

Non-Asymptotic Rates for Manifold, Tangent Space, and Curvature Estimation

Eddie Aamari, Clément Levrard

▶ **To cite this version:**

Eddie Aamari, Clément Levrard. Non-Asymptotic Rates for Manifold, Tangent Space, and Curvature Estimation. Annals of Statistics, 2019, 47 (1), 10.1214/18-AOS1685 . hal-01516032v3

HAL Id: hal-01516032
https://hal.science/hal-01516032v3

Submitted on 2 Feb 2018

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.

NON-ASYMPTOTIC RATES FOR MANIFOLD, TANGENT SPACE AND CURVATURE ESTIMATION

By Eddie Aamari[§,*,†,‡] and Clément Levrard[¶,*,†]

U.C. San Diego[§] *, Université Paris-Diderot*[¶]

Abstract: Given a noisy sample from a submanifold $M \subset \mathbb{R}^D$, we derive optimal rates for the estimation of tangent spaces $T_X M$, the second fundamental form II_X^M, and the submanifold M. After motivating their study, we introduce a quantitative class of $\mathcal{C}^k$-submanifolds in analogy with Hölder classes. The proposed estimators are based on local polynomials and allow to deal simultaneously with the three problems at stake. Minimax lower bounds are derived using a conditional version of Assouad's lemma when the base point X is random.

1. Introduction

A wide variety of data can be thought of as being generated on a shape of low dimensionality compared to possibly high ambient dimension. This point of view led to the development of the so-called topological data analysis, which proved fruitful for instance when dealing with physical parameters subject to constraints, biomolecule conformations, or natural images [35]. This field intends to associate geometric quantities to data without regard of any specific coordinate system or parametrization. If the underlying structure is sufficiently smooth, one can model a point cloud $\mathbb{X}_n = \{X_1, \dots, X_n\}$ as being sampled on a d-dimensional submanifold $M \subset \mathbb{R}^D$. In such a case, geometric and topological intrinsic quantities include (but are not limited to) homology groups [28], persistent homology [15], volume [5], differential quantities [9] or the submanifold itself [20, 1, 26].

The present paper focuses on optimal rates for estimation of quantities up to order two: (0) the submanifold itself, (1) tangent spaces, and (2) second fundamental forms.

Among these three questions, a special attention has been paid to the estimation of the submanifold. In particular, it is a central problem in manifold

[*]Research supported by ANR project TopData ANR-13-BS01-0008

[†]Research supported by Advanced Grant of the European Research Council GUDHI

[‡]Supported by the Conseil régional d'Île-de-France program RDM-IdF

MSC 2010 subject classifications: 62G05, 62C20

Keywords and phrases: geometric inference, minimax, manifold learning

learning. Indeed, there exists a wide bunch of algorithms intended to reconstruct submanifolds from point clouds (Isomap [32], LLE [29], and restricted Delaunay Complexes [6, 12] for instance), but few come with theoretical guarantees [20, 1, 26]. Up to our knowledge, minimax lower bounds were used to prove optimality in only one case [20]. Some of these reconstruction procedures are based on tangent space estimation [6, 1, 12]. Tangent space estimation itself also yields interesting applications in manifold clustering [19, 4]. Estimation of curvature-related quantities naturally arises in shape reconstruction, since curvature can drive the size of a meshing. As a consequence, most of the associated results deal with the case $d = 2$ and $D = 3$, though some of them may be extended to higher dimensions [27, 23]. Several algorithms have been proposed in that case [30, 9, 27, 23], but with no analysis of their performances from a statistical point of view.

To assess the quality of such a geometric estimator, the class of submanifolds over which the procedure is evaluated has to be specified. Up to now, the most commonly used model for submanifolds relied on the reach τ_M, a generalized convexity parameter. Assuming $\tau_M \geq \tau_{min} > 0$ involves both local regularity — a bound on curvature — and global regularity — no arbitrarily pinched area —. This $\mathcal{C}^2$-like assumption has been extensively used in the computational geometry and geometric inference fields [1, 28, 15, 5, 20]. One attempt of a specific investigation for higher orders of regularity $k \geq 3$ has been proposed in [9].

Many works suggest that the regularity of the submanifold has an important impact on convergence rates. This is pretty clear for tangent space estimation, where convergence rates of PCA-based estimators range from $(1/n)^{1/d}$ in the $\mathcal{C}^2$ case [1] to $(1/n)^{\alpha}$ with $1/d < \alpha < 2/d$ in more regular settings [31, 33]. In addition, it seems that PCA-based estimators are outperformed by estimators taking into account higher orders of smoothness [11, 9], for regularities at least $\mathcal{C}^3$. For instance fitting quadratic terms leads to a convergence rate of order $(1/n)^{2/d}$ in [11]. These remarks naturally led us to investigate the properties of local polynomial approximation for regular submanifolds, where "regular" has to be properly defined. Local polynomial fitting for geometric inference was studied in several frameworks such as [9]. In some sense, a part of our work extends these results, by investigating the dependency of convergence rates on the sample size n, but also on the order of regularity k and the ambient and intrinsic dimensions d and D.

1.1. Overview of the Main Results

In this paper, we build a collection of models for $\mathcal{C}^k$-submanifolds ($k \geq 3$) that naturally generalize the commonly used one for $k = 2$ (Section

2). Roughly speaking, these models are defined by their local differential regularity k in the usual sense, and by their minimum reach $\tau_{min} > 0$ that may be thought of as a global regularity parameter (see Section 2.2). On these models, we study the non-asymptotic rates of estimation for tangent space, curvature, and manifold estimation (Section 3). Roughly speaking, if M is a $\mathcal{C}^k_{\tau_{min}}$ submanifold and if $Y_1, \dots, Y_n$ is an n-sample drawn on M uniformly enough, then we can derive the following minimax bounds:

(Theorems 2 and 3)
$$\inf_{\hat{T}} \sup_{\substack{M \in \mathcal{C}^k \\ \tau_M \geq \tau_{min}}} \mathbb{E} \max_{1 \leq j \leq n} \angle\big(T_{Y_j} M, \hat{T}_j\big) \asymp \left(\frac{1}{n}\right)^{\frac{k-1}{d}},$$

where $T_y M$ denotes the tangent space of M at y;

(Theorems 4 and 5)
$$\inf_{\widehat{II}} \sup_{\substack{M \in \mathcal{C}^k \\ \tau_M \geq \tau_{min}}} \mathbb{E} \max_{1 \leq j \leq n} \big\| II^M_{Y_j} - \widehat{II}_j \big\| \asymp \left(\frac{1}{n}\right)^{\frac{k-2}{d}},$$

where II^M_y denotes the second fundamental form of M at y;

(Theorems 6 and 7)
$$\inf_{\hat{M}} \sup_{\substack{M \in \mathcal{C}^k \\ \tau_M \geq \tau_{min}}} \mathbb{E}\, d_H\big(M, \hat{M}\big) \asymp \left(\frac{1}{n}\right)^{\frac{k}{d}},$$

where d_H denotes the Hausdorff distance.

These results shed light on the influence of k, d, and n on these estimation problems, showing for instance that the ambient dimension D plays no role. The estimators proposed for the upper bounds all rely on the analysis of local polynomials, and allow to deal with the three estimation problems in a unified way (Section 5.1). Some of the lower bounds are derived using a new version of Assouad's Lemma (Section 5.2.2).

We also emphasize the influence of the reach τ_M of the manifold M in Theorem 1. Indeed, we show that whatever the local regularity k of M, if we only require $\tau_M \geq 0$, then for any fixed point $y \in M$,

$$\inf_{\hat{T}} \sup_{\substack{M \in \mathcal{C}^k \\ \tau_M \geq 0}} \mathbb{E} \angle\big(T_y M, \hat{T}\big) \geq 1/2, \qquad \inf_{\widehat{II}} \sup_{\substack{M \in \mathcal{C}^k \\ \tau_M \geq 0}} \mathbb{E} \big\| II^M_y - \widehat{II} \big\| \geq c > 0,$$

assessing that the global regularity parameter $\tau_{min} > 0$ is crucial for estimation purpose.

It is worth mentioning that our bounds also allow for perpendicular noise of amplitude $\sigma > 0$. When $\sigma \lesssim (1/n)^{\alpha/d}$ for $1 \leq \alpha$, then our estimators behave as if the corrupted sample $X_1, \dots, X_n$ were exactly drawn on a manifold with regularity α. Hence our estimators turn out to be optimal whenever

$\alpha \geq k$. If $\alpha < k$, the lower bounds suggest that better rates could be obtained with different estimators, by pre-processing data as in [21] for instance.

For the sake of completeness, geometric background and proofs of technical lemmas are given in the Appendix.

2. $\mathcal{C}^k$ Models for Submanifolds

2.1. Notation

Throughout the paper, we consider d-dimensional compact submanifolds $M \subset \mathbb{R}^D$ without boundary. The submanifolds will always be assumed to be at least $\mathcal{C}^2$. For all $p \in M$, T_pM stands for the tangent space of M at p [13, Chapter 0]. We let $II_p^M : T_pM \times T_pM \to T_pM^\perp$ denote the second fundamental form of M at p [13, p. 125]. II_p^M characterizes the curvature of M at p. The standard inner product in $\mathbb{R}^D$ is denoted by $\langle \cdot, \cdot \rangle$ and the Euclidean distance by $\|\cdot\|$. Given a linear subspace $T \subset \mathbb{R}^D$, write $T^\perp$ for its orthogonal space. We write $\mathcal{B}(p, r)$ for the closed Euclidean ball of radius $r > 0$ centered at $p \in \mathbb{R}^D$, and for short $\mathcal{B}_T(p, r) = \mathcal{B}(p, r) \cap T$. For a smooth function $\Phi : \mathbb{R}^D \to \mathbb{R}^D$ and $i \geq 1$, we let $d_x^i \Phi$ denote the ith order differential of Φ at $x \in \mathbb{R}^D$. For a linear map A defined on $T \subset \mathbb{R}^D$, $\|A\|_{\text{op}} = \sup_{v \in T} \frac{\|Av\|}{\|v\|}$ stands for the operator norm. We adopt the same notation $\|\cdot\|_{op}$ for tensors, i.e. multilinear maps. Similarly, if $\{A_x\}_{x \in T'}$ is a family of linear maps, its L^∞ operator norm is denoted by $\|A\|_{op} = \sup_{x \in T'} \|A_x\|_{op}$. When it is well defined, we will write $\pi_B(z)$ for the projection of $z \in \mathbb{R}^D$ onto the closed subset $B \subset \mathbb{R}^D$, that is the nearest neighbor of z in B. The distance between two linear subspaces $U, V \subset \mathbb{R}^D$ of the same dimension is measured by the principal angle $\angle(U, V) = \|\pi_U - \pi_V\|_{\text{op}}$. The Hausdorff distance [20] in $\mathbb{R}^D$ is denoted by d_H. For a probability distribution P, $\mathbb{E}_P$ stands for the expectation with respect to P. We write $P^{\otimes n}$ for the n-times tensor product of P.

Throughout this paper, C_α will denote a generic constant depending on the parameter α. For clarity's sake, C'_α, c_α, or c'_α may also be used when several constants are involved.

2.2. Reach and Regularity of Submanifolds

As introduced in [16], the reach τ_M of a subset $M \subset \mathbb{R}^D$ is the maximal neighborhood radius for which the projection π_M onto M is well defined. More precisely, denoting by $d(\cdot, M)$ the distance to M, the medial axis of M is defined to be the set of points which have at least two nearest neighbors

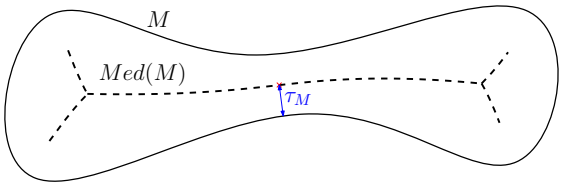


Figure 1: Medial axis and reach of a closed curve in the plane.

on M, that is

$$Med(M) = \left\{ z \in \mathbb{R}^D | \exists p \neq q \in M, \|z-p\| = \|z-q\| = d(z,M) \right\}.$$

The reach is then defined by

$$\tau_M = \inf_{p \in M} d\left(p, Med(M)\right) = \inf_{z \in Med(M)} d\left(z, M\right).$$

It gives a minimal scale of geometric and topological features of M. As a generalized convexity parameter, τ_M is a key parameter in reconstruction [1, 20] and in topological inference [28]. Having $\tau_M \geq \tau_{min} > 0$ prevents M from almost auto-intersecting, and bounds its curvature in the sense that $\left\| II_p^M \right\|_{op} \leq \tau_M^{-1} \leq \tau_{min}^{-1}$ for all $p \in M$ [28, Proposition 6.1].

For $\tau_{min} > 0$, we let $\mathcal{C}^2_{\tau_{min}}$ denote the set of d-dimensional compact connected submanifolds M of $\mathbb{R}^D$ such that $\tau_M \geq \tau_{min} > 0$. A key property of submanifolds $M \in \mathcal{C}^2_{\tau_{min}}$ is the existence of a parametrization closely related to the projection onto tangent spaces. We let $\exp_p : T_pM \to M$ denote the exponential map of M [13, Chapter 3], that is defined by $\exp_p(v) = \gamma_{p,v}(1)$, where $\gamma_{p,v}$ is the unique constant speed geodesic path of M with initial value p and velocity v.

LEMMA 1. *If $M \in \mathcal{C}^2_{\tau_{min}}$, $\exp_p : \mathcal{B}_{T_pM}(0, \tau_{min}/4) \to M$ is one-to-one. Moreover, it can be written as*

$$\begin{aligned} \exp_p \colon \mathcal{B}_{T_pM}(0, \tau_{min}/4) &\longrightarrow M \\ v &\longmapsto p + v + \mathbf{N}_p(v) \end{aligned}$$

with $\mathbf{N}_p$ such that for all $v \in \mathcal{B}_{T_pM}(0, \tau_{min}/4)$,

$$\mathbf{N}_p(0) = 0, \quad d_0\mathbf{N}_p = 0, \quad \left\| d_v\mathbf{N}_p \right\|_{op} \leq L_\perp \|v\|,$$

where $L_\perp = 5/(4\tau_{min})$. Furthermore, for all $p, y \in M$,

$$y - p = \pi_{T_pM}(y-p) + R_2(y-p),$$

where $\|R_2(y-p)\| \leq \frac{\|y-p\|^2}{2\tau_{min}}$.

A proof of Lemma A.1 is given in Section A.1 of the Appendix. In other words, elements of $\mathcal{C}^2_{\tau_{min}}$ have local parametrizations on top of their tangent spaces that are defined on neighborhoods with a minimal radius, and these parametrizations differ from the identity map by at most a quadratic term. The existence of such local parametrizations leads to the following convergence result: if data $Y_1, \dots, Y_n$ are drawn uniformly enough on $M \in \mathcal{C}^2_{\tau_{min}}$, then it is shown in [1, Proposition 14] that a tangent space estimator $\hat{T}$ based on local PCA achieves

$$\mathbb{E} \max_{1 \le j \le n} \angle\big(T_{Y_j} M, \hat{T}_j\big) \le C \left(\frac{1}{n}\right)^{\frac{1}{d}}.$$

When M is smoother, it has been proved in [11] that a convergence rate in $n^{-2/d}$ might be achieved, based on the existence of a local order 3 Taylor expansion of the submanifold on top of its tangent spaces. Thus, a natural extension of the $\mathcal{C}^2_{\tau_{min}}$ model to $\mathcal{C}^k$-submanifolds should ensure that such an expansion exists at order k and satisfies some regularity constraints. To this aim, we introduce the following class of regularity $\mathcal{C}^k_{\tau_{min},\mathbf{L}}$.

Definition 1. *For $k \ge 3$, $\tau_{min} > 0$, and $\mathbf{L} = (L_\perp, L_3, \dots, L_k)$, we let $\mathcal{C}^k_{\tau_{min},\mathbf{L}}$ denote the set of d-dimensional compact connected submanifolds M of $\mathbb{R}^D$ with $\tau_M \ge \tau_{min}$ and such that, for all $p \in M$, there exists a local one-to-one parametrization Ψ_p of the form:*

$$\begin{aligned} \Psi_p \colon \mathcal{B}_{T_pM}(0, r) &\longrightarrow M \\ v &\longmapsto p + v + \mathbf{N}_p(v) \end{aligned}$$

for some $r \ge \frac{1}{4L_\perp}$, with $\mathbf{N}_p \in \mathcal{C}^k\left(\mathcal{B}_{T_pM}(0, r), \mathbb{R}^D\right)$ such that

$$\mathbf{N}_p(0) = 0, \quad d_0\mathbf{N}_p = 0, \quad \left\|d^2_v \mathbf{N}_p\right\|_{op} \le L_\perp,$$

for all $\|v\| \le \frac{1}{4L_\perp}$. Furthermore, we require that

$$\left\|d^i_v \mathbf{N}_p\right\|_{op} \le L_i \text{ for all } 3 \le i \le k.$$

It is important to note that such a family of Ψ_p's exists for any compact $\mathcal{C}^k$-submanifold, if one allows τ^{-1}_{min}, $L_\perp$, L_3,...,L_k to be large enough. Note that the radius $1/(4L_\perp)$ has been chosen for convenience. Other smaller scales would do and we could even parametrize this constant, but without substantial benefits in the results.

The Ψ_p's can be seen as unit parametrizations of M. The conditions on $\mathbf{N}_p(0)$, $d_0\mathbf{N}_p$, and $d^2_v\mathbf{N}_p$ ensure that Ψ_p^{-1} is close to the projection π_{T_pM}.

The bounds on $d_v^i \mathbf{N}_p$ $(3 \leq i \leq k)$ allow to control the coefficients of the polynomial expansion we seek. Indeed, whenever $M \in \mathcal{C}^k_{\tau_{min},\mathbf{L}}$, Lemma 2 shows that for every p in M, and y in $\mathcal{B}\big(p, \frac{\tau_{min} \wedge L_\perp^{-1}}{4}\big) \cap M$,

$$y - p = \pi^*(y-p) + \sum_{i=2}^{k-1} T_i^*(\pi^*(y-p)^{\otimes i}) + R_k(y-p), \tag{1}$$

where π^* denotes the orthogonal projection onto T_pM, the T_i^* are i-linear maps from T_pM to $\mathbb{R}^D$ with $\|T_i^*\|_{op} \leq L_i'$ and R_k satisfies $\|R_k(y-p)\| \leq C\|y-p\|^k$, where the constants C and the L_i''s depend on the parameters τ_{min}, d, k, $L_\perp, \ldots, L_k$.

Note that for $k \geq 3$ the exponential map can happen to be only $\mathcal{C}^{k-2}$ for a $\mathcal{C}^k$-submanifold [24]. Hence, it may not be a good choice of Ψ_p. However, for $k = 2$, taking $\Psi_p = \exp_p$ is sufficient for our purpose. For ease of notation, we may write $\mathcal{C}^2_{\tau_{min},\mathbf{L}}$ although the specification of $\mathbf{L}$ is useless. In this case, we implicitly set by default $\Psi_p = \exp_p$ and $L_\perp = 5/(4\tau_{min})$. As will be shown in Theorem 1, the global assumption $\tau_M \geq \tau_{min} > 0$ cannot be dropped, even when higher order regularity bounds L_i's are fixed.

Let us now describe the statistical model. Every d-dimensional submanifold $M \subset \mathbb{R}^D$ inherits a natural uniform volume measure by restriction of the ambient d-dimensional Hausdorff measure $\mathcal{H}^d$. In what follows, we will consider probability distributions that are almost uniform on some M in $\mathcal{C}^k_{\tau_{min},\mathbf{L}}$, with some bounded noise, as stated below.

Definition 2 (Noise-Free and Tubular Noise Models).
- (Noise-Free Model) For $k \geq 2$, $\tau_{min} > 0$, $\mathbf{L} = (L_\perp, L_3, \ldots, L_k)$ and $f_{min} \leq f_{max}$, we let $\mathcal{P}^k_{\tau_{min},\mathbf{L},f_{min},f_{max}}$ denote the set of distributions P_0 with support $M \in \mathcal{C}^k_{\tau_{min},\mathbf{L}}$ that have a density f with respect to the volume measure on M, and such that for all $y \in M$,

$$0 < f_{min} \leq f(y) \leq f_{max} < \infty.$$

- (Tubular Noise Model) For $0 \leq \sigma < \tau_{min}$, we denote by $\mathcal{P}^k_{\tau_{min},\mathbf{L},f_{min},f_{max}}(\sigma)$ the set of distributions of random variables $X = Y + Z$, where Y has distribution $P_0 \in \mathcal{P}^k_{\tau_{min},\mathbf{L},f_{min},f_{max}}$, and $Z \in T_Y M^\perp$ with $\|Z\| \leq \sigma$ and $\mathbb{E}(Z|Y) = 0$.

For short, we write $\mathcal{P}^k$ and $\mathcal{P}^k(\sigma)$ when there is no ambiguity. We denote by $\mathbb{X}_n$ an i.i.d. n-sample $\{X_1, \ldots, X_n\}$, that is, a sample with distribution $P^{\otimes n}$ for some $P \in \mathcal{P}^k(\sigma)$, so that $X_i = Y_i + Z_i$, where Y has distribution $P_0 \in \mathcal{P}^k$, $Z \in \mathcal{B}_{T_Y M^\perp}(0, \sigma)$ with $\mathbb{E}(Z|Y) = 0$. It is immediate that for

$\sigma < \tau_{min}$, we have $Y = \pi_M(X)$. Note that the tubular noise model $\mathcal{P}^k(\sigma)$ is a slight generalization of that in [21].

In what follows, though M is unknown, all the parameters of the model will be assumed to be known, including the intrinsic dimension d and the order of regularity k. We will also denote by $\mathcal{P}^k_{(x)}$ the subset of elements in $\mathcal{P}^k$ whose support contains a prescribed $x \in \mathbb{R}^D$.

In view of our minimax study on $\mathcal{P}^k$, it is important to ensure by now that $\mathcal{P}^k$ is stable with respect to deformations and dilations.

PROPOSITION 1. *Let $\Phi : \mathbb{R}^D \to \mathbb{R}^D$ be a global $\mathcal{C}^k$-diffeomorphism. If $\|d\Phi - I_D\|_{op}$, $\|d^2\Phi\|_{op}$, ..., $\|d^k\Phi\|_{op}$ are small enough, then for all P in $\mathcal{P}^k_{\tau_{min},\mathbf{L},f_{min},f_{max}}$, the pushforward distribution $P' = \Phi_* P$ belongs to $\mathcal{P}^k_{\tau_{min}/2,2\mathbf{L},f_{min}/2,2f_{max}}$.*

Moreover, if $\Phi = \lambda I_D$ ($\lambda > 0$) is an homogeneous dilation, then $P' \in \mathcal{P}^k_{\lambda\tau_{min},\mathbf{L}_{(\lambda)},f_{min}/\lambda^d,f_{max}/\lambda^d}$, where $\mathbf{L}_{(\lambda)} = (L_\perp/\lambda, L_3/\lambda^2, \ldots, L_k/\lambda^{k-1})$.

Proposition A.4 follows from a geometric reparametrization argument (Proposition A.5 in Appendix A) and a change of variable result for the Hausdorff measure (Lemma A.6 in Appendix A).

2.3. Necessity of a Global Assumption

In the previous Section 2.2, we generalized $\mathcal{C}^2$-like models — stated in terms of reach — to $\mathcal{C}^k$, for $k \geq 3$, by imposing higher order differentiability bounds on parametrizations Ψ_p's. The following Theorem 1 shows that the global assumption $\tau_M \geq \tau_{min} > 0$ is necessary for estimation purpose.

THEOREM 1. *Assume that $\tau_{min} = 0$. If $D \geq d + 3$, then for all $k \geq 3$ and $L_\perp > 0$, provided that $L_3/L_\perp^2, \ldots, L_k/L_\perp^{k-1}, L_\perp^d/f_{min}$ and $f_{max}/L_\perp^d$ are large enough (depending only on d and k), for all $n \geq 1$,*

$$\inf_{\hat{T}} \sup_{P \in \mathcal{P}^k_{(x)}} \mathbb{E}_{P^{\otimes n}} \angle\big(T_x M, \hat{T}\big) \geq \frac{1}{2} > 0,$$

where the infimum is taken over all the estimators $\hat{T} = \hat{T}\big(X_1, \ldots, X_n\big)$.

Moreover, for any $D \geq d+1$, provided that $L_3/L_\perp^2, \ldots, L_k/L_\perp^{k-1}, L_\perp^d/f_{min}$ and $f_{max}/L_\perp^d$ are large enough (depending only on d and k), for all $n \geq 1$,

$$\inf_{\widehat{II}} \sup_{P \in \mathcal{P}^k_{(x)}} \mathbb{E}_{P^{\otimes n}} \left\| II_x^M \circ \pi_{T_x M} - \widehat{II} \right\|_{op} \geq \frac{L_\perp}{4} > 0,$$

where the infimum is taken over all the estimators $\widehat{II} = \widehat{II}\big(X_1, \ldots, X_n\big)$.

The proof of Theorem 1 can be found in Section C.5. In other words, if the class of submanifolds is allowed to have arbitrarily small reach, no estimator can perform uniformly well to estimate neither T_xM nor II_x^M. And this, even though each of the underlying submanifolds have arbitrarily smooth parametrizations. Indeed, if two parts of M can nearly intersect around x at an arbitrarily small scale $\Lambda \to 0$, no estimator can decide whether the direction (resp. curvature) of M at x is that of the first part or the second part (see Figures 8 and 9).

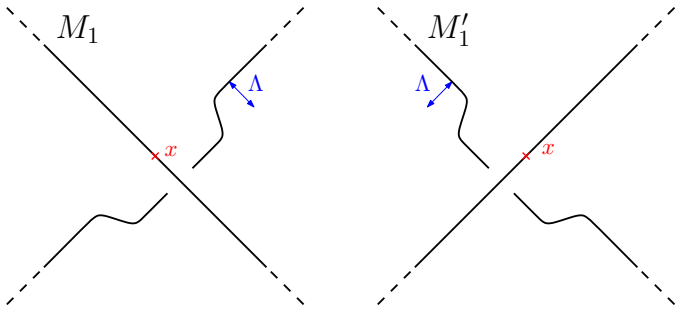


Figure 2: Inconsistency of tangent space estimation for $\tau_{min} = 0$.

