we assume that  $\mathcal{M}_0$  and  $\mu_0$  satisfy the hypotheses 1, 2 and 3.

The following proposition shows that the normalized covariance matrix approximates the tangent spaces of  $\mathcal{M}$ . A similar result appears in [ACLZ17, Lemma 13] in the case where  $\mathcal{M}$  is a submanifold and  $\mu$  is the uniform distribution on  $\mathcal{M}$ .

**Proposition 3.1.** *Let  $x_0 \in \mathcal{M}_0$  and  $r < \lambda(x) \wedge \frac{1}{2\rho}$ . Denote  $T = T_x\mathcal{M}$  and by  $p_T$  the linear projection on  $T$ . Then*

$$\|\bar{\Sigma}_\mu(x) - \frac{1}{d+2}p_T\|_{\mathbb{F}} \leq c_{3.3}r.$$

Proposition 3.1 is a direct consequence of the two following lemmas.

**Lemma 3.2.** [ACLZ17, Lemma 11] *Let  $\Sigma_* = \int_{\bar{\mathcal{B}}_{T_x\mathcal{M}}(0,r)} y^{\otimes 2} \frac{d\mathcal{H}^d(y)}{V_{dr^d}}$ . Then  $\Sigma_* = \frac{1}{d+2}r^2p_T$ .*

**Lemma 3.3.** *Let  $\Sigma_* = \int_{\bar{\mathcal{B}}_{T_x\mathcal{M}}(0,r)} y^{\otimes 2} \frac{d\mathcal{H}^d(y)}{V_{dr^d}}$ . Then*

$$\|\Sigma_\mu(x) - \Sigma_*\|_{\mathbb{F}} \leq c_{3.3}r^3,$$

where  $c_{3.3} = 4\rho + \frac{c_{2.10}}{f_{\min}J_{\min}} + \frac{f_{\max}}{f_{\min}J_{\min}}2^d d\rho + \frac{c_{2.11}}{f_{\min}J_{\min}}$ .

*Proof.* We use the notations of Lemma 2.10. Let  $T = T_x\mathcal{M}$ . We shall consider the following intermediate matrices:

$$\begin{aligned} \Sigma_1 &= \int_{\bar{\mathcal{B}}} ((\overline{\text{exp}}_x^{\mathcal{M}})^{-1}(x'))^{\otimes 2} d\bar{\mu}_x(x') \\ \Sigma_2 &= \int_{\bar{\mathcal{B}}^T} g(0) \cdot y^{\otimes 2} \frac{d\mathcal{H}^d(y)}{|\mu_x|} \\ \Sigma_3 &= \int_{\bar{\mathcal{B}}_T(0,r)} g(0) \cdot y^{\otimes 2} \frac{d\mathcal{H}^d(y)}{|\mu_x|} \end{aligned}$$

Let us write the triangle inequality:

$$\|\Sigma_\mu(x) - \Sigma_*\|_{\mathbb{F}} \leq \underbrace{\|\Sigma_\mu(x) - \Sigma_1\|_{\mathbb{F}}}_{(1)} + \underbrace{\|\Sigma_1 - \Sigma_2\|_{\mathbb{F}}}_{(2)} + \underbrace{\|\Sigma_2 - \Sigma_3\|_{\mathbb{F}}}_{(3)} + \underbrace{\|\Sigma_3 - \Sigma_*\|_{\mathbb{F}}}_{(4)}.$$

*Term (1):* Remind that  $\Sigma_\mu(x) = \int_{\bar{\mathcal{B}}(x,r)} (x - x')^{\otimes 2} \frac{d\mu(x')}{\mu(\bar{\mathcal{B}}(x,r))}$ . We have

$$\|\Sigma_\mu(x) - \Sigma_1\|_{\mathbb{F}} \leq \int_{\bar{\mathcal{B}}(x,r)} \|(x - x')^{\otimes 2} - ((\overline{\text{exp}}_x^{\mathcal{M}})^{-1}(x'))^{\otimes 2}\|_{\mathbb{F}} d\bar{\mu}_x(x')$$

Now, for all  $x' \in \bar{\mathcal{B}}(x,r) \cap \mathcal{M}$ , Lemma B.1 gives  $\|(x - x')^{\otimes 2} - ((\overline{\text{exp}}_x^{\mathcal{M}})^{-1}(x'))^{\otimes 2}\|_{\mathbb{F}} \leq (r+r)\|(x' - x) - (\overline{\text{exp}}_x^{\mathcal{M}})^{-1}(x')\|$ . Moreover, we can use Lemma 2.2 to get  $\|(x' - x) - (\overline{\text{exp}}_x^{\mathcal{M}})^{-1}(x')\| \leq \frac{\rho}{2}d_{\mathcal{M}_0}(x_0, x'_0)^2$ . Indeed, if we write  $x' = \gamma(\delta)$  with  $\gamma$  a geodesic such that  $\gamma(0) = x$  and  $\delta = d_{\mathcal{M}_0}(x_0, x'_0)$ , then  $(\overline{\text{exp}}_x^{\mathcal{M}})^{-1}(x') = \delta\dot{\gamma}(0)$ , and we can write

$$\begin{aligned} \|(x' - x) - (\overline{\text{exp}}_x^{\mathcal{M}})^{-1}(x')\| &= \|x' - (x + \overline{\text{exp}}_x^{\mathcal{M}})^{-1}(x')\| \\ &= \|\gamma(\delta) - (x + \delta\dot{\gamma}(0))\| \leq \frac{\rho}{2}\delta^2. \end{aligned}$$

Now, according to Lemma 2.5,  $d_{\mathcal{M}_0}(x_0, x'_0) \leq 2\|x - x'\| \leq 2r$ , and we obtain

$$\|(x - x')^{\otimes 2} - ((\overline{\text{exp}}_x^{\mathcal{M}})^{-1}(x'))^{\otimes 2}\|_{\mathbb{F}} \leq 2r \frac{\rho}{2} (2r)^2 = 4\rho r^3,$$

from which we deduce that  $\|\Sigma_\mu(x) - \Sigma_1\|_{\mathbb{F}} \leq 4\rho r^3$ .



*Term (2):* Note that  $\Sigma_1$  can be written

$$\Sigma_1 = \int_{\bar{\mathcal{B}}^T} g(y) y^{\otimes 2} \frac{d\mathcal{H}^d(y)}{|\mu_x|},$$

hence we have

$$\|\Sigma_1 - \Sigma_2\|_{\mathbb{F}} \leq \int_{\bar{\mathcal{B}}^T} |g(0) - g(y)| \|y^{\otimes 2}\| \frac{d\mathcal{H}^d(y)}{|\mu_x|}$$

According to Lemma B.1,  $\|y^{\otimes 2}\| = \|y\|^2 \leq r^2$ , and from Lemma 2.10 we get  $|g(y) - g(0)| \leq c_{2.10}r$ . Therefore,

$$\|\Sigma_1 - \Sigma_2\|_{\mathbb{F}} \leq c_{2.10}r^3 \frac{\mathcal{H}^d(\bar{\mathcal{B}}^T)}{|\mu_x|}.$$

To conclude, note that  $|\mu_x| \geq f_{\min} J_{\min} \mathcal{H}^d(\bar{\mathcal{B}}^T)$ , so we obtain  $\|\Sigma_1 - \Sigma_2\|_{\mathbb{F}} \leq \frac{c_{2.10}}{f_{\min} J_{\min}} r^3$ .

*Term (3):* We have

$$\|\Sigma_2 - \Sigma_3\|_{\mathbb{F}} \leq \int_{\bar{\mathcal{B}}_T(0,r) \setminus \bar{\mathcal{B}}^T} \|g(0) \cdot y^{\otimes 2}\|_{\mathbb{F}} \frac{d\mathcal{H}^d(y)}{|\mu_x|}$$

On the one hand,  $\|g(0) \cdot y^{\otimes 2}\|_{\mathbb{F}} \leq g(0)r^2 \leq f_{\max}r^2$ , and we get

$$\|\Sigma_2 - \Sigma_3\|_{\mathbb{F}} \leq f_{\max}r^2 \frac{\mathcal{H}^d(\bar{\mathcal{B}}_T(0,r) \setminus \bar{\mathcal{B}}^T)}{|\mu_x|}.$$

On the other hand, since  $\bar{\mathcal{B}}^T \subseteq \bar{\mathcal{B}}_T(x, c_{2.5}(\rho r)r)$ , we have

$$\mathcal{H}^d(\bar{\mathcal{B}}^T \setminus \bar{\mathcal{B}}_T(0,r)) = (c_{2.5}(\rho r)r)^d V_d - r^d V_d.$$

The inequality  $A^d - 1 \leq d(A-1)A^{d-1}$ , where  $A \geq 1$ , gives  $(c_{2.5}(\rho r)r)^d V_d - r^d V_d \leq V_d r^d d(c_{2.5}(\rho r) - 1)2^{d-1}$ . Combined with the inequalities  $c_{2.5}(\rho r) \leq 1 + 2\rho r$  and  $|\mu_x| \geq f_{\min} J_{\min} V_d r^d$  we get

$$\|\Sigma_2 - \Sigma_3\|_{\mathbb{F}} \leq \frac{f_{\max}}{f_{\min} J_{\min}} 2^d d \rho r^3.$$

*Term (4):* We have

$$\|\Sigma_3 - \Sigma_*\|_{\mathbb{F}} \leq \left| \frac{|\mu_x|}{V_d r^d} - f(x) \right| \int_{\bar{\mathcal{B}}_T(0,r)} \|y^{\otimes 2}\|_{\mathbb{F}} \frac{d\mathcal{H}^d(y)}{|\mu_x|}$$

According to Lemma 2.11 point 2,  $\left| \frac{|\mu_x|}{V_d r^d} - f(x) \right| \leq c_{2.11}r$ . Moreover,  $\int_{\bar{\mathcal{B}}_T(0,r)} \|y^{\otimes 2}\|_{\mathbb{F}} \frac{d\mathcal{H}^d(y)}{|\mu_x|} \leq \frac{1}{f_{\min} J_{\min}} r^2$ . Therefore,

$$\|\Sigma_3 - \Sigma_*\|_{\mathbb{F}} \leq \frac{c_{2.11}}{f_{\min} J_{\min}} r^3. \quad \square$$

**Proposition 3.4.** *Let  $r < \frac{1}{2\rho}$ . Then*

$$W_{p,\gamma}(\check{\mu}, \check{\mu}_0) \leq \gamma(2\mu(\lambda^r)^{\frac{1}{p}} + c_{3.1}r).$$

*Proof.* Define  $\phi : x_0 \in \mathcal{M}_0 \mapsto ((x, \bar{\Sigma}_\mu(x)), (x, \frac{1}{d+2}p_{T_x\mathcal{M}}))$ , and consider the measure  $\pi = \phi_*\mu_0$ . It is a transport plan between  $\check{\mu}$  and  $\check{\mu}_0$ . Therefore,  $W_{p,\gamma}^p(\check{\mu}, \check{\mu}_0) \leq \int \|(x, T) - (x', T')\|_\gamma^p d\pi((x, T), (x', T'))$ , and we can write

$$\begin{aligned} W_{p,\gamma}^p(\check{\mu}, \check{\mu}_0) &\leq \int \|(x, \frac{1}{r^2}\Sigma_\mu(x)) - (x, \frac{1}{d+2}p_{T_x\mathcal{M}})\|_\gamma^p d\mu(x) \\ &= \gamma^p \int \|\frac{1}{r^2}\Sigma_\mu(x) - \frac{1}{d+2}p_{T_x\mathcal{M}}\|_F^p d\mu(x). \end{aligned}$$

We split this last integral onto the sets  $A = \lambda^r$  and  $B = E \setminus \lambda^r$ . Remind that  $\lambda^r$  denotes the sublevel set  $\lambda^{-1}([0, r])$ .

