Non-Asymptotic Rates for Manifold, Tangent Space, and Curvature Estimation
Eddie Aamari, Clément Levrard

To cite this version:

HAL Id: hal-01516032
https://hal.archives-ouvertes.fr/hal-01516032
Submitted on 2 May 2017

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

Eddie Aamari∗†‡
Laboratoire de Mathématiques d’Orsay, INRIA Saclay
Clément Levrard∗†
Université Paris Diderot

Abstract

Given an \( n \)-sample drawn on a submanifold \( M \subset \mathbb{R}^D \), we derive optimal rates for the estimation of tangent spaces \( T_X M \), the second fundamental form \( II_X \), and the submanifold \( M \). After motivating their study, we introduce a quantitative class of \( C^k \)-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 [Was16]. 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 \( X = \{X_1, \ldots, X_n\} \) 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 [NSW08], persistent homology [FLR+14], volume [APR16], differential quantities [CP05] or the submanifold itself [GPPVW12, AL15].

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.

∗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
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 learning. Indeed, there exists a wide bunch of algorithms intended to reconstruct submanifolds from point clouds (Isomap [TSL00], LLE [RS00], and restricted Delaunay Complexes [BG14, CDR05] for instance), but a few come with theoretical guarantees [GPPVW12, AL15]. Up to our knowledge, a minimax lower bound has proved optimality of a reconstruction scheme in only one case [GPPVW12]. Some of these reconstruction procedures are based on tangent space estimation [BG14, AL15, CDR05]. Tangent space estimation itself also yields interesting applications in manifold clustering [GM11, ALZ13]. 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 [MOG11, GWM01]. Several algorithms have been proposed in that case [Rus04, CP05, MOG11, GWM01], 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 $\tau_M$, a generalized convexity parameter. Assuming $\tau_M \geq \tau_{\text{min}} > 0$ involves both local regularity — a bound on curvature — and global regularity — no arbitrarily pinched area. This $C^2$-like assumption has been extensively used in the computational geometry and geometric inference fields [AL15, NSW08, FLR+14, APR16, GPPVW12]. One attempt of a specific investigation for higher orders of regularity $k \geq 3$ has been proposed in [CP05].

However, 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 $C^2$ case [AL15] to $(1/n)^\alpha$ with $1/d < \alpha < 2/d$ in more regular settings [SW12, TVF13]. In addition, it seems that PCA-based estimators are outperformed by estimators taking into account higher orders of smoothness [CC16, CP05], for regularities at least $C^3$. For instance fitting quadratic terms lead to a convergence rate of order $(1/n)^{3/d}$ in [CC16]. 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 [CP05]. 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 Contribution

In this paper, we build a collection of models for $C^k$-submanifolds ($k \geq 3$) that naturally generalize the commonly used one for $k = 2$ (Section 2). We emphasize the necessity of both local and global constraints for estimation. On these models, we study the non-asymptotic rates of estimation for tangent
space, second fundamental form, and submanifold estimation (Section 3). These results shed light on the influence of $k$, $d$, $D$ and $n$ on these estimation problems, showing for instance that the ambient dimension $D$ plays no role. The estimators proposed all rely on the analysis of local polynomials, and allow to deal with the three estimation problems in a unified way (Section 5.1). Minimax lower bounds are derived using standard Bayesian techniques, although a new version of Assouad’s Lemma is used for tangent spaces and second fundamental forms when the base point is random (Section 5.2). For the sake of completeness, geometric background and proofs of technical lemmas are given in the Appendix.

2 $C^k$ 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 $C^2$. For all $p \in M$, $T_p M$ stands for the tangent space of $M$ at $p$ [dC92, Chapter 0]. We let $\Pi^M_p : T_p M \times T_p M \to T_p M^\perp$ denote the second fundamental form of $M$ at $p$ [dC92, p. 125]. $\Pi^M_p$ 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 $B((p,r)$ for the closed Euclidean ball of radius $r > 0$ centered at $p \in \mathbb{R}^D$, and for short $B_T(p,r) = B(p,r) \cap T$. For a smooth function $\Phi : \mathbb{R}^D \to \mathbb{R}^D$ and $i \geq 1$, we let $d^i_x \Phi$ denote the $i$th order differential of $\Phi$ at $x \in \mathbb{R}^D$. For a linear map $A$ defined on $T \subset \mathbb{R}^D$, $\|A\|_{op} = \sup_{v \in T} \|Av\|/\|v\|$ 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, for short, its $L^\infty$ operator norm is denoted by $\|\cdot\|_{op}$, $\sup_{x \in T'} \|A_x\|_{op}$. 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 $\angle(U,V) = \|\pi_U - \pi_V\|_{op}$. The Hausdorff distance $d_{H} \mathbb{P}_{\mathbb{P}}^{V_{2}}$ in $\mathbb{R}^D$ is denoted by $d_{H}$. For a probability distribution $\mathbb{P}$, $E_{\mathbb{P}}$ stands for the expectation with respect to $\mathbb{P}$. We write $\mathbb{P}^\otimes n$ for the $n$-times tensor product of $\mathbb{P}$.

Throughout this paper, $C_\alpha$ will denote a generic constant depending on the parameter $\alpha$. For clarity’s sake, $C'_\alpha$, $c_\alpha$, or $c'_\alpha$ may also be used when several constants are involved.

2.2 Reach and Regularity of Submanifolds

As introduced in [Fed59], the reach $\tau_M$ of a subset $M \subset \mathbb{R}^D$ is the maximal neighborhood radius for which the projection $\pi_M$ 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 on \( M \), that is

\[
\operatorname{Med}(M) = \{ z \in \mathbb{R}^D | \exists p \neq q \in M, \|z - p\| = \|z - q\| = d(z, M) \}.
\]

The reach is then defined by

\[
\tau_M = \inf_{p \in M} d(p, \operatorname{Med}(M)) = \inf_{z \in \operatorname{Med}(M)} d(z, M).
\]

It gives a minimal scale of geometric and topological features of \( M \). As a generalized convexity parameter, \( \tau_M \) is a key parameter in reconstruction [AL15, GPPVW12] and in topological inference [NSW08]. Having \( \tau_M \geq \tau_{\min} > 0 \) prevents \( M \) from almost auto-intersecting, and bounds its curvature in the sense that \( \|H^1_p\|_{op} \leq \tau_{\min}^{-1} \) for all \( p \in M \) [NSW08, Proposition 6.1].

For \( \tau_{\min} > 0 \), we let \( C^2_{\tau_{\min}} \) 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 \( M \in C^2_{\tau_{\min}} \) is the existence of a parametrization closely related to the projection onto tangent spaces. We let \( \exp_p : T_pM \rightarrow M \) denote the geodesic map [dC92, 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 \( M \in C^2_{\tau_{\min}} \), \( \exp_p : B_{T_pM}(0, \tau_{\min}/4) \rightarrow M \) is one-to-one. Moreover, it can be written as

\[
\exp_p : B_{T_pM}(0, \tau_{\min}/4) \rightarrow M
\]

\[
v \mapsto p + v + N_p(v)
\]

with \( N_p \) such that for all \( v \in B_{T_pM}(0, \tau_{\min}/4) \),

\[
N_p(0) = 0, \quad d_0N_p = 0, \quad \|d_vN_p\|_{op} \leq L_{\perp} \|v\|,
\]

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

In other words, elements of \( C^2_{\tau_{\min}} \) 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. In addition, the reach condition provides an order 2 Taylor expansion of the submanifold on top of its tangent spaces. A natural extension to \( C^k \)-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 \( C^k_{\tau_{\min},L} \).
Definition 2. For $k \geq 3$, $\tau_{\min} > 0$, and $L = (L_1, L_3, \ldots, L_k)$, we let $C^k_{\tau_{\min}, L}$ denote the set of $d$-dimensional compact connected submanifolds $M$ of $\mathbb{R}^D$ with $\tau_M \geq \tau_{\min}$ and such that, for all $p \in M$, there exists a local one-to-one parametrization $\Psi_p$ of the form:

$$
\Psi_p : B_{T_p M}(0, r) \rightarrow M
$$

$$
v \mapsto p + v + \mathbf{N}_p(v)
$$

for some $r \geq \frac{1}{8L_1}$, with $\mathbf{N}_p \in C^k(B_{T_p M}(0, r), \mathbb{R}^D)$ such that

$$
\mathbf{N}_p(0) = 0, \quad d_0 \mathbf{N}_p = 0, \quad \|d^2 \mathbf{N}_p\|_{op} \leq L_1,
$$

for all $\|v\| \leq \frac{1}{8L_1}$. Furthermore, we require that

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

Let us precise that the radius $1/(8L_1)$ 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 $\Psi_p$’s can be seen as unit parametrizations of $M$. The conditions on $\mathbf{N}_p(0)$, $d_0 \mathbf{N}_p$, and $d^2 \mathbf{N}_p$ ensure that $\Psi_p^{-1}$ is close to the projection $\pi_{T_p M}$. The bounds on $d^i \mathbf{N}_p$ ($3 \leq i \leq k$) allow to control the coefficients of the polynomial expansion we seek. Indeed, whenever $M \in C^k_{\tau_{\min}, L}$, Lemma 12 shows that for every $p$ in $M$, and $y$ in $B(p, \frac{\tau_{\min} \wedge L_1^{-1}}{4}) \cap M$,

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

(1)

where $\pi^*$ denotes the orthogonal projection onto $T_p M$, the $T_i^*$ are i-linear maps from $T_p M$ 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 lengths $L_i$’s depend on the parameters $\tau_{\min}, d, k, L_1, \ldots, L_k$.

Such $\Psi_p$’s exist for any compact $C^k$-submanifold, if one allows $\tau_{\min}^{-1}, L_1, L_3, \ldots, L_k$ to be large enough. Note that for $k \geq 3$ the exponential map can happen to be only $C^{k-2}$ for a $C^k$-submanifold [Har51]. Hence, it may not be a good choice of $\Psi_p$. However, for $k = 2$, taking $\Psi_p = \exp_p$ is sufficient for our purpose. For ease of notation, we may write $C^2_{\tau_{\min}, L}$ although the specification of $L$ is useless. In this case, we implicitly set by default $\Psi_p = \exp_p$ and $L_1 = 5/(4 \tau_{\min})$.

As will be shown in Theorem 5 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 $M \in C^k_{\tau_{\min}, L}$, as stated below.
Definition 3. For \( k \geq 2 \), \( \tau_{\min} > 0 \), \( \mathbf{L} = (L_\perp, L_3, \ldots, L_k) \) and \( f_{\min} \leq f_{\max} \), we let \( \mathcal{P}^{\tau_{\min}, \mathbf{L}}(f_{\min}, f_{\max}) \) denote the set of distributions \( P \) with support \( M \in \mathcal{C}^k_{\min, \mathbf{L}} \) that have a density \( f \) with respect to the volume measure on \( M \), and such that for all \( x \in M \),

\[
0 < f_{\min} \leq f(x) \leq f_{\max} < \infty.
\]

For short, we write \( \mathcal{P}^k \) when there is no ambiguity. We denote by \( X_n \) an i.i.d. \( n \)-sample \( \{X_1, \ldots, X_n\} \), that is, a sample with distribution \( P^{\otimes n} \) for some \( P \in \mathcal{P}^k \). 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. Here, since we deal with submanifolds, a natural way to model deformations is through ambient diffeomorphisms.

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

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

Proposition 4 follows from a geometric reparametrization argument (Proposition 21) and a change of variable result for the Hausdorff measure (Lemma 22).