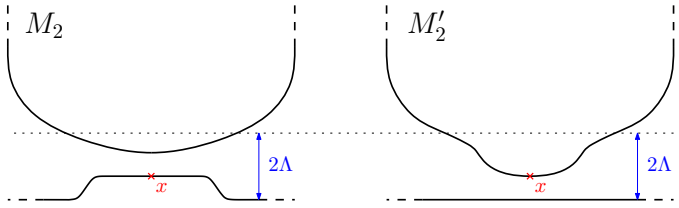


Figure 3: Inconsistency of curvature estimation for $\tau_{min} = 0$.

3. Main Results

Let us now move to the statement of the main results. Given an i.i.d. n-sample $\mathbb{X}_n = \{X_1, \dots, X_n\}$ with unknown common distribution $P \in \mathcal{P}^k(\sigma)$, we detail non-asymptotic rates for the estimation of tangent spaces $T_{Y_j}M$, second fundamental forms $II_{Y_j}^M$, and M itself.

For this, we need one more piece of notation. For $1 \leq j \leq n$, $P_{n-1}^{(j)}$ denotes integration with respect to $1/(n-1)\sum_{i\neq j}\delta_{(X_i - X_j)}$, and $z^{\otimes i}$ denotes the $D \times i$-dimensional vector $(z, \dots, z)$. For a constant $t > 0$ and a bandwidth $h > 0$ to be chosen later, we define the local polynomial estimator

$(\hat{\pi}_j, \hat{T}_{2,j}, \dots, \hat{T}_{k-1,j})$ at X_j to be any element of

$$\underset{\pi, \sup_{2\le i\le k}\|T_i\|_{op}\le t}{\arg\min} P_{n-1}^{(j)}\left[\left\|x-\pi(x)-\sum_{i=2}^{k-1} T_i(\pi(x)^{\otimes i})\right\|^2 \mathbb{1}_{\mathcal{B}(0,h)}(x)\right], \tag{2}$$

where π ranges among all the orthogonal projectors on d-dimensional subspaces, and $T_i : \left(\mathbb{R}^D\right)^i \to \mathbb{R}^D$ among the symmetric tensors of order i such that $\|T_i\|_{op} \le t$. For $k = 2$, the sum over the tensors T_i is empty, and the integrated term reduces to $\|x - \pi(x)\|^2 \mathbb{1}_{\mathcal{B}(0,h)}(x)$. By compactness of the domain of minimization, such a minimizer exists almost surely. In what follows, we will work with a maximum scale $h \le h_0$, with

$$h_0 = \frac{\tau_{min} \wedge L_\perp^{-1}}{8}.$$

The set of d-dimensional orthogonal projectors is not convex, which leads to a more involved optimization problem than usual least squares. In practice, this problem may be solved using tools from optimization on Grassman manifolds [34], or adopting a two-stage procedure such as in [9]: from local PCA, a first d-dimensional space is estimated at each sample point, along with an orthonormal basis of it. Then, the optimization problem (2) is expressed as a minimization problem in terms of the coefficients of $(\pi_j, T_{2,j}, \dots, T_{k,j})$ in this basis under orthogonality constraints. It is worth mentioning that a similar problem is explicitly solved in [11], leading to an optimal tangent space estimation procedure in the case $k = 3$.

The constraint $\|T_i\|_{op} \le t$ involves a parameter t to be calibrated. As will be shown in the following section, it is enough to choose t roughly smaller than $1/h$, but still larger than the unknown norm of the optimal tensors $\|T_i^*\|_{op}$. Hence, for $h \to 0$, the choice $t = h^{-1}$ works to guarantee optimal convergence rates. Such a constraint on the higher order tensors might have been stated under the form of a $\|.\|_{op}$-penalized least squares minimization — as in ridge regression — leading to the same results.

3.1. Tangent Spaces

By definition, the tangent space $T_{Y_j}M$ is the best linear approximation of M nearby Y_j. Thus, it is natural to take the range of the first order term minimizing (2) and write $\hat{T}_j = \operatorname{im}\hat{\pi}_j$. The $\hat{T}_j$'s approximate simultaneously the $T_{Y_j}M$'s with high probability, as stated below.

THEOREM 2. *Assume that* $t \geq C_{k,d,\tau_{min},\mathbf{L}} \geq \sup_{2\leq i\leq k} \|T_i^*\|_{op}$. *Set* $h = \left(C_{d,k}\frac{f_{max}^2 \log n}{f_{min}^3 (n-1)}\right)^{1/d}$, *for* $C_{d,k}$ *large enough, and assume that* $\sigma \leq h/4$. *If* n *is large enough so that* $h \leq h_0$, *then with probability at least* $1 - \left(\frac{1}{n}\right)^{k/d}$,

$$\max_{1\leq j\leq n} \angle\big(T_{Y_j}M, \hat{T}_j\big) \leq C_{d,k,\tau_{min},\mathbf{L}} \sqrt{\frac{f_{max}}{f_{min}}} (h^{k-1} \vee \sigma h^{-1})(1+th).$$

As a consequence, taking $t = h^{-1}$, *for* n *large enough,*

$$\sup_{P\in\mathcal{P}^k(\sigma)} \mathbb{E}_{P^{\otimes n}} \max_{1\leq j\leq n} \angle\big(T_{Y_j}M, \hat{T}_j\big) \leq C \left(\frac{\log n}{n-1}\right)^{\frac{k-1}{d}} \left\{1 \vee \sigma \left(\frac{\log n}{n-1}\right)^{-\frac{k}{d}}\right\},$$

where $C = C_{d,k,\tau_{min},\mathbf{L},f_{min},f_{max}}$.

The proof of Theorem 2 is given in Section 5.1.2. The same bound holds for the estimation of T_yM at a prescribed $y \in M$ in the model $\mathcal{P}^k_{(y)}(\sigma)$. For that, simply take $P_n^{(y)} = 1/n \sum_i \delta_{(X_i - y)}$ as integration in (2).

In the noise-free setting, or when $\sigma \leq h^k$, this result is in line with those of [9] in terms of the sample size dependency $(1/n)^{(k-1)/d}$. Besides, it shows that the convergence rate of our estimator does not depend on the ambient dimension D, even in codimension greater than 2. When $k = 2$, we recover the same rate as [1], where we used local PCA, which is a reformulation of (2). When $k \geq 3$, the procedure (2) outperforms PCA-based estimators of [31] and [33], where convergence rates of the form $(1/n)^\beta$ with $1/d < \beta < 2/d$ are obtained. This bound also recovers the result of [11] in the case $k = 3$, where a similar procedure is used. When the noise level σ is of order h^α, with $1 \leq \alpha \leq k$, Theorem 2 yields a convergence rate in $h^{\alpha-1}$. Since a polynomial decomposition up to order $k_\alpha = \lceil \alpha \rceil$ in (2) results in the same bound, the noise level $\sigma = h^\alpha$ may be thought of as an α-regularity threshold. At last, it may be worth mentioning that the results of Theorem 2 also hold when the assumption $\mathbb{E}(Z|Y) = 0$ is relaxed. Theorem 2 nearly matches the following lower bound.

THEOREM 3. *If* $\tau_{min}L_\perp, \ldots, \tau_{min}^{k-1}L_k, (\tau_{min}^d f_{min})^{-1}$ *and* $\tau_{min}^d f_{max}$ *are large enough (depending only on* d *and* k*), then*

$$\inf_{\hat{T}} \sup_{P\in\mathcal{P}^k(\sigma)} \mathbb{E}_{P^{\otimes n}} \angle\big(T_{\pi_M(X_1)}M, \hat{T}\big) \geq c_{d,k,\tau_{min}} \left\{\left(\frac{1}{n-1}\right)^{\frac{k-1}{d}} \vee \left(\frac{\sigma}{n-1}\right)^{\frac{k-1}{d+k}}\right\},$$

where the infimum is taken over all the estimators $\hat{T} = \hat{T}(X_1, \dots, X_n)$.

A proof of Theorem 3 can be found in Section 5.2.2. When $\sigma \lesssim (1/n)^{k/d}$, the lower bound matches Theorem 2 in the noise-free case, up to a $\log n$ factor. Thus, the rate $(1/n)^{(k-1)/d}$ is optimal for tangent space estimation on the model $\mathcal{P}^k$. The rate $(\log n/n)^{1/d}$ obtained in [1] for $k = 2$ is therefore optimal, as well as the rate $(\log n/n)^{2/d}$ given in [11] for $k = 3$. The rate $(1/n)^{(k-1)/d}$ naturally appears on the the model $\mathcal{P}^k$, as the estimation rate of differential objects of order 1 from k-smooth submanifolds.

When $\sigma \asymp (1/n)^{\alpha/d}$ with $\alpha < k$, the lower bound provided by Theorem 3 is of order $(1/n)^{(k-1)(\alpha+d)/[d(d+k)]}$, hence smaller than the $(1/n)^{\alpha/d}$ rate of Theorem 2. This suggests that the local polynomial estimator (2) is suboptimal whenever $\sigma \gg (1/n)^{k/d}$ on the model $\mathcal{P}^k(\sigma)$.

Here again, the same lower bound holds for the estimation of T_yM at a fixed point y in the model $\mathcal{P}^k_{(y)}(\sigma)$.

3.2. Curvature

The second fundamental form $II^M_{Y_j} : T_{Y_j}M \times T_{Y_j}M \to T_{Y_j}M^\perp \subset \mathbb{R}^D$ is a symmetric bilinear map that encodes completely the curvature of M at Y_j [13, Chap. 6, Proposition 3.1]. Estimating it only from a point cloud $\mathbb{X}_n$ does not trivially make sense, since $II^M_{Y_j}$ has domain $T_{Y_j}M$ which is unknown. To bypass this issue we extend $II^M_{Y_j}$ to $\mathbb{R}^D$. That is, we consider the estimation of $II^M_{Y_j} \circ \pi_{T_{Y_j}M}$ which has full domain $\mathbb{R}^D$. Following the same ideas as in the previous Section 3.1, we use the second order tensor $\hat{T}_{2,j} \circ \hat{\pi}_j$ obtained in (2) to estimate $II^M_{Y_j} \circ \pi_{T_{Y_j}M}$.

THEOREM 4. *Let* $k \geq 3$. *Take* h *as in Theorem 2,* $\sigma \leq h/4$, *and* $t = 1/h$. *If* n *is large enough so that* $h \leq h_0$ *and* $h^{-1} \geq C^{-1}_{k,d,\tau_{min},\mathbf{L}} \geq (\sup_{2\leq i\leq k} \|T_i^*\|_{op})^{-1}$, *then with probability at least* $1 - \left(\frac{1}{n}\right)^{k/d}$,

$$\max_{1\leq j\leq n} \left\| II^M_{Y_j} \circ \pi_{T_{Y_j}M} - \hat{T}_{2,j} \circ \hat{\pi}_j \right\|_{op} \leq C_{d,k,\tau_{min},\mathbf{L}} \sqrt{\frac{f_{max}}{f_{min}}} (h^{k-2} \vee \sigma h^{-2}).$$

In particular, for n *large enough,*

$$\sup_{P\in\mathcal{P}^k(\sigma)} \mathbb{E}_{P^{\otimes n}} \max_{1\leq j\leq n} \left\| II^M_{Y_j} \circ \pi_{T_{Y_j}M} - \hat{T}_{2,j} \circ \hat{\pi}_j \right\|_{op} \leq C_{d,k,\tau_{min},\mathbf{L},f_{min},f_{max}} \left(\frac{\log n}{n-1}\right)^{\frac{k-2}{d}} \left\{ 1 \vee \sigma \left(\frac{\log n}{n-1}\right)^{-\frac{k}{d}} \right\}.$$

The proof of Theorem 4 is given in Section 5.1.3. As in Theorem 2, the case $\sigma \leq h^k$ may be thought of as a noise-free setting, and provides an upper bound of the form h^{k-2}. Interestingly, Theorems 2 and 4 are enough to provide estimators of various notions of curvature. For instance, consider the scalar curvature [13, Section 4.4] at a point Y_j, defined by

$$Sc^M_{Y_j} = \frac{1}{d(d-1)} \sum_{r \neq s} \left[\left\langle II^M_{Y_j}(e_r, e_r), II^M_{Y_j}(e_s, e_s) \right\rangle - \|II^M_{Y_j}(e_r, e_s)\|^2 \right],$$

where $(e_r)_{1 \leq r \leq d}$ is an orthonormal basis of $T_{Y_j} M$. A plugin estimator of $Sc^M_{Y_j}$ is

$$\widehat{Sc}_j = \frac{1}{d(d-1)} \sum_{r \neq s} \left[\left\langle \hat{T}_{2,j}(\hat{e}_r, \hat{e}_r), \hat{T}_{2,j}(\hat{e}_s, \hat{e}_s) \right\rangle - \|\hat{T}_{2,j}(\hat{e}_r, \hat{e}_s)\|^2 \right],$$

where $(\hat{e}_r)_{1 \leq r \leq d}$ is an orthonormal basis of $\hat{T}_{2,j}$. Theorems 2 and 4 yield

$$\mathbb{E}_{P^{\otimes n}} \max_{1 \leq j \leq n} \left| \widehat{Sc}_j - Sc^M_{Y_j} \right| \leq C \left(\frac{\log n}{n-1} \right)^{\frac{k-2}{d}} \left\{ 1 \vee \sigma \left(\frac{\log n}{n-1} \right)^{-\frac{k}{d}} \right\},$$

where $C = C_{d,k,\tau_{min},\mathbf{L},f_{min},f_{max}}$.

The (near-)optimality of the bound stated in Theorem 4 is assessed by the following lower bound.

THEOREM 5. *If $\tau_{min} L_\perp, \ldots, \tau_{min}^{k-1} L_k, (\tau_{min}^d f_{min})^{-1}$ and $\tau_{min}^d f_{max}$ are large enough (depending only on d and k), then*

$$\inf_{\widehat{II}} \sup_{P \in \mathcal{P}^k(\sigma)} \mathbb{E}_{P^{\otimes n}} \left\| II^M_{\pi_M(X_1)} \circ \pi_{T_{\pi_M(X_1)}M} - \widehat{II} \right\|_{op} \geq c_{d,k,\tau_{min}} \left\{ \left(\frac{1}{n-1} \right)^{\frac{k-2}{d}} \vee \left(\frac{\sigma}{n-1} \right)^{\frac{k-2}{d+k}} \right\},$$

where the infimum is taken over all the estimators $\widehat{II} = \widehat{II}(X_1, \ldots, X_n)$.

The proof of Theorem 5 is given in Section 5.2.2. The same remarks as in Section 3.1 hold. If the estimation problem consists in approximating II^M_y at a fixed point y known to belong to M beforehand, we obtain the same rates. The ambient dimension D still plays no role. The shift $k-2$ in the rate of convergence on a $\mathcal{C}^k$-model can be interpreted as the order of derivation of the object of interest, that is 2 for curvature.

Notice that the lower bound (Theorem 5) does not require $k \geq 3$. Hence, we get that for $k = 2$, curvature cannot be estimated uniformly consistently on the $\mathcal{C}^2$-model $\mathcal{P}^2$. This seems natural, since the estimation of a second order quantity should require an additional degree of smoothness.

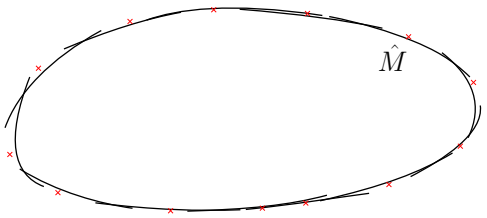


Figure 4: $\hat{M}$ is a union of polynomial patches at sample points.

3.3. Support Estimation

For each $1 \leq j \leq n$, the minimization (2) outputs a series of tensors $(\hat{\pi}_j, \hat{T}_{2,j}, \dots, \hat{T}_{k-1,j})$. This collection of multidimensional monomials can be further exploited as follows. By construction, they fit M at scale h around Y_j, so that

$$\widehat{\Psi}_j(v) = X_j + v + \sum_{i=2}^{k-1} \hat{T}_{i,j}\left(v^{\otimes i}\right)$$

is a good candidate for an approximate parametrization in a neighborhood of Y_j. We do not know the domain $T_{Y_j}M$ of the initial parametrization, though we have at hand an approximation $\hat{T}_j = \operatorname{im}\hat{\pi}_j$ which was proved to be consistent in Section 3.1. As a consequence, we let the support estimator based on local polynomials $\hat{M}$ be

$$\hat{M} = \bigcup_{j=1}^{n} \widehat{\Psi}_j\left(\mathcal{B}_{\hat{T}_j}(0, 7h/8)\right).$$

The set $\hat{M}$ has no reason to be globally smooth, since it consists of a mere union of polynomial patches (Figure 4). However, $\hat{M}$ is provably close to M for the Hausdorff distance.

THEOREM 6. *With the same assumptions as Theorem 4, with probability at least* $1 - 2\left(\frac{1}{n}\right)^{\frac{k}{d}}$, *we have*

$$d_H\big(M, \hat{M}\big) \leq C_{d,k,\tau_{min},\mathbf{L},f_{min},f_{max}}(h^k \vee \sigma).$$

In particular, for n *large enough,*

$$\sup_{P \in \mathcal{P}^k(\sigma)} \mathbb{E}_{P^{\otimes n}} d_H\big(M, \hat{M}\big) \leq C_{d,k,\tau_{min},\mathbf{L},f_{min},f_{max}} \left\{ \left(\frac{\log n}{n-1}\right)^{\frac{k}{d}} \vee \sigma \right\}.$$

A proof of Theorem 6 is given in Section 5.1.4. As in Theorem 2, for a noise level of order h^α, $\alpha \geq 1$, Theorem 6 yields a convergence rate of order $h^{(k\wedge\alpha)/d}$. Thus the noise level σ may also be thought of as a regularity threshold. Contrary to [21, Theorem 2], the case $h/4 < \sigma < \tau_{min}$ is not in the scope of Theorem 6. Moreover, for $1 \leq \alpha < 2d/(d+2)$, [21, Theorem 2] provides a better convergence rate of $h^{2/(d+2)}$. Note however that Theorem 6 is also valid whenever the assumption $\mathbb{E}(Z|Y) = 0$ is relaxed. In this non-centered noise framework, Theorem 6 outperforms [26, Theorem 7] in the case $d \geq 3$, $k = 2$, and $\sigma \leq h^2$.

In the noise-free case or when $\sigma \leq h^k$, for $k = 2$, we recover the rate $(\log n/n)^{2/d}$ obtained in [1, 20, 25] and improve the rate $(\log n/n)^{2/(d+2)}$ in [21, 26]. However, our estimator $\hat{M}$ is an unstructured union of d-dimensional balls in $\mathbb{R}^D$. Consequently, $\hat{M}$ does not recover the topology of M as the estimator of [1] does.

When $k \geq 3$, $\hat{M}$ outperforms reconstruction procedures based on a somewhat piecewise linear interpolation [1, 20, 26], and achieves the faster rate $(\log n/n)^{k/d}$ for the Hausdorff loss. This seems quite natural, since our procedure fits higher order terms. This is done at the price of a probably worse dependency on the dimension d than in [1, 20]. Theorem 6 is now proved to be (almost) minimax optimal.

THEOREM 7. *If $\tau_{min}L_\perp, \ldots, \tau_{min}^{k-1}L_k, (\tau_{min}^d f_{min})^{-1}$ and $\tau_{min}^d f_{max}$ are large enough (depending only on d and k), then for n large enough,*

$$\inf_{\hat{M}} \sup_{P\in\mathcal{P}^k(\sigma)} \mathbb{E}_{P^{\otimes n}} d_H\big(M, \hat{M}\big) \geq c_{d,k,\tau_{min}} \left\{ \left(\frac{1}{n}\right)^{\frac{k}{d}} \vee \left(\frac{\sigma}{n}\right)^{\frac{k}{d+k}} \right\},$$

where the infimum is taken over all the estimators $\hat{M} = \hat{M}(X_1, \ldots, X_n)$.

Theorem 7, whose proof is given in Section 5.2.1, is obtained from Le Cam's Lemma (Theorem C.20). Let us note that it is likely for the extra $\log n$ term appearing in Theorem 6 to actually be present in the minimax rate. Roughly, it is due to the fact that the Hausdorff distance d_H is similar to a L^∞ loss. The $\log n$ term may be obtained in Theorem 7 with the same combinatorial analysis as in [25] for $k = 2$.

As for the estimation of tangent spaces and curvature, Theorem 7 matches the upper bound in Theorem 6 in the noise-free case $\sigma \lesssim (1/n)^{k/d}$. Moreover, for $\sigma < \tau_{min}$, it also generalizes Theorem 1 in [21] to higher orders of regularity ($k \geq 3$). Again, for $\sigma \gg (1/n)^{-k/d}$, the upper bound in Theorem 6 is larger than the lower bound stated in Theorem 7. However our estimator $\hat{M}$ achieves the same convergence rate if the assumption $\mathbb{E}(Z|Y)$ is dropped.

4. Conclusion, Prospects

In this article, we derived non-asymptotic bounds for inference of geometric objects associated with smooth submanifolds $M \subset \mathbb{R}^D$. We focused on tangent spaces, second fundamental forms, and the submanifold itself. We introduced new regularity classes $\mathcal{C}^k_{\tau_{min},\mathbf{L}}$ for submanifolds that extend the case $k = 2$. For each object of interest, the proposed estimator relies on local polynomials that can be computed through a least square minimization. Minimax lower bounds were presented, matching the upper bounds up to $\log n$ factors in the regime of small noise.

The implementation of (2) needs to be investigated. The non-convexity of the criterion comes from that we minimize over the space of orthogonal projectors, which is non-convex. However, that space is pretty well understood, and it seems possible to implement gradient descents on it [34]. Another way to improve our procedure could be to fit orthogonal polynomials instead of monomials. Such a modification may also lead to improved dependency on the dimension d and the regularity k in the bounds for both tangent space and support estimation.

Though the stated lower bounds are valid for quite general tubular noise levels σ, it seems that our estimators based on local polynomials are suboptimal whenever σ is larger than the expected precision for $\mathcal{C}^k$ models in a d-dimensional space (roughly $(1/n)^{k/d}$). In such a setting, it is likely that a preliminary centering procedure is needed, as the one exposed in [21]. Other pre-processings of the data might adapt our estimators to other types of noise. For instance, whenevever outliers are allowed in the model $\mathcal{C}^2$, [1] proposes an iterative denoising procedure based on tangent space estimation. It exploits the fact that tangent space estimation allows to remove a part of outliers, and removing outliers enhances tangent space estimation. An interesting question would be to study how this method can apply with local polynomials.

Another open question is that of exact topology recovery with fast rates for $k \geq 3$. Indeed, $\hat{M}$ converges at rate $(\log n/n)^{k/d}$ but is unstructured. It would be nice to glue the patches of $\hat{M}$ together, for example using interpolation techniques, following the ideas of [18].

5. Proofs

5.1. Upper bounds

5.1.1. Preliminary results on polynomial expansions

To prove Theorem 2, 4 and 6, the following lemmas are needed. First, we relate the existence of parametrizations Ψ_p's mentioned in Definition 1 to a local polynomial decomposition.

LEMMA 2. *For any $M \in \mathcal{C}^k_{\tau_{min},\mathbf{L}}$ and $y \in M$, the following holds.*

(i) For all $v_1, v_2 \in \mathcal{B}_{T_yM}\left(0, \frac{1}{4L_\perp}\right)$,

$$\frac{3}{4}\|v_2 - v_1\| \leq \|\Psi_y(v_2) - \Psi_y(v_1)\| \leq \frac{5}{4}\|v_2 - v_1\|.$$

(ii) For all $h \leq \frac{1}{4L_\perp} \wedge \frac{2\tau_{min}}{5}$,

$$M \cap \mathcal{B}\left(y, \frac{3h}{5}\right) \subset \Psi_y\left(\mathcal{B}_{T_yM}(y,h)\right) \subset M \cap \mathcal{B}\left(y, \frac{5h}{4}\right).$$

(iii) For all $h \leq \frac{\tau_{min}}{2}$,

$$\mathcal{B}_{T_yM}\left(0, \frac{7h}{8}\right) \subset \pi_{T_yM}\left(\mathcal{B}(y,h) \cap M\right).$$

(iv) Denoting by $\pi^* = \pi_{T_yM}$ the orthogonal projection onto T_yM, for all $y \in M$, there exist multilinear maps $T_2^*, \ldots, T_{k-1}^*$ from T_yM to $\mathbb{R}^D$, and R_k such that for all $y' \in \mathcal{B}\left(y, \frac{\tau_{min} \wedge L_\perp^{-1}}{4}\right) \cap M$,

$$\begin{aligned} y' - y = \pi^*(y'-y) + T_2^*(\pi^*(y'-y)^{\otimes 2}) + \ldots + T_{k-1}^*(\pi^*(y'-y)^{\otimes k-1}) \\ + R_k(y'-y), \end{aligned}$$

with

$$\left\|R_k(y'-y)\right\| \leq C\left\|y'-y\right\|^k \quad \text{and} \quad \|T_i^*\|_{op} \leq L_i', \text{ for } 2 \leq i \leq k-1,$$

where L_i' depends on $d, k, \tau_{min}, L_\perp, \ldots, L_i$, and C on d, k, τ_{min}, $L_\perp$,..., L_k. Moreover, for $k \geq 3$, $T_2^* = II_y^M$.

(v) For all $y \in M$, $\left\|II_y^M\right\|_{op} \leq 1/\tau_{min}$. In particular, the sectional curvatures of M satisfy

$$\frac{-2}{\tau_{min}^2} \leq \kappa \leq \frac{1}{\tau_{min}^2}.$$

The proof of Lemma 2 can be found in Section A.2. A direct consequence of Lemma 2 is the following Lemma 3.