On  $A$ , we use the majoration  $\|\frac{1}{r^2}\Sigma_\mu(x) - \frac{1}{d+2}p_{T_x\mathcal{M}}\|_F \leq \|\frac{1}{r^2}\Sigma_\mu(x)\|_F + \|\frac{1}{d+2}p_{T_x\mathcal{M}}\|_F \leq 2$  to obtain

$$\int_A \|\frac{1}{r^2}\Sigma_\mu(x) - \frac{1}{d+2}p_{T_x\mathcal{M}}\|_F^p d\mu(x) \leq 2^p \mu(A).$$

On  $B$ , we use Proposition 3.1 to get

$$\int_B \|\frac{1}{r^2}\Sigma_\mu(x) - \frac{1}{d+2}p_{T_x\mathcal{M}}\|_F^p d\mu(x) \leq (c_{3.1}r)^p.$$

Combining these two inequalities yields  $W_{p,\gamma}^p(\check{\mu}, \check{\mu}_0) \leq \gamma^p(2^p\mu(A) + (c_{3.1}r)^p)$ . Using the inequality  $(a+b)^{\frac{1}{p}} \leq a^{\frac{1}{p}} + b^{\frac{1}{p}}$ , where  $a, b \geq 0$ , we obtain

$$W_{p,\gamma}(\check{\mu}, \check{\mu}_0) \leq \gamma(2\mu(A)^{\frac{1}{p}} + c_{3.1}r). \quad \square$$

### 3.3 Stability of the estimation

In this subsection we study the stability of the operator  $\mu \mapsto \bar{\Sigma}_\mu(\cdot)$  with respect to the  $W_1$  metric on measures. The results of this subsection only rely on the following hypotheses about  $\mu$ :

*Hypothesis 5.*  $\exists a > 0, \forall x \in \text{supp}(\mu), \forall t \in [0, \frac{1}{2\rho})$ ,

$$\mu(\bar{\mathcal{B}}(x, t)) \geq at^d.$$

*Hypothesis 6.*  $\exists b > 0, \forall x \in \text{supp}(\mu), \exists \lambda(x) \geq 0, \forall s, t \in [0, \lambda(x) \wedge \frac{1}{2\rho})$  s.t.  $s \leq t$ ,

$$\mu(\bar{\mathcal{B}}(x, t) \setminus \bar{\mathcal{B}}(x, s)) \leq bt^{d-1}(t-s).$$

*Hypothesis 7.*  $\exists b' > 0, \forall x \in \text{supp}(\mu), \forall s, t \in [0, \frac{1}{2\rho})$  s.t.  $s \leq t$ ,

$$\mu(\bar{\mathcal{B}}(x, t) \setminus \bar{\mathcal{B}}(x, s)) \leq b't^{d-\frac{1}{2}}(t-s)^{\frac{1}{2}}.$$

Note that, as stated in Propositions 2.11 and 2.12, the initial hypotheses 2 and 3 implies the hypotheses 5, 6 and 7 with  $\lambda(x)$  being the normal reach of  $\mathcal{M}$  at  $x$ .

Let  $\mu$  and  $\nu$  be two probability measures,  $x \in \text{supp}(\mu) \cap \text{supp}(\nu)$ , and consider the Frobenius distance  $\|\bar{\Sigma}_\mu(x) - \bar{\Sigma}_\nu(x)\|_F$  between the normalized local covariance matrices. This distance is related to the 1-Wasserstein distance between the localized probability measures  $\bar{\mu}_x$  and  $\bar{\nu}_x$  via the following inequality:

$$\|\bar{\Sigma}_\mu(x) - \bar{\Sigma}_\nu(x)\|_F \leq \frac{2}{r}W_1(\bar{\mu}_x, \bar{\nu}_x).$$

Without any assumption on the measures, it is not true that  $W_1(\bar{\mu}_x, \bar{\nu}_x)$  goes to 0 as  $W_1(\mu, \nu)$  does. However, if we assume that  $\mu$  satisfies the hypotheses 5 and 6, that  $x$  satisfies  $\lambda(x) > 0$  and that  $r$  is chosen such that  $r < \lambda(x)$  and  $r \geq 4(\frac{W_1(\mu, \nu)}{a\wedge 1})^{\frac{1}{d+1}}$ , then we are able to prove (Lemma B.5) that

$$W_1(\bar{\mu}_x, \bar{\nu}_x) \leq c_{B.5} \left( \frac{W_1(\mu, \nu)}{r^{d-1}} \right)^{\frac{1}{2}}.$$

As a consequence, estimating local covariance matrices is robust in Wasserstein distance:

$$\|\bar{\Sigma}_\mu(x) - \bar{\Sigma}_\nu(x)\|_F \leq 2c_{B.5} \left( \frac{W_1(\mu, \nu)}{r^{d+1}} \right)^{\frac{1}{2}}. \quad (7)$$

A stability result of this kind already appears in [MSW19, Theorem 4.3], where  $\mu$  and  $\nu$  are two probability measures on a bounded set  $X$ , and satisfy the following condition:  $\forall x \in X, \forall s, r \leq 0$  s.t.  $s \leq r, \frac{\mu(\bar{B}(x, r))}{\mu(\bar{B}(x, s))} \leq (\frac{r}{s})^d$ . The theorem states that, denoting  $D = \text{diam}(X)$ , for all  $x \in X$ ,

$$W_1(\bar{\mu}_x, \bar{\nu}_x) \leq (1 + 2r) \left[ \frac{W_1(\mu, \nu)^{\frac{1}{2}}}{1 \wedge (\frac{r}{D})^d} + \left(1 + \frac{W_1(\mu, \nu)^{\frac{1}{2}}}{r}\right)^d - 1 \right].$$

When  $r \leq 1$  and  $W_1(\mu, \nu)$  goes to zero, we obtain that  $W_1(\bar{\mu}_x, \bar{\nu}_x)$  is of order  $(1+2r)D^d \left(\frac{W_1(\mu, \nu)}{r^{2d}}\right)^{\frac{1}{2}}$ . Another result in [MMM18, Theorem 3] bounds the distance  $\|\bar{\Sigma}_\mu(x) - \bar{\Sigma}_\nu(x)\|_F$  with the  $\infty$ -Wasserstein distance  $W_\infty(\mu, \nu)$ . Namely, if  $\mu$  and  $\nu$  are fully supported probability measures with densities upper bounded by  $l > 0$  and supports included in  $X \subset \mathbb{R}^d$ , denoting  $D = \text{diam}(X)$ , we have

$$\|\bar{\Sigma}_\mu(x) - \bar{\Sigma}_\nu(x)\|_F \leq l A W_\infty(\mu, \nu),$$

where  $A = \frac{d}{d+2} \frac{(r+D)^{d+1}}{Dr^d} + \frac{(2r+D)(r+D)^d}{r^d} + \frac{2d}{d+2} \frac{(r+D)^{d+2}}{Dr^d}$ .

*Remark 3.1.* Let us show that in general, for  $x \in \text{supp}(\mu) \cap \text{supp}(\nu)$ , it is not true that  $\|\bar{\Sigma}_\mu(x) - \bar{\Sigma}_\nu(x)\|_F$  goes to zero as  $W_1(\mu, \nu)$  goes to zero. Similarly,  $W_{p,\gamma}(\check{\mu}, \check{\nu})$  does not have to go to zero. For example, one can consider  $\epsilon > 0$ , the measures  $\mu = \frac{1}{2}(\delta_0 + \delta_1)$  and  $\nu = \frac{1}{2}(\delta_0 + \delta_{1+\epsilon})$  on  $\mathbb{R}$ , and the scale parameter  $r = 1$ . We have  $\Sigma_\mu(0) = \Sigma_\mu(1) = \frac{1}{2}1^{\otimes 2}$  and  $\Sigma_\nu(0) = \Sigma_\nu(1) = 0$ . Hence  $\check{\mu} = \frac{1}{2}(\delta_{(0, \frac{1}{2}1^{\otimes 2})} + \delta_{(1, \frac{1}{2}1^{\otimes 2})})$ , and  $\check{\nu} = \frac{1}{2}(\delta_{(0,0)} + \delta_{(1+\epsilon,0)})$ . A computation shows that

$$\begin{aligned} W_{p,\gamma}^p(\check{\mu}, \check{\nu}) &= \frac{1}{2} \left\| (0, \frac{1}{2}1^{\otimes 2}) - (0, 0) \right\|_\gamma^p + \frac{1}{2} \left\| (1, \frac{1}{2}1^{\otimes 2}) - (1 + \epsilon, 0) \right\|_\gamma^p \\ &= \frac{1}{2} \left( \left(\frac{\gamma}{2}\right)^p + (\epsilon^2 + \gamma^2 \frac{1}{4})^{\frac{p}{2}} \right) \geq \left(\frac{\gamma}{2}\right)^p. \end{aligned}$$

Hence  $W_{p,\gamma}(\check{\mu}, \check{\nu}) \geq \frac{\gamma}{2} > 0$ . Moreover, we have  $W_1(\mu, \nu) = \frac{1}{2}\epsilon$ . Hence  $W_{p,\gamma}(\check{\mu}, \check{\nu})$  does not go to zero as  $W_1(\mu, \nu)$  does. However, under regularity assumptions on  $\mu$ , the following proposition states that it is the case.

**Proposition 3.5.** *Let  $\mu$  and  $\nu$  be two probability measures on  $E$ . Suppose that  $\mu$  satisfies the hypotheses 5, 6 and 7. Define  $w = W_p(\mu, \nu)$ . Suppose that  $r \leq \frac{1}{2\rho} \wedge 1$  and  $w \leq (a \wedge 1) \left(\frac{r}{4}\right)^{d+1}$ . Then*

$$W_{p,\gamma}(\check{\mu}, \check{\nu}) \leq 2w + \gamma c_{3.5} \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + \gamma (4c_{B.6} \mu(\lambda^r)^{\frac{1}{p}}) \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{4}}$$

with  $c_{3.5} = 4(1 + c_{3.7})$ .