2.3 Necessity of a Global Assumption

In the previous Section 2.2, we generalized \( C^2 \)-like models — stated in terms of reach — to \( C^k \) for \( k \geq 3 \) by imposing higher order differentiability bounds on parametrizations \( \Psi_p \)'s. Though, we did not drop the global assumption \( \tau_M \geq \tau_{\min} > 0 \). Indeed, it appears that such an assumption is necessary for estimation purpose.

Theorem 5. Assume that \( \tau_{\min} = 0 \). If \( D \geq d + 3 \), then for all \( k \geq 3 \) and \( L_\perp > 0 \), provided that \( L_3/L_\perp^2, \ldots, L_k/L_\perp^{k-1}, L_d/L_\perp f_{\min} \) and \( f_{\max}/L_\perp^d \) are large enough (depending only on \( d \) and \( k \)), for all \( n \geq 1 \),

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

where the infimum is taken over all the estimators \( \hat{T} = \hat{T}(X_1, \ldots, X_n) \).
Moreover, for any $D \geq d + 1$, provided that $L_3/L_2, \ldots, L_k/L_{k-1}, L_4/f_{\min}$ and $f_{\max}/L_4$ are large enough (depending only on $d$ and $k$), for all $n \geq 1$,

$$
\inf_{\widehat{II}} \sup_{P \in P_{(\epsilon)}} \mathbb{E}_{P^{\otimes n}} \left\| II^M_x \circ \pi_{T_x M} - \widehat{II} \right\|_{op} \geq \frac{L_4}{4} > 0,
$$

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

In other words, if the class of submanifolds is allowed to have arbitrarily small reach, no estimator can perform uniformly well to estimate neither $T_x M$ nor $II^M_x$. 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 1 and 2).

### 3 Main Results

Let us now move to the description of the main results, that consist of minimax upper and lower bounds for each object of interest. Given an i.i.d. $n$-sample $X_n = \{X_1, \ldots, X_n\}$ with unknown common distribution $P \in \mathcal{P}_k$ having support $M$, we detail non-asymptotic rates for the estimation of tangent spaces $T_{X_j} M$, second fundamental forms $II^M_{X_j}$, and $M$ itself.
For this, we need one more piece of notation. For \(1 \leq j \leq n\), \(P_n^{(j)}\) denotes integration with respect to \(1/(n-1) \sum_{i \neq j} \delta(X_i - X_j)\), and \(\delta^O\) denotes the \(D \times i\)-dimensional vector \((y, \ldots, y)\). For a constant \(t > 0\) and a bandwidth \(h > 0\) to be chosen later, we define the local polynomial estimator \((\hat{\pi}_j, \hat{T}_{2,j}, \ldots, \hat{T}_{k-1,j})\) at \(X_j\) to be any element of

\[
\arg\min_{\pi, \sup_{2 \leq i \leq k} \|T_i\|_{op} \leq t} P_n^{(j)} \left[ \left\| x - \pi(x) - \sum_{i=2}^{k-1} T_i(\pi(x)^{\otimes i}) \right\|^2 \mathbf{1}_{B(0,h)}(x) \right],
\]

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

\[h_0 = \frac{\tau_{\min} \wedge L^{-1}}{8}.\]

Note that the set of \(d\)-dimensional orthogonal projectors is not convex, leading to a more involved optimization problem than usual least squares. In practice, this problem may be solved using tools from optimization on Grassman manifolds [UM14], or adopting a two-stage procedure such as in [CP05]: 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 \((\pi_j, T_{2,j}, \ldots, T_{k,j})\) in this basis under orthogonality constraints. It is worth mentioning that a similar problem is explicitly solved in [CC16], leading to an optimal tangent space estimation procedure in the case \(k = 3\).

The constraint \(\|T_i\|_{op} \leq t\) 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 \(\| \cdot \|_{op}\)-penalized least squares minimization — as in ridge regression — leading to the same results.

### 3.1 Tangent Spaces

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

**Theorem 6.** Assume that \(t \geq C_{k,d,\tau_{\min},L} \geq \sup_{2 \leq i \leq k} \|T^*_i\|_{op}\). Set \(h = \left(C_{d,k} \log(n)f_2^{max}f_{\min}^{-1} \right)^{1/d}\), for \(C_{d,k}\) large enough. If \(n\) is large enough so that \(h \leq h_0\), then with probability...
at least $1 - \left(\frac{1}{n}\right)^{k/d}$,

$$\max_{1 \leq j \leq n} \angle(T_x M, \hat{T}_j) \leq C_{d,k,\tau_{\min}\lambda} \sqrt{\frac{f_{\max}}{f_{\min}}} h^{k-1}(1 + th).$$

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

$$\sup_{P \in P_k} \mathbb{E}_{P \otimes n} \max_{1 \leq j \leq n} \angle(T_x M, \hat{T}_j) \leq C \left(\frac{\log(n)}{n - 1}\right)^{k-1},$$

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

The same bound holds for the estimation of $T_x M$ at a prescribed $x \in M$. For that, simply take $P_n(x) = 1/n \sum_i \delta(x_i - x)$ as integration in (2).

This result is in line with those of [CP05] 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 [AL15], where we used local PCA for estimation, that is a reformulation of (2). When $k \geq 3$, this procedure outperforms PCA-based estimators of [SW12] and [TVF13], where convergence rates of the form $(1/n)^{\alpha}$ is obtained for $1/d < \alpha < 2/d$. This bound also recovers the result of [CC16] in the case $k = 3$, where a similar procedure is used. Moreover, Theorem 6 nearly matches the following lower bound.

**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

$$\inf_{\hat{T}} \sup_{P \in P_k} \mathbb{E}_{P \otimes n} \angle(T_x M, \hat{T}) \geq c_{d,k,\tau_{\min}} \left(\frac{1}{n - 1}\right)^{k-1},$$

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

Hence, up to a log $n$ factor, the rate $n^{-(k-1)/d}$ is optimal for tangent space estimation on the model $P_k$. The rate $(\log n/n)^{1/d}$ obtained in [AL15] for $k = 2$ is therefore optimal, as well as the rate $(\log n/n)^{2/d}$ given in [CC16] for $k = 3$. The rate $n^{-(k-1)/d}$ naturally appears on the the model $P_k$, since it consists of $C^k$-submanifolds, and tangent spaces are differential objects of order 1, yielding the shift $k - 1$. Again, the same lower bound holds for the estimation of $T_x M$ at a fixed point $x$ in the model $P_k^{(x)}$. Interestingly, the tools used to derive the lower bound for $T_x M$ ($x$ fixed) is much less involved than for $T_x M$ ($X_1$ random and depending on the distribution $P$). In the latter case, a conditional Assouad’s Lemma (Lemma 16) is used. We will detail these differences in Section 5.2.

### 3.2 Curvature

The second fundamental form $II^M_{X_j} : T_X M \times T_X M \to T_X M \perp \subset \mathbb{R}^D$ is a symmetric bilinear map that encodes completely the curvature of $M$ at $X_j$.


Let \( k \geq 3 \). Take \( h \) as in Theorem 6 and \( T = 1/h \). If \( n \) is large enough so that \( h \leq h_0 \) and \( h^{-1} \geq C_{k,d,\tau_{\min},L}^{-1} \geq (\sup_{2 \leq i \leq k} ||T_i||_{op})^{-1} \), then with probability at least \( 1 - \left( \frac{1}{n} \right)^{k/d} \),

\[
\max_{1 \leq j \leq n} \left\| I_{X_j}^M \circ \pi_{T X_j} M - \hat{T}_{2,j} \circ \hat{\pi}_j \right\|_{op} \leq C_{d,k,\tau_{\min},L} \sqrt{\frac{f_{\max}}{f_{\min}}} h^{k-2}.
\]

In particular, for \( n \) large enough,

\[
\sup_{P \in \mathcal{P}_n} \mathbb{E}_{P \in \mathcal{P}} \max_{1 \leq j \leq n} \left\| I_{X_j}^M \circ \pi_{T X_j} M - \hat{T}_{2,j} \circ \hat{\pi}_j \right\|_{op} \leq C \left( \frac{\log(n)}{n - 1} \right)^{\frac{k-2}{2}},
\]

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

Interestingly, Theorems 6 and 8 are enough to provide estimators of various notions of curvature. For instance, consider the scalar curvature \([dC92, \text{Section 4.4}]\) at a point \( X_j \), defined by

\[
Sc_{X_j}^M = \frac{1}{d(d-1)} \sum_{r \neq s} \left[ \left\langle I_{X_j}^M(\hat{e}_r,e_r), I_{X_j}^M(e_s,e_s) \right\rangle - \left\| I_{X_j}^M(e_r,e_s) \right\|^2 \right],
\]

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

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

where \( (\hat{e}_r)_{1 \leq r \leq d} \) is an orthonormal basis of \( \hat{T}_{X_j} M \). Theorems 6 and 8 yield

\[
\mathbb{E}_{P \in \mathcal{P}_n} \max_{1 \leq j \leq n} \left| \hat{Sc}_j - Sc_{X_j}^M \right| \leq C_{d,k,\tau_{\min},L,f_{\min},f_{\max}} \left( \frac{\log(n)}{n - 1} \right)^{\frac{k-2}{2}}.
\]

The (near-)optimality of the bound stated in Theorem 8 is assessed by the following lower bound.

**Theorem 9.** If \( \tau_{\min} L^{-1}, \ldots, \tau_{\min}^{k-1} L_k, (\tau_{\min}^{d} f_{\min})^{-1} \) and \( \tau_{\min}^{d} f_{\max} \) are large enough (depending only on \( d \) and \( k \)), then

\[
\inf \sup_{P \in \mathcal{P}_n} \mathbb{E}_{P \in \mathcal{P}_n} \left\| I_{X_j}^M \circ \pi_{T X_j} M - \hat{I} \right\|_{op} \geq C_{d,k,\tau_{\min}} \left( \frac{1}{n - 1} \right)^{\frac{k-2}{2}},
\]

where the infimum is taken over all the estimators \( \hat{I} = \hat{I}(X_1, \ldots, X_n) \).
Figure 3: $\hat{M}$ is a union of polynomial patches at sample points.

The same remarks as in Section 3.1 hold. If the estimation problem consists in approximating $H^x_M$ at a fixed point $x$ known to belong to $M$ beforehand, we obtain the same rate. The ambient dimension $D$ still plays no role. The shift $k - 2$ in the rate of convergence on a $C^k$-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 9) does not require $k \geq 3$. Hence, we get that for $k = 2$, curvature cannot be estimated uniformly consistently on the $C^2$-model $P^2$. This seems natural, since the estimation of a second order quantity should require an additional degree of smoothness.

3.3 Support Estimation

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

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

is a good candidate for an approximate parametrization in a neighborhood of $X_j$. We do not know the domain $T_{X_j}M$ 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 $M$ be

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

The set $\hat{M}$ has no reason to be globally smooth, since it consists of a union of polynomial patches that are not linked together (Figure 3). However, $\hat{M}$ is provably close to $M$ for the Hausdorff distance.

**Theorem 10.** With the same assumptions as Theorem 8 with probability at least $1 - 2 \left( \frac{1}{n} \right)^{\frac{k}{2}}$, we have

$$d_H(M, \hat{M}) \leq C_{d,k,\tau_{\min},L,f_{\min},f_{\max}} h^{\frac{k}{2}}.$$
In particular, for \( n \) large enough,

\[
\sup_{P \in P^k} \mathbb{E}_{P \subseteq \pi} d_H(M, \hat{M}) \leq C \left( \frac{\log(n)}{n - 1} \right)^{\frac{k}{2}},
\]

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

For \( k = 2 \), we recover the rate \((\log n/n)^{2/d}\) obtained in [AL15, GPPVW12, KZ15]. 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 [AL15] does.