LEMMA 3. *Set $h_0 = (\tau_{min} \wedge L_\perp^{-1})/8$ and $h \leq h_0$. Let $M \in \mathcal{C}^k_{\tau_{min},\mathbf{L}}$, $x_0 = y_0 + z_0$, with $y_0 \in M$ and $\|z_0\| \leq \sigma \leq h/4$. Denote by π^* the orthogonal projection onto $T_{y_0}M$, and by $T_2^*, \ldots, T_{k-1}^*$ the multilinear maps given by Lemma 2, iv).*

Then, for any $x = y+z$ such that $y \in M$, $\|z\| \leq \sigma \leq h/4$ and $x \in \mathcal{B}(x_0, h)$, for any orthogonal projection π and multilinear maps $T_2, \ldots, T_{k-1}$, we have

$$x - x_0 - \pi(x - x_0) - \sum_{j=2}^{k-1} T_j(\pi(x - x_0)^{\otimes j}) = \sum_{j=1}^{k} T'_j(\pi^*(y - y_0)^{\otimes j}) + R_k(x - x_0),$$

where T'_j are j-linear maps, and $\|R_k(x - x_0)\| \leq C\left(\sigma \vee h^k\right)(1 + th)$, with $t = \max_{j=2\ldots,k} \|T\|_{op}$ and C depending on d, k, τ_{min}, $L_\perp$,..., L_k. Moreover, we have

$$\begin{aligned} T'_1 &= (\pi^* - \pi), \\ T'_2 &= (\pi^* - \pi) \circ T_2^* + (T_2^* \circ \pi^* - T_2 \circ \pi), \end{aligned}$$

and, if $\pi = \pi^*$ and $T_i = T_i^*$, for $i = 2, \ldots, k-1$, then $T'_j = 0$, for $j = 1, \ldots, k$.

Lemma 3 roughly states that, if π, T_j, $j \geq 2$ are designed to locally approximate $x = y + z$ around $x_0 = y_0 + z_0$, then the approximation error may be expressed as a polynomial expansion in $\pi^*(y - y_0)$.

PROOF OF LEMMA 3. For short assume that $y_0 = 0$. In what follows C will denote a constant depending on d, k, τ_{min}, $L_\perp$,..., L_k. We may write

$$x - x_0 - \pi(x - x_0) - \sum_{j=2}^{k-1} T_j(\pi(x - x_0)^{\otimes j}) = y - \pi(y) - \sum_{j=2}^{k-1} T_j(\pi(y)^{\otimes j}) + R'_k(x - x_0),$$

with $\|R'_k(x - x_0)\| \leq C\sigma(1 + th)$. Since $\sigma \leq h/4$, $y \in \mathcal{B}(0, 3h/2)$, with $h \leq h_0$. Hence Lemma 2 entails

$$y = \pi^*(y) + T_2^*(\pi^*(y)^{\otimes 2}) + \ldots + T_{k-1}^*(\pi^*(y)^{\otimes k-1}) + R''_k(y),$$

with $\|R_k''(y)\| \leq Ch^k$. We deduce that

$$y-\pi(y)-\sum_{j=2}^{k-1} T_j(\pi(y)^{\otimes j}) = (\pi^*-\pi\circ\pi^*)(y)+T_2^*(\pi^*(y)^{\otimes 2})-\pi(T_2^*(\pi^*(y)^{\otimes 2}))$$
$$- T_2(\pi\circ\pi^*(y)^{\otimes 2}) + \sum_{j=3}^{k} T_k'(\pi^*(y)^{\otimes j}) - \pi(R_k''(y)) - R_k'''(y),$$

with $\|R_k'''(y)\| \leq Cth^{k+1}$, since only tensors of order greater than 2 are involved in R_k'''. Since $T_2^* = II_0^M$, $\pi^* \circ T_2^* = 0$, hence the result. □

At last, we need a result relating deviation in terms of polynomial norm and $L^2(P_{0,n-1}^{(j)})$ norm, where $P_0 \in \mathcal{P}^k$, for polynomials taking arguments in $\pi^{*,(j)}(y)$. For clarity's sake, the bounds are given for $j = 1$, and we denote $P_{0,n-1}^{(1)}$ by $P_{0,n-1}$. Without loss of generality, we can assume that $Y_1 = 0$.

Let $\mathbb{R}^k[y_{1:d}]$ denote the set of real-valued polynomial functions in d variables with degree less than k. For $S \in \mathbb{R}^k[y_{1:d}]$, we denote by $\|S\|_2$ the Euclidean norm of its coefficients, and by S_h the polynomial defined by $S_h(y_{1:d}) = S(hy_{1:d})$. With a slight abuse of notation, $S(\pi^*(y))$ will denote $S(e_1^*(\pi^*(y)), \ldots, e_d^*(\pi^*(y)))$, where $e_1^*, \ldots, e_d^*$ form an orthonormal coordinate system of T_0M.

PROPOSITION 2. *Set* $h = \left(K\frac{\log n}{n-1}\right)^{\frac{1}{d}}$. *There exist constants* $\kappa_{k,d}$, $c_{k,d}$ *and* C_d *such that, if* $K \geq (\kappa_{k,d} f_{max}^2/f_{min}^3)$ *and* n *is large enough so that* $h \leq h_0 \leq \tau_{min}/8$, *then with probability at least* $1 - \left(\frac{1}{n}\right)^{\frac{k}{d}+1}$, *we have*

$$\begin{aligned} P_{0,n-1}[S^2(\pi^*(y))\mathbb{1}_{\mathcal{B}(h/2)}(y)] &\geq c_{k,d} h^d f_{min} \|S_h\|_2^2, \\ N(3h/2) &\leq C_d f_{max}(n-1)h^d, \end{aligned}$$

for every $S \in \mathbb{R}^k[y_{1:d}]$, *where* $N(3h/2) = \sum_{j=2}^n \mathbb{1}_{\mathcal{B}(0,3h/2)}(Y_j)$.

The proof of Proposition B.8 is deferred to Section B.2.

5.1.2. Upper Bound for Tangent Space Estimation

PROOF OF THEOREM 2. We recall that for every $j = 1, \ldots, n$, $X_j = Y_j + Z_j$, where $Y_j \in M$ is drawn from P_0 and $\|Z_j\| \leq \sigma \leq h/4$, where $h \leq h_0$ as defined in Lemma 3. Without loss of generality we consider the case $j = 1$, $Y_1 = 0$. From now on we assume that the probability event defined in Proposition B.8 occurs, and denote by $\mathcal{R}_{n-1}(\pi, T_2, \ldots, T_{k-1})$ the empirical criterion

defined by (2). Note that $X_j \in \mathcal{B}(X_1, h)$ entails $Y_j \in \mathcal{B}(0, 3h/2)$. Moreover, since for $t \geq \max_{i=2,\dots,k-1} \|T_i^*\|_{op}$, $\mathcal{R}_{n-1}(\hat{\pi}, \hat{T}_1, \dots, \hat{T}_{k-1}) \leq \mathcal{R}_{n-1}(\pi^*, T_2^*, \dots, T_{k-1}^*)$, we deduce that

$$\mathcal{R}_{n-1}(\hat{\pi}, \hat{T}_1, \dots, \hat{T}_{k-1}) \leq \frac{C_{\tau_{min},\mathbf{L}} \left(\sigma^2 \vee h^{2k}\right) (1+th)^2 N(3h/2)}{n-1},$$

according to Lemma 3. On the other hand, note that if $Y_j \in \mathcal{B}(0, h/2)$, then $X_j \in \mathcal{B}(X_1, h)$. Lemma 3 then yields

$$\begin{aligned}\mathcal{R}_{n-1}(\hat{\pi}, \hat{T}_2, \dots, \hat{T}_{k-1}) \geq & P_{0,n-1} \left(\left\| \sum_{j=1}^{k} \hat{T}_j'(\pi^*(y)^{\otimes j}) \right\|^2 \mathbb{1}_{\mathcal{B}(0,h/2)}(y) \right) \\ & - \frac{C_{\tau_{min},\mathbf{L}} \left(\sigma^2 \vee h^{2k}\right) (1+th)^2 N(3h/2)}{n-1}.\end{aligned}$$

Using Proposition B.8, we can decompose the right-hand side as

$$\begin{aligned}\sum_{r=1}^{D} P_{0,n-1} & \left(\sum_{j=1}^{k} \hat{T'}_j^{(r)}(\pi^*(y)^{\otimes j}) \mathbb{1}_{\mathcal{B}(0,h/2)}(y) \right)^2 \\ & \leq C_{\tau_{min},\mathbf{L}} f_{max} h^d \left(\sigma^2 \vee h^{2k}\right) (1+th)^2,\end{aligned}$$

where for any tensor T, $T^{(r)}$ denotes the r-th coordinate of T and is considered as a real valued r-order polynomial. Then, applying Proposition B.8 to each coordinate leads to

$$c_{d,k} f_{min} \sum_{r=1}^{D} \sum_{j=1}^{k} \left\| \left(T'^{(r)}_j (\pi^*(y)^{\otimes j}) \right)_h \right\|_2^2 \leq C_{\tau_{min},\mathbf{L}} f_{max} h^d \left(\sigma^2 \vee h^{2k}\right) (1+th)^2.$$

It follows that, for $1 \leq j \leq k$,

$$\|\hat{T}_j' \circ \pi^*\|_{op}^2 \leq C_{d,k,\mathbf{L},\tau_{min}} \frac{f_{max}}{f_{min}} (h^{2(k-j)} \vee \sigma^2 h^{-2j})(1+t^2h^2). \tag{3}$$

Noting that, according to [22, Section 2.6.2],

$$\|\hat{T}_1' \circ \pi^*\|_{op} = \|(\pi^* - \hat{\pi})\pi^*\|_{op} = \|\pi_{\hat{T}_1^\perp} \circ \pi^*\| = \angle(T_{Y_1}M, \hat{T}_1),$$

we deduce that

$$\angle\big(T_{Y_1}M, \hat{T}_1\big) \leq C_{d,k,\mathbf{L},\tau_{min}} \sqrt{\frac{f_{max}}{f_{min}}} (h^{(k-1)} \vee \sigma h^{-1})(1+th).$$

Theorem 2 then follows from a straightforward union bound. □

5.1.3. Upper Bound for Curvature Estimation

PROOF OF THEOREM 4. Without loss of generality, the derivation is conducted in the same framework as in the previous Section 5.1.2. In accordance with assumptions of Theorem 4, we assume that $\max_{2\leq i\leq k}\|T_i^*\|_{op}\leq t\leq 1/h$. Since, according to Lemma 3,

$$T_2'(\pi^*(y)^{\otimes 2})=(\pi^*-\hat{\pi})(T_2^*(\pi^*(y)^{\otimes 2}))+(T_2^*\circ\pi^*-\hat{T}_2\circ\hat{\pi})(\pi^*(y)^{\otimes 2}),$$

we deduce that

$$\|T_2^*\circ\pi^*-\hat{T}_2\circ\hat{\pi}\|_{op}\leq\|T_2'\circ\pi^*\|_{op}+\|\hat{\pi}-\pi^*\|_{op}+\|\hat{T}_2\circ\hat{\pi}\circ\pi^*-\hat{T}_2\circ\hat{\pi}\circ\hat{\pi}\|_{op}.$$

Using (3) with $j=1,2$ and $th\leq 1$ leads to

$$\|T_2^*\circ\pi^*-\hat{T}_2\circ\hat{\pi}\|_{op}\leq C_{d,k,\mathbf{L},\tau_{min}}\sqrt{\frac{f_{max}}{f_{min}}}(h^{(k-2)}\vee\sigma h^{-2}).$$

Finally, Lemma 2 states that $II_{Y_1}^M=T_2^*$. Theorem 4 follows from a union bound. □

5.1.4. Upper Bound for Manifold Estimation

PROOF OF THEOREM 6 . Recall that we take $X_i=Y_i+Z_i$, where Y_i has distribution P_0 and $\|Z_j\|\leq\sigma\leq h/4$. We also assume that the probability events of Proposition B.8 occur simultaneously at each Y_i, so that (3) holds for all i, with probability larger than $1-(1/n)^{k/d}$. Without loss of generality set $Y_1=0$. Let $v\in\mathcal{B}_{\hat{T}_1M}(0,7h/8)$ be fixed. Notice that $\pi^*(v)\in\mathcal{B}_{T_0M}(0,7h/8)$. Hence, according to Lemma 2, there exists $y\in\mathcal{B}(0,h)\cap M$ such that $\pi^*(v)=\pi^*(y)$. According to (3), we may write

$$\widehat{\Psi}(v)=Z_1+v+\sum_{j=2}^{k-1}\hat{T}_j(v^{\otimes j})=\pi^*(v)+\sum_{j=2}^{k-1}\hat{T}_j(\pi^*(v)^{\otimes j})+R_k(v),$$

where, since $\|\hat{T}_j\|_{op}\leq 1/h$, $\|R_k(v)\|\leq C_{k,d,\tau_{min},\mathbf{L}}\sqrt{f_{max}/f_{min}}(h^k\vee\sigma)$. Using (3) again leads to

$$\begin{aligned}\pi^*(v)+\sum_{j=2}^{k-1}\hat{T}_j(\pi^*(v)^{\otimes j})&=\pi^*(v)+\sum_{j=2}^{k-1}T_j^*(\pi^*(v)^{\otimes j})+R'(\pi^*(v))\\&=\pi^*(y)+\sum_{j=2}^{k-1}T_j^*(\pi^*(y)^{\otimes j})+R'(\pi^*(y)),\end{aligned}$$

where $\|R'(\pi^*(y))\| \leq C_{k,d,\tau_{min},\mathbf{L}}\sqrt{f_{max}/f_{min}}(h^k \vee \sigma)$. According to Lemma 2, we deduce that $\|\widehat{\Psi}(v) - y\| \leq C_{k,d,\tau_{min},\mathbf{L}}\sqrt{f_{max}/f_{min}}(h^k \vee \sigma)$, hence

$$\sup_{u\in\hat{M}} d(u, M) \leq C_{k,d,\tau_{min},\mathbf{L}}\sqrt{\frac{f_{max}}{f_{min}}}(h^k \vee \sigma). \tag{4}$$

Now we focus on $\sup_{y\in M} d(y, \hat{M})$. For this, we need a lemma ensuring that $\mathbb{Y}_n = \{Y_1, \ldots, Y_n\}$ covers M with high probability.

LEMMA 4. *Let* $h = \left(\frac{C'_d k}{f_{min}}\frac{\log n}{n}\right)^{1/d}$ *with* C'_d *large enough. Then for* n *large enough so that* $h \leq \tau_{min}/4$, *with probability at least* $1 - \left(\frac{1}{n}\right)^{k/d}$,

$$d_H(M, \mathbb{Y}_n) \leq h/2.$$

The proof of Lemma 4 is given in Section B.1. Now we choose h satisfying the conditions of Proposition B.8 and Lemma 4. Let y be in M and assume that $\|y - Y_{j_0}\| \leq h/2$. Then $y \in \mathcal{B}(X_{j_0}, 3h/4)$. According to Lemma 3 and (3), we deduce that $\|\widehat{\Psi}_{j_0}(\hat{\pi}_{j_0}(y - X_{j_0})) - y\| \leq C_{k,d,\tau_{min},\mathbf{L}}\sqrt{f_{max}/f_{min}}(h^k \vee \sigma)$. Hence, from Lemma 4,

$$\sup_{y\in M} d(y, \hat{M}) \leq C_{k,d,\tau_M,\mathbf{L}}\sqrt{\frac{f_{max}}{f_{min}}}(h^k \vee \sigma) \tag{5}$$

with probability at least $1 - 2\left(\frac{1}{n}\right)^{k/d}$. Combining (4) and (5) gives Theorem 6. □

5.2. Minimax Lower Bounds

This section is devoted to describe the main ideas of the proofs of the minimax lower bounds. We prove Theorem 7 on one side, and Theorem 3 and Theorem 5 in a unified way on the other side. The methods used rely on hypothesis comparison [36].

5.2.1. Lower Bound for Manifold Estimation

We recall that for two distributions Q and Q' defined on the same space, the L^1 test affinity $\|Q \wedge Q'\|_1$ is given by

$$\|Q \wedge Q'\|_1 = \int dQ \wedge dQ',$$

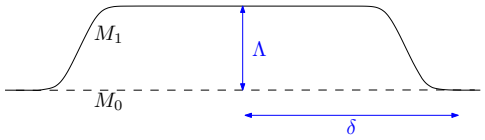


Figure 5: Manifolds M_0 and M_1 of Lemma 5 and Lemma 6. The width δ of the bump is chosen to have $\|P_0^\sigma \wedge P_1^\sigma\|_1^n$ constant. The distance $\Lambda = d_H(M_0, M_1)$ is of order δ^k to ensure that $M_1 \in \mathcal{C}^k$.

where dQ and dQ' denote densities of Q and Q' with respect to any dominating measure.

The first technique we use, involving only two hypotheses, is usually referred to as Le Cam's Lemma [36]. Let $\mathcal{P}$ be a model and $\theta(P)$ be the parameter of interest. Assume that $\theta(P)$ belongs to a pseudo-metric space $(\mathcal{D}, d)$, that is $d(\cdot,\cdot)$ is symmetric and satisfies the triangle inequality. Le Cam's Lemma can be adapted to our framework as follows.

THEOREM 8 (Le Cam's Lemma [36]). *For all pairs P, P' in $\mathcal{P}$,*

$$\inf_{\hat{\theta}} \sup_{P\in\mathcal{P}} \mathbb{E}_{P^{\otimes n}} d(\theta(P), \hat{\theta}) \geq \frac{1}{2} d\left(\theta(P), \theta(P')\right) \left\|P \wedge P'\right\|_1^n,$$

where the infimum is taken over all the estimators $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$.

In this section, we will get interested in $\mathcal{P} = \mathcal{P}^k(\sigma)$ and $\theta(P) = M$, with $d = d_H$. In order to derive Theorem 7, we build two different pairs (P_0, P_1), (P_0^σ, P_1^σ) of hypotheses in the model $\mathcal{P}^k(\sigma)$. Each pair will exploit a different property of the model $\mathcal{P}^k(\sigma)$.

The first pair (P_0, P_1) of hypotheses (Lemma 5) is built in the model $\mathcal{P}^k \subset \mathcal{P}^k(\sigma)$, and exploits the geometric difficulty of manifold reconstruction, even if no noise is present. These hypotheses, depicted in Figure 5, consist of bumped versions of one another.

LEMMA 5. *Under the assumptions of Theorem 7, there exist $P_0, P_1 \in \mathcal{P}^k$ with associated submanifolds M_0, M_1 such that*

$$d_H(M_0, M_1) \geq c_{k,d,\tau_{min}} \left(\frac{1}{n}\right)^{\frac{k}{d}}, \quad \textit{and} \quad \|P_0 \wedge P_1\|_1^n \geq c_0.$$

The proof of Lemma 5 is to be found in Section C.4.1.

The second pair (P_0^σ, P_1^σ) of hypotheses (Lemma 6) has a similar construction than (P_0, P_1). Roughly speaking, they are the uniform distributions on

the offsets of radii $\sigma/2$ of M_0 and M_1 of Figure 5. Here, the hypotheses are built in $\mathcal{P}^k(\sigma)$, and fully exploit the statistical difficulty of manifold reconstruction induced by noise.

LEMMA 6. *Under the assumptions of Theorem 7, there exist* $P_0^\sigma, P_1^\sigma \in \mathcal{P}^k(\sigma)$ *with associated submanifolds* M_0^σ, M_1^σ *such that*

$$d_H(M_0^\sigma, M_1^\sigma) \geq c_{k,d,\tau_{min}} \left(\frac{\sigma}{n}\right)^{\frac{k}{d+k}}, \quad \textit{and} \quad \|P_0^\sigma \wedge P_1^\sigma\|_1^n \geq c_0.$$

The proof of Lemma 6 is to be found in Section C.4.2. We are now in position to prove Theorem 7.

PROOF OF THEOREM 7. Let us apply Theorem C.20 with $\mathcal{P} = \mathcal{P}^k(\sigma)$, $\theta(P) = M$ and $d = d_H$. Taking $P = P_0$ and $P' = P_1$ of Lemma 5, these distributions both belong to $\mathcal{P}^k \subset \mathcal{P}^k(\sigma)$, so that Theorem C.20 yields

$$\begin{aligned}\inf_{\hat{M}} \sup_{P \in \mathcal{P}^k(\sigma)} \mathbb{E}_{P^{\otimes n}} d_H\big(M, \hat{M}\big) &\geq d_H(M_0, M_1)\, \|P_0 \wedge P_1\|_1^n \\ &\geq c_{k,d,\tau_{min}} \left(\frac{1}{n}\right)^{\frac{k}{d}} \times c_0.\end{aligned}$$

Similarly, setting hypotheses $P = P_0^\sigma$ and $P' = P_1^\sigma$ of Lemma 6 yields

$$\begin{aligned}\inf_{\hat{M}} \sup_{P \in \mathcal{P}^k(\sigma)} \mathbb{E}_{P^{\otimes n}} d_H\big(M, \hat{M}\big) &\geq d_H(M_0^\sigma, M_1^\sigma)\, \|P_0^\sigma \wedge P_1^\sigma\|_1^n \\ &\geq c_{k,d,\tau_{min}} \left(\frac{\sigma}{n}\right)^{\frac{k}{k+d}} \times c_0,\end{aligned}$$

which concludes the proof. □

5.2.2. Lower Bounds for Tangent Space and Curvature Estimation

Let us now move to the proof of Theorem 3 and 5, that consist of lower bounds for the estimation of $T_{X_1}M$ and $II_{X_1}^M$ with random base point X_1. In both cases, the loss can be cast as

$$\begin{aligned}E_{P^{\otimes n}}\, d(\theta_{X_1}(P), \hat{\theta}) &= \mathbb{E}_{P^{\otimes n-1}} \Big[E_P\, d(\theta_{X_1}(P), \hat{\theta})\Big] \\ &= \mathbb{E}_{P^{\otimes n-1}} \Big[\Big\| d\big(\theta_\cdot(P), \hat{\theta}\big)\Big\|_{L^1(P)}\Big],\end{aligned}$$

where $\hat{\theta} = \hat{\theta}(X, X')$, with $X = X_1$ driving the parameter of interest, and $X' = (X_2, \ldots, X_n) = X_{2:n}$. Since $\|.\|_{L^1(P)}$ obviously depends on P, the

technique exposed in the previous section does not apply anymore. However, a slight adaptation of Assouad's Lemma [36] with an extra conditioning on $X = X_1$ carries out for our purpose. Let us now detail a general framework where the method applies.

We let $\mathcal{X}, \mathcal{X}'$ denote measured spaces. For a probability distribution Q on $\mathcal{X} \times \mathcal{X}'$, we let (X, X') be a random variable with distribution Q. The marginals of Q on $\mathcal{X}$ and $\mathcal{X}'$ are denoted by μ and ν respectively. Let $(\mathcal{D}, d)$ be a pseudo-metric space. For $Q \in \mathcal{Q}$, we let $\theta_\cdot(Q) : \mathcal{X} \to \mathcal{D}$ be defined μ-almost surely, where μ is the marginal distribution of Q on $\mathcal{X}$. The parameter of interest is $\theta_X(Q)$, and the associated minimax risk over $\mathcal{Q}$ is

$$\inf_{\hat{\theta}} \sup_{Q \in \mathcal{Q}} \mathbb{E}_Q \left[d\big(\theta_X(Q), \hat{\theta}(X, X')\big) \right], \tag{6}$$

where the infimum is taken over all the estimators $\hat{\theta} : \mathcal{X} \times \mathcal{X}' \to \mathcal{D}$.

Given a set of probability distributions $\mathcal{Q}$ on $\mathcal{X} \times \mathcal{X}'$, write $\overline{Conv}(\mathcal{Q})$ for the set of mixture probability distributions with components in $\mathcal{Q}$. For all $\tau = (\tau_1, \dots, \tau_m) \in \{0,1\}^m$, τ^k denotes the m-tuple that differs from τ only at the kth position. We are now in position to state the conditional version of Assouad's Lemma that allows to lower bound the minimax risk (6).

LEMMA 7 (Conditional Assouad). *Let $m \geq 1$ be an integer and let $\{\mathcal{Q}_\tau\}_{\tau \in \{0,1\}^m}$ be a family of 2^m submodels $\mathcal{Q}_\tau \subset \mathcal{Q}$. Let $\{U_k \times U_k'\}_{1 \leq k \leq m}$ be a family of pairwise disjoint subsets of $\mathcal{X} \times \mathcal{X}'$, and $\mathcal{D}_{\tau,k}$ be subsets of $\mathcal{D}$. Assume that for all $\tau \in \{0,1\}^m$ and $1 \leq k \leq m$,*

- *for all $Q_\tau \in \mathcal{Q}_\tau$, $\theta_X(Q_\tau) \in \mathcal{D}_{\tau,k}$ on the event $\{X \in U_k\}$;*
- *for all $\theta \in \mathcal{D}_{\tau,k}$ and $\theta' \in \mathcal{D}_{\tau^k,k}$, $d(\theta, \theta') \geq \Delta$.*

For all $\tau \in \{0,1\}^m$, let $\overline{Q}_\tau \in \overline{Conv}(\mathcal{Q}_\tau)$, and write $\bar{\mu}_\tau$ and $\bar{\nu}_\tau$ for the marginal distributions of $\overline{Q}_\tau$ on $\mathcal{X}$ and $\mathcal{X}'$ respectively. Assume that if (X, X') has distribution $\overline{Q}_\tau$, X and X' are independent conditionally on the event $\{(X, X') \in U_k \times U_k'\}$, and that

$$\min_{\substack{\tau \in \{0,1\}^m \\ 1 \leq k \leq m}} \left\{ \left(\int_{U_k} d\bar{\mu}_\tau \wedge d\bar{\mu}_{\tau^k} \right) \left(\int_{U_k'} d\bar{\nu}_\tau \wedge d\bar{\nu}_{\tau^k} \right) \right\} \geq 1 - \alpha.$$

Then,

$$\inf_{\hat{\theta}} \sup_{Q \in \mathcal{Q}} \mathbb{E}_Q \left[d\big(\theta_X(Q), \hat{\theta}(X, X')\big) \right] \geq m \frac{\Delta}{2} (1 - \alpha),$$

where the infimum is taken over all the estimators $\hat{\theta} : \mathcal{X} \times \mathcal{X}' \to \mathcal{D}$.

Note that for a model of the form $\mathcal{Q} = \{\delta_{x_0} \otimes P, P \in \mathcal{P}\}$ with fixed $x_0 \in \mathcal{X}$, one recovers the classical Assouad's Lemma [36] taking $U_k = \mathcal{X}$ and $U_k' = \mathcal{X}'$. Indeed, when $X = x$ is deterministic, the parameter of interest $\theta_X(Q) = \theta(Q)$ can be seen as non-random.