*Proof.* According to Lemma 3.6, we have

$$W_{p,\gamma}(\check{\mu}, \check{\nu}) \leq 2^{\frac{p-1}{p}} \left(1 + \frac{2\gamma}{r}\right) w + 2^{\frac{p-1}{p}} \frac{2\gamma}{r} \left( \int W_1^p(\bar{\mu}_x, \bar{\nu}_y) d\pi(x, y) \right)^{\frac{1}{p}}.$$

Moreover, Lemma 3.7 gives

$$\left( \int W_1(\bar{\mu}_x, \bar{\nu}_y) d\pi(x, y) \right)^{\frac{1}{p}} \leq 2^{\frac{p-1}{p}} \left( c_{B.6} r^{\frac{1}{2}} \mu(\lambda^r)^{\frac{1}{p}} \alpha^{\frac{1}{2}} + c_{3.7} \alpha \right)$$

Combining these inequalities yields

$$W_{p,\gamma}(\check{\mu}, \check{\nu}) \leq 2^{\frac{p-1}{p}} w + 2^{\frac{p-1}{p}} \frac{2\gamma}{r} (w + 2^{\frac{p-1}{p}} c_{3.7} \alpha) + (2^{\frac{p-1}{p}})^2 \frac{2\gamma}{r} c_{B.6} r^{\frac{1}{2}} \mu(\lambda^r)^{\frac{1}{p}} \alpha^{\frac{1}{2}},$$

and since  $r \leq 1$  we conclude with  $w = (\frac{w}{r^{d-1}})^{\frac{1}{2}} r^{\frac{d-1}{2}} w^{\frac{1}{2}} \leq (\frac{w}{r^{d-1}})^{\frac{1}{2}} = \alpha$ :

$$\begin{aligned} W_{p,\gamma}(\check{\mu}, \check{\nu}) &\leq 2^{\frac{p-1}{p}} w + 2^{\frac{p-1}{p}} \frac{2\gamma}{r} (1 + 2^{\frac{p-1}{p}} c_{3.7}) \alpha + (2^{\frac{p-1}{p}})^2 \frac{2\gamma}{r} c_{B.6} r^{\frac{1}{2}} \mu(\lambda^r)^{\frac{1}{p}} \alpha^{\frac{1}{2}} \\ &\leq 2^{\frac{p-1}{p}} w + 2^{\frac{p-1}{p}} 2\gamma (1 + 2^{\frac{p-1}{p}} c_{3.7}) \frac{\alpha}{r} + (2^{\frac{p-1}{p}})^2 2\gamma c_{B.6} \mu(\lambda^r)^{\frac{1}{p}} \left(\frac{\alpha}{r}\right)^{\frac{1}{2}}. \end{aligned}$$

Writing  $2^{\frac{p-1}{p}} \leq 2^{\frac{1}{2}} \leq 2$  and  $\alpha = (\frac{w}{r^{d-1}})^{\frac{1}{2}}$  yields

$$W_{p,\gamma}(\check{\mu}, \check{\nu}) \leq 2w + 4\gamma(1 + 2c_{3.7}) \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + 4\gamma c_{B.6} \mu(\lambda^r)^{\frac{1}{p}} \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{4}}. \quad \square$$

Let us interpret the inequality  $W_{p,\gamma}(\check{\mu}, \check{\nu}) \leq 2w + \gamma c_{3.5} (\frac{w}{r^{d+1}})^{\frac{1}{2}} + \gamma (4c_{B.6} \mu(\lambda^r)^{\frac{1}{p}}) (\frac{w}{r^{d+1}})^{\frac{1}{4}}$ . The first term  $2w$  is to be seen as the initial error between the measures  $\mu$  and  $\nu$ . The second term  $\gamma c_{3.5} (\frac{w}{r^{d+1}})^{\frac{1}{2}}$  corresponds to the local errors  $W_1(\bar{\mu}_x, \bar{\nu}_y)$  when comparing the normalized covariance matrices. The third term  $\gamma (4c_{B.6} \mu(\lambda^r)^{\frac{1}{p}}) (\frac{w}{r^{d+1}})^{\frac{1}{4}}$  stands for the error on points  $x$  such that  $\lambda(x) \leq r$ , where the stability is weaker. As a consequence of this proposition, the application  $\mu \mapsto \check{\mu}$ , seen as an application between spaces of measures endowed with the Wasserstein metric, is continuous on the set of measures  $\mu$  which satisfy 5, 6 and 7 with  $\frac{1}{2\rho} \geq r$ .

We now state the lemmas used in the proof of Proposition 3.5.

**Lemma 3.6.** *Let  $\pi$  be an optimal transport plan for  $W_p(\mu, \nu)$ . Then*

$$W_{p,\gamma}(\check{\mu}, \check{\nu}) \leq 2^{\frac{p-1}{p}} (1 + \frac{2\gamma}{r}) W_p(\mu, \nu) + 2^{\frac{p-1}{p}} \frac{2\gamma}{r} \left( \int W_1^p(\bar{\mu}_x, \bar{\nu}_y) d\pi(x, y) \right)^{\frac{1}{p}}.$$

*Proof.* We first prove the following fact: for every  $x \in \text{supp}(\mu)$  and  $y \in \text{supp}(\nu)$ ,

$$\|\Sigma_\mu(x) - \Sigma_\nu(y)\|_F \leq 2r(\|x - y\| + W_1(\bar{\mu}_x, \bar{\nu}_y)). \quad (8)$$

Let  $\rho$  be any transport plan between  $\bar{\mu}_x$  and  $\bar{\nu}_y$ . We have

$$\begin{aligned} \Sigma_\mu(x) - \Sigma_\nu(y) &= \int (x - y)^{\otimes 2} d\bar{\mu}_x(x') - \int (y - y')^{\otimes 2} d\bar{\nu}_y(y') \\ &= \int ((x - x')^{\otimes 2} - (y - y')^{\otimes 2}) d\rho(x', y') \end{aligned}$$

For any  $x' \in \bar{\mathcal{B}}(x, r)$  and  $y' \in \bar{\mathcal{B}}(y, r)$ , using Lemma B.1, we have  $\|(x - x')^{\otimes 2} - (y - y')^{\otimes 2}\|_F \leq (r + r)(\|x - y\| + \|x' - y'\|)$ . Hence,

$$\|\Sigma_\mu(x) - \Sigma_\nu(y)\|_F \leq \int 2r(\|x - y\| + \|x' - y'\|) d\rho(x', y'),$$

and we deduce that  $\|\Sigma_\mu(x) - \Sigma_\nu(y)\|_F \leq 2r(\|x - y\| + W_1(\bar{\mu}_x, \bar{\nu}_y))$ .

Now, the transport plan  $\pi$  begin given, we build a transport plan  $\tilde{\pi}$  for  $(\check{\mu}, \check{\nu})$  as follows: for every  $\phi : (E \times \mathcal{M}(E))^2 \rightarrow \mathbb{R}$  with compact support,

$$\int \phi(x, A, y, B) d\tilde{\pi}(x, A, y, B) = \int \phi(x, \bar{\Sigma}_\mu(x), y, \bar{\Sigma}_\nu(y)) d\pi(x, y).$$

We have

$$\begin{aligned}
W_{p,\gamma}^p(\check{\mu}, \check{\nu}) &\leq \int \| (x, A) - (y, B) \|_\gamma^p d\check{\pi}(x, A, y, B) \\
&= \int (\|x - y\|^2 + \gamma^2 \|\bar{\Sigma}_\mu(x) - \bar{\Sigma}_\nu(y)\|_{\mathbb{F}}^2)^{\frac{p}{2}} d\pi(x, y) \\
&\leq \int (\|x - y\| + \gamma \|\bar{\Sigma}_\mu(x) - \bar{\Sigma}_\nu(y)\|_{\mathbb{F}})^p d\pi(x, y)
\end{aligned}$$

Besides, Equation 8 gives  $\|\bar{\Sigma}_\mu(x) - \bar{\Sigma}_\nu(y)\|_{\mathbb{F}} \leq \frac{1}{r^2} \|\Sigma_\mu(x) - \Sigma_\nu(y)\|_{\mathbb{F}} \leq \frac{2}{r} (\|x - y\| + W_1(\bar{\mu}_x, \bar{\nu}_y))$ . We can use the inequality  $(a + b)^p \leq 2^{p-1}(a^p + b^p)$ , where  $a, b \geq 0$ , to deduce

$$\begin{aligned}
(\|x - y\| + \gamma \|\bar{\Sigma}_\mu(x) - \bar{\Sigma}_\nu(y)\|_{\mathbb{F}})^p &\leq (\|x - y\| + \gamma \frac{2}{r} (\|x - y\| + W_1(\bar{\mu}_x, \bar{\nu}_y)))^p \\
&\leq 2^{p-1} \left( (1 + \frac{2\gamma}{r}) \|x - y\| \right)^p + 2^{p-1} \left( \frac{2\gamma}{r} W_1(\bar{\mu}_x, \bar{\nu}_y) \right)^p
\end{aligned}$$

To conclude,

$$\begin{aligned}
W_{p,\gamma}^p(\check{\mu}, \check{\nu}) &\leq 2^{p-1} \int \left( (1 + \frac{2\gamma}{r}) \|x - y\| \right)^p + \left( \frac{2\gamma}{r} W_1(\bar{\mu}_x, \bar{\nu}_y) \right)^p d\pi(x, y) \\
&= 2^{p-1} \left( 1 + \frac{2\gamma}{r} \right)^p W_p^p(\mu, \nu) + 2^{p-1} \left( \frac{2\gamma}{r} \right)^p \int W_1^p(\bar{\mu}_x, \bar{\nu}_y) d\pi(x, y),
\end{aligned}$$

hence

$$W_{p,\gamma}(\check{\mu}, \check{\nu}) \leq 2^{\frac{p-1}{p}} \left( 1 + \frac{2\gamma}{r} \right) W_p(\mu, \nu) + 2^{\frac{p-1}{p}} \frac{2\gamma}{r} \left( \int W_1^p(\bar{\mu}_x, \bar{\nu}_y) d\pi(x, y) \right)^{\frac{1}{p}} \quad \square$$

**Lemma 3.7.** *Let  $w = W_p(\mu, \nu)$ . Suppose that  $r \leq \frac{1}{2\rho}$  and  $w \leq (a \wedge 1) \left( \frac{r}{4} \right)^{d+1}$ . Let  $\pi$  be an optimal transport plan for  $W_p(\mu, \nu)$ . Then*

$$\begin{aligned}
&\left( \int W_1^p(\bar{\mu}_x, \bar{\nu}_y) d\pi(x, y) \right)^{\frac{1}{p}} \\
&\leq 2^{\frac{p-1}{p}} \left( c_{B.6} r^{\frac{1}{2}} \mu(\lambda^r)^{\frac{1}{p}} \alpha^{\frac{1}{2}} + (2r^d + c_{B.4} r^{\frac{d+1}{2}} + c_{B.5}) \alpha + (1 + c_{B.3}) w \right)
\end{aligned}$$

where  $\alpha = \left( \frac{w}{r^{d-1}} \right)^{\frac{1}{2}}$ . If we suppose that  $r \leq 1$ , then

$$\left( \int W_1^p(\bar{\mu}_x, \bar{\nu}_y) d\pi(x, y) \right)^{\frac{1}{p}} \leq 2^{\frac{p-1}{p}} \left( c_{B.6} r^{\frac{1}{2}} \mu(\lambda^r)^{\frac{1}{p}} \alpha^{\frac{1}{2}} + c_{3.7} \alpha \right)$$

with  $c_{3.7} = 3 + c_{B.3} + c_{B.4} + c_{B.5}$ .

*Proof.* In order to ease the reading of the proof, we define  $w = W_p(\mu, \nu)$  and  $\alpha = \left( \frac{w}{r^{d-1}} \right)^{\frac{1}{2}}$ . We cut the integral as follows:

$$\int W_1^p(\bar{\mu}_x, \bar{\nu}_y) d\pi(x, y) = \int_A + \int_B + \int_C W_1^p(\bar{\mu}_x, \bar{\nu}_y) d\pi(x, y)$$

where  $A = \{(x, y), \|x - y\| \geq \alpha\}$ ,  $B = \{(x, y), \|x - y\| < \alpha \text{ and } \lambda(x) > r\}$  and  $C = \{(x, y), \|x - y\| < \alpha \text{ and } \lambda(x) \leq r\}$ .

*Term A:* We use the following loose majoration:

$$\begin{aligned}
W_1(\bar{\mu}_x, \bar{\nu}_y) &\leq W_1(\bar{\mu}_x, \delta_x) + W_1(\delta_x, \delta_y) + W_1(\delta_y, \bar{\nu}_y) \\
&\leq r + \|x - y\| + r
\end{aligned}$$

to obtain  $W_1^p(\overline{\mu_x}, \overline{\nu_y}) \leq 2^{p-1}((2r)^p + \|x - y\|^p)$  and

$$\begin{aligned} \int_A W_1^p(\overline{\mu_x}, \overline{\nu_y}) d\pi(x, y) &\leq \int_A 2^{p-1}((2r)^p + \|x - y\|^p) d\pi(x, y) \\ &\leq 2^{p-1}(2r)^p \pi(A) + \int 2^{p-1} \|x - y\|^p d\pi(x, y) \\ &= 2^{p-1}(2r)^p \pi(A) + 2^{p-1} w^p. \end{aligned}$$

But  $\pi(A) = \pi(\{(x, y), \|x - y\| > \alpha\}) = \pi(\{(x, y), \|x - y\|^p > \alpha^p\}) \leq (\frac{w}{\alpha})^p$  by Markov inequality. Therefore,

$$\int_A W_1^p(\overline{\mu_x}, \overline{\nu_y}) d\pi(x, y) \leq 2^{p-1}(2r)^p \left(\frac{w}{\alpha}\right)^p + 2^{p-1} w^p = 2^{p-1}(2r^d \alpha)^p + 2^{p-1} w^p,$$

where we used  $r \frac{w}{\alpha} = r^d \alpha$ .