When \( k \geq 3 \), \( \hat{M} \) outperforms reconstruction procedures based on a somewhat piecewise linear interpolation [AL15, GPPVW12], 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 [AL15, GPPVW12]. Theorem 10 is now proved to be (almost) minimax optimal.

**Theorem 11.** If \( \tau_{\min,L}, \tau_{\min,L}^{-1} \) and \( \tau_{\min,f_{\min}} \) are large enough (depending only on \( d \) and \( k \)), then for \( n \) large enough,

\[
\inf_{\hat{M}} \sup_{P \in P^k} \mathbb{E}_{P \subseteq \pi} d_H(M, \hat{M}) \geq c_{d,k,\tau_{\min}} \left( \frac{1}{n} \right)^{\frac{k}{2}},
\]

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

Theorem 11 is obtained from Le Cam’s Lemma (Theorem 15). Let us note that it is likely for the extra \( \log n \) term appearing in Theorem 10 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 \( L^\infty \) loss. The \( \log n \) term may be obtained in Theorem 11 with the same combinatorial analysis as in [KZ15] for \( k = 2 \).

### 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 \( C_{\tau_{\min},L}^k \) for submanifolds that extend naturally 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.

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 [UM14]. 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.

As a first attempt to a minimax study over models of higher order regularity \(C^k\) \((k \geq 3)\) for submanifolds, we chose not to include noise. This is a limitation of the model \(P^k\), and one could argue that the methods described are not robust. However, with outliers in the model \(C^2\), \(\text{[AL15]}\) 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 recovering with fast rates for \(k \geq 3\). Indeed, \(\hat{M}\) converges at rate \((\log(n/n)/n)^{k/d}\) but is unstructured. It would be nice to glue the patches of \(\hat{M}\) together, for example using interpolation techniques, following the ideas of \(\text{[FIK+15]}\).

5 Proofs

5.1 Local Polynomials

We now turn to the proof of the upper bounds of Section 3. First, to relate the existence of parametrizations \(\Psi_p\)’s to a local polynomial decomposition, the following lemma is needed.

Lemma 12. For any \(M \in C^k_{\tau_{\text{min}}}\) and \(x \in M\), the following holds.

(i) For all \(v_1, v_2 \in B_{T_x M} \left(0, \frac{1}{4L_{x,4}}\right)\),
\[
\frac{3}{4} \|v_2 - v_1\| \leq \|\Psi_x(v_2) - \Psi_x(v_1)\| \leq \frac{5}{4} \|v_2 - v_1\|.
\]

(ii) For all \(h \leq \frac{1}{4L_{x,4}} \wedge \frac{2\tau_{\text{min}}}{5}\),
\[
M \cap B \left(x, \frac{3h}{5}\right) \subset \Psi_x \left(B_{T_x M} \left(x, h\right)\right) \subset M \cap B \left(x, \frac{5h}{4}\right).
\]

(iii) For all \(h \leq \frac{\tau_{\text{min}}}{2}\),
\[
B_{T_x M} \left(0, \frac{7h}{8}\right) \subset \pi_{T_x M} \left(B \left(x, h\right) \cap M\right).
\]

(iv) Denoting \(\pi^* = \pi_{T_x M}\) the orthogonal projection onto \(T_x M\), for all \(x \in M\), there exist multilinear maps \(T^*_2, \ldots, T^*_k\) from \(T_x M\) to \(\mathbb{R}^D\), and \(R_k\) such that for all \(y \in B \left(x, \frac{\tau_{\text{min}}^k L_{x,4}^{-1}}{4}\right) \cap M\),
\[
y - x = \pi^* (y - x) + T^*_2 (\pi^* (y - x) \otimes 2) + \ldots + T^*_k (\pi^* (y - x) \otimes k - 1) + R_k (y - x),
\]
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 \( d, k, \tau_{\min}, L_1, \ldots, L_i \), and \( C \) on \( d, k, \tau_{\min}, L_1, \ldots, L_k \). Moreover, for \( k \geq 3 \), \( T^*_2 = \Pi_{x}^M \).

(v) For all \( x \in M \), \( \|\Pi_{x}^M\|_{op} \leq 1/\tau_{\min} \). In particular, the sectional curvatures of \( M \) satisfy
\[ -2/\tau_{\min} \leq \kappa \leq 1/\tau_{\min}. \]

The proof of Lemma 12 can be found in Section A.2.

We are now in position to analyze local polynomial estimators. For clarity’s sake, the bounds are given for \( j = 1 \), where we denote by \( \hat{\pi}, \hat{T}_i \) (\( 2 \leq i \leq k - 1 \)) the fitted polynomials of (2), and \( P_{n-1} = P^{(1)}_{n-1} \). The results of Theorems 6 and 8 then follow from a straightforward union bound. We also set \( k \geq 3 \), the case \( k = 2 \) proceeding from the same derivation, omitting the higher order tensors. Without loss of generality, we can assume that \( X_1 = 0 \) and that \( T_0M \) is spanned by the first \( d \) vectors of the canonical basis, so that \( \pi^*(x) = (x_1, \ldots, x_d, 0, \ldots, 0) = (x_1, d, 0, \ldots, 0) \).

Recall that \( h_0 = (\tau_{\min} \wedge \Pi_{x}^M)/8 \). According to Lemma 12 if \( M \in \mathcal{C}_{\tau_{\min}, L}^\ast \), for any \( x \in M \) such that \( \|x\| \leq h_0 \), we may write
\[ x = \pi^*(x) + T^*_2(\pi^*(x)^{\otimes 2}) + \ldots + T^*_{k-1}(\pi^*(x)^{\otimes k-1}) + R_k(x), \]
where \( \|R_k(x)\| \leq C_{\tau_{\min}, L}\|x\|^k \). Every coordinate of \((\hat{T}_i - T^*_i)(\pi^*(x))\) may be thought of as a polynomial map in \( x_{1:d} \). Thus, proximity between \( \hat{T}_i \) and \( T^*_i \) will be first stated in terms of polynomial norm.

Let \( \mathbb{R}^k[x_{1:d}] \) denote the set of real-valued polynomial functions in \( d \) variables with degree less than \( k \). For \( Q \in \mathbb{R}^k[x_{1:d}] \), we denote by \( \|Q\|_2 \) the Euclidean norm of its coefficients, and by \( Q_h \) the polynomial defined by \( Q_h(x_{1:d}) = Q(hx_{1:d}) \). The following result relates the \( L^2(P_{n-1}) \) norm involved in (2) to polynomial norms.

**Proposition 13.** Set \( h = \left( \frac{K \log(n)}{n-1} \right)^\frac{1}{2} \). There exist constants \( \kappa_{k,d}, c_{k,d} \) and \( C_d \) such that, if \( K \geq (\kappa_{k,d}f_{\text{max}}^2/f_{\text{min}}^3) \) and \( n \) is large enough so that \( h \leq h_0 \leq \tau_{\min}/4 \), then with probability at least 1 \(-\left(\frac{1}{n}\right)^{\frac{3}{2} + 1}\), we have
\[ P_{n-1}[Q^2(\pi^*(x))\mathbb{1}_{\mathcal{B}(h)}(x)] \geq c_{k,d}h^d f_{\text{min}}\|Q_h\|_2^2, \]
\[ \leq C_d f_{\text{max}}(n - 1) h^d, \]
for every \( Q \in \mathbb{R}^k[x_{1:d}] \), where \( N(h) = \sum_{j=2}^{n} \mathbb{1}_{\mathcal{B}(0,h)}(X_j) \).
The proof of Proposition 13 is deferred to Section 5.2. From now on we assume that the probability event defined in Proposition 13 occurs. For short, with a slight abuse of notation, we denote by $T_{p,q}(x)$ the sum $T_{p}(x^\otimes p) + \ldots + T_{q}(x^\otimes q)$, and by $R_{n-1}(\pi,T_{2},\ldots,T_{k-1})$ the empirical criterion defined by (2).

Since for $t \geq \max_{i=2,\ldots,k-1} ||T_{i}||_{op}$,

$$R_{n-1}(\hat{\pi},\hat{T}_{1},\ldots,\hat{T}_{k-1}) \leq R_{n-1}(\pi^{*},T_{2}^{*},\ldots,T_{k-1}^{*}) \leq C_{\tau_{min},L}h^{2k}N(h)/(n-1)$$

according to (1), we may write

$$C_{\tau_{min},L}h^{2k}N(h)/(n-1) \geq R_{n-1}(\hat{\pi},\hat{T}_{2},\ldots,\hat{T}_{k-1})$$

$$= P_{n-1}\left(\left\| (\pi^{*} - \hat{\pi})(x) + (T_{2,k-1}^{*} \circ \pi^{*} - \hat{T}_{2,k-1} \circ \hat{\pi})(x) + R_{k}(x) \right\|^{2} \mathbf{1}_{B(0,h)}(x) \right),$$

with $||R_{k}(x)|| \leq C_{\tau_{min},L}h^{2k}$. It follows that

$$P_{n-1}\left(\left\| (\pi^{*} - \hat{\pi})(x) + (T_{2,k-1}^{*} \circ \pi^{*} - \hat{T}_{2,k-1} \circ \hat{\pi})(x) + R_{k}(x) \right\|^{2} \mathbf{1}_{B(0,h)}(x) \right)$$

$$\leq C_{\tau_{min},L}h^{2k}N(h)/(n-1)$$

$$\leq C_{\tau_{min},L}d_{f}f_{max}h^{d+2k}.$$

On the other hand, using (1) again yields, for $x \in B(0,h) \cap M$,

$$(\pi^{*} - \hat{\pi})(x) + (T_{2,k-1}^{*} \circ \pi^{*} - \hat{T}_{2,k-1} \circ \hat{\pi})(x)$$

$$= T_{1}(\pi^{*})(x) + T_{2}(\pi^{*} \otimes 2) + T_{3,k}(\pi^{*} \otimes 2) + \hat{\pi}(R_{k}(x)) + R_{k}'(x),$$

with $||R_{k}(x)|| \leq C_{\tau_{min},L}h^{k}$, $||R_{k}'(x)|| \leq tC_{\tau_{min},L}h^{k+1}$ since only tensors of order greater than 2 are involved in $R_{k}'$, and

$$T_{1}(\pi^{*})(x) = (\pi^{*} - \hat{\pi})\pi^{*}(x)$$

$$T_{2}(\pi^{*} \otimes 2) = (\pi^{*} - \hat{\pi})(T_{2}(\pi^{*} \otimes 2)) + (T_{2} \circ \pi^{*} - \hat{T}_{2} \circ \hat{\pi})(\pi^{*} \otimes 2).$$

Hence,

$$P_{n-1}\left(\left\| T_{1}(\pi^{*})(x) + T_{2}(\pi^{*} \otimes 2) + T_{3,k}(\pi^{*}) \right\|^{2} \mathbf{1}_{B(0,h)}(x) \right)$$

$$\leq C_{\tau_{min},L}d_{f}f_{max}h^{d+2k}(1 + ht). \quad (3)$$

The left-hand side of (3) may be decomposed coordinate-wise as

$$P_{n-1}\left(\left\| T_{1}(\pi^{*})(x) + T_{2}(\pi^{*} \otimes 2) + T_{3,k}(\pi^{*}) \right\|^{2} \mathbf{1}_{B(0,h)}(x) \right)$$

$$= \sum_{j=1}^{D} P_{n-1}\left(\left\| T_{1}^{(j)}(\pi^{*})(x) + T_{2}^{(j)}(\pi^{*} \otimes 2) + T_{3,k}^{(j)}(\pi^{*}) \right\|^{2} \mathbf{1}_{B(0,h)}(x) \right),$$

15
where for any tensor $T$, $T^{(j)}$ denotes the $j$-th coordinate of $T$ and is considered as a real valued $j$-order polynomial. Then, for every $j$, Proposition 13 leads to

$$P_{n-1} \left( \left( T^{(j)}_1 (x) + T^{(j)}_2 (x) + T^{(j)}_{3k}(x) \right)^2 \mathbf{I}_{B(0,h)}(x) \right) \geq c_{d,k} f_{\min} h^d \left\| \left( T^{(j)}_1 (x) + T^{(j)}_2 (x) + T^{(j)}_{3k}(x) \right) \right\|_2^2$$

$$= c_{d,k} f_{\min} h^d \sum_{i=1}^k \left\| \left( T^{(j)}_i (x) \right) \right\|_2^2.$$

Summing all contributions leads to

$$c_{d,k} f_{\min} \sum_{j=1}^D \sum_{i=1}^k \left\| \left( T^{(j)}_i (x) \right) \right\|_2^2 \leq C_{d,k,L,\tau_{\min}} f_{\max} h^2 (1 + t^2 h^2).$$

This entails

$$\|T^i\|_f \leq C_{d,k,L,\tau_{\min}} \frac{f_{\max}}{f_{\min}} h^2 (1 + t^2 h^2), \quad (4)$$

for $1 \leq i \leq k$, as well as

$$\|\pi^* - \hat{\pi}\| = \|T_{2,k-1} \circ \pi^* - \hat{T}_{2,k-1} \circ \hat{\pi}\| \leq C_{d,k,L,\tau_{\min}} \sqrt{\frac{f_{\max}}{f_{\min}}} h (1 + th), \quad (5)$$

for $x \in B(0,h) \cap M$, according to [1].