In this section, we will get interested in $\mathcal{Q} = \mathcal{P}^k(\sigma)^{\otimes n}$, and $\theta_X(Q) = \theta_{X_1}(Q)$ being alternatively $T_{X_1}M$ and $II^M_{X_1}$. Similarly to Section 5.2.1, we build two different families of submodels, each of them will exploit a different kind of difficulty for tangent space and curvature estimation.

The first family, described in Lemma 8, highlights the geometric difficulty of the estimation problems, even when the noise level σ is small, or even zero. Let us emphasize that the estimation error is integrated with respect to the distribution of X_1. Hence, considering mixture hypotheses is natural, since building manifolds with different tangent spaces (or curvature) necessarily leads to distributions that are locally singular. Here, as in Section 5.2.1, the considered hypotheses are composed of bumped manifolds (see Figure 7). We defer the proof of Lemma 8 to Section C.3.1.

LEMMA 8. *Assume that the conditions of Theorem 3 or 5 hold. Given $i \in \{1,2\}$, there exists a family of 2^m submodels $\{\mathcal{P}_\tau^{(i)}\}_{\tau\in\{0,1\}^m} \subset \mathcal{P}^k$, together with pairwise disjoint subsets $\{U_k \times U_k'\}_{1\le k\le m}$ of $\mathbb{R}^D \times \big(\mathbb{R}^D\big)^{n-1}$ such that the following holds for all $\tau \in \{0,1\}^m$ and $1 \le k \le m$.*

For any distribution $P_\tau^{(i)} \in \mathcal{P}_\tau^{(i)}$ with support $M_\tau^{(i)} = Supp\big(P_\tau^{(i)}\big)$, if $(X_1,\dots,X_n)$ has distribution $\big(P_\tau^{(i)}\big)^{\otimes n}$, then on the event $\{X_1 \in U_k\}$, we have:

- *if $\tau_k = 0$,*

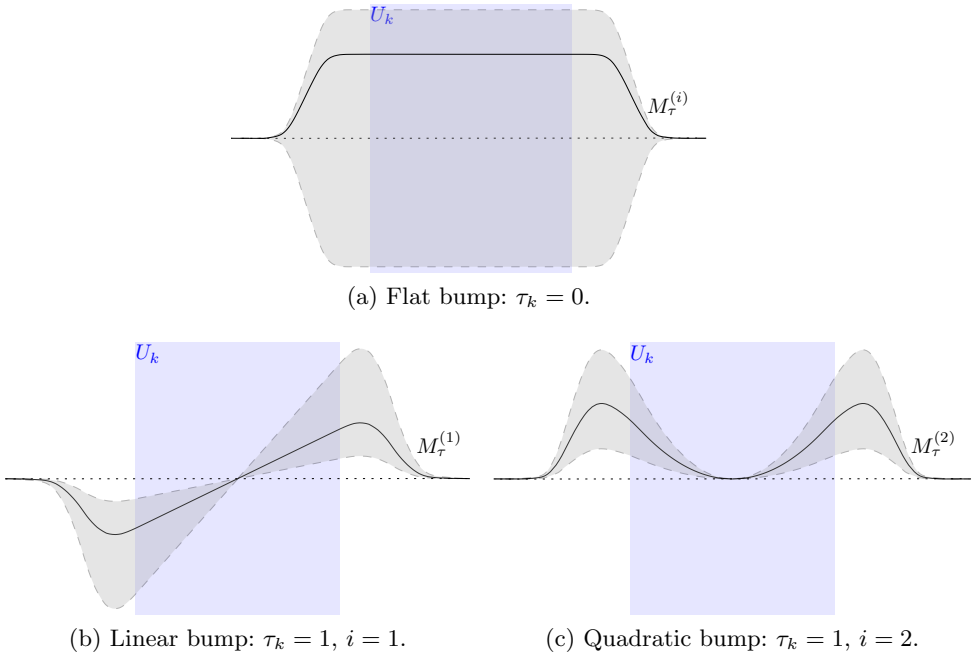
$$T_{X_1}M_\tau^{(i)} = \mathbb{R}^d \times \{0\}^{D-d} \quad , \quad \left\| II_{X_1}^{M_\tau^{(i)}} \circ \pi_{T_{X_1}M_\tau^{(i)}} \right\|_{op} = 0,$$

- *if $\tau_k = 1$,*
  - *for $i = 1$: $\angle\left(T_{X_1}M_\tau^{(1)}, \mathbb{R}^d \times \{0\}^{D-d}\right) \ge c_{k,d,\tau_{min}} \left(\frac{1}{n-1}\right)^{\frac{k-1}{d}}$,*
  - *for $i = 2$: $\left\| II_{X_1}^{M_\tau^{(2)}} \circ \pi_{T_{X_1}M_\tau^{(2)}} \right\|_{op} \ge c_{k,d,\tau_{min}} \left(\frac{1}{n-1}\right)^{\frac{k-2}{d}}$.*

Furthermore, there exists $\bar{Q}_{\tau,n}^{(i)} \in \overline{Conv}\big(\big(\mathcal{P}_\tau^{(i)}\big)^{\otimes n}\big)$ such that if $(Z_1,\dots,Z_n) = (Z_1, Z_{2:n})$ has distribution $\bar{Q}_{\tau,n}^{(i)}$, Z_1 and $Z_{2:n}$ are independent conditionally on the event $\{(Z_1, Z_{2:n}) \in U_k \times U_k'\}$. The marginal distributions of $\bar{Q}_{\tau,n}^{(i)}$ on

$\mathbb{R}^D \times \left(\mathbb{R}^D\right)^{n-1}$ *are* $\bar{Q}^{(i)}_{\tau,1}$ *and* $\bar{Q}^{(i)}_{\tau,n-1}$*, and we have*

$$\int_{U'_k} d\bar{Q}^{(i)}_{\tau,n-1} \wedge d\bar{Q}^{(i)}_{\tau^k,n-1} \geq c_0, \text{ and } \quad m \cdot \int_{U_k} d\bar{Q}^{(i)}_{\tau,1} \wedge d\bar{Q}^{(i)}_{\tau^k,1} \geq c_d.$$



(a) Flat bump: $\tau_k = 0$.

(b) Linear bump: $\tau_k = 1$, $i = 1$.

(c) Quadratic bump: $\tau_k = 1$, $i = 2$.

Figure 6: Distributions of Lemma 8 in the neighborhood of U_k $(1 \leq k \leq m)$. Black curves correspond to the support $M_\tau^{(i)}$ of a distribution of $\mathcal{P}_\tau^{(i)} \subset \mathcal{P}^k$. The area shaded in grey depicts the mixture distribution $\bar{Q}^{(i)}_{\tau,1} \in \overline{Conv}\left(\mathcal{P}_\tau^{(i)}\right)$.

The second family, described in Lemma 9, testifies of the statistical difficulty of the estimation problem when the noise level σ is large enough. The construction is very similar to Lemma 8 (see Figure 7). Though, in this case, the magnitude of the noise drives the statistical difficulty, as opposed to the sampling scale in Lemma 8. Note that in this case, considering mixture distributions is not necessary since the ample-enough noise make bumps that are absolutely continuous with respect to each other. The proof of Lemma 9 can be found in Section C.3.2.

LEMMA 9. *Assume that the conditions of Theorem 3 or 5 hold, and that* $\sigma \geq C_{k,d,\tau_{min}} \left(1/(n-1)\right)^{k/d}$ *for* $C_{k,d,\tau_{min}} > 0$ *large enough. Given* $i \in \{1, 2\}$*,*

there exists a collection of 2^m *distributions* $\{\mathbf{P}_\tau^{(i),\sigma}\}_{\tau\in\{0,1\}^m} \subset \mathcal{P}^k(\sigma)$ *with associated submanifolds* $\{M_\tau^{(i),\sigma}\}_{\tau\in\{0,1\}^m}$, *together with pairwise disjoint subsets* $\{U_k^\sigma\}_{1\le k\le m}$ *of* $\mathbb{R}^D$ *such that the following holds for all* $\tau\in\{0,1\}^m$ *and* $1\le k\le m$.

If $x\in U_k^\sigma$ *and* $y=\pi_{M_\tau^{(i),\sigma}}(x)$, *we have*

- *if* $\tau_k=0$,
$$T_yM_\tau^{(i),\sigma}=\mathbb{R}^d\times\{0\}^{D-d} \quad , \quad \left\|II_y^{M_\tau^{(i),\sigma}}\circ\pi_{T_yM_\tau^{(i),\sigma}}\right\|_{op}=0,$$
- *if* $\tau_k=1$,
  - *for* $i=1$: $\angle\left(T_yM_\tau^{(1),\sigma},\mathbb{R}^d\times\{0\}^{D-d}\right)\ge c_{k,d,\tau_{min}}\left(\frac{\sigma}{n-1}\right)^{\frac{k-1}{k+d}}$,
  - *for* $i=2$: $\left\|II_y^{M_\tau^{(2),\sigma}}\circ\pi_{T_yM_\tau^{(2),\sigma}}\right\|_{op}\ge c'_{k,d,\tau_{min}}\left(\frac{\sigma}{n-1}\right)^{\frac{k-2}{k+d}}$.

Furthermore,
$$\int_{(\mathbb{R}^D)^{n-1}}\left(\mathbf{P}_\tau^{(i),\sigma}\right)^{\otimes n-1}\wedge\left(\mathbf{P}_{\tau^k}^{(i),\sigma}\right)^{\otimes n-1}\ge c_0, \text{ and } \quad m\cdot\int_{U_k^\sigma}\mathbf{P}_\tau^{(i),\sigma}\wedge\mathbf{P}_{\tau^k}^{(i),\sigma}\ge c_d.$$

Proof of Theorem 3. Let us apply Lemma C.11 with $\mathcal{X}=\mathbb{R}^D$, $\mathcal{X}'=\left(\mathbb{R}^D\right)^{n-1}$, $\mathcal{Q}=\left(\mathcal{P}^k(\sigma)\right)^{\otimes n}$, $X=X_1$, $X'=(X_2,\dots,X_n)=X_{2:n}$, $\theta_X(Q)=T_XM$, and the angle between linear subspaces as the distance d.

If $\sigma<C_{k,d,\tau_{min}}\left(1/(n-1)\right)^{k/d}$, for $C_{k,d,\tau_{min}}>0$ defined in Lemma 9, then, applying Lemma C.11 to the family $\{\bar{Q}_{\tau,n}^{(1)}\}_\tau$ together with the disjoint sets $U_k\times U'_k$ of Lemma 8, we get
$$\begin{aligned}\inf_{\hat{T}}\sup_{P\in\mathcal{P}^k(\sigma)}\mathbb{E}_{P^{\otimes n}}\angle\left(T_{\pi_M(X_1)}M,\hat{T}\right)&\ge m\cdot c_{k,d,\tau_{min}}\left(\frac{1}{n-1}\right)^{\frac{k-1}{d}}\cdot c_0\cdot c_d\\&=c'_{d,k,\tau_{min}}\left\{\left(\frac{1}{n-1}\right)^{\frac{k-1}{d}}\vee\left(\frac{\sigma}{n-1}\right)^{\frac{k-1}{d+k}}\right\},\end{aligned}$$
where the second line uses that $\sigma<C_{k,d,\tau_{min}}\left(1/(n-1)\right)^{k/d}$.

If $\sigma\ge C_{k,d,\tau_{min}}\left(1/(n-1)\right)^{k/d}$, then Lemma 9 holds, and considering the family $\left\{\left(\mathbf{P}_\tau^{(1),\sigma}\right)^{\otimes n}\right\}_\tau$, together with the disjoint sets $U_k^\sigma\times\left(\mathbb{R}^D\right)^{n-1}$, Lemma

C.11 gives

$$\inf_{\hat{T}} \sup_{P\in\mathcal{P}^k(\sigma)} \mathbb{E}_{P^{\otimes n}} \angle\big(T_{\pi_M(X_1)}M, \hat{T}\big) \geq m \cdot c_{k,d,\tau_{min}} \left(\frac{\sigma}{n-1}\right)^{\frac{k-1}{k+d}} \cdot c_0 \cdot c_d$$
$$= c''_{d,k,\tau_{min}} \left\{ \left(\frac{1}{n-1}\right)^{\frac{k-1}{d}} \vee \left(\frac{\sigma}{n-1}\right)^{\frac{k-1}{d+k}} \right\},$$

hence the result.

□

Proof of Theorem 5. The proof follows the exact same lines as that of Theorem 3 just above. Namely, consider the same setting with $\theta_X(Q) = II^M_{\pi_M(X)}$. If $\sigma \geq C_{k,d,\tau_{min}} \left(1/(n-1)\right)^{k/d}$, apply Lemma C.11 with the family $\left\{\bar{Q}^{(2)}_{\tau,n}\right\}_\tau$ of Lemma 8. If $\sigma > C_{k,d,\tau_{min}} \left(1/(n-1)\right)^{k/d}$, Lemma C.11 can be applied to $\left\{\left(\mathbf{P}^{(2),\sigma}_\tau\right)^{\otimes n}\right\}_\tau$ in Lemma 9. This yields the announced rate. □

Acknowledgements

We would like to thank Frédéric Chazal and Pascal Massart for their constant encouragements, suggestions and stimulating discussions. We also thank the anonymous reviewers for valuable comments and suggestions.

REFERENCES

[1] Aamari, E. and Levrard, C. (2015). Stability and Minimax Optimality of Tangential Delaunay Complexes for Manifold Reconstruction. *ArXiv e-prints*.

[2] Aamari, E. and Levrard, C. (2017). Non-asymptotic rates for manifold, tangent space and curvature estimation.

[3] Alexander, S. B. and Bishop, R. L. (2006). Gauss equation and injectivity radii for subspaces in spaces of curvature bounded above. *Geom. Dedicata* **117** 65–84. MR2231159 (2007c:53110)

[4] Arias-Castro, E., Lerman, G. and Zhang, T. (2013). Spectral Clustering Based on Local PCA. *ArXiv e-prints*.

[5] Arias-Castro, E., Pateiro-López, B. and Rodríguez-Casal, A. (2016). Minimax Estimation of the Volume of a Set with Smooth Boundary. *ArXiv e-prints*.

[6] Boissonnat, J.-D. and Ghosh, A. (2014). Manifold reconstruction using tangential Delaunay complexes. *Discrete Comput. Geom.* **51** 221–267. MR3148657

[7] Boucheron, S., Lugosi, G. and Massart, P. (2013). *Concentration inequalities*. Oxford University Press, Oxford A nonasymptotic theory of independence, With a foreword by Michel Ledoux. MR3185193

[8] Bousquet, O. (2002). A Bennett concentration inequality and its application to suprema of empirical processes. *C. R. Math. Acad. Sci. Paris* **334** 495–500. MR1890640 (2003f:60039)

[9] Cazals, F. and Pouget, M. (2005). Estimating differential quantities using polynomial fitting of osculating jets. *Comput. Aided Geom. Design* **22** 121–146. MR2116098

[10] Chazal, F., Glisse, M., Labruère, C. and Michel, B. (2013). Optimal rates of convergence for persistence diagrams in Topological Data Analysis. *ArXiv e-prints*.

[11] Cheng, S.-W. and Chiu, M.-K. (2016). Tangent estimation from point samples. *Discrete Comput. Geom.* **56** 505–557. MR3544007

[12] Cheng, S.-W., Dey, T. K. and Ramos, E. A. (2005). Manifold reconstruction from point samples. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms* 1018–1027. ACM, New York. MR2298361

[13] do Carmo, M. P. (1992). *Riemannian geometry. Mathematics: Theory & Applications*. Birkhäuser Boston, Inc., Boston, MA Translated from the second Portuguese edition by Francis Flaherty. MR1138207 (92i:53001)

[14] Dyer, R., Vegter, G. and Wintraecken, M. (2015). Riemannian simplices and triangulations. *Geom. Dedicata* **179** 91–138. MR3424659

[15] Fasy, B. T., Lecci, F., Rinaldo, A., Wasserman, L., Balakrishnan, S. and Singh, A. (2014). Confidence sets for persistence diagrams. *Ann. Statist.* **42** 2301–2339. MR3269981

[16] Federer, H. (1959). Curvature measures. *Trans. Amer. Math. Soc.* **93** 418–491. MR0110078 (22 ##961)

[17] Federer, H. (1969). *Geometric measure theory. Die Grundlehren der mathematischen Wissenschaften, Band 153*. Springer-Verlag New York Inc., New York. MR0257325 (41 ##1976)

[18] Fefferman, C., Ivanov, S., Kurylev, Y., Lassas, M. and Narayanan, H. (2015). Reconstruction and interpolation of manifolds I: The geometric Whitney problem. *ArXiv e-prints*.

[19] Gashler, M. S. and Martinez, T. (2011). Tangent Space Guided Intelligent Neighbor Finding. In *Proceedings of the IEEE International Joint Conference on Neural Networks IJCNN'11* 2617–2624. IEEE Press.

[20] Genovese, C. R., Perone-Pacifico, M., Verdinelli, I. and Wasserman, L. (2012). Manifold estimation and singular deconvolution under Hausdorff loss. *Ann. Statist.* **40** 941–963. MR2985939

[21] Genovese, C. R., Perone-Pacifico, M., Verdinelli, I. and Wasserman, L. (2012). Minimax manifold estimation. *J. Mach. Learn. Res.* **13** 1263–1291. MR2930639

[22] Golub, G. H. and Van Loan, C. F. (1996). *Matrix computations*, third ed. *Johns Hopkins Studies in the Mathematical Sciences*. Johns Hopkins University Press, Baltimore, MD. MR1417720 (97g:65006)

[23] Gumhold, S., Wang, X. and MacLeod, R. (2001). Feature Extraction from Point Clouds. In *10th International Meshing Roundtable* 293–305. Sandia National Laboratories,.

[24] Hartman, P. (1951). On geodesic coordinates. *Amer. J. Math.* **73** 949–954. MR0046087

[25] Kim, A. K. H. and Zhou, H. H. (2015). Tight minimax rates for manifold estimation under Hausdorff loss. *Electron. J. Stat.* **9** 1562–1582. MR3376117

[26] Maggioni, M., Minsker, S. and Strawn, N. (2016). Multiscale dictionary learning: non-asymptotic bounds and robustness. *J. Mach. Learn. Res.* **17** Paper No. 2, 51. MR3482922

[27] Merigot, Q., Ovsjanikov, M. and Guibas, L. J. (2011). Voronoi-Based Curvature and Feature Estimation from Point Clouds. *IEEE Transactions on Visualization and Computer Graphics* **17** 743-756.

[28] Niyogi, P., Smale, S. and Weinberger, S. (2008). Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput. Geom.* **39** 419–441. MR2383768 (2009b:60038)

[29] Roweis, S. T. and Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *SCIENCE* **290** 2323–2326.

[30] Rusinkiewicz, S. (2004). Estimating Curvatures and Their Derivatives on Triangle Meshes. In *Proceedings of the 3D Data Processing, Visualization, and Transmission, 2Nd International Symposium. 3DPVT '04* 486–493. IEEE Computer Society, Washington, DC, USA.

[31] Singer, A. and Wu, H. T. (2012). Vector diffusion maps and the connection Laplacian. *Comm. Pure Appl. Math.* **65** 1067–1144. MR2928092

[32] Tenenbaum, J. B., de Silva, V. and Langford, J. C. (2000). A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* **290** 2319.

[33] Tyagi, H., Vural, E. f. and Frossard, P. (2013). Tangent space estimation for smooth embeddings of Riemannian manifolds. *Inf. Inference* **2** 69–114. MR3311444

[34] Usevich, K. and Markovsky, I. (2014). Optimization on a Grassmann manifold with application to system identification. *Automatica* **50** 1656 - 1662.

[35] Wasserman, L. (2016). Topological Data Analysis. *ArXiv e-prints*.

[36] Yu, B. (1997). Assouad, fano, and le cam. In *Festschrift for Lucien Le Cam* 423–435. Springer.

Appendix: Geometric background and proofs of intermediate results

Appendix A: Properties and Stability of the Models

A.1. Property of the Exponential Map in $\mathcal{C}^2_{\tau_{min}}$

Here we show the following Lemma 1, reproduced as Lemma A.1.

Lemma A.1. *If $M \in \mathcal{C}^2_{\tau_{min}}$, $\exp_p : \mathcal{B}_{T_pM}(0, \tau_{min}/4) \to M$ is one-to-one. Moreover, it can be written as*

$$\begin{aligned} \exp_p \colon \mathcal{B}_{T_pM}(0, \tau_{min}/4) &\longrightarrow M \\ v &\longmapsto p + v + \mathbf{N}_p(v) \end{aligned}$$

with $\mathbf{N}_p$ such that for all $v \in \mathcal{B}_{T_pM}(0, \tau_{min}/4)$,

$$\mathbf{N}_p(0) = 0, \quad d_0\mathbf{N}_p = 0, \quad \|d_v\mathbf{N}_p\|_{op} \leq L_\perp \|v\|,$$

where $L_\perp = 5/(4\tau_{min})$. Furthermore, for all $p, y \in M$,

$$y - p = \pi_{T_pM}(y - p) + R_2(y - p),$$

where $\|R_2(y-p)\| \leq \frac{\|y-p\|^2}{2\tau_{min}}$.

Proof of Lemma A.1. Proposition 6.1 in [28] states that for all $x \in M$, $\left\|II^M_x\right\|_{op} \leq 1/\tau_{min}$. In particular, Gauss equation ([13, Proposition 3.1 (a), p.135]) yields that the sectional curvatures of M satisfy $-2/\tau^2_{min} \leq \kappa \leq 1/\tau^2_{min}$. Using Corollary 1.4 of [3], we get that the injectivity radius of M is at least $\pi\tau_{min} \geq \tau_{min}/4$. Therefore, $\exp_p : \mathcal{B}_{T_pM}(0, \tau_{min}/4) \to M$ is one-to-one.

Let us write $\mathbf{N}_p(v) = \exp_p(v) - p - v$. We clearly have $\mathbf{N}_p(0) = 0$ and $d_0\mathbf{N}_p = 0$. Let now $v \in \mathcal{B}_{T_pM}(0, \tau_{min}/4)$ be fixed. We have $d_v\mathbf{N}_p = d_v \exp_p - Id_{T_pM}$. For $0 \leq t \leq \|v\|$, we write $\gamma(t) = \exp_p(tv/\|v\|)$ for the arc-length parametrized geodesic from p to $\exp_p(v)$, and P_t for the parallel translation along γ. From Lemma 18 of [14],

$$\left\| d_{t\frac{v}{\|v\|}} \exp_p - P_t \right\|_{op} \leq \frac{2}{\tau^2_{min}} \frac{t^2}{2} \leq \frac{t}{4\tau_{min}}.$$

We now derive an upper bound for $\left\|P_t - Id_{T_pM}\right\|_{op}$. For this, fix two unit vectors $u \in \mathbb{R}^D$ and $w \in T_pM$, and write $g(t) = \langle P_t(w) - w, u\rangle$. Letting $\bar{\nabla}$ denote the ambient derivative in $\mathbb{R}^D$, by definition of parallel translation,

$$\begin{aligned}\left|g'(t)\right| &= \left|\langle \bar{\nabla}_{\gamma'(t)} P_t(w) - w, u\rangle\right| \\ &= \left|\langle II^M_{\gamma(t)}\big(\gamma'(t), P_t(w)\big), u\rangle\right| \\ &\leq 1/\tau_{min}.\end{aligned}$$

Since $g(0) = 0$, we get $\left\|P_t - Id_{T_pM}\right\|_{op} \leq t/\tau_{min}$. Finally, the triangle inequality leads to

$$\begin{aligned}\left\|d_v \mathbf{N}_p\right\|_{op} &= \left\|d_v \exp - Id_{T_pM}\right\|_{op} \\ &\leq \left\|d_v \exp - P_{\|v\|}\right\|_{op} + \left\|P_{\|v\|} - Id_{T_pM}\right\|_{op} \\ &\leq \frac{5\left\|v\right\|}{4\tau_{min}}.\end{aligned}$$

We conclude with the property of the projection $\pi^* = \pi_{T_pM}$. Indeed, defining $R_2(y-p) = (y-p) - \pi^*(y-p)$, Lemma 4.7 in [16] gives

$$\begin{aligned}\left\|R_2(y-p)\right\| &= d(y-p, T_pM) \\ &\leq \frac{\left\|y-p\right\|^2}{2\tau_{min}}.\end{aligned}$$

□

A.2. Geometric Properties of the Models $\mathcal{C}^k$

LEMMA A.2. *For any $M \in \mathcal{C}^k_{\tau_{min},\mathbf{L}}$ and $x \in M$, the following holds.*

(i) For all $v_1, v_2 \in \mathcal{B}_{T_xM}\left(0, \frac{1}{4L_\perp}\right)$,

$$\frac{3}{4}\left\|v_2 - v_1\right\| \leq \left\|\Psi_x(v_2) - \Psi_x(v_1)\right\| \leq \frac{5}{4}\left\|v_2 - v_1\right\|.$$

(ii) For all $h \leq \frac{1}{4L_\perp} \wedge \frac{2\tau_{min}}{5}$,

$$M \cap \mathcal{B}\left(x, \frac{3h}{5}\right) \subset \Psi_x\left(\mathcal{B}_{T_xM}\left(x, h\right)\right) \subset M \cap \mathcal{B}\left(x, \frac{5h}{4}\right).$$

(iii) For all $h \leq \frac{\tau_{min}}{2}$,

$$\mathcal{B}_{T_xM}\left(0, \frac{7h}{8}\right) \subset \pi_{T_xM}\left(\mathcal{B}(x,h) \cap M\right).$$

(iv) *Denoting by* $\pi^* = \pi_{T_xM}$ *the orthogonal projection onto* T_xM*, for all* $x \in M$*, there exist multilinear maps* $T_2^*, \dots, T_{k-1}^*$ *from* T_xM *to* $\mathbb{R}^D$*, and* R_k *such that for all* $y \in \mathcal{B}\left(x, \frac{\tau_{min} \wedge L_\perp^{-1}}{4}\right) \cap M$*,*

$$y - x = \pi^*(y-x) + T_2^*(\pi^*(y-x)^{\otimes 2}) + \ldots + T_{k-1}^*(\pi^*(y-x)^{\otimes k-1}) + R_k(y-x),$$

with

$$\|R_k(y-x)\| \leq C\|y-x\|^k \quad \text{and} \quad \|T_i^*\|_{op} \leq L_i', \text{ for } 2 \leq i \leq k-1,$$

where L_i' *depends on* $d, k, \tau_{min}, L_\perp, \dots, L_i$*, and* C *on* d*,* k*,* τ_{min}*,* $L_\perp$*,* $\dots$*,* L_k*. Moreover, for* $k \geq 3$*,* $T_2^* = II_x^M$*.*

(v) *For all* $x \in M$*,* $\left\|II_x^M\right\|_{op} \leq 1/\tau_{min}$*. In particular, the sectional curvatures of* M *satisfy*

$$\frac{-2}{\tau_{min}^2} \leq \kappa \leq \frac{1}{\tau_{min}^2}.$$

Proof of Lemma A.2. (i) Simply notice that from the reverse triangle inequality,

$$\left| \frac{\|\Psi_x(v_2) - \Psi_x(v_1)\|}{\|v_2 - v_1\|} - 1 \right| \leq \frac{\|N_x(v_2) - N_x(v_1)\|}{\|v_2 - v_1\|} \leq L_\perp(\|v_1\| \vee \|v_2\|) \leq \frac{1}{4}.$$

(ii) The right-hand side inclusion follows straightforwardly from (i). Let us focus on the left-hand side inclusion. For this, consider the map defined by $G = \pi_{T_xM} \circ \Psi_x$ on the domain $\mathcal{B}_{T_xM}(0,h)$. For all $v \in \mathcal{B}_{T_xM}(0,h)$, we have

$$\|d_vG - Id_{T_xM}\|_{op} = \|\pi_{T_xM} \circ d_v\mathbf{N}_x\|_{op} \leq \|d_v\mathbf{N}_x\|_{op} \leq L_\perp\|v\| \leq \frac{1}{4} < 1.$$

Hence, G is a diffeomorphism onto its image and it satisfies $\|G(v)\| \geq 3\|v\|/4$. It follows that

$$\mathcal{B}_{T_xM}\left(0, \frac{3h}{4}\right) \subset G\left(\mathcal{B}_{T_xM}(0,h)\right) = \pi_{T_xM}\left(\Psi_x\left(\mathcal{B}_{T_xM}(0,h)\right)\right).$$

Now, according to Lemma A.1, for all $y \in \mathcal{B}\left(x, \frac{3h}{5}\right) \cap M$,

$$\|\pi_{T_xM}(y-x)\| \leq \|y-x\| + \frac{\|y-x\|^2}{2\tau_{min}} \leq \left(1 + \frac{1}{4}\right)\|y-x\| \leq \frac{3h}{4},$$

from what we deduce $\pi_{T_xM}\left(\mathcal{B}\left(x,\frac{3h}{5}\right)\cap M\right)\subset\mathcal{B}_{T_xM}\left(0,\frac{3h}{4}\right)$. As a consequence,

$$\pi_{T_xM}\left(\mathcal{B}\left(x,\frac{3h}{5}\right)\cap M\right)\subset\pi_{T_xM}\left(\Psi_x\left(\mathcal{B}_{T_xM}\left(0,h\right)\right)\right),$$

which yields the announced inclusion since π_{T_xM} is one to one on $\mathcal{B}\left(x,\frac{5h}{4}\right)\cap M$ from Lemma 3 in [4], and

$$\left(\mathcal{B}\left(x,\frac{3h}{5}\right)\cap M\right)\subset\Psi_x\left(\mathcal{B}_{T_xM}\left(0,h\right)\right)\subset\mathcal{B}\left(x,\frac{5h}{4}\right)\cap M.$$

(iii) Straightforward application of Lemma 3 in [4].