*Term B:* On the event  $B$ , we write

$$W_1(\overline{\mu_x}, \overline{\nu_y}) \leq W_1(\overline{\mu_x}, \overline{\mu_y}) + W_1(\overline{\mu_y}, \overline{\nu_y}).$$

Since  $\lambda(x) > r$ , Lemma B.3 and Lemma B.5 give  $W_1(\overline{\mu_x}, \overline{\mu_y}) \leq c_{B.3} \|x - y\|$  and  $W_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{B.5} \alpha$ . We deduce that

$$\begin{aligned} \int_B W_1^p(\overline{\mu_x}, \overline{\nu_y}) d\pi(x, y) &\leq 2^{p-1} \int_B (c_{B.3} \|x - y\|)^p + (c_{B.5} \alpha)^p d\pi(x, y) \\ &\leq 2^{p-1} (c_{B.3} w)^p + 2^{p-1} (c_{B.5} \alpha)^p. \end{aligned}$$

*Term C:* We proceed as for Term B, but using Lemmas B.4 and B.6 instead of Lemmas B.3 and B.5. This yields

$$\begin{aligned} W_1(\overline{\mu_x}, \overline{\nu_y}) &\leq W_1(\overline{\mu_x}, \overline{\mu_y}) + W_1(\overline{\mu_y}, \overline{\nu_y}) \\ &\leq c_{B.4} r^{\frac{1}{2}} \|x - y\|^{\frac{1}{2}} + c_{B.6} r^{\frac{1}{2}} \alpha^{\frac{1}{2}}, \end{aligned}$$

and we deduce that

$$\int_C W_1^p(\overline{\mu_x}, \overline{\nu_y}) d\pi(x, y) \leq \int_C 2^{p-1} (c_{B.4} r^{\frac{1}{2}} \|x - y\|^{\frac{1}{2}})^p d\pi(x, y) + 2^{p-1} \pi(C) (c_{B.6} r^{\frac{1}{2}} \alpha^{\frac{1}{2}})^p.$$

We have  $\int_C \|x - y\|^{\frac{p}{2}} d\pi(x, y) \leq \int_{E \times E} \|x - y\|^{\frac{p}{2}} d\pi(x, y)$ , and by Jensen's inequality,  $\int_{E \times E} \|x - y\|^{\frac{p}{2}} d\pi(x, y) \leq (w^p)^{\frac{1}{2}}$ . Moreover, by definition,  $\pi(C) \leq \mu(\lambda^r)$ . Therefore,

$$\int_C W_1^p(\overline{\mu_x}, \overline{\nu_y}) d\pi(x, y) \leq 2^{p-1} (c_{B.4} r^{\frac{1}{2}} w^{\frac{1}{2}})^p + 2^{p-1} \mu(\lambda^r) (c_{B.6} r^{\frac{1}{2}} \alpha^{\frac{1}{2}})^p.$$

To conclude the proof, we write

$$\begin{aligned} \int W_1^p(\overline{\mu_x}, \overline{\nu_y}) d\pi(x, y) &= \int_A + \int_B + \int_C W_1^p(\overline{\mu_x}, \overline{\nu_y}) d\pi(x, y) \\ &\leq 2^{p-1} (2r^d \alpha)^p + 2^{p-1} w^p + 2^{p-1} (c_{B.3} w)^p + 2^{p-1} (c_{B.5} \alpha)^p \\ &\quad + 2^{p-1} (c_{B.4} r^{\frac{1}{2}} w^{\frac{1}{2}})^p + 2^{p-1} \mu(\lambda^r) (c_{B.6} r^{\frac{1}{2}} \alpha^{\frac{1}{2}})^p, \end{aligned}$$

and we use the inequality  $(a + b)^{\frac{1}{p}} \leq a^{\frac{1}{p}} + b^{\frac{1}{p}}$ , where  $a, b \geq 0$ , to get

$$\begin{aligned} &\left( \int W_1^p(\overline{\mu_x}, \overline{\nu_y}) d\pi(x, y) \right)^{\frac{1}{p}} \\ &\leq 2^{\frac{p-1}{p}} \left( 2r^d \alpha + w + c_{B.3} w + c_{B.5} \alpha + c_{B.4} r^{\frac{1}{2}} w^{\frac{1}{2}} + \mu(\lambda^r)^{\frac{1}{p}} c_{B.6} r^{\frac{1}{2}} \alpha^{\frac{1}{2}} \right) \\ &\leq 2^{\frac{p-1}{p}} \left( c_{B.6} r^{\frac{1}{2}} \mu(\lambda^r)^{\frac{1}{p}} \alpha^{\frac{1}{2}} + (2r^d + c_{B.4} r^{\frac{d+1}{2}} + c_{B.5}) \alpha + (1 + c_{B.3}) w \right) \end{aligned}$$

where we used on the the last line  $c_{B.4}r^{\frac{1}{2}}w^{\frac{1}{2}} = c_{B.4}r^{\frac{d+1}{2}}\alpha$ .

If we suppose  $r \leq 1$ , we can use the inequalities  $r^d \leq r^{\frac{d+1}{2}} \leq 1$  and  $w = \alpha r^{\frac{d-1}{2}}w^{\frac{1}{2}} \leq \alpha$  to obtain the simplified expression

$$\left( \int W_1(\overline{\mu}_x, \overline{\nu}_y) d\pi(x, y) \right)^{\frac{1}{p}} \leq 2^{\frac{p-1}{p}} \left( c_{B.6}r^{\frac{1}{2}}\mu(\lambda^r)^{\frac{1}{p}}\alpha^{\frac{1}{2}} + (3 + c_{B.3} + c_{B.4} + c_{B.5})\alpha \right) \quad \square$$

*Remark 3.2.* On Term  $C$ , we could have used the inequality  $W_1(\overline{\mu}_x, \overline{\nu}_y) \leq r + \|x - y\| + r$  to obtain

$$\begin{aligned} \int_C W_1^p(\overline{\mu}_x, \overline{\nu}_y) d\pi(x, y) &\leq 2^{p-1} \int_C (2r)^p + \|x - y\|^p d\pi(x, y) \\ &\leq 2^{p-1}(2r)^p \pi(C) + 2^{p-1}w^p. \end{aligned}$$

Following the rest of the proof, and under the assumption  $r \leq 1$ , we eventually obtain

$$\left( \int W_1(\overline{\mu}_x, \overline{\nu}_y) d\pi(x, y) \right)^{\frac{1}{p}} \leq 2^{\frac{p-1}{p}} \left( 2r\mu(\lambda^r)^{\frac{1}{p}} + c'_{3.7}\alpha \right)$$

with  $c'_{3.7} = 4 + c_{B.3} + c_{B.5}$ .

## 4 Topological inference with $\check{\nu}$

Combining the results of the last section, we derive an approximation theorem for  $\check{\mu}_0$ , based on  $\check{\nu}$ . As a consequence, we show that  $\check{\nu}$  can be used to infer the homotopy type of  $\check{\mathcal{M}}$ , and that it can be used to estimate the persistent homology of  $\check{\mu}_0$ .

### 4.1 An approximation theorem

Let us recall the definitions of Subsection 3.1:  $\check{\mu}_0 = (u_*\mu_0)(x_0) \otimes \delta_{\frac{1}{d+2}pT_x\mathcal{M}}$  and  $\check{\nu} = \nu(x) \otimes \delta_{\overline{\Sigma}_\nu(x)}$ . Moreover, the distance  $W_{p,\gamma}(\check{\nu}, \check{\mu}_0)$  corresponds to the  $p$ -Wasserstein distance in  $E \times M(E)$  with respect to the metric  $\|\cdot\|_\gamma$ .

**Theorem 4.1.** *Assume that  $\mathcal{M}_0$  and  $\mu$  satisfy the hypotheses 1, 2, 3. Let  $\nu$  be any probability measure. Denote  $W_p(\mu, \nu) = w$ . Suppose that  $r \leq \frac{1}{2^p} \wedge 1$  and  $w \leq (a \wedge 1)\left(\frac{r}{4}\right)^{d+1}$ . Then*

$$W_{p,\gamma}(\check{\nu}, \check{\mu}_0) \leq \gamma c_{4.1}\mu(\lambda^r)^{\frac{1}{p}} + \gamma c_{3.1}r + \gamma c_{3.5}\left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + 2w$$

where  $c_{4.1} = 2(1 + 2c_{B.6})$ .

*Proof.* It is a direct consequence of Propositions 3.4 and 3.5. □

In order to simplify the results of this section, we shall use the following weaker result. It shows that  $\check{\mu}_0$  is well approximated by  $\check{\nu}$ , as long as the the parameter  $r$  is well chosen.

**Corollary 4.2.** *Let  $r > 0$ . Assume that  $\mathcal{M}_0$  and  $\mu_0$  satisfy the hypotheses 1, 2, 3 and Hypothesis 4 with  $r_0 \geq r$ . Let  $\nu$  be any probability measure. Denote  $W_p(\mu, \nu) = w$ . Suppose that  $r \leq \frac{1}{2^p} \wedge 1$  and  $w \leq (a \wedge 1)\left(\frac{r}{4}\right)^{d+2}$ . Then*

$$W_{p,\gamma}(\check{\nu}, \check{\mu}_0) \leq (1 + \gamma c_{4.2})r^{\frac{1}{p}}$$

with  $c_{4.2} = c_{4.1}(c_4)^{\frac{1}{p}} + c_{3.5} + c_{3.1}$ .