### 5.1.1 Bounds for Tangent Space Estimation

Noting that

$$\|T^i\|_f = \|\pi^* - \hat{\pi}\| = \|\pi^* \circ \hat{T}_i\| = \angle(T_0 M, \hat{T}_i)$$

from [GVL96] Section 2.6.2, and using (4) for $i = 1$ yields Theorem 6.

### 5.1.2 Bounds for Curvature Estimation

In accordance with assumptions of Theorem 8, we assume that $\max_{2 \leq i \leq k} \|T^*_i\|_f \leq t \leq 1/h$. Since

$$T^*_2 (\pi^* \circ \hat{T}_2 \circ \hat{\pi}) = (\pi^* \circ \hat{\pi}) (T^*_2 (\pi^* \circ \hat{T}_2 \circ \hat{\pi})) + (T^*_2 \circ \pi^* - \hat{T}_2 \circ \hat{\pi}) (\pi^* \circ \hat{T}_2 \circ \hat{\pi}),$$

we deduce that

$$\|T^*_2 - \hat{T}_2 \circ \hat{\pi}\| \leq \|T^*_2\|_f + \|\pi^* - \hat{\pi}\| + \|\hat{T}_2 - \hat{T}_2 \circ \hat{\pi}\|_f.$$  

Using (4) with $i = 1, 2$ and $th \leq 1$ leads to

$$\|T^*_2 \circ \pi^* - \hat{T}_2 \circ \hat{\pi}\|_f \leq C_{d,k,L,\tau_{\min}} \sqrt{\frac{f_{\max}}{f_{\min}}} h^{-2}.$$

Finally, Lemma 12 states that $H_{X_1}^M = T^*_2$, hence Theorem 8 is proved.
5.1.3 Bounds for Reconstruction

Let \( v \in B_{T_0}(0, 7h/8) \) be fixed. Notice that \( \pi^*(v) \in B_{T_0}(0, 7h/8) \). Hence, according to Lemma 12, there exists \( x \in B(0, h) \cap M \) such that \( \pi^*(v) = \pi^*(x) \).

We may write

\[
\hat{\Psi}(v) = v + k-1 \sum_{i=2}^{k-1} \hat{T}_i(v^{\otimes i}) = \pi^*(v) + k-1 \sum_{i=2}^{k-1} \hat{T}_i(\pi^*(v)^{\otimes i}) + R_k(v),
\]

where, since \( \|\hat{T}_i\|_{op} \leq 1/h, \|R_k(v)\| \leq C_{k,d,\tau_{min},L} \sqrt{f_{max}/f_{min} h^k} \) according to (4). Then, according to (5),

\[
\pi^*(v) + k-1 \sum_{i=2}^{k-1} \hat{T}_i(\pi^*(v)^{\otimes i}) = \pi^*(x) + k-1 \sum_{i=2}^{k-1} T_i(\pi^*(x)^{\otimes i}) + R'(\pi^*(x)),
\]

where \( \|R'(\pi^*(x))\| \leq C_{k,d,\tau_{min},L} \sqrt{f_{max}/f_{min} h^{k+1}} \). According to Lemma 12, we deduce that \( \|\hat{\Psi}(v) - x\| \leq C_{k,d,\tau_{min},L} \sqrt{f_{max}/f_{min} h^k} \), hence

\[
\sup_{u \in M} d(u, M) \leq C_{k,d,\tau_{min},L} \sqrt{f_{max}/f_{min} h^k}.
\]

(6)

Now we focus on \( \sup_{x \in M} d(x, M) \). For this, we need a lemma ensuring that \( X_n = \{X_1, \ldots, X_n\} \) covers \( M \) with high probability.

**Lemma 14.** Let \( h = \left( \frac{C'_d k \log n}{f_{min}} \right)^{1/d} \) with \( C'_d \) large enough. Then for \( n \) large enough so that \( h \leq \tau_{min}/8 \), with probability at least \( 1 - \left( \frac{1}{n} \right)^{k/d} \),

\[
d_H(M, X_n) \leq h.
\]

The proof of Lemma 14 is given in Section B.1. Now we choose \( h \) satisfying the conditions of Proposition 13 and Lemma 14. Let \( x \) be in \( M \) and assume that \( \|x - X_{j_0}\| \leq h \). According to (5) and (4), we deduce that \( \|\hat{\Psi}_{j_0}(\pi_{j_0}(x)) - x\| \leq C_{k,d,\tau_{min},L} \sqrt{f_{max}/f_{min} h^k} \). Hence, from Lemma 14

\[
\sup_{x \in M} d(x, M) \leq C_{k,d,\tau_{M},L} \sqrt{f_{max}/f_{min} h^k}.
\]

(7)

with probability at least \( 1 - 2 \left( \frac{1}{n} \right)^{k/d} \). Combining (6) and (7) gives Theorem 10.
5.2 Minimax Lower Bounds

This section is devoted to describe the main ideas of the proofs of the minimax lower bounds, Theorems 7, 9 and 11. The methods we use rely on hypothesis comparison \[Yu97\]. We recall that for two distributions \( Q \) and \( Q' \) defined on the same space, the total variation distance \( \text{TV}(Q, Q') \) and the \( L^1 \) test affinity \( \|Q \wedge Q'\|_1 \) are given by

\[
\text{TV}(Q, Q') = \frac{1}{2} \int |dQ - dQ'|,
\]

\[
\|Q \wedge Q'\|_1 = \int dQ \wedge dQ',
\]

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

5.2.1 Le Cam’s Lemma and Consequences

The first technique we use, involving only two hypotheses, is usually referred to as Le Cam’s Lemma. Let \( P \) be a model and \( \theta(P) \) be the parameter of interest. Assume that \( \theta(P) \) belongs to a pseudo-metric space \((D, d)\), 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 15** (Le Cam’s Lemma [Yu97]). For all \( P, P' \) in the model \( P \),

\[
\inf_{\hat{\theta}} \sup_{P \in P} E_{P \otimes n} d(\theta(P), \hat{\theta}) \geq \frac{1}{2} d(\theta(P), \theta(P')) \|P^\otimes n \wedge P'^\otimes n\|_1,
\]

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

Moreover, \( \|P^\otimes n \wedge P'^\otimes n\|_1 \geq \|P \wedge P'\|_1^n = (1 - \text{TV}(P, P'))^n \).

We derive Theorem 11 as well as Theorems 7 and 9 with fixed base point \( x \), \( \theta(P) \) being \( \text{supp}(P) = M, T_x M \) and \( H_x^M \circ \pi_{T_x M} \) respectively. The hypotheses \( P, P' \) are built in Section 5.2.3. Such constructions are not substantially new in minimax geometric inference [GPPVW12]. Therefore, we do not detail it further.

5.2.2 Conditional Assouad’s Lemma

Now, consider the estimation of the differential quantities \( T_{X_1} M \) and \( H_{X_1}^M \) with random base point \( X_1 \). In both cases, the loss can be cast as

\[
E_{P^\otimes n} d(\theta_{X_1}(P), \hat{\theta}) = E_{P^\otimes n-1} \left[ E_{P} d(\theta_{X_1}(P), \hat{\theta}) \right]
\]

\[
= E_{P^\otimes n-1} \left[ \left\| d(\theta(P), \hat{\theta}) \right\|_{L^1(P)} \right],
\]

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 \( \left\| d(\theta(P), \hat{\theta}) \right\|_{L^1(P)} \) obviously depends on \( P \), the technique exposed in the previous section does not apply anymore. However, a slight
adaptation of Assouad’s Lemma [Yu97] 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 \( \mu \) and \( \nu \) respectively. Let \((\mathcal{D}, d)\) be a pseudo-metric space. For \( Q \in \mathcal{Q} \), we let \( \theta_0(Q) : \mathcal{X} \to \mathcal{D} \) be defined \( \mu \)-almost surely, where \( \mu \) is the marginal distribution of \( Q \) on \( \mathcal{X} \). The parameter of interest is \( \theta_\mathcal{X}(Q) \), and the associated minimax risk over \( \mathcal{Q} \) is

\[
\inf_{\theta} \sup_{Q \in \mathcal{Q}} \mathbb{E}_{Q} \left[ d(\theta_\mathcal{X}(Q), \hat{\theta}(X, X')) \right],
\]

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 \( Q \) on \( \mathcal{X} \times \mathcal{X}' \), write \( \text{Conv}(Q) \) for the set of mixture probability distributions with components in \( Q \). For all \( \tau = (\tau_1, \ldots, \tau_m) \in \{0,1\}^m \), \( \tau^k \) denotes the \( m \)-tuple that differs from \( \tau \) only at the \( k \)th position. We are now in position to state the conditional version of Assouad’s Lemma that allows to lower bound the minimax risk (8).

**Lemma 16 (Conditional Assouad).** Let \( m \geq 1 \) be an integer and let \( \{Q_\tau\}_{\tau \in \{0,1\}^m} \) be a family of \( 2^m \) submodels \( Q_\tau \subset \mathcal{Q} \). 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 \( Q_\tau \in \mathcal{Q}_\tau \), \( \theta_\mathcal{X}(Q_\tau) \in \mathcal{D}_{\tau,k} \) on the event \( \{X \in U_k\} \);

- for all \( \theta \in \mathcal{D}_{\tau,k} \) and \( \theta' \in \mathcal{D}_{\tau',k} \), \( d(\theta, \theta') \geq \Delta \).