(iv) Notice that Lemma A.1 gives the existence of such an expansion for $k=2$. Hence, we can assume $k\geq 3$. Taking $h=\frac{\tau_{min}\wedge L_\perp^{-1}}{4}$, we showed in the proof of (ii) that the map G is a diffeomorphism onto its image, with $\|d_vG-Id_{T_xM}\|_{op}\leq\frac{1}{4}<1$. Additionally, the chain rule yields $\left\|d_v^iG\right\|_{op}\leq\left\|d_v^i\Psi_x\right\|_{op}\leq L_i$ for all $2\leq i\leq k$. Therefore, from Lemma A.3, the differentials of G^{-1} up to order k are uniformly bounded. As a consequence, we get the announced expansion writing

$$y-x=\Psi_x\circ G^{-1}\left(\pi^*(y-x)\right),$$

and using the Taylor expansions of order k of Ψ_x and G^{-1}.
Let us now check that $T_2^*=II_x^M$. Since, by construction, T_2^* is the second order term of the Taylor expansion of $\Psi_x\circ G^{-1}$ at zero, a straightforward computation yields

$$\begin{aligned}T_2^*&=(I_D-\pi_{T_xM})\circ d_0^2\Psi_x\\&=\pi_{T_xM^\perp}\circ d_0^2\Psi_x.\end{aligned}$$

Let $v\in T_xM$ be fixed. Letting $\gamma(t)=\Psi_x(tv)$ for $|t|$ small enough, it is clear that $\gamma''(0)=d_0^2\Psi(v^{\otimes 2})$. Moreover, by definition of the second fundamental form [13, Proposition 2.1, p.127], since $\gamma(0)=x$ and $\gamma'(0)=v$, we have

$$II_x^M(v^{\otimes 2})=\pi_{T_xM^\perp}(\gamma''(0)).$$

Hence

$$\begin{aligned}T_2^*(v^{\otimes 2})&=\pi_{T_xM^\perp}\circ d_0^2\Psi_x(v^{\otimes 2})\\&=\pi_{T_xM^\perp}(\gamma''(0))\\&=II_x^M(v^{\otimes 2}),\end{aligned}$$

which concludes the proof.

(v) The first statement is a rephrasing of Proposition 6.1 in [28]. It yields the bound on sectional curvature, using the Gauss equation [13, Proposition 3.1 (a), p.135].

□

In the proof of Lemma A.2 (iv), we used a technical lemma of differential calculus that we now prove. It states quantitatively that if G is $\mathcal{C}^k$-close to the identity map, then it is a diffeomorphism onto its image and the differentials of its inverse G^{-1} are controlled.

LEMMA A.3. *Let $k \geq 2$ and U be an open subset of $\mathbb{R}^d$. Let $G : U \to \mathbb{R}^d$ be $\mathcal{C}^k$. Assume that $\|I_d - dG\|_{op} \leq \varepsilon < 1$, and that for all $2 \leq i \leq k$, $\|d^i G\|_{op} \leq L_i$ for some $L_i > 0$. Then G is a $\mathcal{C}^k$-diffeomorphism onto its image, and for all $2 \leq i \leq k$,*

$$\|I_d - dG^{-1}\|_{op} \leq \frac{\varepsilon}{1-\varepsilon} \quad \text{and} \quad \|d^i G^{-1}\|_{op} \leq L'_{i,\varepsilon,L_2,\dots,L_i} < \infty, \text{ for } 2 \leq i \leq k.$$

PROOF OF LEMMA A.3. For all $x \in U$, $\|d_x G - I_d\|_{op} < 1$, so G is one to one, and for all $y = G(x) \in G(U)$,

$$\begin{aligned} \|I_d - d_y G^{-1}\|_{op} &= \|I_d - (d_x G)^{-1}\|_{op} \\ &\leq \|(d_x G)^{-1}\|_{op} \|I_d - d_x G\|_{op} \\ &\leq \frac{\|I_d - d_x G\|_{op}}{1 - \|I_d - d_x G\|_{op}} \\ &\leq \frac{\varepsilon}{1-\varepsilon}. \end{aligned}$$

For $2 \leq i \leq k$ and $1 \leq j \leq i$, write $\Pi_i^{(j)}$ for the set of partitions of $\{1, \dots, i\}$ with j blocks. Differentiating i times the identity $G \circ G^{-1} = Id_{G(U)}$, Faa di Bruno's formula yields that, for all $y = G(x) \in G(U)$ and all unit vectors $h_1, \dots, h_i \in \mathbb{R}^D$,

$$0 = d_y \left(G \circ G^{-1}\right).(h_\alpha)_{1 \leq \alpha \leq i} = \sum_{j=1}^{i} \sum_{\pi \in \Pi_i^{(j)}} d_x^j G. \left(\left(d_y^{|I|} G^{-1}. (h_\alpha)_{\alpha \in I} \right)_{I \in \pi} \right).$$

Isolating the term for $j = 1$ entails

$$\begin{aligned}
&\left\| d_x G. \left(d_y^i G^{-1}. (h_\alpha)_{1 \le \alpha \le i} \right) \right\|_{op} \\
&= \left\| - \sum_{j=2}^{i} \sum_{\pi \in \Pi_i^{(j)}} d_x^j G. \left(\left(d_y^{|I|} G^{-1}. (h_\alpha)_{\alpha \in I} \right)_{I \in \pi} \right) \right\|_{op} \\
&\le \sum_{j=2}^{i} \sum_{\pi \in \Pi_i^{(j)}} \left\| d^j G \right\|_{op} \prod_{I \in \pi} \left\| d^{|I|} G^{-1} \right\|_{op}.
\end{aligned}$$

Using the first order Lipschitz bound on G^{-1}, we get

$$\left\| d^i G^{-1} \right\|_{op} \le \frac{1+\varepsilon}{1-\varepsilon} \sum_{j=2}^{i} L_j \sum_{\pi \in \Pi_i^{(j)}} \prod_{I \in \pi} \left\| d^{|I|} G^{-1} \right\|_{op}.$$

The result follows by induction on i. □

A.3. Proof of Proposition 1

This section is devoted to prove Proposition 1 (reproduced below as Proposition A.4), that asserts the stability of the model with respect to ambient diffeomorphisms.

PROPOSITION A.4. *Let $\Phi : \mathbb{R}^D \to \mathbb{R}^D$ be a global $\mathcal{C}^k$-diffeomorphism. If $\|d\Phi - I_D\|_{op}$, $\left\|d^2\Phi\right\|_{op}$, ..., $\left\|d^k\Phi\right\|_{op}$ are small enough, then for all P in $\mathcal{P}^k_{\tau_{min},\mathbf{L},f_{min},f_{max}}$, the pushforward distribution $P' = \Phi_* P$ belongs to $\mathcal{P}^k_{\tau_{min}/2,2\mathbf{L},f_{min}/2,2f_{max}}$.*

Moreover, if $\Phi = \lambda I_D$ ($\lambda > 0$) is an homogeneous dilation, then $P' \in \mathcal{P}^k_{\lambda\tau_{min},\mathbf{L}_{(\lambda)},f_{min}/\lambda^d,f_{max}/\lambda^d}$, where $\mathbf{L}_{(\lambda)} = (L_\perp/\lambda, L_3/\lambda^2, \ldots, L_k/\lambda^{k-1})$.

PROOF OF PROPOSITION A.4. The second part is straightforward since the dilation λM has reach $\tau_{\lambda M} = \lambda\tau_M$, and can be parametrized locally by $\tilde{\Psi}_{\lambda p}(v) = \lambda\Psi_p(v/\lambda) = \lambda p + v + \lambda \mathbf{N}_p(v/\lambda)$, yielding the differential bounds $\mathbf{L}_{(\lambda)}$. Bounds on the density follow from homogeneity of the d-dimensional Hausdorff measure.

The first part follows combining Proposition A.5 and Lemma A.6. □

Proposition A.5 asserts the stability of the geometric model, that is, the reach bound and the existence of a smooth parametrization when a submanifold is perturbed.

PROPOSITION A.5. *Let $\Phi : \mathbb{R}^D \to \mathbb{R}^D$ be a global $\mathcal{C}^k$-diffeomorphism. If $\|d\Phi - I_D\|_{op}$, $\left\|d^2\Phi\right\|_{op}$, ... , $\left\|d^k\Phi\right\|_{op}$ are small enough, then for all M in $\mathcal{C}^k_{\tau_{min},\mathbf{L}}$, the image $M' = \Phi(M)$ belongs to $\mathcal{C}^k_{\tau_{min}/2,2L_\perp,2L_3,\dots,2L_k}$.*

PROOF OF PROPOSITION A.5. To bound $\tau_{M'}$ from below, we use the stability of the reach with respect to $\mathcal{C}^2$ diffeomorphisms. Namely, from Theorem 4.19 in [16],

$$\begin{aligned}\tau_{M'} = \tau_{\Phi(M)} &\geq \frac{(1 - \|I_D - d\Phi\|_{op})^2}{\frac{1+\|I_D - d\Phi\|_{op}}{\tau_M} + \|d^2\Phi\|_{op}} \\ &\geq \tau_{min} \frac{(1 - \|I_D - d\Phi\|_{op})^2}{1 + \|I_D - d\Phi\|_{op} + \tau_{min}\|d^2\Phi\|_{op}} \geq \frac{\tau_{min}}{2}\end{aligned}$$

for $\|I_D - d\Phi\|_{op}$ and $\left\|d^2\Phi\right\|_{op}$ small enough. This shows the stability for $k = 2$, as well as that of the reach assumption for $k \geq 3$.

By now, take $k \geq 3$. We focus on the existence of a good parametrization of M' around a fixed point $p' = \Phi(p) \in M'$. For $v' \in T_{p'}M' = d_p\Phi(T_pM)$, let us define

$$\begin{aligned}\Psi'_{p'}(v') &= \Phi\left(\Psi_p\left(d_{p'}\Phi^{-1}.v'\right)\right) \\ &= p' + v' + \mathbf{N}'_{p'}(v'),\end{aligned}$$

where $\mathbf{N}'_{p'}(v') = \left\{\Phi\left(\Psi_p\left(d_{p'}\Phi^{-1}.v'\right)\right) - p' - v'\right\}$.

$$\begin{array}{ccc} M & \xrightarrow{\quad\Phi\quad} & M' \\ {\scriptstyle\Psi_p}\uparrow & & \uparrow{\scriptstyle\Psi'_{p'}} \\ T_pM & \xrightarrow[d_p\Phi]{} & T_{p'}M' \end{array}$$

The maps $\Psi'_{p'}(v')$ and $\mathbf{N}'_{p'}(v')$ are well defined whenever $\left\|d_{p'}\Phi^{-1}.v'\right\| \leq \frac{1}{4L_\perp}$, so in particular if $\|v'\| \leq \frac{1}{4(2L_\perp)} \leq \frac{1-\|I_D-d\Phi\|_{op}}{4L_\perp}$ and $\|I_D - d\Phi\|_{op} \leq \frac{1}{2}$. One easily checks that $\mathbf{N}'_{p'}(0) = 0$, $d_0\mathbf{N}'_{p'} = 0$ and writing $c(v') = p + d_{p'}\Phi^{-1}.v' + \mathbf{N}_{p'}\left(d_{p'}\Phi^{-1}.v'\right)$, for all unit vector $w' \in T_{p'}M'$,

$$
\begin{aligned}
\left\|d^2_{v'}\mathbf{N}'_{p'}(w'^{\otimes 2})\right\| &= \left\|d^2_{c(v')}\Phi\left(\left\{d_{d_{p'}\Phi^{-1}.v'}\Psi_p \circ d_{p'}\Phi^{-1}.w'\right\}^{\otimes 2}\right)\right. \\
&\quad \left. + d_{c(v')}\Phi \circ d^2_{d_{p'}\Phi^{-1}.v'}\Psi_p\left(\left\{d_{p'}\Phi^{-1}.w'\right\}^{\otimes 2}\right)\right\| \\
&= \left\|d^2_{c(v')}\Phi\left(\left\{d_{d_{p'}\Phi^{-1}.v'}\Psi_p \circ d_{p'}\Phi^{-1}.w'\right\}^{\otimes 2}\right)\right. \\
&\quad + \left(d_{c(v')}\Phi - Id\right) \circ d^2_{d_{p'}\Phi^{-1}.v'}\Psi_p\left(\left\{d_{p'}\Phi^{-1}.w'\right\}^{\otimes 2}\right) \\
&\quad \left. + d^2_{d_{p'}\Phi^{-1}.v'}\Psi_p\left(\left\{d_{p'}\Phi^{-1}.w'\right\}^{\otimes 2}\right)\right\| \\
&\leq \left\|d^2\Phi\right\|_{op}\left(1 + L_\perp\left\|d_{p'}\Phi^{-1}.v'\right\|\right)^2\left\|d_{p'}\Phi^{-1}.w'\right\|^2 \\
&\quad + \left\|I_D - d\Phi\right\|_{op} L_\perp\left\|d_{p'}\Phi^{-1}.w'\right\|^2 \\
&\quad + L_\perp\left\|d_{p'}\Phi^{-1}.w'\right\|^2 \\
&\leq \left\|d^2\Phi\right\|_{op}(1+1/4)^2\left\|d_{p'}\Phi^{-1}\right\|^2_{op} \\
&\quad + \left\|I_D - d\Phi\right\|_{op} L_\perp\left\|d\Phi^{-1}\right\|^2_{op} \\
&\quad + L_\perp\left\|d_{p'}\Phi^{-1}\right\|^2_{op}.
\end{aligned}
$$

Writing further $\left\|d\Phi^{-1}\right\|_{op} \leq (1 - \left\|I_D - d\Phi\right\|_{op})^{-1} \leq 1 + 2\left\|I_D - \Phi\right\|_{op}$ for $\left\|I_D - d\Phi\right\|_{op}$ small enough depending only on $L_\perp$, it is clear that the right-hand side of the latter inequality goes below $2L_\perp$ for $\left\|I_D - d\Phi\right\|_{op}$ and $\left\|d^2\Phi\right\|_{op}$ small enough. Hence, for $\left\|I_D - d\Phi\right\|_{op}$ and $\left\|d^2\Phi\right\|_{op}$ small enough depending only on $L_\perp$, $\|d^2_{v'}\mathbf{N}'_{p'}\|_{op} \leq 2L_\perp$ for all $\|v'\| \leq \frac{1}{4(2L_\perp)}$. From the chain rule, the same argument applies for the order $3 \leq i \leq k$ differential of $\mathbf{N}'_{p'}$. □

Lemma A.6 deals with the condition on the density in the models $\mathcal{P}^k$. It gives a change of variable formula for pushforward of measure on submanifolds, ensuring a control on densities with respect to intrinsic volume measure.

LEMMA A.6 (Change of variable for the Hausdorff measure). *Let P be a probability distribution on $M \subset \mathbb{R}^D$ with density f with respect to the d-dimensional Hausdorff measure $\mathcal{H}^d$. Let $\Phi : \mathbb{R}^D \to \mathbb{R}^D$ be a global diffeomorphism such that $\left\|I_D - d\Phi\right\|_{\mathrm{op}} < 1/3$. Let $P' = \Phi_* P$ be the pushforward of P by Φ. Then P' has a density g with respect to $\mathcal{H}^d$. This density can be*

chosen to be, for all $z \in \Phi(M)$,

$$g(z) = \frac{f\left(\Phi^{-1}(z)\right)}{\sqrt{\det\left(\pi_{T_{\Phi^{-1}(z)}M} \circ d_{\Phi^{-1}(z)}\Phi^T \circ d_{\Phi^{-1}(z)}\Phi\big|_{T_{\Phi^{-1}(z)}M}\right)}}.$$

In particular, if $f_{min} \le f \le f_{max}$ *on* M, *then for all* $z \in \Phi(M)$,

$$\left(1 - 3d/2\,\|I_D - d\Phi\|_{\mathrm{op}}\right) f_{min} \le g(z) \le f_{max}\left(1 + 3(2^{d/2}-1)\,\|I_D - d\Phi\|_{\mathrm{op}}\right).$$

Proof of Lemma A.6. Let $p \in M$ be fixed and $A \subset \mathcal{B}(p,r) \cap M$ for r small enough. For a differentiable map $h : \mathbb{R}^d \to \mathbb{R}^D$ and for all $x \in \mathbb{R}^d$, we let $J_h(x)$ denote the d-dimensional Jacobian $J_h(x) = \sqrt{\det\left(d_x h^T\, d_x h\right)}$. The area formula ([17, Theorem 3.2.5]) states that if h is one-to-one,

$$\int_A u\left(h(x)\right) J_h(x) \lambda^d(dx) = \int_{h(A)} u(y) \mathcal{H}^d(dy),$$

whenever $u : \mathbb{R}^D \to \mathbb{R}$ is Borel, where λ^d is the Lebesgue measure on $\mathbb{R}^d$. By definition of the pushforward, and since $dP = f d\mathcal{H}^d$,

$$\int_{\Phi(A)} dP'(z) = \int_A f(y) \mathcal{H}^d(dy).$$

Writing $\Psi_p = \exp_p : T_pM \to \mathbb{R}^D$ for the exponential map of M at p, we have

$$\int_A f(y) \mathcal{H}^d(dy) = \int_{{\Psi_p}^{-1}(A)} f(\Psi_p(x)) J_{\Psi_p}(x) \lambda^d(dx).$$

Rewriting the right hand term, we apply the area formula again with $h = \Phi \circ \Psi_p$,

$$\begin{aligned}
&\int_{{\Psi_p}^{-1}(A)} f(\Psi_p(x)) J_{\Psi_p}(x) \lambda^d(dy) \\
&\qquad = \int_{{\Psi_p}^{-1}(A)} f\left(\Phi^{-1}\left(h(x)\right)\right) \frac{J_{\Psi_p}(h^{-1}\left(h(x)\right))}{J_{\Phi\circ\Psi_p}(h^{-1}\left(h(x)\right))} J_{\Phi\circ\Psi_p}(x) \lambda^d(dx) \\
&\qquad = \int_{\Phi(A)} f\left(\Phi^{-1}(z)\right) \frac{J_{\Psi_p}(h^{-1}(z))}{J_{\Phi\circ\Psi_p}(h^{-1}(z))} \mathcal{H}^d(dz).
\end{aligned}$$

Since this is true for all $A \subset \mathcal{B}(p,r) \cap M$, P' has a density g with respect to $\mathcal{H}^d$, with

$$g(z) = f\left(\Phi^{-1}(z)\right) \frac{J_{\Psi_{\Phi^{-1}(z)}}(\Psi^{-1}_{\Phi^{-1}(z)} \circ \Phi^{-1}(z))}{J_{\Phi\circ\Psi_{\Phi^{-1}(z)}}(\Psi^{-1}_{\Phi^{-1}(z)} \circ \Phi^{-1}(z))}.$$

Writing $p = \Phi^{-1}(z)$, it is clear that $\Psi^{-1}_{\Phi^{-1}(z)} \circ \Phi^{-1}(z) = \Psi_p^{-1}(p) = 0 \in T_pM$. Since $d_0 \exp_p : T_pM \to \mathbb{R}^D$ is the inclusion map, we get the first statement.

We now let B and π_T denote $d_p\Phi$ and π_{T_pM} respectively. For any unit vector $v \in T_pM$,

$$\begin{aligned}
\left| \left\| \pi_T B^T B v \right\| - \|v\| \right| &\leq \left\| \pi_T \left(B^T B - I_D \right) v \right\| \\
&\leq \left\| B^T B - I_D \right\|_{\text{op}} \\
&\leq \left(2 + \|I_D - B\|_{\text{op}} \right) \|I_D - B\|_{\text{op}} \\
&\leq 3 \|I_D - B\|_{\text{op}} .
\end{aligned}$$

Therefore, $1 - 3\|I_D - B\|_{\text{op}} \leq \left\| \pi_T B^T B_{|T_pM} \right\|_{\text{op}} \leq 1 + 3\|I_D - B\|_{\text{op}}$. Hence,

$$\sqrt{\det\left(\pi_T B^T B_{|T_pM} \right)} \leq \left(1 + 3\|I_D - B\|_{\text{op}} \right)^{d/2} \leq \frac{1}{1 - \frac{3d}{2}\|I_D - B\|_{\text{op}}},$$

and

$$\sqrt{\det\left(\pi_T B^T B_{|T_pM} \right)} \geq \left(1 - 3\|I_D - B\|_{\text{op}} \right)^{d/2} \geq \frac{1}{1 + 3(2^{d/2} - 1)\|I_D - B\|_{\text{op}}},$$

which yields the result. □

Appendix B: Some Probabilistic Tools

B.1. Volume and Covering Rate

The first lemma of this section gives some details about the covering rate of a manifold with bounded reach.

LEMMA B.7. *Let $P_0 \in \mathcal{P}^k$ have support $M \subset \mathbb{R}^D$. Then for all $r \leq \tau_{min}/4$ and x in M,*

$$c_d f_{min} r^d \leq p_x(r) \leq C_d f_{max} r^d,$$

for some $c_d, C_d > 0$, with $p_x(r) = P_0\big(\mathcal{B}(x,r)\big)$.

Moreover, letting $h = \left(\frac{C'_d k}{f_{min}} \frac{\log n}{n}\right)^{1/d}$ with C'_d large enough, the following holds. For n large enough so that $h \leq \tau_{min}/4$, with probability at least $1 - \left(\frac{1}{n}\right)^{k/d}$,

$$d_H\left(M, \mathbb{Y}_n\right) \leq h/2.$$

PROOF OF LEMMA B.7. Denoting by $\mathcal{B}_M(x,r)$ the geodesic ball of radius r centered at x, Proposition 25 of [1] yields

$$\mathcal{B}_M(x,r) \subset \mathcal{B}(x,r) \cap M \subset \mathcal{B}_M(x, 6r/5).$$

Hence, the bounds on the Jacobian of the exponential map given by Proposition 27 of [1] yield

$$c_d r^d \leq Vol\big(\mathcal{B}(x,r) \cap M\big) \leq C_d r^d,$$

for some $c_d, C_d > 0$. Now, since P has a density $f_{min} \leq f \leq f_{max}$ with respect to the volume measure of M, we get the first result.

Now we notice that since $p_x(r) \geq c_d f_{min} r^d$, Theorem 3.3 in [10] entails, for $s \leq \tau_{min}/8$,

$$\mathbb{P}\left(d_H\big(M, \mathbb{X}_n\big) \geq s\right) \leq \frac{4^d}{c_d f_{min} s^d} \exp\left(-\frac{c_d f_{min}}{2^d} n s^d\right).$$

Hence, taking $s = h/2$, and $h = \left(\frac{C'_d k}{f_{min}} \frac{\log n}{n}\right)^{1/d}$ with C'_d so that $C'_d \geq \frac{8^d}{c_d k} \vee \frac{2^d(1+k/d)}{c_d k}$ yields the result. Since $k \geq 1$, taking $C'_d = \frac{8^d}{c_d}$ is sufficient. □

B.2. Concentration Bounds for Local Polynomials

This section is devoted to the proof of the following proposition.