*Proof.* According to Theorem 4.1, we have

$$W_{p,\gamma}(\check{\nu}, \check{\mu}_0) \leq \gamma c_{4.1} \mu(\lambda^r)^{\frac{1}{p}} + \gamma c_{3.1} r + \gamma c_{3.5} \left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} + 2w.$$

Note that the assumption  $w \leq (a \wedge 1) \left(\frac{r}{4}\right)^{d+2}$  yields  $\left(\frac{w}{r^{d+1}}\right)^{\frac{1}{2}} \leq \frac{r}{4} \leq r$ . Besides,  $r \leq 1$  yields  $w \leq \left(\frac{r}{4}\right)^{d+2} \leq \frac{r}{4} \leq r$ . Finally, Hypothesis 4 gives  $\mu(\lambda^r) \leq c_4 r$ , and we deduce the result thanks to a rough majoration:

$$\begin{aligned} W_{p,\gamma}(\check{\nu}, \check{\mu}_0) &\leq \gamma c_{4.1} (c_4 r)^{\frac{1}{p}} + \gamma c_{3.1} r + \gamma c_{3.5} r + 2r \\ &\leq (\gamma c_{4.1} (c_4)^{\frac{1}{p}} + \gamma c_{3.1} + \gamma c_{3.5} + 2) r^{\frac{1}{p}}. \end{aligned} \quad \square$$

## 4.2 Homotopy type estimation with the DTM

In this subsection, we use the DTM, as defined in Subsection 1.3, to infer the homotopy type of  $\check{\mathcal{M}}$  from  $\check{\nu}$ . In order to apply Theorem 1.3 in our setting, we have to consider the reach of the submanifold  $\check{\mathcal{M}}$ . For every  $\gamma > 0$ , we denote by  $\tau_\gamma$  the reach of  $\check{\mathcal{M}}$ , where the space  $E \times M(E)$  is endowed with the norm  $\|\cdot\|_\gamma$ . Alternatively,  $\tau_\gamma$  can be defined as the reach of  $\check{\mathcal{M}}_\gamma$  where the space  $E \times M(E)$  is endowed with the usual Euclidean norm  $\|(x, A)\|^2 = \|x\|^2 + \|A\|_F^2$ , and  $\check{\mathcal{M}}_\gamma$  is defined as  $\check{\mathcal{M}}_\gamma = \{(x, \gamma A), (x, A) \in \check{\mathcal{M}}\}$ . Consider the application  $i_\gamma : E \times M(E) \rightarrow E \times M(E)$  defined as  $i_\gamma(x, A) = (x, \gamma A)$ . We have  $\check{\mathcal{M}}_\gamma = i_\gamma(\check{\mathcal{M}})$ . Moreover, observe that, for every probability measures  $\alpha, \beta$  on  $E \times M(E)$ ,

$$W_{2,\gamma}(\alpha, \beta) = W_2((i_\gamma)_* \alpha, (i_\gamma)_* \beta),$$

where  $W_2$  denotes the 2-Wasserstein distance on  $E \times M(E)$  endowed with the usual Euclidean norm  $\|\cdot\|$ . Corollary 4.2 then reformulates as  $W_2((i_\gamma)_* \check{\mu}_0, (i_\gamma)_* \check{\nu}) \leq (1 + \gamma c_{4.2}) r^{\frac{1}{2}}$ , and Theorem 1.3 yields the following corollary:

**Corollary 4.3.** *Let  $m \in (0, 1)$ . Assume that  $\mathcal{M}_0$  and  $\mu_0$  satisfy the hypotheses 1, 2, 3, and Hypothesis 4 with  $r_0 \geq m$ . Let  $\nu$  be any probability measure on  $E$ . Denote  $w = W_2(\mu, \nu)$ . Choose  $r > 0$  such that  $r \leq \frac{1}{2\rho} \wedge 1$  and  $w \leq (a \wedge 1) \left(\frac{r}{4}\right)^{d+2}$ .*

*Let  $\gamma > 0$ , and denote by  $\tau_\gamma$  the reach of  $\check{\mathcal{M}}^\gamma$ . Suppose that  $(1 + \gamma c_{4.2}) r^{\frac{1}{2}} \leq m^{\frac{1}{2}} \left(\frac{\tau_\gamma}{9} - \left(\frac{m}{a}\right)^{\frac{1}{d}}\right)$ . Define  $\epsilon = \left(\frac{m}{a}\right)^{\frac{1}{d}} + m^{-\frac{1}{2}} w$  and choose  $t \in [4\epsilon, R - 4\epsilon]$ . Then  $(d_{(i_\gamma)_* \check{\nu}, m})^t$  and  $\check{\mathcal{M}}$  are homotopic equivalent.*

**Illustration.** Say that  $\mathcal{M}$  is the lemniscate of Bernoulli, as in the introduction (Figure 2). Suppose that  $\mu$  is the uniform distribution on  $\mathcal{M}$ , and  $\nu$  is the empirical measure on a 500-sample of  $\mathcal{M}$ . Figure 22 represents the value of the DTM on  $\text{supp}(\check{\nu})$ . Observe that the outliers, i.e. points for which the local covariance matrix is not well estimated, have large DTM values.



Figure 21: The set  $\text{supp}(\check{\nu}) \subset \mathbb{R}^2 \times M(\mathbb{R}^2)$ . Parameter  $r = 0.03$ . Figure 22: The set  $\text{supp}(\check{\nu})$  with colors indicating the value of  $d_{\check{\nu}}(x)$ . Parameters  $\gamma = 2$  and  $m = 0.01$ .



### 4.3 Persistent homology with DTM-filtrations

In this subsection, we apply the DTM-filtrations to  $\check{\nu}$ , as defined in subsection 1.3.

**Corollary 4.4.** *Let  $m \in (0, 1)$ . Assume that  $\mathcal{M}_0$  and  $\mu_0$  satisfy the hypotheses 1, 2, 3, and Hypothesis 4 with  $r_0 \geq m$ . Let  $\nu$  be any probability measure. Denote  $W_2(\mu, \nu) = w$ . Suppose that  $r \leq \frac{1}{2\rho} \wedge 1$  and  $w \leq (a \wedge 1)(\frac{r}{4})^{d+2}$ . Then*

$$d_i(W[(i^\gamma)_*\check{\mu}_0], W[(i^\gamma)_*\check{\nu}]) \leq c_{1.6}(1 + \gamma c_{4.2})^{\frac{1}{2}} m^{-\frac{1}{2}} r^{\frac{1}{4}} + 2c_{1.5} m^{\frac{1}{d}}.$$

*Proof.* Let  $\check{w} = W_2((i^\gamma)_*\check{\mu}_0, (i^\gamma)_*\check{\nu})$ . Corollary 1.6 gives

$$d_i(W[\mu], W[\nu]) \leq c_{1.6} \left(\frac{\check{w}}{m}\right)^{\frac{1}{2}} + 2c_{1.5} m^{\frac{1}{d}},$$

and Corollary 4.2 gives  $(1 + \gamma c_{4.2})r^{\frac{1}{p}}$ . Combining these inequalities yields the result.  $\square$

**Illustration.** Say that  $\mu$  is the uniform measure on the union of five circles, as represented in Figure 23, and that  $\nu$  is the empirical measure on the point cloud drawn in Figure 24. It consists in 300 points per circle, and 100 points of clutter noise. Observe that the barcodes of the DTM-filtration  $W[(i^\gamma)_*\check{\mu}_0]$ , represented in Figure 25, reveals the homology of the disjoint union of five circles—which is the set  $\mathcal{M}_0$ . The same holds for the barcodes of  $W[(i^\gamma)_*\check{\nu}]$  (Figure 26). Only bars of length larger than 0.1 are displayed in red, and only bars of length larger than 0.01 are displayed in green.

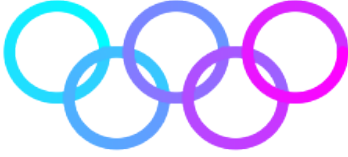


Figure 23: The set  $\text{supp}(\mu)$

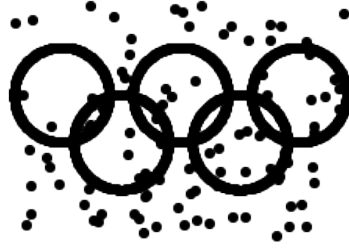


Figure 24: The set  $\text{supp}(\nu)$

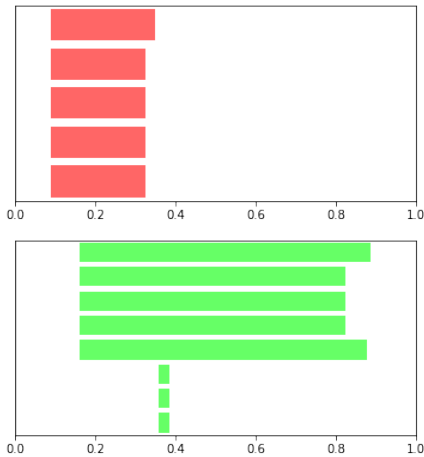


Figure 25: Persistence barcode of the 0- and 1-homology of the DTM-filtration on  $\check{\mu}_0$ . Parameters  $\gamma = 1$  and  $m = 0.01$

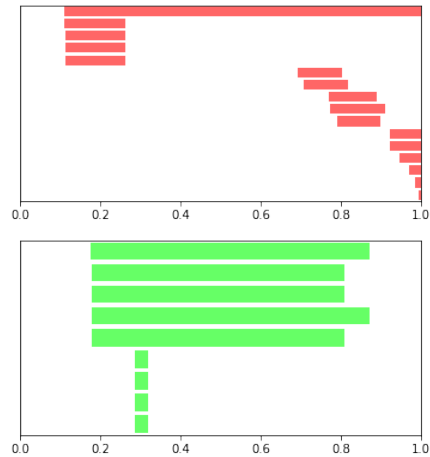


Figure 26: Persistence barcode of the 0- and 1-homology of the DTM-filtration on  $\check{\nu}$ . Parameters  $\gamma = 1$ ,  $m = 0.01$  and  $r = 0.03$

## 5 Conclusion

In this paper we described a method to estimate the tangent bundle of a manifold  $\mathcal{M}_0$  immersed in the Euclidean space, based on a sample of its image. This estimation is stable in Wasserstein distance. Using the DTM, we are able to estimate the homotopy type of  $\mathcal{M}_0$ . Moreover, via the DTM-filtrations, we can define a filtration of the space  $E \times M(E)$ , whose persistence module contains information about the homology of  $\mathcal{M}_0$ .

The robust estimation of tangent bundles of manifolds opens the way to the estimation of other topological invariants than homology groups—such as characteristic classes—a problem that will be addressed in a further work.

## A Supplementary material for Section 1

*Proof of Lemma 1.5.* By definition,  $\delta_{\mu,t}(x) = \inf\{r \geq 0, \mu(\overline{\mathcal{B}}(x,r)) > t\}$ , and  $d_\mu(x) = \frac{1}{m} \int_0^m \delta_{\mu,t}(x) dt$ . Using Hypothesis 5, we get  $\delta_{\mu,t}(x) \leq (\frac{t}{a})^{\frac{1}{d}}$ , and a simple computation yields  $d_\mu(x) \leq a^{-(1+\frac{1}{d})} m^{\frac{1}{d}}$ .  $\square$

*Proof of Corollary 1.6.* Let  $\pi$  be an optimal transport plan for  $w = W_2(\mu, \nu)$ . Denote  $\alpha = w^{\frac{1}{2}}$  and  $D = \text{diam}(\text{supp}(\mu))$ . Define  $\pi'$  to be  $\pi$  restricted to the set  $\{x, y \in E, \|x - y\| < \alpha\}$ . We denote its marginals  $\mu'$  and  $\nu'$ . By Markov inequality,  $1 - |\pi'| \leq \frac{w^2}{\alpha^2} = w$ . Consider the probability measures  $\overline{\mu'}$  and  $\overline{\nu'}$ . We have:

- $W_2(\mu, \overline{\mu'}) = 2D\alpha$ ,
- $W_2(\overline{\mu'}, \overline{\nu'}) \leq \alpha$ ,
- $W_2(\nu, \overline{\nu'}) \leq 2(1 + D)\alpha$ .

The first inequality is an application of Lemma B.2:

$$W_2(\mu, \overline{\mu'}) \leq 2(1 - |\mu'|)^{\frac{1}{2}} D = 2(1 - |\pi'|)^{\frac{1}{2}} D \leq 2w^{\frac{1}{2}} D.$$

We obtain the second one as follows:

$$W_2^2(\overline{\mu'}, \overline{\nu'}) = \int \|x - y\|^2 d\overline{\pi'}(x, y) = \int \|x - y\| \frac{d\pi'(x, y)}{|\pi'|} \leq \frac{1}{|\pi'|} \int \|x - y\| d\pi(x, y).$$

Now, Jensen inequality leads to  $W_2(\overline{\mu'}, \overline{\nu'}) \leq \frac{w}{|\pi'|^{\frac{1}{2}}}$ . Since  $1 - |\pi'| \leq w$ , we have  $\frac{w}{|\pi'|^{\frac{1}{2}}} \leq \frac{w}{1-w}$ , and  $w \leq \frac{1}{4}$  yields  $\frac{w}{1-w} \leq \alpha$ . Finally, we obtain the third inequality with the triangular inequality:

$$W_1(\nu, \overline{\nu'}) \leq W_1(\nu, \mu) + W_1(\mu, \overline{\mu'}) + W_1(\overline{\mu'}, \overline{\nu'}).$$

Now, note that we also have

- $c(\overline{\mu'}) \leq c(\mu) + m^{-\frac{1}{2}} 2D\alpha$ ,
- $c(\overline{\nu'}) \leq c(\mu) + (m^{-\frac{1}{2}} + m^{-\frac{1}{2}} 2D + 1)\alpha$ .