For all \( \tau \in \{0,1\}^m \), let \( \overline{Q}_\tau \in \text{Conv}(Q_\tau) \), and write \( \overline{\mu}_\tau \) and \( \overline{\nu}_\tau \) 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

\[
\min_{\tau \in \{0,1\}^m} \min_{1 \leq k \leq m} \left\{ \left( \int_{U_k} d\overline{\mu}_\tau \wedge d\overline{\mu}_{\tau^k} \right) \left( \int_{U'_k} d\overline{\nu}_\tau \wedge d\overline{\nu}_{\tau^k} \right) \right\} \geq 1 - \alpha.
\]

Then,

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

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

Notice that for a model of the form \( Q = \{\delta_{x_0} \otimes P, P \in \mathcal{P}\} \) with fixed \( x_0 \in \mathcal{X} \), one recovers the classical Assouad’s Lemma [Yu97] taking \( U_k = \mathcal{X} \) and \( U'_k = \mathcal{X}' \). Indeed, when \( X = x \) a.s, the parameter of interest \( \theta_\mathcal{X}(Q) = \theta(Q) \) can be seen as non-random.
5.2.3 Construction of Hypotheses

In order to apply Le Cam’s Lemma (Theorem 15) or the conditional Assouad’s Lemma (Lemma 16), we describe in this section the construction of the hypotheses involved in the different contexts of estimation. For this, the strategy consists in building distributions that are stochastically close — i.e. with a large test affinity — for which the associated parameters of interest are as different as possible. Before continuing to the precise construction, let us make two remarks about the lower bounds with random point $X_1$. First, the associated minimax risks (Theorems 7 and Theorem 9) involve the integration with respect to $X_1$. Hence, as for regression with $L^p$ loss, multiple locations of bumps are required to yield the right rate. Second, building manifolds with different tangent spaces (resp curvature) would lead to locally singular distributions. Therefore it is natural to consider mixture distributions to get non-trivial bounds.

Let $M_0^{(0)}$ be a $d$-dimensional $C^\infty$-submanifold of $\mathbb{R}^D$ with reach greater than 1 and such that it contains $B_{\mathbb{R}^d \times \{0\}}(\tau_{\min}, 0, 1/2)$. $M_0^{(0)}$ can be built for example by flattening smoothly a unit $d$-sphere. Since $M_0^{(0)}$ is $C^\infty$, the uniform probability distribution $P_0^{(0)}$ on $M_0^{(0)}$ belongs to $\mathcal{P}_1^{k,\mathbb{R}^d}(1/V_0^{(0)}; 1/V_0^{(0)})$, for some $L^{(0)}$ 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 homothety. By construction, from Proposition 14 we have $\tau_{\min} \geq 2\tau_{\min}, B_{\mathbb{R}^d \times \{0\}}(\tau_{\min}) \subset M_0$, and the uniform probability distribution $P_0$ on $M_0$ belongs to the model $\mathcal{P}_2^{k,\mathbb{R}^d,L}(f_{\min}, f_{\max})$ whenever $L \geq L^{(0)}/(2\tau_{\min}), \ldots, L_k \geq L_k^{(0)}/(2\tau_{\min})^{k-1}$, and provided that $f_{\min} \leq ((2\tau_{\min})^{d}V_0^{(0)})^{-1} \leq f_{\max}$. Note that $L_k^{(0)}, \ldots, L_k^{(0)}, \text{Vol}(M_0^{(0)})$ depend only on $d$ and $k$. For this reason, all the lower bounds will be valid for $\tau_{\min}L_1^{(0)}, \ldots, \tau_k^{d-1}L_k^{(0)}/(2\tau_{\min})^{k-1}$ and $\tau^{d-1}f_{\min}$ large enough to exceed the thresholds $L_1^{(0)}, \ldots, L_k^{(0)}$, $\tau^{d-1}V_0^{(0)}$, and $\tau^{d-1}V_0^{(0)}$ respectively.

For $0 < \delta \leq \tau_{\min}/2$, let $x_1, \ldots, x_m \in M_0 \cap B(0, \tau_{\min})$ be such that for all $k \neq k', \|x_k - x_{k'}\| \geq \delta$. A classical packing argument (see [Mas07], p. 71) shows that one can take up $m = [c_d/\delta^d]$ for some $c_d > 0$. We let $e \in \mathbb{R}^D$ denote any unit vector orthogonal to $\mathbb{R}^d \times \{0\}^{D-d}$.

Let $\phi : \mathbb{R}^D \to [0, 1]$ be a smooth scalar map such that $\phi|_{B(0,1)} = 1$ and $\phi|_{B(0,1)^c} = 0$. Let $A_+ > 0$ and $1 \geq A_+ > A_- > 0$ be real numbers to be chosen later. Let $\Lambda = (\Lambda_1, \ldots, \Lambda_m)$ with entries $-\Lambda_+ \leq \Lambda_k \leq A_+$, and $\Lambda = (\Lambda_1, \ldots, \Lambda_m)$ with entries $A_- \leq \Lambda_k \leq A_+$. For $z \in \mathbb{R}^D$, we write $z = (z_1, \ldots, z_D)$ for its coordinates in the canonical basis. For all $\tau = (\tau_1, \ldots, \tau_m) \in \{0, 1\}^m$, define the bump map as

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

An analogous deformation map was considered in [AL15]. We let $P_{\tau, \Lambda, \cdot}^{i}$ denote the pushforward distribution of $P_0$ by $\Phi_{\tau, \Lambda, \cdot}^{i}$, and write $M_{\tau, \Lambda, \cdot}^{i}$ for its...
support. Roughly speaking, $M^\Lambda,A,i$ consists of $m$ bumps at the $x_k$'s having different shapes (Figure 4). If $\tau_k = 0$, the bump at $x_k$ is a symmetric plateau function and has height $\Lambda_k$. If $\tau_k = 1$, it fits the graph of the polynomial $A_k(x - x_k)$ locally. The following Lemma 17 gives differential bounds and geometric properties of $\Phi^\Lambda,A,i$. It follows straightforwardly from chain rule, similarly to Lemma 11 in [AL15].

**Lemma 17.** There exists $c_{\phi,i} < 1$ such that if $A_+ \leq c_{\phi,i} \delta^{i-1}$ and $\Lambda_+ \leq c_{\phi,i} \delta$, then $\Phi^\Lambda,A,i$ is a global $C^\infty$-diffeomorphism of $\mathbb{R}^D$ such that for all $1 \leq k \leq m$, $\Phi^\Lambda,A,i(B(x_k, \delta)) = B(x_k, \delta)$. Moreover,

$$
\|I_D - d\Phi^\Lambda,A,i\|_{op} \leq C_{\phi,i} \left\{ \frac{A_+}{\delta^{i-1}} \right\} \lor \left\{ \frac{\Lambda_+}{\delta} \right\},
$$

and for $j \geq 2$,

$$
\|d^j \Phi^\Lambda,A,i\|_{op} \leq C_{\phi,i,j} \left\{ \frac{A_+}{\delta^{j-1}} \right\} \lor \left\{ \frac{\Lambda_+}{\delta^j} \right\}.
$$

Finally, we define the mixture distribution $\bar{Q}^{(i)}_{\tau,n}$ on $(\mathbb{R}^D)^n$ by

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

(10)
Then the sets $U_k$ for $P_{c > 3}$, using Lemma 17 hold, and let $P_{\tau, n}$ be the distributions $P_{\tau, n}$ on $M_0$. Then $(Z_1, \ldots, Z_n)$ has distribution $Q^{(i)}_{\tau, n}$.

We now state useful probabilistic and geometric properties of $Q^{(i)}_{\tau, n}$, in view of using Theorem 16. For this, let us denote by $P^{(i)}_{\tau}$ the set composed of all the distributions $P^{(i)}_{\tau, A}$ for $A_- \leq A_1, \ldots, A_m \leq A_+$ and $A_+ \leq A_1, \ldots, A_m \leq A_+$. Again, we omit the dependency on $A_-, A_+$ and $A_+$.

**Lemma 18.** Assume that the conditions of Lemma 17 hold, and let

$$U_k = B_{\mathbb{R}^d}((0)^d + B_{\text{span}(e)}(0, \tau_{\min}/2),$$

where for $B, B' \subset \mathbb{R}^D$, $B + B'$ denotes their Minkowski sum, and

$$U'_k = (\mathbb{R}^D \setminus \{B_{\mathbb{R}^d}((0)^d + B_{\text{span}(e)}(0, \tau_{\min}/2)\})^{n-1}.$$

Then the sets $U_k \times U'_k$ are pairwise disjoint, $Q^{(i)}_{\tau, n} \subset \text{Conv}(P^{(i)}_{\tau, n} \otimes n)$, and if $(Z_1, A_1, Z_{2n}) = (Z_1, Z_{2n})$ has distribution $Q^{(i)}_{\tau, n}$, $Z_1$ and $Z_{2n}$ are independent conditionally on the event $(\{Z_1, Z_{2n}\} \in U_k \times U'_k)$.

Moreover, if $(X_1, \ldots, X_n)$ has distribution $P^{(i)}_{\tau, A} \otimes n$ (with fixed $A$ and $A$), then on the event $\{X_1 \in U_k\}$, we have:

- if $\tau_k = 0$,

$$T_{X_1} M_{\tau, A}^{(i)} = \mathbb{R}^d \otimes \{0\}^{D-d}, \quad \|\pi_{X_1}^{M_{\tau, A}^{(i)}} \|_{\text{op}} = 0,$$

and $d_H(M_0, M_{\tau, A}^{(i)}) = |\Lambda_k|$.

- if $\tau_k = 1$,

$$- \text{ for } i = 1: T_{X_1} M_{\tau, A}^{(i)} = \mathbb{R}^d \otimes \{0\}^{D-d}, \quad \|\pi_{X_1}^{M_{\tau, A}^{(i)}} \|_{\text{op}} \geq A_- / 2;$$

$$- \text{ for } i = 2: T_{X_1} M_{\tau, A}^{(i)} = \mathbb{R}^d \otimes \{0\}^{D-d}, \quad \|\pi_{X_1}^{M_{\tau, A}^{(i)}} \|_{\text{op}} \geq A_- / 2.$$

To apply Theorem 16 to the $Q^{(i)}_{\tau, n}$, with $A = \mathbb{R}^D$, $A' = (\mathbb{R}^D)^{n-1}$, it remains to bound the test affinities between their marginals on $\mathcal{X}$ and $\mathcal{X}'$. By construction (10), these are respectively $\tilde{Q}^{(i)}_{\tau, 1}$ and $\tilde{Q}^{(i)}_{\tau, n-1}$.

**Lemma 19.** Assume that the conditions of Lemma 17 and Lemma 18 hold. If in addition, $cA_+(\delta/4) \leq \Lambda_+ \leq CA_+(\delta/4)$ for some absolute constants $C \geq c > 3/4$, and $A_- = A_+/2$, then,

$$\int_{U_k} d\tilde{Q}^{(i)}_{\tau, 1} \wedge d\tilde{Q}^{(i)}_{\tau, n} \geq C \left(\frac{\delta}{\tau_{\min}}\right)^d,$$
\[ \int_{U_i} d\bar{Q}^{(i)}_{\tau,n-1} \wedge d\bar{Q}^{(i)}_{\tau,k,n-1} = \left(1 - c'_d \left(\frac{\delta}{\tau_{\min}}\right)^d\right)^{n-1}. \]

Now, to derive Theorem 7, set \( i = 1 \), take \( A_+ = 2A_- = \varepsilon \delta^{k-1} \), and \( \Lambda_+ = \delta A_+/4 \) for \( \varepsilon = \varepsilon_{\phi,k,d,\tau_{\min}} \) small enough so that \( P^{(1)}_\tau \subset P^k_{\tau_{\min},L}(f_{\min},f_{\max}) \), according to Lemma 17 and Proposition 4. Hence, applying Lemma 16 together with Lemma 18 and Lemma 19, recalling that \( m \) can be taken of order \( c_d/\delta^d \), we get, for all estimators \( \hat{T} \),

\[ \sup_{P \in P^k} \mathbb{E}_{P^{\otimes n}} \mathcal{L}(T_X,M,\hat{T}) \geq c_d,k,\varepsilon m A_+ - \frac{\delta}{\tau_{\min}} \left(1 - c'_d \left(\frac{\delta}{\tau_{\min}}\right)^d\right)^{n-1} \]

Taking \( (\delta/\tau_{\min})^d = 1/(n-1) \) yields the result.