PROPOSITION B.8. *Set* $h = \left(K\frac{\log n}{n-1}\right)^{\frac{1}{d}}$. *There exist constants* $\kappa_{k,d}$, $c_{k,d}$ *and* C_d *such that, if* $K \geq (\kappa_{k,d} f_{max}^2/f_{min}^3)$ *and* n *is large enough so that* $3h/2 \leq h_0 \leq \tau_{min}/4$, *then with probability at least* $1 - \left(\frac{1}{n}\right)^{\frac{k}{d}+1}$, *we have*

$$\begin{array}{rcl} P_{0,n-1}[S^2(\pi^*(x))\mathbb{1}_{\mathcal{B}(h/2)}(x)] & \geq & c_{k,d} h^d f_{min} \|S_h\|_2^2, \\ N(3h/2) & \leq & C_d f_{max}(n-1)h^d, \end{array}$$

for every $S \in \mathbb{R}^k[x_{1:d}]$, *where* $N(h) = \sum_{j=2}^n \mathbb{1}_{\mathcal{B}(0,h)}(Y_j)$.

A first step is to ensure that empirical expectations of order k polynomials are close to their deterministic counterparts.

PROPOSITION B.9. *Let* $b \leq \tau_{min}/8$. *For any* $y_0 \in M$, *we have*

$$\mathbb{P}\left[\sup_{u_1,\dots,u_k,\varepsilon\in\{0,1\}^k} \left|(P_0 - P_{0,n-1})\prod_{j=1}^p \left(\frac{\langle u_j, y\rangle}{b}\right)^{\varepsilon_j} \mathbb{1}_{\mathcal{B}(y_0,b)}(y)\right| \right.$$
$$\left. \geq p_{y_0}(b)\left(\frac{4k\sqrt{2\pi}}{\sqrt{(n-1)p_{y_0}(b)}} + \sqrt{\frac{2t}{(n-1)p_{y_0}(b)}} + \frac{2}{3(n-1)p_{y_0}(b)}\right)\right] \leq e^{-t},$$

where $P_{0,n-1}$ *denotes the empirical distribution of* $n-1$ *i.i.d. random variables* Y_i *drawn from* P_0.

PROOF OF PROPOSITION B.9. Without loss of generality we choose $y_0 = 0$ and shorten notation to $\mathcal{B}(b)$ and $p(b)$. Let $\mathcal{Z}$ denote the empirical process on the left-hand side of Proposition B.9. Denote also by $f_{u,\varepsilon}$ the map $\prod_{j=1}^k \left(\frac{\langle u_j, y\rangle}{b}\right)^{\varepsilon_j} \mathbb{1}_{\mathcal{B}(b)}(y)$, and let $\mathcal{F}$ denote the set of such maps, for u_j in $\mathcal{B}(1)$ and ε in $\{0,1\}^k$.

Since $\|f_{u,\varepsilon}\|_\infty \leq 1$ and $Pf_{u,\varepsilon}^2 \leq p(b)$, the Talagrand-Bousquet inequality ([8, Theorem 2.3]) yields

$$\mathcal{Z} \leq 4\mathbb{E}\mathcal{Z} + \sqrt{\frac{2p(b)t}{n-1}} + \frac{2t}{3(n-1)},$$

with probability larger than $1 - e^{-t}$. It remains to bound $\mathbb{E}\mathcal{Z}$ from above.

LEMMA B.10. *We may write*

$$\mathbb{E}\mathcal{Z} \leq \frac{\sqrt{2\pi p(b)}}{\sqrt{n-1}} k.$$

PROOF OF LEMMA B.10. Let σ_i and g_i denote some independent Rademacher and Gaussian variables. For convenience, we denote by $\mathbb{E}_A$ the expectation with respect to the random variable A. Using symmetrization inequalities we may write

$$\begin{aligned}
\mathbb{E}\mathcal{Z} &= \mathbb{E}_Y \sup_{u,\varepsilon} \left| (P_0 - P_{0,n-1}) \prod_{j=1}^{k} \left(\frac{\langle u_j, y \rangle}{b} \right)^{\varepsilon_j} \mathbb{1}_{\mathcal{B}(b)}(y) \right| \\
&\leq \frac{2}{n-1} \mathbb{E}_Y \mathbb{E}_\sigma \sup_{u,\varepsilon} \sum_{i=1}^{n-1} \sigma_i \prod_{j=1}^{k} \left(\frac{\langle u_j, Y_i \rangle}{b} \right)^{\varepsilon_j} \mathbb{1}_{\mathcal{B}(b)}(Y_i) \\
&\leq \frac{\sqrt{2\pi}}{n-1} \mathbb{E}_Y \mathbb{E}_g \sup_{u,\varepsilon} \sum_{i=1}^{n-1} g_i \prod_{j=1}^{k} \left(\frac{\langle u_j, Y_i \rangle}{b} \right)^{\varepsilon_j} \mathbb{1}_{\mathcal{B}(b)}(Y_i).
\end{aligned}$$

Now let $\mathcal{Y}_{u,\varepsilon}$ denote the Gaussian process $\sum_{i=1}^{n-1} g_i \prod_{j=1}^{k} \left(\frac{\langle u_j, Y_i \rangle}{b} \right)^{\varepsilon_j} \mathbb{1}_{\mathcal{B}(b)}(Y_i)$. Since, for any y in $\mathcal{B}(b)$, u,v in $\mathcal{B}(1)^k$, and ε, ε' in $\{0,1\}^k$, we have

$$\begin{aligned}
&\left| \prod_{j=1}^{k} \left(\frac{\langle y, u_j \rangle}{b} \right)^{\varepsilon_j} - \prod_{j=1}^{k} \left(\frac{\langle y, v_j \rangle}{b} \right)^{\varepsilon'_j} \right| \\
&\quad \leq \left| \sum_{r=1}^{k} \left(\prod_{j=1}^{k+1-r} \left(\frac{\langle y, u_j \rangle}{b} \right)^{\varepsilon_j} \prod_{j=k+2-r}^{k} \left(\frac{\langle y, v_j \rangle}{b} \right)^{\varepsilon'_j} \right. \right. \\
&\qquad \left. \left. - \prod_{j=1}^{k-r} \left(\frac{\langle y, u_j \rangle}{b} \right)^{\varepsilon_j} \prod_{j=k+1-r}^{k} \left(\frac{\langle y, v_j \rangle}{b} \right)^{\varepsilon'_j} \right) \right| \\
&\quad \leq \sum_{r=1}^{k} \left| \prod_{j=1}^{k-r} \left(\frac{\langle y, u_j \rangle}{b} \right)^{\varepsilon_j} \prod_{j=k+2-r}^{k} \left(\frac{\langle y, v_j \rangle}{b} \right)^{\varepsilon'_j} \left[\left(\frac{\langle u_{k+1-r}, y \rangle}{b} \right)^{\varepsilon_{k+1-r}} \right. \right. \\
&\qquad \left. \left. - \left(\frac{\langle v_{k+1-r}, y \rangle}{b} \right)^{\varepsilon'_{k+1-r}} \right] \right| \\
&\quad \leq \sum_{r=1}^{k} \left| \frac{\langle \varepsilon_r u_r - \varepsilon'_r v_r, y \rangle}{b} \right|,
\end{aligned}$$

we deduce that

$$\begin{aligned}\mathbb{E}_g(\mathcal{Y}_{u,\varepsilon}-\mathcal{Y}_{v,\varepsilon'})^2 &\leq k\sum_{i=1}^{n-1}\sum_{r=1}^{k}\left(\frac{\langle\varepsilon_r u_r,Y_i\rangle}{b}-\frac{\langle\varepsilon'_r v_r,Y_i\rangle}{b}\right)^2\mathbb{1}_{\mathcal{B}(b)}(Y_i)\\ &\leq \mathbb{E}_g(\Theta_{u,\varepsilon}-\Theta_{v,\varepsilon'})^2,\end{aligned}$$

where $\Theta_{u,\varepsilon}=\sqrt{k}\sum_{i=1}^{n-1}\sum_{r=1}^{k} g_{i,r}\frac{\langle\varepsilon_r u_r,Y_i\rangle}{b}\mathbb{1}_{\mathcal{B}(b)}(Y_i)$. According to Slepian's Lemma [7, Theorem 13.3], it follows that

$$\begin{aligned}\mathbb{E}_g\sup_{u,\varepsilon}\mathcal{Y}_g &\leq \mathbb{E}_g\sup_{u,\varepsilon}\Theta_{u,\varepsilon}\\ &\leq \sqrt{k}\mathbb{E}_g\sup_{u,\varepsilon}\sum_{r=1}^{k}\frac{\left\langle\varepsilon_r u_r,\sum_{i=1}^{n-1} g_{i,r}\mathbb{1}_{\mathcal{B}(b)}(Y_i)Y_i\right\rangle}{b}\\ &\leq \sqrt{k}\mathbb{E}_g\sup_{u,\varepsilon}\sqrt{k\sum_{r=1}^{k}\frac{\left\langle\varepsilon_r u_r,\sum_{i=1}^{n-1} g_{i,r}\mathbb{1}_{\mathcal{B}(b)}(Y_i)Y_i\right\rangle^2}{b^2}}.\end{aligned}$$

We deduce that

$$\begin{aligned}\mathbb{E}_g\sup_{u,\varepsilon}Y_g &\leq \mathbb{E}_g\sup_{u,\varepsilon}\Theta_g\\ &\leq k\sqrt{\mathbb{E}_g\sup_{\|u\|=1,\varepsilon\in\{0,1\}}\frac{\left\langle\varepsilon u,\sum_{i=1}^{n-1} g_i\mathbb{1}_{\mathcal{B}(b)}(Y_i)Y_i\right\rangle^2}{b^2}}\\ &\leq k\sqrt{\mathbb{E}_g\left\|\sum_{i=1}^{n-1}\frac{g_i Y_i}{b}\mathbb{1}_{\mathcal{B}(b)}(Y_i)\right\|^2}\\ &\leq k\sqrt{N(b)}.\end{aligned}$$

Then we can deduce that $\mathbb{E}_X\mathbb{E}_g\sup_{u,\varepsilon}Y_g\leq k\sqrt{p(b)}$. □

Combining Lemma B.10 with Talagrand-Bousquet's inequality gives the result of Proposition B.9. □

We are now in position to prove Proposition B.8.

Proof of Proposition B.8. If $h/2\leq\tau_{min}/4$, then, according to Lemma B.7, $p(h/2)\geq c_d f_{min}h^d$, hence, if $h=\left(K\frac{\log(n)}{n-1}\right)^{\frac{1}{d}}$, $(n-1)p(h/2)\geq Kc_d f_{min}\log(n)$.

Choosing $b = h/2$ and $t = (k/d+1)\log(n) + \log(2)$ in Proposition B.9 and $K = K'/f_{min}$, with $K' > 1$ leads to

$$\mathbb{P}\left[\sup_{u_1,\dots,u_k,\varepsilon\in\{0,1\}^k}\left|(P_0 - P_{0,n-1})\prod_{j=1}^{k}\left(2\frac{\langle u_j, y\rangle}{h}\right)^{\varepsilon_j}\mathbb{1}_{\mathcal{B}(y_0,h/2)}(y)\right| \geq \frac{c_{d,k}f_{max}}{\sqrt{K'}}h^d\right] \leq \frac{1}{2}\left(\frac{1}{n}\right)^{\frac{k}{d}+1}.$$

On the complement of the probability event mentioned just above, for a polynomial $S = \sum_{\alpha\in[0,k]^d||\alpha|\leq k} a_\alpha y_{1:d}^\alpha$, we have

$$\begin{aligned}(P_{0,n-1} - P_0)S^2(y_{1:d})\mathbb{1}_{\mathcal{B}(h/2)}(y) &\geq -\sum_{\alpha,\beta}\frac{c_{d,k}f_{max}}{\sqrt{K'}}|a_\alpha a_\beta|h^{d+|\alpha|+|\beta|}\\ &\geq -\frac{c_{d,k}f_{max}}{\sqrt{K'}}h^d\|S_h\|_2^2.\end{aligned}$$

On the other hand, we may write, for all $r > 0$,

$$\int_{\mathcal{B}(0,r)} S^2(y_{1:d})dy_1\dots dy_d \geq C_{d,k}r^d\|S_r\|_2^2,$$

for some constant $C_{d,k}$. It follows that

$$P_0S^2(y_{1:d})\mathbb{1}_{\mathcal{B}(h/2)}(y) \geq P_0S^2(y_{1:d})\mathbb{1}_{B(7h/16)}(y_{1:d}) \geq c_{k,d}h^d f_{min}\|S_h\|_2^2,$$

according to Lemma A.2. Then we may choose $K' = \kappa_{k,d}(f_{max}/f_{min})^2$, with $\kappa_{k,d}$ large enough so that

$$P_{0,n-1}S^2(x_{1:d})\mathbb{1}_{\mathcal{B}(h/2)}(y) \geq c_{k,d}f_{min}h^d\|S_h\|_2^2.$$

The second inequality of Proposition B.8 is derived the same way from Proposition B.9, choosing $\varepsilon = (0,\dots,0)$, $b = 3h/2$ and $h \leq \tau_{min}/8$ so that $b \leq \tau_{min}/4$. □

Appendix C: Minimax Lower Bounds

C.1. Conditional Assouad's Lemma

This section is dedicated to the proof of Lemma 7, reproduced below as Lemma C.11.

LEMMA C.11 (Conditional Assouad). *Let $m \geq 1$ be an integer and let $\{\mathcal{Q}_\tau\}_{\tau\in\{0,1\}^m}$ be a family of 2^m submodels $\mathcal{Q}_\tau \subset \mathcal{Q}$. Let $\{U_k \times U_k'\}_{1\leq k\leq m}$ be a family of pairwise disjoint subsets of $\mathcal{X}\times\mathcal{X}'$, and $\mathcal{D}_{\tau,k}$ be subsets of $\mathcal{D}$. Assume that for all $\tau\in\{0,1\}^m$ and $1\leq k\leq m$,*

- *for all $Q_\tau \in \mathcal{Q}_\tau$, $\theta_X(Q_\tau)\in\mathcal{D}_{\tau,k}$ on the event $\{X\in U_k\}$;*
- *for all $\theta\in\mathcal{D}_{\tau,k}$ and $\theta'\in\mathcal{D}_{\tau^k,k}$, $d(\theta,\theta')\geq\Delta$.*

For all $\tau\in\{0,1\}^m$, let $\overline{Q}_\tau \in \overline{Conv}(\mathcal{Q}_\tau)$, and write $\bar{\mu}_\tau$ and $\bar{\nu}_\tau$ for the marginal distributions of $\overline{Q}_\tau$ on $\mathcal{X}$ and $\mathcal{X}'$ respectively. Assume that if (X,X') has distribution $\overline{Q}_\tau$, X and X' are independent conditionally on the event $\{(X,X')\in U_k\times U_k'\}$, and that

$$\min_{\substack{\tau\in\{0,1\}^m\\1\leq k\leq m}}\left\{\left(\int_{U_k} d\bar{\mu}_\tau\wedge d\bar{\mu}_{\tau^k}\right)\left(\int_{U_k'} d\bar{\nu}_\tau\wedge d\bar{\nu}_{\tau^k}\right)\right\}\geq 1-\alpha.$$

Then,

$$\inf_{\hat{\theta}}\sup_{Q\in\mathcal{Q}}\mathbb{E}_Q\left[d\big(\theta_X(Q),\hat{\theta}(X,X')\big)\right]\geq m\frac{\Delta}{2}(1-\alpha),$$

where the infimum is taken over all the estimators $\hat{\theta}:\mathcal{X}\times\mathcal{X}'\to\mathcal{D}$.

PROOF OF LEMMA C.11. The proof follows that of Lemma 2 in [36]. Let $\hat{\theta}=\hat{\theta}(X,X')$ be fixed. For any family of 2^m distributions $\{Q_\tau\}_\tau\in\{\mathcal{Q}_\tau\}_\tau$,

since the $U_k \times U_k'$'s are pairwise disjoint,

$$\begin{aligned}
&\sup_{Q\in\mathcal{Q}} \mathbb{E}_Q\left[d\big(\theta_X(Q),\hat\theta(X,X')\big)\right] \\
&\qquad \geq \max_\tau \mathbb{E}_{Q_\tau} d(\hat\theta,\theta_X(Q_\tau)) \\
&\qquad \geq \max_\tau \mathbb{E}_{Q_\tau} \sum_{k=1}^m d\big(\hat\theta,\theta_X(Q_\tau)\big)\mathbb{1}_{U_k\times U_k'}(X,X') \\
&\qquad \geq 2^{-m}\sum_\tau \sum_{k=1}^m \mathbb{E}_{Q_\tau} d\big(\hat\theta,\theta_X(Q_\tau)\big)\mathbb{1}_{U_k\times U_k'}(X,X') \\
&\qquad \geq 2^{-m}\sum_\tau \sum_{k=1}^m \mathbb{E}_{Q_\tau} d\big(\hat\theta,\mathcal{D}_{\tau,k}\big)\mathbb{1}_{U_k\times U_k'}(X,X') \\
&\qquad = \sum_{k=1}^m 2^{-(m+1)} \sum_\tau \Big(\mathbb{E}_{Q_\tau} d\big(\hat\theta,\mathcal{D}_{\tau,k}\big)\mathbb{1}_{U_k\times U_k'}(X,X') \\
&\qquad\qquad\qquad + \mathbb{E}_{Q_{\tau^k}} d\big(\hat\theta,\mathcal{D}_{\tau^k,k}\big)\mathbb{1}_{U_k\times U_k'}(X,X')\Big).
\end{aligned}$$

Since the previous inequality holds for all $Q_\tau \in \mathcal{Q}_\tau$, it extends to $\overline{Q}_\tau \in \overline{Conv}(\mathcal{Q}_\tau)$ by linearity. Let us now lower bound each of the terms of the sum for fixed $\tau \in \{0,1\}^m$ and $1 \leq k \leq m$. By assumption, if (X,X') has distribution $\overline{Q}_\tau$, then conditionally on $\{(X,X') \in U_k \times U_k'\}$, X and X' are independent. Therefore,

$$\begin{aligned}
&\mathbb{E}_{\overline{Q}_\tau} d\big(\hat\theta,\mathcal{D}_{\tau,k}\big)\mathbb{1}_{U_k\times U_k'}(X,X') + \mathbb{E}_{\overline{Q}_{\tau^k}} d\big(\hat\theta,\mathcal{D}_{\tau^k,k}\big)\mathbb{1}_{U_k\times U_k'}(X,X') \\
&\quad \geq \mathbb{E}_{\overline{Q}_\tau} d\big(\hat\theta,\mathcal{D}_{\tau,k}\big)\mathbb{1}_{U_k}(X)\mathbb{1}_{U_k'}(X') + \mathbb{E}_{\overline{Q}_{\tau^k}} d\big(\hat\theta,\mathcal{D}_{\tau^k,k}\big)\mathbb{1}_{U_k}(X)\mathbb{1}_{U_k'}(X') \\
&\quad = \mathbb{E}_{\bar\nu_\tau}\left[\mathbb{E}_{\bar\mu_\tau}\left(d\big(\hat\theta,\mathcal{D}_{\tau,k}\big)\mathbb{1}_{U_k}(X)\right)\mathbb{1}_{U_k'}(X')\right] \\
&\qquad + \mathbb{E}_{\bar\nu_{\tau^k}}\left[\mathbb{E}_{\bar\mu_{\tau^k}}\left(d\big(\hat\theta,\mathcal{D}_{\tau^k,k}\big)\mathbb{1}_{U_k}(X)\right)\mathbb{1}_{U_k'}(X')\right] \\
&\quad = \int_{U_k}\int_{U_k'} d(\hat\theta,\mathcal{D}_{\tau,k})d\bar\mu_\tau(x)d\bar\nu_\tau(x') + \int_{U_k}\int_{U_k'} d(\hat\theta,\mathcal{D}_{\tau^k,k})d\bar\mu_{\tau^k}(x)d\bar\nu_{\tau^k}(x') \\
&\quad \geq \int_{U_k}\int_{U_k'} \Big(d(\hat\theta,\mathcal{D}_{\tau,k}) + d(\hat\theta,\mathcal{D}_{\tau^k,k})\Big)\, d\bar\mu_\tau \wedge d\bar\mu_{\tau^k}(x)d\bar\nu_\tau \wedge d\bar\nu_{\tau^k}(x') \\
&\quad \geq \Delta\left(\int_{U_k} d\bar\mu_\tau \wedge d\bar\mu_{\tau^k}\right)\left(\int_{U_k'} d\bar\nu_\tau \wedge d\bar\nu_{\tau^k}\right) \\
&\quad \geq \Delta(1-\alpha),
\end{aligned}$$

where we used that $d(\hat\theta,\mathcal{D}_{\tau,k}) + d(\hat\theta,\mathcal{D}_{\tau^k,k}) \geq \Delta$. The result follows by summing the above bound $|\{1,\dots,m\}\times\{0,1\}^m| = m2^m$ times. □

C.2. Construction of Generic Hypotheses

Let $M_0^{(0)}$ be a d-dimensional $\mathcal{C}^\infty$-submanifold of $\mathbb{R}^D$ with reach greater than 1 and such that it contains $\mathcal{B}_{\mathbb{R}^d\times\{0\}^{D-d}}(0,1/2)$. $M_0^{(0)}$ can be built for example by flattening smoothly a unit d-sphere in $\mathbb{R}^{d+1}\times\{0\}^{D-d-1}$. Since $M_0^{(0)}$ is $\mathcal{C}^\infty$, the uniform probability distribution $P_0^{(0)}$ on $M_0^{(0)}$ belongs to $\mathcal{P}^k_{1,\mathbf{L}^{(0)},1/V_0^{(0)},1/V_0^{(0)}}$, for some $\mathbf{L}^{(0)}$ and $V_0^{(0)} = Vol(M_0^{(0)})$.

Let now $M_0 = (2\tau_{min})M_0^{(0)}$ be the submanifold obtained from $M_0^{(0)}$ by homothecy. By construction, and from Proposition A.4, we have

$$\tau_{M_0} \geq 2\tau_{min}, \qquad \mathcal{B}_{\mathbb{R}^d\times\{0\}^{D-d}}(0,\tau_{min}) \subset M_0, \qquad Vol(M_0) = C_d\tau_{min}^d,$$

and the uniform probability distribution P_0 on M_0 satisfies

$$P_0 \in \mathcal{P}^k_{2\tau_{min},\mathbf{L}/2,2f_{min},f_{max}/2},$$

whenever $L_\perp/2 \geq L_\perp^{(0)}/(2\tau_{min})$, ..., $L_k/2 \geq L_k^{(0)}/(2\tau_{min})^{k-1}$, and provided that $2f_{min} \leq \left((2\tau_{min})^d V_0^{(0)}\right)^{-1} \leq f_{max}/2$. Note that $L_\perp^{(0)}, \ldots, L_k^{(0)}$, $Vol(M_0^{(0)})$ depend only on d and k. For this reason, all the lower bounds will be valid for $\tau_{min}L_\perp, \ldots, \tau_{min}^{k-1}L_k, (\tau_{min}^d f_{min})^{-1}$ and $\tau_{min}^d f_{max}$ large enough to exceed the thresholds $L_\perp^{(0)}/2, \ldots, L_k^{(0)}/2^{k-1}$, $2^d V_0^{(0)}$ and $(2^d V_0^{(0)})^{-1}$ respectively.

For $0 < \delta \leq \tau_{min}/4$, let $x_1, \ldots, x_m \in M_0 \cap \mathcal{B}(0,\tau_{min}/4)$ be a family of points such that

$$\text{for} \quad 1 \leq k \neq k' \leq m, \quad \|x_k - x_{k'}\| \geq \delta.$$

For instance, considering the family $\left\{\left(l_1\delta, \ldots, l_d\delta, 0, \ldots, 0\right)\right\}_{l_i\in\mathbb{Z},|l_i|\leq\lfloor\tau_{min}/(4\delta)\rfloor}$,

$$m \geq c_d\left(\frac{\tau_{min}}{\delta}\right)^d,$$

for some $c_d > 0$.

We let $e \in \mathbb{R}^D$ denote the $(d+1)$th vector of the canonical basis. In particular, we have the orthogonal decomposition of the ambient space

$$\mathbb{R}^D = \left(\mathbb{R}^d \times \{0\}^{D-d}\right) + span(e) + \left(\{0\}^{d+1} \times \mathbb{R}^{D-d-1}\right).$$

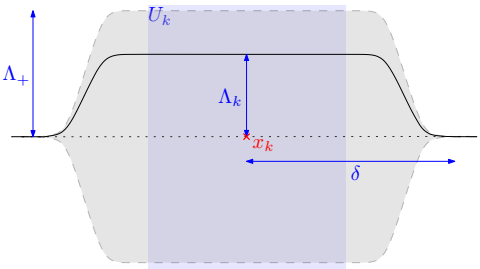
Let $\phi : \mathbb{R}^D \to [0,1]$ be a smooth scalar map such that $\phi|_{\mathcal{B}\left(0,\frac{1}{2}\right)} = 1$ and $\phi|_{\mathcal{B}(0,1)^c} = 0$.