The first inequality follows from Theorem 1.2:

$$c(\overline{\mu'}) = \sup_{x \in \text{supp}(\overline{\mu'})} d_{\overline{\mu'}}(x) \leq \sup_{x \in \text{supp}(\mu')} d_{\mu'}(x) + m^{-\frac{1}{2}} W_2(\overline{\mu'}, \mu),$$

and we conclude with  $W_2(\mu, \overline{\mu'}) = 2D\alpha$ . In order to prove the second, we also use Theorem 1.2:

$$c(\overline{\nu'}) = \sup_{x \in \text{supp}(\overline{\nu'})} d_{\overline{\nu'}}(x) \leq \sup_{x \in \text{supp}(\nu')} d_{\nu'}(x) + m^{-\frac{1}{2}} W_2(\overline{\mu'}, \overline{\nu'})$$

Since  $\pi'$  has support included in  $\{x, y \in E, \|x - y\| < \alpha\}$ , we can use Proposition 1.1 to obtain

$$\sup_{x \in \text{supp}(\overline{\nu'})} d_{\overline{\mu'}}(x) \leq \sup_{x \in \text{supp}(\overline{\mu'})} d_{\overline{\mu'}}(x) + \alpha = c(\mu') + \alpha$$

and we deduce

$$c(\overline{\nu'}) \leq c(\mu') + \alpha + m^{-\frac{1}{2}} W_2(\overline{\mu'}, \overline{\nu'}) \leq c(\mu) + (m^{-\frac{1}{2}} + m^{-\frac{1}{2}} 2D + 1)\alpha.$$

To conclude, Theorem 1.4 gives:

$$\begin{aligned} d_i(W[\mu], W[\nu]) &\leq m^{-1} W_1(\mu, \overline{\mu'}) + m^{-1} W_1(\overline{\mu'}, \overline{\nu'}) + m^{-1} W_1(\nu, \overline{\nu'}) + c(\overline{\mu'}) + c(\overline{\nu'}) \\ &\leq (m^{-\frac{1}{2}}(4D + 1) + 4(D + 1))\alpha + 2c(\mu). \end{aligned}$$

Since  $m \leq 1$ , we can simplify this expression into

$$d_i(W[\mu], W[\nu]) \leq m^{-\frac{1}{2}}(8D + 5)\alpha + 2c(\mu).$$

We conclude the proof using  $c(\mu) \leq c_{1.5} m^{\frac{1}{d}}$  (Lemma 1.5).  $\square$

## B Supplementary material for Section 3

In this subsection, we suppose that  $\mu$  and  $\nu$  are any probability measures on  $E$ .

**Lemma B.1.** *For every  $x, y \in E$ ,  $\|x^{\otimes 2} - y^{\otimes 2}\|_{\mathbb{F}} \leq (\|x\| + \|y\|)\|x - y\|$ .*

*Proof.* We apply the triangular inequality to  $x^t x - y^t y = (x - y)^t x + y^t (x - y)$ :

$$\begin{aligned} \|x^t x - y^t y\|_{\mathbb{F}} &\leq \|(x - y)^t x\|_{\mathbb{F}} + \|y^t (x - y)\|_{\mathbb{F}} \leq \|x - y\| \|x\| + \|y\| \|x - y\| \\ &= (\|x\| + \|y\|)\|x - y\|. \end{aligned} \quad \square$$

**Lemma B.2.** *Let  $\mu'$  be a submeasure of  $\mu$  with  $|\mu'| > 0$ , and consider the corresponding probability measure  $\overline{\mu'}$ . Suppose that  $\text{supp}(\mu)$  is included in a ball  $\overline{\mathcal{B}}(x, r)$ . Then*

$$W_p(\mu, \overline{\mu'}) \leq 2(1 - |\mu'|)^{\frac{1}{p}} r.$$

*More generally, let  $\mu$  be any measure of positive mass (potentially with  $|\mu| \neq 1$ ), and let  $\mu'$  be a submeasure of  $\mu$  with  $|\mu'| > 0$ . Suppose that  $\text{supp}(\mu)$  is included in a ball  $\overline{\mathcal{B}}(x, r)$ . Then*

$$W_p(\overline{\mu}, \overline{\mu'}) \leq 2\left(1 - \frac{|\mu'|}{|\mu|}\right)^{\frac{1}{p}} r.$$

*Proof.* We start with the first inequality. Consider the intermediate probability measure  $\omega = \mu' + (1 - |\mu'|)\delta_x$ . We shall use the triangular inequality  $W_1(\mu, \overline{\mu'}) \leq W_1(\mu, \omega) + W_1(\omega, \overline{\mu'})$ . We can write

- $\mu = \mu' + (\mu - \mu')$ ,
- $\omega = \mu' + (1 - |\mu'|)\delta_x$ ,
- $\overline{\mu'} = \mu' + (\overline{\mu'} - \mu')$ .

Observe that  $\mu$  and  $\omega$  admits  $\mu'$  as a common submeasure of mass  $|\mu'|$ . Therefore we can build a transport plan between  $\mu$  and  $\omega$  where only a mass  $1 - |\mu'|$  of  $\mu$  is moved to  $x$ . In other words,

$$W_p(\mu, \omega) \leq (1 - |\mu'|)^{\frac{1}{p}} r.$$

Similarly, one shows that  $W_p(\omega, \overline{\mu'}) \leq (1 - |\mu'|)^{\frac{1}{p}} r$ .

Now let us prove the second inequality. Since  $\mu'$  is a submeasure of  $\mu$  of mass  $|\mu'|$ , then  $\frac{1}{|\mu|}\mu'$  is a submeasure of  $\overline{\mu} = \frac{1}{|\mu|}\mu$  of mass  $\frac{1}{|\mu|}|\mu'|$ . We then apply the first inequality.  $\square$

**Lemma B.3.** *Let  $x \in \text{supp}(\mu)$ . Suppose that  $x$  satisfies the Hypotheses 5 and 6 with  $\lambda(x) \wedge \frac{1}{2\rho} > r$ . Let  $y \in E$  such that  $\|x - y\| < \frac{r}{4}$ . Then  $|\mu_x|, |\mu_y| > 0$ , and*

$$W_1(\overline{\mu_x}, \overline{\mu_y}) \leq c_{B.3} \|x - y\|$$

with  $c_{B.3} = 2(1 + 4^{\frac{5^{d-1}}{3^d}}) \frac{b}{a}$ .

*Proof.* It is clear that  $|\mu_y| > 0$  since  $\mu(\overline{\mathcal{B}}(y, r)) \geq \mu(\overline{\mathcal{B}}(x, r - \|x - y\|))$  and  $x \in \text{supp}(\mu)$ .

Let us show the inequality  $W_1(\overline{\mu_x}, \overline{\mu_y}) \leq c_{B.3} \|x - y\|$  by studying the measure  $\mu$  on the intersection  $\overline{\mathcal{B}}(x, r) \cap \overline{\mathcal{B}}(y, r)$ . Let  $\mu_{x,y}$  be the restriction of  $\mu$  to  $\overline{\mathcal{B}}(x, r) \cap \overline{\mathcal{B}}(y, r)$ , and  $\overline{\mu_{x,y}}$  the corresponding probability measure. The triangular inequality gives:

$$W_1(\overline{\mu_x}, \overline{\mu_y}) \leq \underbrace{W_1(\overline{\mu_x}, \overline{\mu_{x,y}})}_{(1)} + \underbrace{W_1(\overline{\mu_{x,y}}, \overline{\mu_y})}_{(2)}.$$

*Term (1):* Let us show that  $W_1(\overline{\mu_x}, \overline{\mu_{x,y}}) \leq 2 \frac{b}{a} \|x - y\|$ . Note that  $\overline{\mu_{x,y}}$  is a submeasure of  $\overline{\mu_x}$ . According to Lemma B.2, we have

$$W_1(\overline{\mu_x}, \overline{\mu_{x,y}}) \leq 2(1 - \frac{|\mu_{x,y}|}{|\mu_x|})r = 2 \frac{|\mu_x| - |\mu_{x,y}|}{|\mu_x|} r.$$

We know from Hypothesis 5 that  $|\mu_x| \geq ar^d$ . On the other hand,

$$\begin{aligned} |\mu_x| - |\mu_{x,y}| &= \mu(\overline{\mathcal{B}}(x, r)) - \mu(\overline{\mathcal{B}}(x, r) \cap \overline{\mathcal{B}}(y, r)) \\ &\leq \mu(\overline{\mathcal{B}}(x, r)) - \mu(\overline{\mathcal{B}}(x, r - \|x - y\|)), \end{aligned}$$

hence we can apply Hypothesis 6 to get  $|\mu_x| - |\mu_{x,y}| \leq br^{d-1} \|x - y\|$ . We finally obtain

$$W_1(\overline{\mu_x}, \overline{\mu_{x,y}}) \leq 2 \frac{br^{d-1} \|x - y\|}{ar^d} r = 2 \frac{b}{a} \|x - y\|.$$

*Term (2):* Similarly, Lemma B.2 yields

$$W_1(\overline{\mu_y}, \overline{\mu_{x,y}}) \leq 2 \frac{|\mu_y| - |\mu_{x,y}|}{|\mu_y|} r.$$

Let us show that we still have  $|\mu_y| \geq a'r^d$  and  $|\mu_y| - |\mu_{x,y}| \leq b'r^{d-1} \|x - y\|$  with the constants  $a' = (\frac{3}{4})^d a$  and  $b' = 2(\frac{5}{4})^{d-1} b$ . The first inequality comes from Hypothesis 5:

$$\mu(\overline{\mathcal{B}}(y, r)) \geq \mu(\overline{\mathcal{B}}(x, r - \|x - y\|)) \geq a(r - \|x - y\|)^d$$

and  $\|x - y\| \leq \frac{r}{4}$ . The second inequality comes from Hypothesis 6:

$$\begin{aligned} \mu(\overline{\mathcal{B}}(y, r)) - \mu(\overline{\mathcal{B}}(x, r) \cap \overline{\mathcal{B}}(y, r)) &\leq \mu(\overline{\mathcal{B}}(x, r + \|x - y\|)) - \mu(\overline{\mathcal{B}}(x, r - \|x - y\|)) \\ &\leq b(r + \|x - y\|)^{d-1} 2\|x - y\| \end{aligned}$$

and  $\|x - y\| \leq \frac{r}{4}$ . To conclude,

$$W_1(\overline{\mu_y}, \overline{\mu_{x,y}}) \leq 2 \frac{2(\frac{5}{4})^{d-1} r^{d-1} a \|x - y\|}{2(\frac{3}{4})^d b r^d} r = 8 \frac{5^{d-1} b}{3^d a} \|x - y\|. \quad \square$$

**Lemma B.4.** *Let  $x \in \text{supp}(\mu)$ . Suppose that  $x$  satisfies the Hypotheses 5 and 7 at  $x$  with  $\frac{1}{2\rho} > r$ . Let  $y \in E$  such that  $\|x - y\| < \frac{r}{4}$ . Then  $|\mu_x|, |\mu_y| > 0$ , and*

$$W_1(\overline{\mu_x}, \overline{\mu_y}) \leq c_{B.4} r^{\frac{1}{2}} \|x - y\|^{\frac{1}{2}}$$

with  $c_{B.4} = (2 + \frac{2^{\frac{5}{2}} 5^{d-\frac{1}{2}}}{3^d}) \frac{b'}{a}$ .