Similarly, to derive Theorem 9, set \( i = 2 \), take \( A_+ = 2A_- = \varepsilon' \delta^{k-2} \), and \( \Lambda_+ = \delta^2 A_+/4^2 \) with \( \varepsilon' = \varepsilon'_{\phi,k,d,\tau_{\min}} \) small enough so that \( P^{(2)}_\tau \subset P^k_{\tau_{\min},L}(f_{\min},f_{\max}) \). With \( (\delta/\tau_{\min})^d = 1/(n-1) \), the same derivation as above leads to the result.

Finally, for Theorem 11, simply take \( m = 1 \), \( \tau = 0 \) and \( \Lambda_1 = \delta A_1/4 = \varepsilon \delta^k \) for \( \varepsilon = \varepsilon_{\phi,k,d,\tau_{\min}} \) as above. We may conclude using Theorem 15 with \( P_0 \) and \( P^0_{\Lambda_1,A_1,(i)} \). Indeed, using \( d_H(M_0,M^0_{\Lambda_1,A_1,(i)}) \geq |A_1| = \varepsilon \delta^k \) from Lemma 18 and noticing that the total variation distance between the two distributions is \( P_0(B(x_1,\delta)) = c_d(\delta/\tau_{\min})^d \), since they differ only outside \( B(x_1,\delta) \), we get the result.

**Acknowledgements**

We would like to thank Frédéric Chazal and Pascal Massart for their constant encouragements, suggestions and stimulating discussions.
A Properties and Stability of the Models

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

Here we show the Lemma 1. Proposition 6.1 in [NSW08] states that for all $x \in M$, $\|II^M_x\|_{\text{op}} \leq 1/\tau_{\text{min}}$. In particular, Gauss equation ([dC92, Proposition 3.1 (a), p.135]) yields that the sectional curvatures of $M$ satisfy $-2/\tau_{\text{min}}^2 \leq \kappa \leq 1/\tau_{\text{min}}^2$. Using Corollary 1.4 of [AB06], we get that the injectivity radius of $M$ is at least $\pi \tau_{\text{min}} \geq \tau_{\text{min}}/4$. Therefore, $\exp_p : B_T(M,0,\tau_{\text{min}}/4) \to M$ is one-to-one.

Let us write $N_p(v) = \exp_p(v) - p - v$. We clearly have $N_p(0) = 0$ and $d_0 N_p = 0$. Let now $v \in B_T(M,0,\tau_{\text{min}}/4)$ be fixed. We have $d_v N_p = d_v \exp_p - Id_{T_pM}$. For $0 \leq t \leq \|v\|$, we write $\gamma(t) = \exp_{\gamma(t)} tv/\|v\|$ for the arc-length parametrized geodesic from $p$ to $\exp_{\gamma(t)} v$, and $P_t$ for the parallel translation along $\gamma$. From Lemma 18 of [DVW15],

$$\left\|d_{\gamma(t)} \exp_p - P_t\right\|_{\text{op}} \leq \frac{2}{\tau_{\text{min}}^2} \frac{t^2}{2} \leq \frac{t}{4\tau_{\text{min}}}.$$  

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

$$|g'(t)| = \left|\langle \nabla_{\gamma'(t)} P_t(w), w, u \rangle \right| = \left|\langle II^M_{\gamma(t)} (\gamma'(t), P_t(w)), u \rangle \right| \leq 1/\tau_{\text{min}}.$$  

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

$$\|d_v N_p\|_{\text{op}} = \left\|d_v \exp_p - Id_{T_pM}\right\|_{\text{op}} \leq \left\|d_v \exp_p - P_{\|v\|}\right\|_{\text{op}} + \|P_{\|v\|} - Id_{T_pM}\|_{\text{op}} \leq \frac{5 \|v\|}{4\tau_{\text{min}}}.$$  

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 [Fed59] gives

$$\|R_2(y - p)\| = d(y - p, T_pM) \leq \frac{\|y - p\|^2}{2\tau_{\text{min}}}.$$  

A.2 Geometric Properties of the Models $C^k$

We now move to the proof of Lemma [12]
Proof of Lemma 12. (i) Simply notice that from the reverse triangle inequality,
\[
\left| \frac{||\Psi_x(v_2) - \Psi_x(v_1)||}{||v_2 - v_1||} - 1 \right| \leq \frac{||N_x(v_2) - N_x(v_1)||}{||v_2 - v_1||} \leq L_{\perp} (||v_1|| \lor ||v_2||) \leq \frac{1}{4}.
\]

(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 \( G = \pi_{T_x M} \circ \Psi_x \) on the domain \( B_{T_x M} (0, h) \). For all \( v \in B_{T_x M} (0, h) \), we have
\[
\| d_v G - Id_{T_x M} \|_{\text{op}} = \| \pi_{T_x M} \circ d_v N_x \|_{\text{op}} \leq \| d_v N_x \|_{\text{op}} \leq L_{\perp} \| v \| \leq \frac{1}{4} < 1.
\]
Hence, \( G \) is a diffeomorphism onto its image and it satisfies \( \| G(v) \| \geq 3 \| v \| / 4 \). It follows that
\[
B_{T_x M} \left( 0, \frac{3h}{4} \right) \subset G (B_{T_x M} (0, h)) = \pi_{T_x M} (\Psi_x (B_{T_x M} (0, h))).
\]

Now, according to Lemma 1, for all \( y \in B \left( x, \frac{3h}{5} \right) \cap M \),
\[
\| \pi_{T_x M} (y - x) \| \leq \| y - x \| + \frac{\| y - x \|^2}{2 \tau_{\text{min}}} \leq \left( 1 + \frac{1}{4} \right) \| y - x \| \leq \frac{3h}{4},
\]
from what we deduce \( \pi_{T_x M} (B \left( x, \frac{3h}{5} \right) \cap M) \subset B_{T_x M} (0, \frac{3h}{4}) \). As a consequence,
\[
\pi_{T_x M} \left( B \left( x, \frac{3h}{5} \right) \cap M \right) \subset \pi_{T_x M} (\Psi_x (B_{T_x M} (0, h))),
\]
which yields the announced inclusion since \( \pi_{T_x M} \) is one to one on \( B \left( x, \frac{3h}{5} \right) \cap M \) from Lemma 3 in [ALZ13], and
\[
\left( B \left( x, \frac{3h}{5} \right) \cap M \right) \subset \Psi_x (B_{T_x M} (0, h)) \subset B \left( x, \frac{5h}{4} \right) \cap M.
\]

(iii) Straightforward application of Lemma 3 in [ALZ13].

(iv) Notice that Lemma 1 gives the existence of such an expansion for \( k = 2 \). Hence, we can assume \( k \geq 3 \). Taking \( h = \frac{\tau_{\text{min}} \land L_{\perp}}{4} \), we showed in the proof of (ii) that the map \( G \) is a diffeomorphism onto its image, with \( \| d_v G - Id_{T_x M} \|_{\text{op}} \leq \frac{1}{4} < 1 \). Additionally, the chain rule yields \( \| d_i^v G \|_{\text{op}} \leq \| d_i^v \Psi_x \|_{\text{op}} \leq L_i \) for all \( 2 \leq i \leq k \). Therefore, from Lemma 20 the differentials of \( G^{-1} \) up to order \( k \) are uniformly bounded. As a consequence, we get the announced expansion writing
\[
y - x = \Psi_x \circ G^{-1} (\pi^{*} (y - x)),
\]
and using the Taylor expansions of order \( k \) of \( \Psi_x \) and \( G^{-1} \).

Let us now check that \( T^*_2 = I_{\mathcal{M}}^M \). First, 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

\[
T^*_2 = (I_D - \pi_{T_x M}) \circ d^2_0 \Psi_x \\
= \pi_{T_x M^\perp} \circ d^2_0 \Psi_x.
\]

Let \( v \in T_x M \) be fixed. Letting \( \gamma(t) = \Psi_x(tv) \) for \( |t| \) small enough, it is clear that \( \gamma''(0) = d^2_0 \Psi(v \otimes 2) \). Moreover, by definition of the second fundamental form \([\text{dC92}, \text{Proposition 2.1, p.127}]\), since \( \gamma(0) = x \) and \( \gamma'(0) = v \), we have

\[
I^M_x (v \otimes 2) = \pi_{T_x M^\perp} (\gamma''(0)).
\]

Hence

\[
T^*_2 (v \otimes 2) = \pi_{T_x M^\perp} \circ d^2_0 \Psi_x (v \otimes 2) \\
= \pi_{T_x M^\perp} (\gamma''(0)) \\
= I^M_x (v \otimes 2),
\]

which concludes the proof.

(v) The first statement is a rephrasing of Proposition 6.1 in \([\text{NSW08}]\). It yields the bound on sectional curvature, using the Gauss equation \([\text{dC92}, \text{Proposition 3.1 (a), p.135}]\).

In the proof of Lemma \( 12 \) (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 20.** 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 \( 2 \leq i \leq k \),

\[
\|I_d - dG^{-1}\|_{op} \leq \frac{\varepsilon}{1 - \varepsilon} \quad \text{and} \quad \|d^i G^{-1}\|_{op} \leq L_i' \varepsilon \leq L_1 \varepsilon \cdots L_k \varepsilon < \infty \text{ for } 2 \leq i \leq k.
\]

**Proof of Lemma 20.** For all \( x \in U \), \( \|d_x G - I_d\|_{op} < 1 \), so \( G \) is one to one, and for all \( y = G(x) \in G(U) \),

\[
\|I_d - d_y G^{-1}\|_{op} = \|I_d - (d_x G)^{-1}\|_{op} \\
\leq \|(d_x G)^{-1}\|_{op} \|I_d - d_x G\|_{op} \\
\leq \frac{\|I_d - d_x G\|_{op}}{1 - \|I_d - d_x G\|_{op}} \\
\leq \frac{\varepsilon}{1 - \varepsilon}.
\]
For $2 \leq i \leq k$ and $1 \leq j \leq i$, write $\Pi_{i}^{(j)}$ for the set of partitions of $\{1, \ldots, i\}$ with $j$ blocks. Differentiating $i$ times the identity $G \circ G^{-1} = I_{d(G(U))}$, Faa di Bruno’s formula yields that, for all $y = G(x) \in G(U)$ and all unit vectors $h_1, \ldots, h_i \in \mathbb{R}^D$,

$$0 = d_{y} (G \circ G^{-1} \cdot (h_{\alpha})_{1 \leq \alpha \leq i}) = \sum_{j=1}^{i} \sum_{\pi \in \Pi_{i}^{(j)}} d_{x}^{j} G \cdot \left( (d_{y}^{j} G^{-1} \cdot (h_{\alpha})_{\alpha \in I})_{I \in \pi} \right).$$

Isolating the term for $j = 1$ entails

$$\left\| d_{x} \Phi \cdot (d_{y}^{1} G^{-1} \cdot (h_{\alpha})_{1 \leq \alpha \leq i}) \right\|_{op} = \left\| - \sum_{j=2}^{i} \sum_{\pi \in \Pi_{i}^{(j)}} d_{x}^{j} G \cdot \left( (d_{y}^{j} G^{-1} \cdot (h_{\alpha})_{\alpha \in I})_{I \in \pi} \right) \right\|_{op} \leq \sum_{j=2}^{i} \sum_{\pi \in \Pi_{i}^{(j)}} \left\| d^{j} G \right\|_{op} \prod_{I \in \pi} \left\| d^{I} G^{-1} \right\|_{op}.$$

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

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

The result follows by induction on $i$. \hfill \Box

### A.3 Stability of the Model

This section is devoted to prove the stability of the model with respect to ambient diffeomorphisms (Proposition 4).