Let $\Lambda_+ > 0$ and $1 \geq A_+ > A_- > 0$ be real numbers to be chosen later. Let $\mathbf{\Lambda} = (\Lambda_1, \ldots, \Lambda_m)$ with entries $-\Lambda_+ \leq \Lambda_k \leq \Lambda_+$, and $\mathbf{A} = (A_1, \ldots, A_m)$

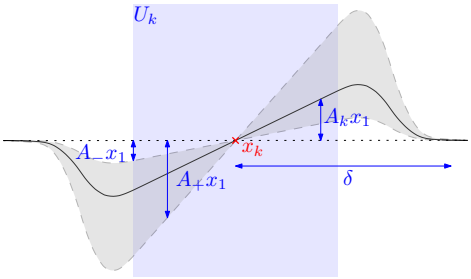
with entries $A_- \le A_k \le A_+$. For $z \in \mathbb{R}^D$, we write $z = (z_1, \dots, z_D)$ for its coordinates in the canonical basis. For all $\tau = (\tau_1, \dots, \tau_m) \in \{0,1\}^m$, define the bump map as

$$\Phi_\tau^{\mathbf{\Lambda},\mathbf{A},i}(x) = x + \sum_{k=1}^{m} \phi\left(\frac{x - x_k}{\delta}\right) \left\{\tau_k A_k (x - x_k)_1^i + (1 - \tau_k)\Lambda_k\right\} e. \tag{7}$$

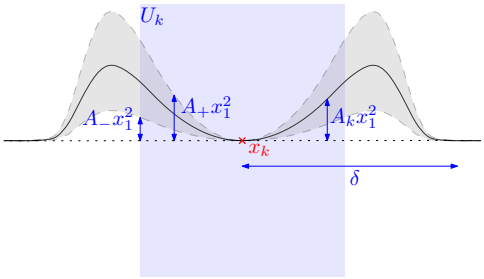
An analogous deformation map was considered in [1]. We let $P_\tau^{\mathbf{\Lambda},\mathbf{A},(i)}$ denote the pushforward distribution of P_0 by $\Phi_\tau^{\mathbf{\Lambda},\mathbf{A},(i)}$, and write $M_\tau^{\mathbf{\Lambda},\mathbf{A},(i)}$ for its support. Roughly speaking, $M_\tau^{\mathbf{\Lambda},\mathbf{A},i}$ consists of m bumps at the x_k's having different shapes (Figure 7). If $\tau_k = 0$, the bump at x_k is a symmetric plateau function and has height Λ_k. If $\tau_k = 1$, it fits the graph of the polynomial $A_k(x - x_k)_1^i$ locally. The following Lemma C.12 gives differential bounds and



(a) Flat bump: $\tau_k = 0$.



(b) Linear bump: $\tau_k = 1$, $i = 1$.



(c) Quadratic bump: $\tau_k = 1$, $i = 2$.

Figure 7: The three shapes of the bump map $\Phi_\tau^{\mathbf{\Lambda},\mathbf{A},i}$ around x_k.

geometric properties of $\Phi_\tau^{\mathbf{\Lambda},\mathbf{A},i}$.

LEMMA C.12. *There exists $c_{\phi,i} < 1$ such that if $A_+ \le c_{\phi,i}\delta^{i-1}$ and $\Lambda_+ \le c_{\phi,i}\delta$, then $\Phi_\tau^{\mathbf{\Lambda},\mathbf{A},i}$ is a global $\mathcal{C}^\infty$-diffeomorphism of $\mathbb{R}^D$ such that for*

all $1 \leq k \leq m$, $\Phi_\tau^{\mathbf{\Lambda},\mathbf{A},i}(\mathcal{B}(x_k,\delta)) = \mathcal{B}(x_k,\delta)$. *Moreover,*

$$\left\|I_D - d\Phi_\tau^{\mathbf{\Lambda},\mathbf{A},i}\right\|_{op} \leq C_i \left\{\frac{A_+}{\delta^{1-i}}\right\} \vee \left\{\frac{\Lambda_+}{\delta}\right\},$$

and for $j \geq 2$,

$$\left\|d^j \Phi_\tau^{\mathbf{\Lambda},\mathbf{A},i}\right\|_{op} \leq C_{i,j} \left\{\frac{A_+}{\delta^{j-i}}\right\} \vee \left\{\frac{\Lambda_+}{\delta^j}\right\}.$$

Proof of Lemma C.12. Follows straightforwardly from chain rule, similarly to Lemma 11 in [1]. □

Lemma C.13. *If* $\tau_{min}L_\perp, \ldots, \tau_{min}^{k-1}L_k, (\tau_{min}^d f_{min})^{-1}$ *and* $\tau_{min}^d f_{max}$ *are large enough (depending only on* d *and* k*), then provided that* $\Lambda_+ \vee A_+\delta^i \leq c_{k,d,\tau_{min}}\delta^k$, *for all* $\tau \in \{0,1\}^m$, $P_\tau^{\mathbf{\Lambda},\mathbf{A},i} \in \mathcal{P}^k_{\tau_{min},\mathbf{L},f_{min},f_{max}}$

Proof of Lemma C.13. Follows using the stability of the model Lemma A.4 applied to the distribution $P_0 \in \mathcal{P}^k_{2\tau_{min},\mathbf{L}/2,2f_{min},f_{max}/2}$ and the map $\Phi_\tau^{\mathbf{\Lambda},\mathbf{A},i}$, of which differential bounds are asserted by Lemma C.12. □

C.3. Hypotheses for Tangent Space and Curvature

C.3.1. Proof of Lemma 8

This section is devoted to the proof of Lemma 8, for which we first derive two slightly more general results, with parameters to be tuned later. The proof is split into two intermediate results Lemma C.14 and Lemma C.15.

Let us write $\bar{Q}_{\tau,n}^{(i)}$ for the mixture distribution on $(\mathbb{R}^D)^n$ defined by

$$\bar{Q}_{\tau,n}^{(i)} = \int_{[-\Lambda_+,\Lambda_+]^m} \int_{[A_-,A_+]^m} \left(P_\tau^{\mathbf{\Lambda},\mathbf{A},(i)}\right)^{\otimes n} \frac{d\mathbf{A}}{(A_+ - A_-)^m} \frac{d\mathbf{\Lambda}}{(2\Lambda_+)^m}. \tag{8}$$

Although the probability distribution $\bar{Q}_{\tau,n}^{(i)}$ depends on A_-, A_+ and Λ_+, we omit this dependency for the sake of compactness. Another way to define $\bar{Q}_{\tau,n}^{(i)}$ is the following: draw uniformly $\mathbf{\Lambda}$ in $[-\Lambda_+,\Lambda_+]^m$ and $\mathbf{A}$ in $[A_-,A_+]^m$, and given $(\mathbf{\Lambda},\mathbf{A})$, take $Z_i = \Phi_\tau^{\mathbf{\Lambda},\mathbf{A},i}(Y_i)$, where $Y_1,\ldots,Y_n$ is an i.i.d. n-sample with common distribution P_0 on M_0. Then $(Z_1,\ldots,Z_n)$ has distribution $\bar{Q}_{\tau,n}^{(i)}$.

Lemma C.14. *Assume that the conditions of Lemma C.12 hold, and let*

$$U_k = \mathcal{B}_{\mathbb{R}^d\times\{0\}^{D-d}}(x_k,\delta/2) + \mathcal{B}_{span(e)}(0,\tau_{min}/2),$$

and

$$U_k' = \left(\mathbb{R}^D \setminus \left\{\mathcal{B}_{\mathbb{R}^d\times\{0\}^{D-d}}(x_k,\delta) + \mathcal{B}_{span(e)}(0,\tau_{min}/2)\right\}\right)^{n-1}.$$

Then the sets $U_k \times U_k'$ are pairwise disjoint, $\bar{Q}_{\tau,n}^{(i)} \in \overline{Conv}\big((\mathcal{P}_\tau^{(i)})^{\otimes n}\big)$, and if $(Z_1,\dots,Z_n) = (Z_1, Z_{2:n})$ has distribution $\bar{Q}_{\tau,n}^{(i)}$, Z_1 and $Z_{2:d}$ are independent conditionally on the event $\{(Z_1, Z_{2:n}) \in U_k \times U_k'\}$.

Moreover, if $(X_1,\dots,X_n)$ has distribution $\big(P_\tau^{\mathbf{\Lambda},\mathbf{A},(i)}\big)^{\otimes n}$ (with fixed $\mathbf{A}$ and $\mathbf{\Lambda}$), then on the event $\{X_1 \in U_k\}$, we have:

- *if $\tau_k = 0$,*

$$T_{X_1}M_\tau^{\mathbf{\Lambda},\mathbf{A},(i)} = \mathbb{R}^d \times \{0\}^{D-d} \quad , \quad \left\| II_{X_1}^{M_\tau^{\mathbf{\Lambda},\mathbf{A},(i)}} \circ \pi_{T_{X_1}M_\tau^{\mathbf{\Lambda},\mathbf{A},(i)}} \right\|_{op} = 0$$

  and $d_H\big(M_0, M_\tau^{\mathbf{\Lambda},\mathbf{A},(i)}\big) \geq |\Lambda_k|$.
- *if $\tau_k = 1$,*
  - *for $i = 1$: $\angle\left(T_{X_1}M_\tau^{\mathbf{\Lambda},\mathbf{A},(1)}, \mathbb{R}^d \times \{0\}^{D-d}\right) \geq A_-/2$.*
  - *for $i = 2$: $\left\| II_{X_1}^{M_\tau^{\mathbf{\Lambda},\mathbf{A},(2)}} \circ \pi_{T_{X_1}M_\tau^{\mathbf{\Lambda},\mathbf{A},(2)}} \right\|_{op} \geq A_-/2$.*

Proof of Lemma C.14. It is clear from the definition (8) that $\bar{Q}_{\tau,n}^{(i)} \in \overline{Conv}\big((\mathcal{P}_\tau^{(i)})^{\otimes n}\big)$. By construction of the $\Phi_\tau^{\mathbf{\Lambda},\mathbf{A},i}$'s, these maps leave the sets

$$\mathcal{B}_{\mathbb{R}^d\times\{0\}^{D-d}}(x_k,\delta) + \mathcal{B}_{span(e)}(0,\tau_{min}/2)$$

unchanged for all $\mathbf{\Lambda}, \mathbf{L}$. Therefore, on the event $\{(Z_1, Z_{2:n}) \in U_k \times U_k'\}$, one can write Z_1 only as a function of X_1, Λ_k, A_k, and $Z_{2:n}$ as a function of the rest of the X_j's,Λ_k's and A_k's. Therefore, Z_1 and $Z_{2:n}$ are independent.

We now focus on the geometric statements. For this, we fix a deterministic point $z = \Phi_\tau^{\mathbf{\Lambda},\mathbf{A},(i)}(x_0) \in U_k \cap M_\tau^{\mathbf{\Lambda},\mathbf{A},(i)}$. By construction, one necessarily has $x_0 \in M_0 \cap \mathcal{B}(x_k, \delta/2)$.

- If $\tau_k = 0$, locally around x_0, $\Phi_\tau^{\mathbf{\Lambda},\mathbf{A},(1)}$ is the translation of vector $\Lambda_k e$. Therefore, since M_0 satisfies $T_{x_0}M_0 = \mathbb{R}^d \times \{0\}^{D-d}$ and $II_{x_0}^{M_0} = 0$, we have

$$T_z M_\tau^{\mathbf{\Lambda},\mathbf{A},(i)} = \mathbb{R}^d \times \{0\}^{D-d} \quad \text{and} \quad \left\| II_z^{M_\tau^{\mathbf{\Lambda},\mathbf{A},(i)}} \circ \pi_{T_z M_\tau^{\mathbf{\Lambda},\mathbf{A},(i)}} \right\|_{op} = 0.$$

- if $\tau_k = 1$,

– for $i = 1$: locally around x_0, $\Phi_\tau^{\mathbf{\Lambda},\mathbf{A},(1)}$ can be written as $x \mapsto x + A_k(x - x_k)_1 e$. Hence, $T_z M_\tau^{\mathbf{\Lambda},\mathbf{A},(i)}$ contains the direction $(1, A_k)$ in the plane $span(e_1, e)$ spanned by the first vector of the canonical basis and e. As a consequence, since e is orthogonal to $\mathbb{R}^d \times \{0\}^{D-d}$,

$$\angle\left(T_z M_\tau^{\mathbf{\Lambda},\mathbf{A},(1)}, \mathbb{R}^d \times \{0\}^{D-d}\right) \geq \left(1 + 1/A_k^2\right)^{-1/2} \geq A_k/2 \geq A_-/2.$$

– for $i = 2$: locally around x_0, $\Phi_\tau^{\mathbf{\Lambda},\mathbf{A},(2)}$ can be written as $x \mapsto x + A_k(x - x_k)_1^2 e$. Hence, $M_\tau^{\mathbf{\Lambda},\mathbf{A},(2)}$ contains an arc of parabola of equation $y = A_k(x - x_k)_1^2$ in the plane $span(e_1, e)$. As a consequence,

$$\left\| II_z^{M_\tau^{\mathbf{\Lambda},\mathbf{A},(2)}} \circ \pi_{T_z M_\tau^{\mathbf{\Lambda},\mathbf{A},(2)}} \right\|_{op} \geq A_k/2 \geq A_-/2.$$

□

LEMMA C.15. *Assume that the conditions of Lemma C.12 and Lemma C.14 hold. If in addition,* $cA_+(\delta/4)^i \leq \Lambda_+ \leq CA_+(\delta/4)^i$ *for some absolute constants* $C \geq c > 3/4$, *and* $A_- = A_+/2$, *then,*

$$\int_{U_k} d\bar{Q}_{\tau,1}^{(i)} \wedge d\bar{Q}_{\tau^k,1}^{(i)} \geq \frac{c_{d,i}}{C} \left(\frac{\delta}{\tau_{min}}\right)^d,$$

and

$$\int_{U_k'} d\bar{Q}_{\tau,n-1}^{(i)} \wedge d\bar{Q}_{\tau^k,n-1}^{(i)} = \left(1 - c_d' \left(\frac{\delta}{\tau_{min}}\right)^d\right)^{n-1}.$$

PROOF OF LEMMA C.15. First note that all the involved distributions have support in $\mathbb{R}^d \times span(e) \times \{0\}^{D-(d+1)}$. Therefore, we use the canonical coordinate system of $\mathbb{R}^d \times span(e)$, centered at x_k, and we denote the components by $(x_1, x_2, \ldots, x_d, y) = (x_1, x_{2:d}, y)$. Without loss of generality, assume that $\tau_k = 0$ (if not, flip τ and τ^k). Recall that ϕ has been chosen to be constant and equal to 1 on the ball $\mathcal{B}(0, 1/2)$.

By definition (8), on the event $\{Z \in U_k\}$, a random variable Z having distribution $\bar{Q}_{\tau,1}^{(i)}$ can be represented by $Z = X + \phi\left(\frac{X - x_k}{\delta}\right)\Lambda_k e = X + \Lambda_k e$ where X and Λ_k are independent and have respective distributions P_0 (the uniform distribution on M_0) and the uniform distribution on $[-\Lambda_+, \Lambda_+]$.

Therefore, on U_k, $\bar{Q}^{(i)}_{\tau,1}$ has a density with respect to the Lebesgue measure λ_{d+1} on $\mathbb{R}^d \times span(e)$ that can be written as

$$\bar{q}^{(i)}_{\tau,1}(x_1, x_{2:d}, y) = \frac{\mathbb{1}_{[-\Lambda_+,\Lambda_+]}(y)}{2Vol(M_0)\Lambda_+}.$$

Analogously, nearby x_k a random variable Z having distribution $\bar{Q}^{(i)}_{\tau^k,1}$ can be represented by $Z = X + A_k(X - x_k)^i_1 e$ where A_k has uniform distribution on $[A_-, A_+]$. Therefore, a straightforward change of variable yields the density

$$\bar{q}^{(i)}_{\tau^k,1}(x_1, x_{2:d}, y) = \frac{\mathbb{1}_{[A_- x_1^i, A_+ x_1^i]}(y)}{Vol(M_0)\,(A_+ - A_-)\,x_1^i}.$$

We recall that $Vol(M_0) = (2\tau_{min})^d Vol\big(M_0^{(0)}\big) = c'_d \tau^d_{min}$. Let us now tackle the right-hand side inequality, writing

$$\begin{aligned}
&\int_{U_k} d\bar{Q}^{(i)}_{\tau,1} \wedge d\bar{Q}^{(i)}_{\tau^k,1} \\
&\quad = \int_{\mathcal{B}(x_k,\delta/2)} \left(\frac{\mathbb{1}_{[-\Lambda_+,\Lambda_+]}(y)}{2Vol(M_0)\Lambda_+}\right) \wedge \left(\frac{\mathbb{1}_{[A_- x_1^i, A_+ x_1^i]}(y)}{Vol(M_0)\,(A_+ - A_-)\,x_1^i}\right) dy dx_1 dx_{2:d} \\
&\quad \geq \int_{\mathcal{B}_{\mathbb{R}^{d-1}}(0,\frac{\delta}{4})} \int_{-\delta/4}^{\delta/4} \int_{\mathbb{R}} \left(\frac{\mathbb{1}_{[-\Lambda_+,\Lambda_+]}(y)}{2\Lambda_+}\right) \wedge \left(\frac{\mathbb{1}_{[A_- x_1^i, A_+ x_1^i]}(y)}{A_+ x_1^i/2}\right) \frac{dy dx_1 dx_{2:d}}{Vol(M_0)}.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\int_{U_k} d\bar{Q}^{(i)}_{\tau,1} \wedge d\bar{Q}^{(i)}_{\tau^k,1} \\
&\quad \geq \frac{c_d}{\tau^d_{min}} \delta^{d-1} \int_0^{\delta/4} \int_{A_+ x_1^i/2}^{\Lambda_+ \wedge (A_+ x_1^i)} \frac{1}{2\Lambda_+} \wedge \frac{2}{A_+ x_1^i} dy dx_1 \\
&\quad \geq \frac{c_d}{\tau^d_{min}} \delta^{d-1} \int_0^{\delta/4} \int_{A_+ x_1^i/2}^{(c\wedge 1)(A_+ x_1^i)} \frac{(2c \wedge 1/2)}{2\Lambda_+} dy dx_1 \\
&\quad = \frac{c_d}{\tau^d_{min}} \delta^{d-1} (2c \wedge 1/2)\,(c \wedge 1 - 1/2)\,\frac{A_+}{\Lambda_+} \frac{(\delta/4)^{i+1}}{i+1} \\
&\quad \geq \frac{c_{d,i}}{C} \left(\frac{\delta}{\tau_{min}}\right)^d.
\end{aligned}$$

For the integral on U'_k, notice that by definition, $\bar{Q}^{(i)}_{\tau,n-1}$ and $\bar{Q}^{(i)}_{\tau^k,n-1}$ coincide on U'_k since they are respectively the image distributions of P_0 by

functions that are equal on that set. Moreover, these two functions leave $\mathbb{R}^D \setminus \left\{ \mathcal{B}_{\mathbb{R}^d \times \{0\}^{D-d}}(x_k, \delta) + \mathcal{B}_{span(e)}(0, \tau_{min}/2) \right\}$ unchanged. Therefore,

$$\begin{aligned}
\int_{U_k'} & d\bar{Q}^{(i)}_{\tau,n-1} \wedge d\bar{Q}^{(i)}_{\tau^k,n-1} \\
&= P_0^{\otimes n-1}\left(U_k'\right) \\
&= \left(1 - P_0\left(\mathcal{B}_{\mathbb{R}^d \times \{0\}^{D-d}}(x_k, \delta) + \mathcal{B}_{span(e)}(0, \tau_{min}/2)\right)\right)^{n-1} \\
&= \left(1 - \omega_d \delta^d / Vol(M_0)\right)^{n-1},
\end{aligned}$$

hence the result. □

Proof of Lemma 8. The properties of $\left\{\bar{Q}^{(i)}_{\tau,n}\right\}_\tau$ and $\{U_k \times U_k'\}_k$ given by Lemma C.14 and Lemma C.15 yield the result, setting $\Lambda_+ = A_+\delta^i/4$, $A_+ = 2A_- = \varepsilon\delta^{k-i}$ for $\varepsilon = \varepsilon_{k,d,\tau_{min}}$, and δ such that $c_d'\left(\frac{\delta}{\tau_{min}}\right)^d = \frac{1}{n-1}$. □

C.3.2. Proof of Lemma 9

This section details the construction leading to Lemma 9 that we restate in Lemma C.16.

Lemma C.16. *Assume that $\tau_{min}L_\perp, \ldots, \tau_{min}^{k-1}L_k, (\tau_{min}^d f_{min})^{-1}$, $\tau_{min}^d f_{max}$ are large enough (depending only on d and k), and $\sigma \geq C_{k,d,\tau_{min}}\left(1/(n-1)\right)^{k/d}$ for $C_{k,d,\tau_{min}} > 0$ large enough. Given $i \in \{1,2\}$, there exists a collection of 2^m distributions $\left\{\mathbf{P}_\tau^{(i),\sigma}\right\}_{\tau\in\{0,1\}^m} \subset \mathcal{P}^k(\sigma)$ with associated submanifolds $\left\{M_\tau^{(i),\sigma}\right\}_{\tau\in\{0,1\}^m}$, together with pairwise disjoint subsets $\{U_k^\sigma\}_{1\leq k\leq m}$ of $\mathbb{R}^D$ such that the following holds for all $\tau \in \{0,1\}^m$ and $1 \leq k \leq m$.*

If $x \in U_k^\sigma$ and $y = \pi_{M_\tau^{(i),\sigma}}(x)$, we have

- *if $\tau_k = 0$,*

$$T_y M_\tau^{(i),\sigma} = \mathbb{R}^d \times \{0\}^{D-d} \quad , \quad \left\| II_y^{M_\tau^{(i),\sigma}} \circ \pi_{T_y M_\tau^{(i),\sigma}} \right\|_{op} = 0,$$

- *if $\tau_k = 1$,*
  - *for $i = 1$: $\angle\left(T_y M_\tau^{(1),\sigma}, \mathbb{R}^d \times \{0\}^{D-d}\right) \geq c_{k,d,\tau_{min}}\left(\frac{\sigma}{n-1}\right)^{\frac{k-1}{k+d}}$,*
  - *for $i = 2$: $\left\| II_y^{M_\tau^{(2),\sigma}} \circ \pi_{T_y M_\tau^{(2),\sigma}} \right\|_{op} \geq c'_{k,d,\tau_{min}}\left(\frac{\sigma}{n-1}\right)^{\frac{k-2}{k+d}}$.*

Furthermore,

$$\int_{(\mathbb{R}^D)^{n-1}} \left(\mathbf{P}_\tau^{(i),\sigma}\right)^{\otimes n-1} \wedge \left(\mathbf{P}_{\tau^k}^{(i),\sigma}\right)^{\otimes n-1} \geq c_0, \quad \textit{and} \quad m \cdot \int_{U_k^\sigma} \mathbf{P}_\tau^{(i),\sigma} \wedge \mathbf{P}_{\tau^k}^{(i),\sigma} \geq c_d.$$

Proof of Lemma C.16. Following the notation of Section C.2, for $i \in \{1,2\}$, $\tau \in \{0,1\}^m$, $\delta \leq \tau_{min}/4$ and $A > 0$, consider

$$\Phi_\tau^{A,i}(x) = x + \sum_{k=1}^m \phi\left(\frac{x - x_k}{\delta}\right) \left\{\tau_k A (x - x_k)_1^i\right\} e. \tag{9}$$

Note that (9) is a particular case of (7). Clearly from the definition, $\Phi_\tau^{A,i}$ and $\Phi_{\tau^k}^{A,i}$ coincide outside $\mathcal{B}(x_k, \delta)$, $(\Phi(x) - x) \in span(e)$ for all $x \in \mathbb{R}^D$, and $\|I_D - \Phi\|_\infty \leq A\delta^i$. Let us define $M_\tau^{A,i} = \Phi_\tau^{A,i}(M_0)$. From Lemma C.13, we have $M_\tau^{A,i} \in \mathcal{C}^k_{\tau_{min},\mathbf{L}}$ provided that $\tau_{min} L_\perp, \ldots, \tau_{min}^{k-1} L_k$ are large enough, and that $\delta \leq \tau_{min}/2$, with $A/\delta^{k-i} \leq \varepsilon$ for $\varepsilon = \varepsilon_{k,d,\tau_{min},i}$ small enough.

Furthermore, let us write

$$U_k^\sigma = \mathcal{B}_{\mathbb{R}^d \times \{0\}^{D-d}}(x_k, \delta/2) + \mathcal{B}_{\{0\}^d \times \mathbb{R}^{D-d}}(x_k, \sigma/2).$$

Then the family $\{U_k^\sigma\}_{1 \leq k \leq m}$ is pairwise disjoint. Also, since $\tau_k = 0$ implies that $M_\tau^{A,i}$ coincides with M_0 on $\mathcal{B}(x_k, \delta)$, we get that if $x \in U_k^\sigma$ and $y = \pi_{M_\tau^{A,i}}(x)$,

$$T_y M_\tau^{A,i} = \mathbb{R}^d \times \{0\}^{D-d} \quad , \quad \left\| II_y^{M_\tau^{A,i}} \circ \pi_{T_y M_\tau^{A,i}} \right\|_{op} = 0.$$

Furthermore, by construction of the bump function $\Phi_\tau^{A,i}$, if $x \in U_k^\sigma$ and $\tau_k = 1$, then

$$\angle\left(T_y M_\tau^{A,i}, \mathbb{R}^d \times \{0\}^{D-d}\right) \geq \frac{A}{2},$$

and

$$\left\| II_y^{M_\tau^{A,i}} \circ \pi_{T_y M_\tau^{A,i}} \right\|_{op} \geq \frac{A}{2}.$$

Now, let us write

$$\mathcal{O}_\tau^{A,i} = \left\{ y + \xi \,\middle|\, y \in M_\tau^{A,i}, \xi \in \left(T_y M_\tau^{A,i}\right)^\perp, \|\xi\| \leq \sigma/2 \right\}$$

for the offset of $M_\tau^{\Lambda,A,i}$ of radius $\sigma/2$. The sets $\left\{\mathcal{O}_\tau^{A,i}\right\}_\tau$ are closed subsets of $\mathbb{R}^D$ with non-empty interiors. Let $\mathbf{P}_\tau^{A,i}$ denote the uniform distribution on $\mathcal{O}_\tau^{A,i}$. Finally, let us denote by $P_\tau^{A,i} = \left(\pi_{M_\tau^{A,i}}\right)_* \mathbf{P}_\tau^{A,i}$ the pushforward

distributions of $\mathbf{P}_\tau^{A,i}$ by the projection maps $\pi_{M_\tau^{A,i}}$. From Lemma 19 in [26], $P_\tau^{A,i}$ has a density $f_\tau^{A,i}$ with respect to the volume measure on $M_\tau^{A,i}$, and this density satisfies

$$Vol\left(M_\tau^{A,i}\right) f_\tau^{A,i} \leq \left(\frac{\tau_{min} + \sigma/2}{\tau_{min} - \sigma/2}\right)^d \leq \left(\frac{5}{3}\right)^d,$$

and

$$Vol\left(M_\tau^{A,i}\right) f_\tau^{A,i} \geq \left(\frac{\tau_{min} - \sigma/2}{\tau_{min} + \sigma/2}\right)^d \geq \left(\frac{3}{5}\right)^d.$$

Since, by construction, $Vol(M_0) = c_d \tau_{min}^d$, and $c_d' \leq Vol\big(M_\tau^{\Lambda,A,i}\big)/Vol(M_0) \leq C_d'$ whenever $A/\delta^{i-1} \leq \varepsilon'_{d,\tau_{min},i}$, we get that $P_\tau^{A,i}$ belongs to the model $\mathcal{P}^k$ provided that $(\tau_{min}^d f_{min})^{-1}$ and $\tau_{min}^d f_{max}$ are large enough. This proves that under these conditions, the family $\left\{\mathbf{P}_\tau^{A,i}\right\}_{\tau\in\{0,1\}^m}$ is included in the model $\mathcal{P}^k(\sigma)$.