*Proof.* The proof is similar to Lemma B.5 with slight modifications. We still consider

$$W_1(\overline{\mu_x}, \overline{\mu_y}) \leq \underbrace{W_1(\overline{\mu_x}, \overline{\mu_{x,y}})}_{(1)} + \underbrace{W_1(\overline{\mu_{x,y}}, \overline{\mu_y})}_{(2)}.$$

*Term (1):* We have  $W_1(\overline{\mu_x}, \overline{\mu_{x,y}}) \leq 2 \frac{|\mu_x| - |\mu_{x,y}|}{|\mu_x|} r$ . Hypothesis 5 still gives  $|\mu_x| \geq ar^d$ . But Hypothesis 7 now yields

$$\begin{aligned} |\mu_x| - |\mu_{x,y}| &\leq \mu(\overline{\mathcal{B}}(x, r)) - \mu(\overline{\mathcal{B}}(x, r - \|x - y\|)) \\ &\leq b' r^{d-\frac{1}{2}} \|x - y\|^{\frac{1}{2}}. \end{aligned}$$

We finally obtain  $W_1(\overline{\mu_x}, \overline{\mu_{x,y}}) \leq 2 \frac{b'}{a} r^{\frac{1}{2}} \|x - y\|^{\frac{1}{2}}$ .

*Term (2):* In order to bound  $W_1(\overline{\mu_y}, \overline{\mu_{x,y}}) \leq 2 \frac{|\mu_y| - |\mu_{x,y}|}{|\mu_y|} r$ , Hypothesis 5 still gives  $|\mu_x| \geq (\frac{3}{4})^d ar^d$ , and Hypothesis 7 yields

$$\begin{aligned} |\mu_y| - |\mu_{x,y}| &\leq \mu(\overline{\mathcal{B}}(x, r + \|x - y\|)) - \mu(\overline{\mathcal{B}}(x, r - \|x - y\|)) \\ &\leq b'(r + \|x - y\|)^{d-\frac{1}{2}} (2\|x - y\|)^{\frac{1}{2}}, \end{aligned}$$

which is not greater than  $b'(\frac{5}{4}r)^{d-\frac{1}{2}} (2\|x - y\|)^{\frac{1}{2}}$ .

We finally get  $W_1(\overline{\mu_y}, \overline{\mu_{x,y}}) \leq 2 \frac{b'(\frac{5}{4}r)^{d-\frac{1}{2}} (2\|x - y\|)^{\frac{1}{2}}}{(\frac{3}{4})^d ar^d} r \leq \frac{2^{\frac{5}{2}} 5^{d-\frac{1}{2}} b'}{3^d a} r^{\frac{1}{2}} \|x - y\|^{\frac{1}{2}}$ .  $\square$

**Lemma B.5.** *Let  $w = W_p(\mu, \nu)$ . Let  $y \in E$ . Suppose that there exists  $x \in \text{supp}(\mu)$  such that  $\|x - y\| \leq \alpha$  with  $\alpha = (\frac{w}{r^{d-1}})^{\frac{1}{2}}$ , and that  $\mu$  satisfies the Hypotheses 5 and 6 at  $x$  with  $\lambda(x) \wedge \frac{1}{2p} > r$ . Assume that  $w \leq (a \wedge 1)(\frac{r}{4})^{d+1}$ . Then*

$$W_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{B.5} \alpha$$

with  $c_{B.5} = \frac{2^{d-1}}{a} + 2 \frac{12 \cdot 5^{d-1} b + 1}{3^d a} + 2^{d+3} \frac{(\frac{3}{2})^{d-1} b + 1}{a}$ .

*Proof.* Let  $\pi$  be an optimal transport for  $W_p(\mu, \nu)$ . Define  $\pi_y$  to be the restriction of the measure  $\pi$  to the set  $\overline{\mathcal{B}}(y, r) \times \overline{\mathcal{B}}(y, r) \subset E \times E$ . Its marginals  $p_{1*}\pi_y$  and  $p_{2*}\pi_y$  are submeasures of  $\mu_y$  and  $\nu_y$ . We shall use the triangular inequality:

$$W_1(\overline{\mu_y}, \overline{\nu_y}) \leq \underbrace{W_1(\overline{\mu_y}, \overline{p_{1*}\pi_y})}_{(1)} + \underbrace{W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y})}_{(2)} + \underbrace{W_1(\overline{p_{2*}\pi_y}, \overline{\nu_y})}_{(3)}$$

Before examining each of these terms, note that we have

$$|\pi_y| = |p_{1*}\pi_y| = |p_{2*}\pi_y| \geq \mu(\overline{\mathcal{B}}(y, r - \alpha)) - \frac{w}{\alpha} \quad (9)$$

$$|\nu_y| \leq \mu(\overline{\mathcal{B}}(y, r + \alpha)) + \frac{w}{\alpha} \quad (10)$$

$$|\nu_y| \geq \mu(\overline{\mathcal{B}}(y, r - \alpha)) - \frac{w}{\alpha} \quad (11)$$

The first equation can be proven as follows:

$$\begin{aligned} \mu(\overline{\mathcal{B}}(y, r - \alpha)) &= \pi(\overline{\mathcal{B}}(y, r - \alpha) \times E) \\ &= \pi(\overline{\mathcal{B}}(y, r - \alpha) \times \overline{\mathcal{B}}(y, r)) + \pi(\overline{\mathcal{B}}(y, r - \alpha) \times \overline{\mathcal{B}}(y, r)^c) \end{aligned}$$

On the one hand,  $\pi(\overline{\mathcal{B}}(y, r - \alpha) \times \overline{\mathcal{B}}(y, r)) \leq \pi(\overline{\mathcal{B}}(y, r) \times \overline{\mathcal{B}}(y, r)) \leq |\pi_y|$ . On the other hand, Markov inequality yields

$$\pi(\overline{\mathcal{B}}(y, r - \alpha) \times \overline{\mathcal{B}}(y, r)^c) \leq \pi(\{(z, z'), \|z - z'\| \geq \alpha\}) \leq \frac{1}{\alpha} \int \|z - z'\| d\pi(z, z'),$$

and Jensen inequality gives

$$\frac{1}{\alpha} \int \|z - z'\| d\pi(z, z') \leq \frac{1}{\alpha} \left( \int \|z - z'\|^p d\pi(z, z') \right)^{\frac{1}{p}} = \frac{w}{\alpha}.$$

We deduce that  $\mu(\overline{\mathcal{B}}(y, r - \alpha)) \leq |\pi_y| + \frac{w}{\alpha}$ , which gives Equation 9. Equations 10 and 11 can be proven similarly.

In addition, note that the assumption  $w \leq (a \wedge 1) \left(\frac{r}{4}\right)^{d+1}$  yields

$$\alpha \leq \frac{r}{4} \tag{12}$$

$$\frac{w}{\alpha} \leq \frac{a}{2} \left(\frac{r}{2}\right)^d \tag{13}$$

We now study the terms (1), (2) and (3).

*Term (2):* Since  $\overline{\pi_y} = \frac{\pi_y}{|\pi_y|}$  is a transport plan between  $\overline{p_{1*}\pi_y}$  and  $\overline{p_{2*}\pi_y}$ , we have

$$W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \leq \int \|z - z'\| \frac{d\pi_y(z, z')}{|\pi_y|} \leq \frac{1}{|\pi_y|} \int \|z - z'\| d\pi(z, z').$$

Moreover, Jensen inequality yields  $\int \|z - z'\| d\pi(z, z') \leq w$ . Hence

$$W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \leq \frac{w}{|\pi_y|}.$$

Let us prove that  $|\pi_y| \geq \frac{a}{2} \left(\frac{r}{2}\right)^d$ . According to Equation 9,  $|\pi_y| \geq \mu(\overline{\mathcal{B}}(y, r - \alpha)) - \frac{w}{\alpha}$ . Now, remark that  $\mu(\overline{\mathcal{B}}(y, r - \alpha)) \geq \frac{a}{2^d} r^d$ . Indeed, using Hypothesis 5,

$$\mu(\overline{\mathcal{B}}(y, r - \alpha)) \geq \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x - y\|)) \geq a(r - \alpha - \|x - y\|)^d,$$

and we conclude with  $\|x - y\| \leq \alpha \leq \frac{r}{4}$ . Now, using Equation 13, we get

$$\begin{aligned} |\pi_y| &\geq \mu(\overline{\mathcal{B}}(y, r - \alpha)) - \frac{w}{\alpha} \\ &\geq a \left(\frac{r}{2}\right)^d - \frac{a}{2} \left(\frac{r}{2}\right)^d \geq \frac{a}{2} \left(\frac{r}{2}\right)^d \end{aligned}$$

Finally, since  $\alpha = \left(\frac{w}{r^{d-1}}\right)^{\frac{1}{2}}$  and  $\alpha \leq \frac{r}{4}$ , we obtain

$$W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \leq \frac{w}{|\pi_y|} \leq \frac{w}{\frac{a}{2} \left(\frac{r}{2}\right)^d} = \frac{2^{d+1}}{a} \alpha^2 \frac{1}{r} \leq \frac{2^{d-1}}{a} \alpha.$$

*Term (1):* According to Lemma B.2, we have

$$W_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \leq 2 \frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} r.$$

We can use Equation 9 to get

$$\begin{aligned} |\mu_y| - |p_{1*}\pi_y| &\leq \mu(\overline{\mathcal{B}}(y, r)) - \mu(\overline{\mathcal{B}}(y, r - \alpha)) + \frac{w}{\alpha} \\ &\leq \mu(\overline{\mathcal{B}}(x, r + \|x - y\|)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x - y\|)) + \frac{w}{\alpha}. \end{aligned}$$

By Hypothesis 6,

$$\mu(\overline{\mathcal{B}}(x, r + \|x - y\|)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x - y\|)) \leq b(r + \|x - y\|)^{d-1} (2\|x - y\| + \alpha),$$

which is not greater than  $b(\frac{5}{4}r)^{d-1}3\alpha$  since  $\|x - y\| \leq \alpha \leq \frac{r}{4}$ . Moreover,  $\frac{w}{\alpha} = r^{d-1}\alpha$ , and we obtain

$$|\mu_y| - |p_{1*}\pi_y| \leq (3(\frac{5}{4})^{d-1}b + 1)r^{d-1}\alpha,$$

Finally, thanks to Hypothesis 5, we write

$$\begin{aligned} |\mu_y| &= \mu(\overline{\mathcal{B}}(y, r)) \geq \mu(\overline{\mathcal{B}}(x, r - \|x - y\|)) \\ &\geq a(r - \|x - y\|)^d \geq a(\frac{3}{4})^d r^d \end{aligned}$$

and we obtain

$$\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} \leq \frac{((3(\frac{5}{4})^{d-1}b + 1)r^{d-1})}{a(\frac{3}{4})^d r^d} \alpha = \frac{1}{r} \cdot \frac{12 \cdot 5^{d-1}b + 1}{3^d a} \alpha.$$

We deduce

$$W_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \leq 2 \frac{12 \cdot 5^{d-1}b + 1}{3^d a} \alpha.$$

*Term (3):* It is similar to Term (1). First, one shows that

$$W_1(\overline{\nu_y}, \overline{p_{2*}\pi_y}) \leq 2 \frac{|\nu_y| - |p_{2*}\pi_y|}{|\nu_y|} r.$$