The second part is pretty straightforward since the dilation $\lambda M$ has reach $\tau_{\lambda M} = \lambda \tau_{M}$, and can be parametrized locally by $\tilde{\Psi}_{\lambda p}(v) = \lambda \Psi_{p}(v/\lambda) = \lambda p + v + \lambda N_{p}(v/\lambda)$, yielding the differential bounds $L_{(\lambda)}$. Bounds on the density follow from homogeneity of the $d$-dimensional Hausdorff measure.

For the first part, we split the proof into two intermediate results. Proposition 21 deals with the stability of the geometric model, that is, the reach bound and the existence of a smooth parametrization when a submanifold is perturbed. Lemma 22 deals with the condition on the density in the models $P^{k}$. It gives a change of variable formula for pushforward of measure on submanifolds, ensuring a control on densities with respect to intrinsic volume measure.

**Proposition 21.** Let $\Phi : \mathbb{R}^D \rightarrow \mathbb{R}^D$ be a global $C^{k}$-diffeomorphism. If $\|d\Phi - I_{D}\|_{op}$, $\|d^{2}\Phi\|_{op}$, $\ldots$, $\|d^{k}\Phi\|_{op}$ are small enough, then for all $M$ in $C_{\tau_{\min}, L}$, the image $M' = \Phi(M)$ belongs to $C_{\tau_{\min}/2, 2L, 2L, \ldots, 2L_k}$. 

27
Proof of Proposition 27. To bound \( \tau_{M'} \) from below, we use the stability of the reach with respect to \( C^2 \) diffeomorphisms. Namely, from Theorem 4.19 in [Fed59],

\[
\tau_{M'} = \tau_{\Phi(M)} \geq \frac{(1 - \|I_D - d\Phi\|_{op})^2}{1 + \|I_D - d\Phi\|_{op} + \|d^2\Phi\|_{op}}
\]

\[
\geq \tau_{\min} \frac{(1 - \|I_D - d\Phi\|_{op})^2}{1 + \|I_D - d\Phi\|_{op} + \tau_{\min} \|d^2\Phi\|_{op}} \geq \frac{\tau_{\min}}{2}
\]

for \( \|I_D - d\Phi\|_{op} \) and \( \|d^2\Phi\|_{op} \) 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 \( v' \in T_{p'} M' = d_p \Phi(T_p M) \), let us define

\[
\Psi_{p'}(v') = \Phi \left( \Psi_p \left( d_{p'} \Phi^{-1}.v' \right) \right)
\]

\[
= p' + v' + N_{p'}(v'),
\]

where \( N_{p'}(v') = \left\{ \Phi \left( \Psi_p \left( d_{p'} \Phi^{-1}.v' \right) \right) - p' - v' \right\} \).

The maps \( \Psi_{p'}(v') \) and \( N_{p'}(v') \) are well defined whenever \( \|d_{p'} \Phi^{-1}.v'\| \leq \frac{1}{8L_{\perp}} \), so in particular if \( \|v'\| \leq \frac{1}{8(2L_{\perp})} \leq \frac{1-\|I_D - d\Phi\|_{op}}{8L_{\perp}} \) and \( \|I_D - d\Phi\|_{op} \leq \frac{1}{2} \). One easily checks that \( N_{p'}(0) = 0 \), \( d_0 N_{p'} = 0 \) and writing \( c(v') = p + d_{p'} \Phi^{-1}.v' + N_{p'}(d_{p'} \Phi^{-1}.v') \), for all unit vector \( w' \in T_{p'} M' \),
$$\|d^2_{\omega} N^r_p(w^{\otimes 2})\| = \|d^2_{\omega}(\Phi \left( \{d_{\omega} \Phi^{-1} \circ d_{\omega} \Phi^{-1}\} \otimes 2 \right)) + d_{\omega}(\Phi \circ d_{\omega} \Phi^{-1}) \left( \{d_{\omega} \Phi^{-1}\} \otimes 2 \right)\|
\leq \|d^2 \Phi\|_{op} (1 + L_1 \|d_{\omega} \Phi^{-1}\|)^2 \|d_{\omega} \Phi^{-1}\|^2
\leq \|d^2 \Phi\|_{op} (1 + 1/8)^2 \|d_{\omega} \Phi^{-1}\|^2_{op}
\leq \|d^2 \Phi\|_{op} (1 + 1/8)^2 \|d_{\omega} \Phi^{-1}\|^2_{op} + L_1 \|d_{\omega} \Phi^{-1}\|_{op}^2$$

Writing further $\|d \Phi^{-1}\|_{op} \leq (1 - \|I_D - d \Phi\|_{op})^{-1} \leq 1 + 2 \|I_D - d \Phi\|_{op}$ for $\|I_D - d \Phi\|_{op}$ small enough only on $L_1$, it is clear that the right-hand side of the latter inequality goes below $2L_1$ for $\|I_D - d \Phi\|_{op}$ and $\|d^2 \Phi\|_{op}$ small enough. Hence, for $\|I_D - d \Phi\|_{op}$ and $\|d^2 \Phi\|_{op}$ small enough depending only on $L_1$, $\|d_{\omega} \Phi_{op} \|_{op}^2 \leq 2L_1$ for all $\|\omega\| \leq \frac{1}{8(2L_1)}$. From the chain rule, the same argument applies for the order $3 \leq i \leq k$ differential of $N^r_p$.

**Lemma 22** (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 $\|I_D - d \Phi\|_{op} < 1/3$. Let $P' = \Phi \ast P$ be the pushforward of $P$ by $\Phi$. Then $P'$ has a density $g$ with respect to $\mathcal{H}^d$. This density can be chosen to be, for all $z \in \Phi(M)$,

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

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

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

**Proof of Lemma 22.** Let $p \in M$ be fixed and $A \subset B(p, r) \cap M$ 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 $J_h(x) = \sqrt{\det (d_x h^T d_x h)}$. The area formula ([Fed69, Theorem 3.2.5]) states that if $h$ is one-to-one,

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

whenever $u : \mathbb{R}^D \to \mathbb{R}$ is Borel, where $\lambda^d$ 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 $\Psi_p = \exp_p : T_p M \to \mathbb{R}^D$ 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 $h = \Phi \circ \Psi_p$,

$$\int_{\Psi_p^{-1}(A)} f(\Psi_p(x)) J_{\Psi_p}(x) \lambda^d(dx)$$

$$= \int_{\Psi_p^{-1}(A)} f (\Phi^{-1}(h(x))) \frac{J_{\Psi_p}(h^{-1}(h(x)))}{J_{\Phi \circ \Psi_p}(h^{-1}(h(x)))} J_{\Phi \circ \Psi_p}(x) \lambda^d(dx)$$

$$= \int_{\Phi(A)} f (\Phi^{-1}(z)) \frac{J_{\Psi_p}(h^{-1}(z))}{J_{\Phi \circ \Psi_p}(h^{-1}(z))} \mathcal{H}^d(dz).$$

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

$$g(z) = f (\Phi^{-1}(z)) \frac{J_{\Psi_p^{-1}(z)}(\Psi_p^{-1}(z) \circ \Phi^{-1}(z))}{J_{\Phi \circ \Psi_p^{-1}(z)}(\Psi_p^{-1}(z) \circ \Phi^{-1}(z))}.$$

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

We now let $B$ and $\pi_T$ denote $d_p \Phi$ and $\pi_{T_p M}$ respectively. For any unit vector $v \in T_p M$,

$$\|\pi_T B^T B v\| \leq \|\pi_T (B^T B - I_D) v\|$$

$$\leq \|B^T B - I_D\|_{op}$$

$$\leq \left( 2 + \|I_D - B\|_{op} \right) \|I_D - B\|_{op}$$

$$\leq 3 \|I_D - B\|_{op}.$$  

Therefore, $1 - 3 \|I_D - B\|_{op} \leq \|\pi_T B^T B|_{T_p M}\|_{op} \leq 1 + 3 \|I_D - B\|_{op}$. Hence,

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

30
and
\[
\sqrt{\det (\pi_T B^T B )} \geq \left( 1 - 3 \| I_D - B \|_\text{op} \right)^{d/2} \geq \frac{1}{1 + 3(2^{d/2} - 1) \| I_D - B \|_\text{op}},
\]
which yields the result.

\[\Box\]

## 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 23.** Let \( P \in \mathcal{P}^k \) have support \( M \subset \mathbb{R}^D \). Then for all \( r \leq \tau_{\text{min}}/4 \) and \( x \in M \),

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

for some \( c_d, C_d > 0 \), with \( p_x(r) = P(B(x,r)) \).

Moreover, letting \( h = \left( \frac{C_k \log n}{f_{\text{min}}} \right)^{1/d} \) with \( C_k \) large enough, the following holds. For \( n \) large enough so that \( h \leq \tau_{\text{min}}/8 \), with probability at least \( 1 - \left( \frac{1}{n} \right)^{k/d} \),

\[
d_H (M, X_n) \leq h.
\]

Lemma 23 includes Lemma 14, hence we only prove Lemma 23.

**Proof of Lemma 23.** Denoting by \( B_M (x, r) \) the geodesic ball of radius \( r \) centered at \( x \), Proposition 25 of [AL15] yields

\[
B_M (x, r) \subset B(x, r) \cap M \subset B_M (x, 6r/5).
\]

Hence, the bounds on the Jacobian of the exponential map given by Proposition 27 of [AL15] yield

\[
c_d r^d \leq \text{Vol} (B(x, r) \cap M) \leq C_d r^d,
\]

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

Now we notice that since \( p_x(r) \geq c_d f_{\text{min}} r^d \), Theorem 3.3 in [CGLM13] entails, for \( h \leq \tau_{\text{min}}/8 \),

\[
\mathbb{P} \left( d_H (M, X_n) \geq h \right) \leq \frac{4^d}{c_d f_{\text{min}} h^d} \exp \left( - \frac{c_d f_{\text{min}} h^d}{2^d} \right).
\]

Hence, taking \( h = \left( \frac{C_k \log n}{f_{\text{min}}} \right)^{1/d} \) with \( C_k \) so that \( C_k \geq 4^d \), \( \left( \frac{1+k/d}{c_k} \right) \) yields the result. Since \( k \geq 1 \), taking \( C_k = \frac{4^d}{c_k} \) is sufficient. \[\Box\]
B.2 Concentration Bounds for Local Polynomials

This section is devoted to the proof of Proposition 13. A first step is to ensure that empirical expectations order \( k \) polynomials are close to their deterministic counterparts.

**Proposition 24.** For any \( x \in M \), we have

\[
\mathbb{P} \left[ \sup_{u_1, \ldots, u_k, \varepsilon \in \{0,1\}^k} \left| (P - P_{n-1}) \prod_{j=1}^{p} \left( \frac{\langle u_j, y \rangle}{h} \right)^{\varepsilon_j} \mathbb{1}_{B(x,h)}(y) \right| \right] \geq p_x(h) \left( \frac{4k\sqrt{2\pi}}{\sqrt{n-1}p_x(h)} + \frac{2t}{(n-1)p_x(h)} + \frac{2}{3(n-1)p_x(h)} \right) \leq e^{-t},
\]

where \( P_{n-1} \) denotes the empirical distribution of \( n-1 \) i.i.d. random variables \( X_i \) drawn from \( P \).