Let us now focus on the bounds on the L^1 test affinities. Let $\tau \in \{0,1\}^m$ and $1 \leq k \leq m$ be fixed, and assume, without loss of generality, that $\tau_k = 0$ (if not, flip the role of τ and τ^k). First, note that

$$\int_{(\mathbb{R}^D)^{n-1}} \left(\mathbf{P}_\tau^{A,i}\right)^{\otimes n-1} \wedge \left(\mathbf{P}_{\tau^k}^{A,i}\right)^{\otimes n-1} \geq \left(\int_{\mathbb{R}^D} \mathbf{P}_\tau^{A,i} \wedge \mathbf{P}_{\tau^k}^{A,i}\right)^{n-1}.$$

Furthermore, since $\mathbf{P}_\tau^{A,i}$ and $\mathbf{P}_{\tau^k}^{A,i}$ are the uniform distributions on $\mathcal{O}_\tau^{A,i}$ and $\mathcal{O}_\tau^{A,i}$,

$$\begin{aligned}\int_{\mathbb{R}^D} \mathbf{P}_\tau^{A,i} \wedge \mathbf{P}_{\tau^k}^{A,i} &= 1 - \frac{1}{2}\int_{\mathbb{R}^D} \left|\mathbf{P}_\tau^{A,i} - \mathbf{P}_{\tau^k}^{A,i}\right| \\ &= 1 - \frac{1}{2}\int_{\mathbb{R}^D} \left|\frac{\mathbb{1}_{\mathcal{O}_\tau^{A,i}}(a)}{Vol\left(\mathcal{O}_\tau^{A,i}\right)} - \frac{\mathbb{1}_{\mathcal{O}_\tau^{A,i}}(a)}{Vol\left(\mathcal{O}_{\tau^k}^{A,i}\right)}\right| d\mathcal{H}^D(a).\end{aligned}$$

Furthermore,

$$\begin{aligned}\frac{1}{2}\int_{\mathbb{R}^D}\left|\frac{\mathbb{1}_{\mathcal{O}_\tau^{A,i}}(a)}{Vol\left(\mathcal{O}_\tau^{A,i}\right)}-\frac{\mathbb{1}_{\mathcal{O}_{\tau^k}^{A,i}}(a)}{Vol\left(\mathcal{O}_{\tau^k}^{A,i}\right)}\right|d\mathcal{H}^D(a)&\\ =\frac{1}{2}Vol\left(\mathcal{O}_\tau^{A,i}\cap\mathcal{O}_{\tau^k}^{A,i}\right)\left|\frac{1}{Vol\left(\mathcal{O}_\tau^{A,i}\right)}-\frac{1}{Vol\left(\mathcal{O}_{\tau^k}^{A,i}\right)}\right|&\\ +\frac{1}{2}\left(\frac{Vol\left(\mathcal{O}_\tau^{A,i}\setminus\mathcal{O}_{\tau^k}^{A,i}\right)}{Vol\left(\mathcal{O}_\tau^{A,i}\right)}+\frac{Vol\left(\mathcal{O}_{\tau^k}^{A,i}\setminus\mathcal{O}_\tau^{A,i}\right)}{Vol\left(\mathcal{O}_{\tau^k}^{A,i}\right)}\right)&\\ \leq\frac{3}{2}\frac{Vol\left(\mathcal{O}_\tau^{A,i}\setminus\mathcal{O}_{\tau^k}^{A,i}\right)\vee Vol\left(\mathcal{O}_{\tau^k}^{A,i}\setminus\mathcal{O}_\tau^{A,i}\right)}{Vol\left(\mathcal{O}_\tau^{A,i}\right)\wedge Vol\left(\mathcal{O}_{\tau^k}^{A,i}\right)}.&\end{aligned}$$

To get a lower bound on the denominator, note that for $\delta\leq\tau_{min}/2$, $M_\tau^{A,i}$ and $M_{\tau^k}^{A,i}$ both contain

$$\mathcal{B}_{\mathbb{R}^d\times\{0\}^{D-d}}(0,\tau_{min})\setminus\mathcal{B}_{\mathbb{R}^d\times\{0\}^{D-d}}(0,\tau_{min}/4),$$

so that $\mathcal{O}_\tau^{A,i}$ and $\mathcal{O}_{\tau^k}^{A,i}$ both contain

$$\left(\mathcal{B}_{\mathbb{R}^d\times\{0\}^{D-d}}(0,\tau_{min})\setminus\mathcal{B}_{\mathbb{R}^d\times\{0\}^{D-d}}(0,\tau_{min}/4)\right)+\mathcal{B}_{\{0\}^d\times\mathbb{R}^{D-d}}(0,\sigma/2).$$

As a consequence, $Vol\left(\mathcal{O}_\tau^{A,i}\right)\wedge Vol\left(\mathcal{O}_{\tau^k}^{A,i}\right)\geq c_d\omega_d\tau_{min}^d\omega_{D-d}(\sigma/2)^{D-d}$, where ω_ℓ denote the volume of a ℓ-dimensional unit Euclidean ball.

We now derive an upper bound on $Vol\big(\mathcal{O}_\tau^{A,i}\setminus\mathcal{O}_{\tau^k}^{A,i}\big)$. To this aim, let us consider $a_0=y+\xi\in\mathcal{O}_\tau^{A,i}\setminus\mathcal{O}_{\tau^k}^{A,i}$, with $y\in M_\tau^{A,i}$ and $\xi\in\big(T_yM_\tau^{A,i}\big)^\perp$. Since $\Phi_\tau^{A,i}$ and $\Phi_{\tau^k}^{A,i}$ coincide outside $\mathcal{B}(x_k,\delta)$, so do $M_\tau^{A,i}$ and $M_{\tau^k}^{A,i}$. Hence, one necessarily has $y\in\mathcal{B}(x_k,\delta)$. Thus, $\big(T_yM_\tau^{A,i}\big)^\perp=T_yM_0^\perp=span(e)+\{0\}^{d+1}\times\mathbb{R}^{D-d-1}$, so we can write $\xi=se+z$ with $s\in\mathbb{R}$ and $z\in\{0\}^{d+1}\times\mathbb{R}^{D-d-1}$. By definition of $\mathcal{O}_\tau^{A,i}$, $\|\xi\|=\sqrt{s^2+\|z\|^2}\leq\sigma/2$, which yields $\|z\|\leq\sigma/2$ and $|s|\leq\sqrt{(\sigma/2)^2-\|z\|^2}$. Furthermore, y_0 does not belong to $\mathcal{O}_{\tau^k}^{A,i}$, which translates to

$$\begin{aligned}\sigma/2<d\big(a_0,M_{\tau^k}^{A,i}\big)&\leq\left\|y_0+se+z-\Phi_{\tau^k}^{A,i}(y_0)\right\|\\ &=\sqrt{\left|s+\left\langle e,y_0-\Phi_{\tau^k}^{A,i}(y_0)\right\rangle\right|^2+\|z\|^2},\end{aligned}$$

from what we get $|s| \geq \sqrt{(\sigma/2)^2 - \|z\|^2} - \left\|I_D - \Phi_{\tau^k}^{A,i}\right\|_\infty$. We just proved that $\mathcal{O}_\tau^{A,i} \setminus \mathcal{O}_{\tau^k}^{A,i}$ is a subset of

$$\mathcal{B}_d(x_k,\delta) + \left\{ se + z \,\middle|\, (s,z) \in \mathbb{R} \times \mathbb{R}^{D-d-1}, \|z\| \leq \sigma/2 \text{ and } \sqrt{(\sigma/2)^2 - \|z\|^2} - \left\|I_D - \Phi_{\tau^k}^{A,i}\right\|_\infty \leq |s| \leq \sqrt{(\sigma/2)^2 - \|z\|^2} \right\}.$$

Hence,

$$Vol\left(\mathcal{O}_\tau^{A,i} \setminus \mathcal{O}_{\tau^k}^{A,i}\right) \leq \omega_d \delta^d \times 2\left\|I_D - \Phi_{\tau^k}^{A,i}\right\|_\infty \times \omega_{D-d-1}(\sigma/2)^{D-d-1}. \tag{10}$$

Similar arguments lead to

$$Vol\left(\mathcal{O}_{\tau^k}^{A,i} \setminus \mathcal{O}_\tau^{A,i}\right) \leq \omega_d \delta^d \times 2\left\|I_D - \Phi_\tau^{A,i}\right\|_\infty \times \omega_{D-d-1}(\sigma/2)^{D-d-1}. \tag{11}$$

Since $\left\|I_D - \Phi_\tau^{A,i}\right\|_\infty \vee \left\|I_D - \Phi_{\tau^k}^{A,i}\right\|_\infty \leq A\delta^i$, summing up bounds (10) and (11) yields

$$\begin{aligned}\int_{\mathbb{R}^D} \mathbf{P}_\tau^{A,i} \wedge \mathbf{P}_{\tau^k}^{A,i} &\geq 1 - 3\frac{\omega_d \omega_{D-d-1} A\delta^i \cdot \delta^d (\sigma/2)^{D-d-1}}{\omega_d \tau_{min}^d \omega_{D-d}(\sigma/2)^{D-d}} \\ &\geq 1 - 3\frac{A\delta^i}{\sigma}\left(\frac{\delta}{\tau_{min}}\right)^d.\end{aligned}$$

To derive the last bound, we notice that since $U_k^\sigma \subset \mathcal{O}_\tau^{A,i} = Supp(\mathbf{P}_\tau^{A,i})$, we have

$$\begin{aligned}\int_{U_k^\sigma} \mathbf{P}_\tau^{A,i} \wedge \mathbf{P}_{\tau^k}^{A,i} &\geq \frac{Vol\left(U_k^\sigma \cap \mathcal{O}_{\tau^k}^{A,i}\right)}{Vol\left(\mathcal{O}_\tau^{A,i}\right) \wedge Vol\left(\mathcal{O}_{\tau^k}^{A,i}\right)} \\ &\geq \frac{Vol\left(U_k^\sigma\right) - Vol\left(U_k^\sigma \setminus \mathcal{O}_{\tau^k}^{A,i}\right)}{Vol\left(\mathcal{O}_\tau^{A,i}\right) \wedge Vol\left(\mathcal{O}_{\tau^k}^{A,i}\right)} \\ &\geq \frac{Vol\left(U_k^\sigma\right) - Vol\left(\mathcal{O}_\tau^{A,i} \setminus \mathcal{O}_{\tau^k}^{A,i}\right)}{Vol\left(\mathcal{O}_\tau^{A,i}\right) \wedge Vol\left(\mathcal{O}_{\tau^k}^{A,i}\right)} \\ &\geq \frac{\omega_d(\delta/2)^d \omega_{D-d}(\sigma/2)^{D-d} - \omega_d \delta^d A\delta^i \omega_{D-d-1}(\sigma/2)^{D-d-1}}{\omega_d \tau_{min}^d \omega_{D-d}(\sigma/2)^{D-d}}.\end{aligned}$$

Hence, whenever $A\delta^i \leq c_d\sigma$ for c_d small enough, we get

$$\int_{U_k^\sigma} \mathbf{P}_\tau^{A,i} \wedge \mathbf{P}_{\tau^k}^{A,i} \geq c_d' \left(\frac{\delta}{\tau_{min}}\right)^d.$$

Since m can be chosen such that $m \geq c_d(\tau_{min}/\delta)^d$, we get the last bound.

Eventually, writting $\mathbf{P}_\tau^{(i),\sigma} = \mathbf{P}_\tau^{A,i}$ for the particular parameters $A = \varepsilon\delta^{k-i}$, for $\varepsilon = \varepsilon_{k,d,\tau_{min}}$ small enough, and δ such that $\frac{3A\delta^i}{\sigma}\left(\frac{\delta}{\tau_{min}}\right)^d = \frac{1}{n-1}$ yields the result. Such a choice of parameter δ does meet the condition $A\delta^i = \varepsilon\delta^k \leq c_d\sigma$, provided that $\sigma \geq \frac{c_d}{\varepsilon}\left(\frac{1}{n-1}\right)^{k/d}$. □

C.4. Hypotheses for Manifold Estimation

C.4.1. Proof of Lemma 5

Let us prove Lemma 5, stated here as Lemma C.17.

LEMMA C.17. *If $\tau_{min}L_\perp, \ldots, \tau_{min}^{k-1}L_k, (\tau_{min}^d f_{min})^{-1}$ and $\tau_{min}^d f_{max}$ are large enough (depending only on d and k), there exist $P_0, P_1 \in \mathcal{P}^k$ with associated submanifolds M_0, M_1 such that*

$$d_H(M_0, M_1) \geq c_{k,d,\tau_{min}}\left(\frac{1}{n}\right)^{\frac{k}{d}}, \quad \text{and} \quad \|P_0 \wedge P_1\|_1^n \geq c_0.$$

PROOF OF LEMMA C.17. Following the notation of Section C.2, for $\delta \leq \tau_{min}/4$ and $\Lambda > 0$, consider

$$\Phi_\tau^\Lambda(x) = x + \phi\left(\frac{x}{\delta}\right)\Lambda \cdot e,$$

which is a particular case of (7). Define $M^\Lambda = \Phi^\Lambda(M_0)$, and $P^\Lambda = \Phi_*^\Lambda P_0$. Under the conditions of Lemma C.13, P_0 and P^Λ belong to $\mathcal{P}^k$, and by construction, $d_H(M_0, M^\Lambda) = \Lambda$. In addition, since P_0 and P^Λ coincide outside $\mathcal{B}(0,\delta)$,

$$\int_{\mathbb{R}^D} dP_0 \wedge dP^\Lambda = P_0\big(\mathcal{B}(0,\delta)\big) = \omega_d\left(\frac{\delta}{\tau_{min}}\right)^d.$$

Setting $P_1 = P^\Lambda$ with $\omega_d\left(\frac{\delta}{\tau_{min}}\right)^d = \frac{1}{n}$ and $\Lambda = c_{k,d,\tau_{min}}\delta^k$ for $c_{k,d,\tau_{min}} > 0$ small enough yields the result. □

C.4.2. Proof of Lemma 6

Here comes the proof of Lemma 6, stated here as Lemma C.17.

LEMMA C.18. *If $\tau_{min}L_\perp, \dots, \tau_{min}^{k-1}L_k, (\tau_{min}^d f_{min})^{-1}$ and $\tau_{min}^d f_{max}$ are large enough (depending only on d and k), there exist $P_0^\sigma, P_1^\sigma \in \mathcal{P}^k(\sigma)$ with associated submanifolds M_0^σ, M_1^σ such that*

$$d_H(M_0^\sigma, M_1^\sigma) \geq c_{k,d,\tau_{min}} \left(\frac{\sigma}{n}\right)^{\frac{k}{d+k}}, \quad \text{and} \quad \|P_0^\sigma \wedge P_1^\sigma\|_1^n \geq c_0.$$

PROOF OF LEMMA C.18. The proof follows the lines of that of Lemma C.16. Indeed, with the notation of Section C.2, for $\delta \leq \tau_{min}/4$ and $0 < \Lambda \leq c_{k,d,\tau_{min}}\delta^k$ for $c_{k,d,\tau_{min}} > 0$ small enough, consider

$$\Phi_\tau^\Lambda(x) = x + \phi\left(\frac{x}{\delta}\right)\Lambda \cdot e.$$

Define $M^\Lambda = \Phi^\Lambda(M_0)$. Write $\mathcal{O}_0$, $\mathcal{O}^\Lambda$ for the offsets of radii $\sigma/2$ of M_0, M^Λ, and and $\mathbf{P}_0, \mathbf{P}^\Lambda$ for the uniform distributions on these sets.

By construction, we have $d_H(M_0, M^\Lambda) = \Lambda$, and as in the proof of Lemma C.16, we get

$$\int_{\mathbb{R}^D} \mathbf{P}_0 \wedge \mathbf{P}^\Lambda \geq 1 - 3\frac{\Lambda}{\sigma}\left(\frac{\delta}{\tau_{min}}\right)^d.$$

Denoting $P_0^\sigma = \mathbf{P}_0$ and $P_1^\sigma = \mathbf{P}^\Lambda$ with $\Lambda = \varepsilon_{k,d,\tau_{min}}\delta^k$ and δ such that $3\frac{\Lambda}{\sigma}\left(\frac{\delta}{\tau_{min}}\right)^d$ yields the result.

□

C.5. Minimax Inconsistency Results

This section is devoted to the proof of Theorem 1, reproduced here as Theorem C.19.

THEOREM C.19. *Assume that $\tau_{min} = 0$. If $D \geq d+3$, then, for all $k \geq 2$ and $L_\perp > 0$, provided that $L_3/L_\perp^2, \dots, L_k/L_\perp^{k-1}, L_\perp^d/f_{min}$ and $f_{max}/L_\perp^d$ are large enough (depending only on d and k), for all $n \geq 1$,*

$$\inf_{\hat{T}} \sup_{P \in \mathcal{P}_{(x)}^k} \mathbb{E}_{P^{\otimes n}} \angle\big(T_x M, \hat{T}\big) \geq \frac{1}{2} > 0,$$

where the infimum is taken over all the estimators $\hat{T} = \hat{T}\big(X_1, \dots, X_n\big)$.

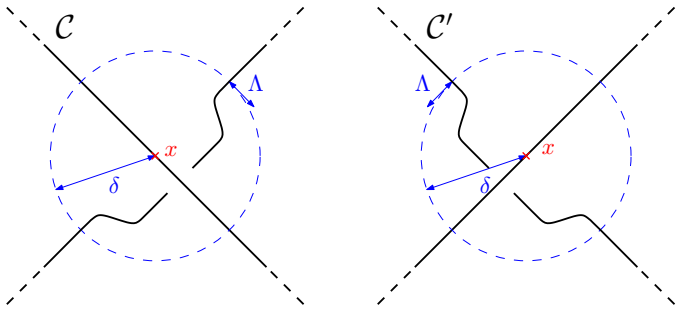


Figure 8: Hypotheses for minimax lower bound on tangent space estimation with $\tau_{min} = 0$.

Moreover, for any $D \geq d+1$, provided that $L_3/L_\perp^2, \dots, L_k/L_\perp^{k-1}, L_\perp^d/f_{min}$ and $f_{max}/L_\perp^d$ are large enough (depending only on d and k), for all $n \geq 1$,

$$\inf_{\widehat{II}} \sup_{P \in \mathcal{P}^k_{(x)}} \mathbb{E}_{P^{\otimes n}} \left\| II_x^M \circ \pi_{T_xM} - \widehat{II} \right\|_{op} \geq \frac{L_\perp}{4} > 0,$$

where the infimum is taken over all the estimators $\widehat{II} = \widehat{II}(X_1, \dots, X_n)$.

We will make use of Le Cam's Lemma, which we recall here.

THEOREM C.20 (Le Cam's Lemma [36]). *For all pairs P, P' in $\mathcal{P}$,*

$$\inf_{\hat{\theta}} \sup_{P \in \mathcal{P}} \mathbb{E}_{P^{\otimes n}} d(\theta(P), \hat{\theta}) \geq \frac{1}{2} d\left(\theta(P), \theta(P')\right) \left\| P \wedge P' \right\|_1^n,$$

where the infimum is taken over all the estimators $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$.

PROOF OF THEOREM C.19. For $\delta \geq \Lambda > 0$, let $\mathcal{C}, \mathcal{C}' \subset \mathbb{R}^3$ be closed curves of the Euclidean space as in Figure 8, and such that outside the figure, $\mathcal{C}$ and $\mathcal{C}'$ coincide and are $\mathcal{C}^\infty$. The bumped parts are obtained with a smooth diffeomorphism similar to (7) and centered at x. Here, δ and Λ can be chosen arbitrarily small.

Let $\mathcal{S}^{d-1} \subset \mathbb{R}^d$ be a $d-1$-sphere of radius $1/L_\perp$. Consider the Cartesian products $M_1 = \mathcal{C} \times \mathcal{S}^{d-1}$ and $M_1' = \mathcal{C}' \times \mathcal{S}^{d-1}$. M_1 and M_1' are subsets of $\mathbb{R}^{d+3} \subset \mathbb{R}^D$. Finally, let P_1 and P_1' denote the uniform distributions on M and M'. Note that M, M' can be built by homothecy of ratio $\lambda = 1/L_\perp$ from some unitary scaled $M_1^{(0)}, M_1'^{(0)}$, similarly to Section 5.3.2 in [2], yielding, from Proposition A.4, that P_1, P_1' belong to $\mathcal{P}^k_{(x)}$ provided that

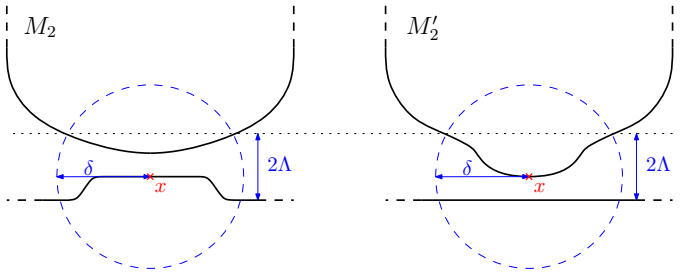


Figure 9: Hypotheses for minimax lower bound on curvature estimation with $\tau_{min} = 0$.

$L_3/L_\perp^2, \dots, L_k/L_\perp^{k-1}, L_\perp^d/f_{min}$ and $f_{max}/L_\perp^d$ are large enough (depending only on d and k), and that Λ, δ and Λ^k/δ are small enough. From Le Cam's Lemma C.20, we have for all $n \geq 1$,

$$\inf_{\hat{T}} \sup_{P \in \mathcal{P}^k_{(x)}} \mathbb{E}_{P^{\otimes n}} \angle\big(T_x M, \hat{T}\big) \geq \frac{1}{2} \angle\big(T_x M_1, T_x M_1'\big) \left\| P_1 \wedge P_1' \right\|_1^n .$$

By construction, $\angle\big(T_x M_1, T_x M_1'\big) = 1$, and since $\mathcal{C}$ and $\mathcal{C}'$ coincide outside $\mathcal{B}_{\mathbb{R}^3}(0, \delta)$,

$$\begin{aligned} \left\| P_1 \wedge P_1' \right\|_1 &= 1 - Vol\left(\left(\mathcal{B}_{\mathbb{R}^3}(0,\delta) \cap \mathcal{C} \right) \times \mathcal{S}^{d-1} \right) / Vol\left(\mathcal{C} \times \mathcal{S}^{d-1} \right) \\ &= 1 - Length\left(\mathcal{B}_{\mathbb{R}^3}(0,\delta) \cap \mathcal{C} \right) / Length(\mathcal{C}) \\ &\geq 1 - c_{L_\perp} \delta. \end{aligned}$$

Hence, at fixed $n \geq 1$, letting Λ, δ go to 0 with Λ^k/δ small enough, we get the announced bound.

We now tackle the lower bound on curvature estimation with the same strategy. Let $M_2, M_2' \subset \mathbb{R}^D$ be d-dimensional submanifolds as in Figure 9: they both contain x, the part on the top of M_2 is a half d-sphere of radius $2/L_\perp$, the bottom part of M_2' is a piece of a d-plane, and the bumped parts are obtained with a smooth diffeomorphism similar to (7), centered at x. Outside $\mathcal{B}(x, \delta)$, M_2 and M_2' coincide and connect smoothly the upper and lower parts. Let P_2, P_2' be the probability distributions obtained by the pushforward given by the bump maps. Under the same conditions on the parameters as previously, P_2 and P_2' belong to $\mathcal{P}^k_{(x)}$ according to Proposition

A.4. Hence from Le Cam's Lemma C.20 we deduce

$$\inf_{\widehat{II}} \sup_{P\in\mathcal{P}^k_{(x)}} \mathbb{E}_{P^{\otimes n}} \left\| II^M_x \circ \pi_{T_xM} - \widehat{II} \right\|_{op} \geq \frac{1}{2} \left\| II^{M_2}_x \circ \pi_{T_xM_2} - II^{M_2'}_x \circ \pi_{T_xM_2'} \right\|_{op} \left\| P_2 \wedge P_2' \right\|_1^n .$$

But by construction, $\left\| II^{M_2}_x \circ \pi_{T_xM_2} \right\|_{op} = 0$, and since M_2' is a part of a sphere of radius $2/L_\perp$ nearby x, $\left\| II^{M_2'}_x \circ \pi_{T_xM_2'} \right\|_{op} = L_\perp/2$. Hence,

$$\left\| II^{M_2}_x \circ \pi_{T_xM_2} - II^{M_2'}_x \circ \pi_{T_xM_2'} \right\|_{op} \geq L_\perp/2.$$

Moreover, since P_2 and P_2' coincide on $\mathbb{R}^D \setminus \mathcal{B}(x,\delta)$,

$$\left\| P_2 \wedge P_2' \right\|_1 = 1 - P_2(\mathcal{B}(x,\delta)) \geq 1 - c_{d,L_\perp}\delta^d.$$

At $n \geq 1$ fixed, letting Λ, δ go to 0 with Λ^k/δ small enough, we get the desired result.

□

Department of Mathematics
University of California San Diego
9500 Gilman Dr. La Jolla
CA 92093
United States
E-mail: eaamari@ucsd.edu
URL: http://www.math.ucsd.edu/ eaamari/

Laboratoire de probabilités et modèles aléatoires
Bâtiment Sophie Germain
Université Paris-Diderot
75013 Paris
France
E-mail: levrard@math.univ-paris-diderot.fr
URL: http://www.normalesup.org/ levrard/