Using Equations 9 and 10 we get

$$\begin{aligned} |\nu_y| - |p_{2*}\pi_y| &\leq \mu(\overline{\mathcal{B}}(y, r + \alpha)) + \frac{w}{\alpha} - \mu(\overline{\mathcal{B}}(y, r - \alpha)) + \frac{w}{\alpha} \\ &\leq \mu(\overline{\mathcal{B}}(x, r + \|x - y\| + \alpha)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x - y\|)) + 2\frac{w}{\alpha}. \end{aligned}$$

By Hypothesis 6, we have

$$\begin{aligned} &\mu(\overline{\mathcal{B}}(x, r + \|x - y\| + \alpha)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x - y\|)) \\ &\leq b(r + \|x - y\| + \alpha)^{d-1}(2\|x - y\| + 2\alpha) \end{aligned}$$

which is not greater than  $b(\frac{3}{2}r)^{d-1}4\alpha$  since  $\|x - y\| \leq \alpha \leq \frac{r}{4}$ . Moreover,  $\frac{w}{\alpha} = r^{d-1}\alpha$ , and we obtain

$$|\nu_y| - |p_{2*}\pi_y| \leq (4(\frac{3}{2})^{d-1}b + 2)r^{d-1}\alpha.$$

We have seen that

$$|\nu_y| \geq \mu(\overline{\mathcal{B}}(y, r - \alpha)) - \frac{w}{\alpha} \geq \frac{a}{2}(\frac{r}{2})^d.$$

Hence

$$\frac{|\nu_y| - |p_{2*}\pi_y|}{|\nu_y|} \leq \frac{(4(\frac{3}{2})^{d-1}b + 2)r^{d-1}}{\frac{a}{2}(\frac{r}{2})^d} \alpha = \frac{1}{r} \cdot 2^{d+2} \frac{(\frac{3}{2})^{d-1}b + 1}{a} \alpha,$$

and we finally obtain

$$W_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \leq 2^{d+3} \frac{(\frac{3}{2})^{d-1}b + 1}{a} \alpha.$$

To conclude, summing up these three terms gives  $W_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{B.5}\alpha$  with  $c_{B.5} = \frac{2^{d-1}}{a} + 2\frac{12 \cdot 5^{d-1}b + 1}{3^d a} + 2^{d+3} \frac{(\frac{3}{2})^{d-1}b + 1}{a}$ .  $\square$

**Lemma B.6.** Let  $w = W_p(\mu, \nu)$ . Let  $y \in E$ . Suppose that there exists  $x \in \text{supp}(\mu)$  such that  $\|x - y\| \leq \alpha$  with  $\alpha = (\frac{w}{r^{d-1}})^{\frac{1}{2}}$ , and that  $\mu$  satisfies the Hypotheses 5 and 7 at  $x$  with  $\frac{1}{2\rho} > r$ . Assume that  $w \leq (a \wedge 1)(\frac{r}{4})^{d+1}$ . Then

$$W_1(\overline{\mu_y}, \overline{\nu_y}) \leq c_{B.6} r^{\frac{1}{2}} \alpha^{\frac{1}{2}}$$

$$\text{with } c_{B.6} = \frac{2^{d-2}}{a} + \frac{4 \cdot 3^{\frac{1}{2}} 5^{d-\frac{1}{2}} b' + 4^{d-\frac{1}{2}}}{3^d a} + 2 \cdot 4^d \frac{2b'(\frac{3}{2})^{d-\frac{1}{2}+1}}{3^d a}.$$

*Proof.* The proof is similar as Lemma B.5. Let us highlight the modifications. Since  $\alpha \leq \frac{r}{4}$  and  $\frac{w}{\alpha} = r^{d-1}\alpha$ , we have the inequalities

$$\alpha^{\frac{1}{2}} \leq \frac{1}{2} r^{\frac{1}{2}}$$

$$\frac{w}{\alpha} \leq \frac{1}{2} r^{d-\frac{1}{2}} \alpha^{\frac{1}{2}}$$

We still write the triangular inequality:

$$W_1(\overline{\mu_y}, \overline{\nu_y}) \leq \underbrace{W_1(\overline{\mu_y}, \overline{p_{1*}\pi_y})}_{(1)} + \underbrace{W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y})}_{(2)} + \underbrace{W_1(\overline{p_{2*}\pi_y}, \overline{\nu_y})}_{(3)}$$

where  $\pi$  is an optimal transport plan for  $W_p(\mu, \nu)$ .

*Term (2):* The argument to obtain  $W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \leq \frac{2^{d-1}}{a} \alpha$  is unchanged, and we use  $\alpha^{\frac{1}{2}} \leq \frac{1}{2} r^{\frac{1}{2}}$  to get

$$W_1(\overline{p_{1*}\pi_y}, \overline{p_{2*}\pi_y}) \leq \frac{2^{d-2}}{a} \alpha^{\frac{1}{2}} r^{\frac{1}{2}}.$$

*Term (1):* Using Hypothesis 7, we have

$$\begin{aligned} & \mu(\overline{\mathcal{B}}(x, r + \|x - y\|)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x - y\|)) \\ & \leq b'(r + \|x - y\|)^{d-\frac{1}{2}} (2\|x - y\| + \alpha)^{\frac{1}{2}} \\ & \leq b'(\frac{5}{4}r)^{d-\frac{1}{2}} 3^{\frac{1}{2}} \alpha^{\frac{1}{2}}. \end{aligned}$$

And since  $\frac{w}{\alpha} \leq \frac{1}{2} r^{d-\frac{1}{2}} \alpha^{\frac{1}{2}}$ , we get

$$\begin{aligned} |\mu_y| - |p_{1*}\pi_y| & \leq \mu(\overline{\mathcal{B}}(x, r + \|x - y\|)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x - y\|)) + \frac{w}{\alpha} \\ & \leq (b'(\frac{5}{4})^{d-\frac{1}{2}} 3^{\frac{1}{2}} + \frac{1}{2}) r^{d-\frac{1}{2}} \alpha^{\frac{1}{2}}. \end{aligned}$$

Finally, we use

$$\begin{aligned} |\mu_y| & = \mu(\overline{\mathcal{B}}(y, r)) \geq \mu(\overline{\mathcal{B}}(x, r - \|x - y\|)) \\ & \geq a(r - \|x - y\|)^d \geq a(\frac{3}{4})^d r^d \end{aligned}$$

to obtain

$$\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} \leq \frac{((b'(\frac{5}{4})^{d-\frac{1}{2}} 3^{\frac{1}{2}} + \frac{1}{2}) r^{d-\frac{1}{2}})}{a(\frac{3}{4})^d r^d} \alpha^{\frac{1}{2}} = \frac{1}{r^{\frac{1}{2}}} \cdot \frac{2 \cdot 3^{\frac{1}{2}} 5^{d-\frac{1}{2}} b' + 4^{d-\frac{1}{2}}}{3^d a} \alpha^{\frac{1}{2}}$$

and we deduce

$$W_1(\overline{\mu_y}, \overline{p_{1*}\pi_y}) \leq 2 \frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} r \leq \frac{4 \cdot 3^{\frac{1}{2}} 5^{d-\frac{1}{2}} b' + 4^{d-\frac{1}{2}}}{3^d a} r^{\frac{1}{2}} \alpha^{\frac{1}{2}}.$$



Term (3): We use Hypothesis 7 to get

$$\begin{aligned} & \mu(\overline{\mathcal{B}}(x, r + \|x - y\| + \alpha)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x - y\|)) \\ & \leq b'(r + \|x - y\| + \alpha)^{d-\frac{1}{2}}(2\|x - y\| + 2\alpha)^{\frac{1}{2}} \\ & \leq 2b'\left(\frac{3}{2}r\right)^{d-\frac{1}{2}}\alpha^{\frac{1}{2}}. \end{aligned}$$

And since  $\frac{w}{\alpha} \leq \frac{1}{2}r^{d-\frac{1}{2}}\alpha^{\frac{1}{2}}$ , we get

$$\begin{aligned} |\nu_y| - |p_{2*}\pi_y| & \leq \mu(\overline{\mathcal{B}}(x, r + \|x - y\| + \alpha)) - \mu(\overline{\mathcal{B}}(x, r - \alpha - \|x - y\|)) + 2\frac{w}{\alpha} \\ & \leq (2b'\left(\frac{3}{2}\right)^{d-\frac{1}{2}} + 1)r^{d-\frac{1}{2}}\alpha^{\frac{1}{2}}. \end{aligned}$$

Finally, we use

$$\begin{aligned} |\mu_y| = \mu(\overline{\mathcal{B}}(y, r)) & \geq \mu(\overline{\mathcal{B}}(x, r - \|x - y\|)) \\ & \geq a(r - \|x - y\|)^d \geq a\left(\frac{3}{4}\right)^d r^d \end{aligned}$$

to obtain

$$\frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} \leq \frac{(2b'\left(\frac{3}{2}\right)^{d-\frac{1}{2}} + 1)r^{d-\frac{1}{2}}}{a\left(\frac{3}{4}\right)^d r^d} \alpha^{\frac{1}{2}} = \frac{1}{r^{\frac{1}{2}}} \cdot 4^d \frac{2b'\left(\frac{3}{2}\right)^{d-\frac{1}{2}} + 1}{3^d a} \alpha^{\frac{1}{2}}$$

and we deduce

$$W_1(\mu_y, p_{1*}\pi_y) \leq 2 \frac{|\mu_y| - |p_{1*}\pi_y|}{|\mu_y|} r \leq 2 \cdot 4^d \frac{2b'\left(\frac{3}{2}\right)^{d-\frac{1}{2}} + 1}{3^d a} r^{\frac{1}{2}} \alpha^{\frac{1}{2}}. \quad \square$$

## References

- [Aam18] Eddie Aamari. *Vitesses de convergence en inférence géométrique*. PhD thesis, Université Paris-Saclay, 2018.
- [ACG<sup>+</sup>18] Hirokazu Anai, Frédéric Chazal, Marc Glisse, Yuichi Ike, Hiroya Inakoshi, Raphaël Tinarrage, and Yuhei Umeda. Dtm-based filtrations. In *Symposium on Computational Geometry*, 2018.
- [ACLZ17] Ery Arias-Castro, Gilad Lerman, and Teng Zhang. Spectral clustering based on local pca. *The Journal of Machine Learning Research*, 18(1):253–309, 2017.
- [AKC<sup>+</sup>19] Eddie Aamari, Jisu Kim, Frédéric Chazal, Bertrand Michel, Alessandro Rinaldo, and Larry Wasserman. Estimating the Reach of a Manifold. *Electronic journal of statistics*, April 2019.
- [CCSM11] F. Chazal, D. Cohen-Steiner, and Q. Mérigot. Geometric inference for probability measures. *Journal on Found. of Comp. Mathematics*, 11(6):733–751, 2011.
- [CdSGO16] Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Oudot. *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics, 2016.
- [dC92] M.P. do Carmo. *Riemannian Geometry*. Mathematics (Boston, Mass.). Birkhäuser, 1992.
- [Fed59] Herbert Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93(3):418–491, 1959.

- [Fed14] Herbert Federer. *Geometric measure theory*. Springer, 2014.
- [MMM18] Diego H Díaz Martínez, Facundo Mémoli, and Washington Mio. The shape of data and probability measures. *Applied and Computational Harmonic Analysis*, 2018.
- [MSW19] Facundo Memoli, Zane Smith, and Zhengchao Wan. The Wasserstein transform. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 4496–4504, Long Beach, California, USA, 09–15 Jun 2019. PMLR.
- [NSW08] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete & Computational Geometry*, 39(1-3):419–441, 2008.