**Proof of Proposition 24.** Without loss of generality we choose \( x = 0 \) and shorten notation to \( B(h) \) and \( p(h) \). Let \( Z \) denote the empirical process on the left-hand side of Proposition 24. Denote also by \( f_{u,\varepsilon} \) the map \( \prod_{j=1}^{k} \left( \langle u_j, y \rangle \right)^{\varepsilon_j} \mathbb{1}_{B(h)}(y) \), and let \( \mathcal{F} \) denote the set of such maps, for \( u_j \) in \( B(1) \) and \( \varepsilon \) in \( \{0,1\}^k \).

Since \( \|f_{u,\varepsilon}\|_{\infty} \leq 1 \) and \( Pf_{u,\varepsilon} \leq p(h) \), the Talagrand-Bousquet inequality ([Bou02, Theorem 2.3]) yields

\[
Z \leq 4\mathbb{E}Z + \sqrt{\frac{2p(h)t}{n-1} + \frac{2t}{3(n-1)}},
\]

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

**Lemma 25.** We may write

\[
\mathbb{E}Z \leq \frac{\sqrt{2\pi p(h)}}{\sqrt{n-1} k}.
\]

**Proof of Lemma 25.** Let \( \sigma_i \) 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

\[
\mathbb{E}Z = \mathbb{E}_X \sup_{u,\varepsilon} \left| (P - P_{n-1}) \prod_{j=1}^{k} \left( \frac{\langle u_j, y \rangle}{h} \right)^{\varepsilon_j} \mathbb{1}_{B(h)}(y) \right| \leq \frac{2}{n-1} \mathbb{E}_X \mathbb{E}_\sigma \sup_{u,\varepsilon} \sum_{i=1}^{n-1} \sigma_i \prod_{j=1}^{k} \left( \frac{\langle u_j, X_i \rangle}{h} \right)^{\varepsilon_j} \mathbb{1}_{B(h)}(X_i) \leq \frac{\sqrt{2\pi}}{n-1} \mathbb{E}_X \mathbb{E}_g \sup_{u,\varepsilon} \sum_{i=1}^{n-1} g_i \prod_{j=1}^{k} \left( \frac{\langle u_j, X_i \rangle}{h} \right)^{\varepsilon_j} \mathbb{1}_{B(h)}(X_i).
\]
Now let $Y_g$ denote the Gaussian process $\sum_{i=1}^{n-1} g_I \prod_{j=1}^k \left( \frac{\langle u_j, X_i \rangle}{h} \right)^{\varepsilon_j} 1_{B(h)}(X_i)$. Since, for any $x$ in $B(h)$, $u, v$ in $B(1)^k$, and $\varepsilon, \varepsilon'$ in $\{0, 1\}^k$, we have

$$\left| \prod_{j=1}^k \left( \frac{\langle x, u_j \rangle}{h} \right)^{\varepsilon_j} - \prod_{j=1}^k \left( \frac{\langle x, v_j \rangle}{h} \right)^{\varepsilon_j} \right| \leq \sum_{r=1}^k \prod_{j=1}^{k-r} \left( \frac{\langle x, u_j \rangle}{h} \right)^{\varepsilon_j} \prod_{j=k+2-r}^{k} \left( \frac{\langle x, v_j \rangle}{h} \right)^{\varepsilon_j} \left[ \left( \frac{\langle u_{k+1-r}, x \rangle}{h} \right)^{\varepsilon_{k+1-r}} - \left( \frac{\langle v_{k+1-r}, x \rangle}{h} \right)^{\varepsilon_{k+1-r}} \right]$$

$$\leq \sum_{r=1}^k \left| \left\langle \varepsilon_r u_r - \varepsilon'_r v_r, x \right\rangle \right|.$$  

We deduce that

$$\mathbb{E}_g (Y_{u, \varepsilon} - Y_{v, \varepsilon'})^2 \leq k \sum_{i=1}^{n-1} \left( \frac{\left\langle \varepsilon_r u_r, X_i \right\rangle}{h} - \frac{\left\langle \varepsilon'_r v_r, X_i \right\rangle}{h} \right)^2 1_{B(h)}(X_i)$$

$$\leq \mathbb{E}_g (\Theta_{u, \varepsilon} - \Theta_{v, \varepsilon'})^2,$$

where $\Theta_{u, \varepsilon} = \sqrt{k} \sum_{i=1}^{n-1} \sum_{r=1}^k g_i r \frac{\left\langle \varepsilon_r u_r, X_i \right\rangle}{h} 1_{B(h)}(X_i)$. According to Slepian’s Lemma [BLM13, Theorem 13.3], it follows that

$$\mathbb{E}_g \sup_{u, \varepsilon} Y_g \leq \mathbb{E}_g \sup_{u, \varepsilon} \Theta_g$$

$$\leq \sqrt{k} \mathbb{E}_g \sup_{u, \varepsilon} \left\| \sum_{r=1}^k \left( \varepsilon_r u_r, \sum_{i=1}^{n-1} g_i r 1_{B(h)}(X_i) \right) \right\|^2$$

$$\leq \sqrt{k} \mathbb{E}_g \sup_{u, \varepsilon} \sqrt{\sum_{r=1}^k \left( \varepsilon_r u_r, \sum_{i=1}^{n-1} g_i r 1_{B(h)}(X_i) \right)^2 \frac{h^2}{h^2}}.$$
We deduce that
\[ E_g \sup_{u, \varepsilon} Y_g \leq E_g \sup_{u, \varepsilon} \Theta_g \]
\[ \leq k \sqrt{E_g \left\| \sum_{i=1}^{n-1} g_i \mathbb{1}_{B(h)}(X_i) \right\|^2} \]
\[ \leq k \sqrt{E_g \left\| \sum_{i=1}^{n-1} g_i X_i \mathbb{1}_{B(h)}(X_i) \right\|^2} \]
\[ \leq k \sqrt{N(h)}. \]

Then we can deduce that \( E X \sup_{u, \varepsilon} Y_g \leq k \sqrt{p(h)} \), hence the result. \( \square \)

Combining Lemma 25 with Talagrand-Bousquet’s inequality gives the result of Proposition 24.

We are now in position to prove Proposition 13.

**Proof of Proposition 13.** If \( h \leq \tau_{\min}/8 \), then, according to Lemma 23, \( p(h) \geq c_d f_{\min} h^d \), hence, if \( h = \left( \frac{K \log(n)}{n-1} \right)^{\frac{1}{d}} \), \( (n-1)p(h) \geq K c_d f_{\min} \log(n) \). Choosing \( t = (k/d + 1) \log(n) \) in Proposition 24 and \( K = K'/f_{\min} \), with \( K' > 1 \) leads to

\[ \mathbb{P} \left[ \sup_{u_{1}, \ldots, u_{k}, \varepsilon \in \{0,1\}^k} \left\| (P - P_{n-1}) \prod_{j=1}^{k} \left( \frac{\langle u_j, y \rangle}{h} \right)^{\varepsilon_j} \mathbb{1}_{B(x, h)}(y) \right\| \geq \frac{c_d f_{\max} h^d}{\sqrt{K'}} \right] \leq \left( \frac{1}{n} \right)^{\frac{1}{d} + 1}. \]

On the complement of the probability event mentioned just above, for a polynomial \( Q = \sum_{\alpha \in [0,k]^d} a_{\alpha} x_{1,d}^{\alpha} \), we have

\[ (P_{n-1} - P) Q^2(x_{1,d}) \mathbb{1}_{B(h)}(x) \geq - \sum_{\alpha, \beta} \frac{c_d f_{\max} h^{d+|\alpha|+|\beta|}}{\sqrt{K'}} |a_{\alpha} a_{\beta}| h^{d+|\alpha|+|\beta|} \]
\[ \geq - \frac{c_d f_{\max} h^d}{\sqrt{K'}} \|Q_h\|_2^2. \]

On the other hand, we may write, for all \( r > 0 \),

\[ \int_{B(0,r)} Q^2(x_{1,d})dx_1 \ldots dx_d \geq C_{d,k} r^d \|Q_r\|_2^2, \]

for some constant \( C_{d,k} \). It follows that

\[ PQ^2(x_{1,d}) \mathbb{1}_{B(h)}(x) \geq PQ^2(x_{1,d}) \mathbb{1}_{B(7h/8)}(x_{1,d}) \geq c_k h^d f_{\min} \|Q_h\|_2^2. \]

34
according to Lemma 12. Then we may choose $K' = \kappa_{k,d}(f_{\text{max}}/f_{\text{min}})^2$, with $\kappa_{k,d}$ large enough so that

$$P_{n-1}Q^2(x_1:d)\mathbb{E}(h)(x) \geq c_{k,d}f_{\text{min}}h^d\|Q_h\|_2^2.$$ 

\[\square\]

C Minimax Lower Bounds

C.1 Proof of the Conditional Assouad’s Lemma

This section is dedicated to the proof of Lemma 16. The proof follows that of Lemma 2 in [Yu97]. Let $\hat{\theta} = \hat{\theta}(X,X')$ be fixed. For any family of $2^m$ distributions $\{Q_\tau\}_{\tau \in \{Q_\tau\}_{\tau}}$, since the $U_k \times U'_k$'s are pairwise disjoint,

$$\sup_{Q \in \Omega} \mathbb{E}_Q \left[d(\theta_X(Q), \hat{\theta}(X,X'))\right] \geq \max_{\tau} \mathbb{E}_Q, d(\hat{\theta}, \theta_X(Q_\tau))$$

$$\geq \max_{\tau} \mathbb{E}_Q, \sum_{k=1}^{m} d(\hat{\theta}, \theta_X(Q_\tau)) \mathbb{I}_{U_k \times U'_k}(X,X')$$

$$\geq 2^{-m} \sum_{\tau} \sum_{k=1}^{m} \mathbb{E}_Q, d(\hat{\theta}, \theta_X(Q_\tau)) \mathbb{I}_{U_k \times U'_k}(X,X')$$

$$\geq 2^{-m} \sum_{\tau} \sum_{k=1}^{m} \mathbb{E}_Q, d(\hat{\theta}, D_{\tau,k}) \mathbb{I}_{U_k \times U'_k}(X,X')$$

$$= \sum_{k=1}^{m} 2^{-(m+1)} \sum_{\tau} \left( \mathbb{E}_Q, d(\hat{\theta}, D_{\tau,k}) \mathbb{I}_{U_k \times U'_k}(X,X') \right).$$

Since the previous inequality holds for all $Q_\tau \in Q_\tau$, it extends to $\overline{Q}_\tau \in \overline{\text{Conv}}(Q_\tau)$ 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.
fore,
\[
\mathbb{E}_{\mathcal{Q}_r} d(\hat{\theta}, D_{\tau,k}) \mathbb{1}_{U_k \times U_k'}(X, X') + \mathbb{E}_{\mathcal{Q}_r} d(\hat{\theta}, D_{\tau,k}) \mathbb{1}_{U_k \times U_k'}(X, X') \\
\geq \mathbb{E}_{\mathcal{Q}_r} d(\hat{\theta}, D_{\tau,k}) \mathbb{1}_{U_k}(X) \mathbb{1}_{U_k'}(X') + \mathbb{E}_{\mathcal{Q}_r} d(\hat{\theta}, D_{\tau,k}) \mathbb{1}_{U_k}(X) \mathbb{1}_{U_k'}(X') \\
= \mathbb{E}_{\nu_r} \left[ d(\hat{\theta}, D_{\tau,k}) \mathbb{Q}_r(X) \right] \mathbb{1}_{U_k'}(X') \\
+ \mathbb{E}_{\nu_r} \left[ d(\hat{\theta}, D_{\tau,k}) \mathbb{1}_{U_k}(X) \right] \mathbb{1}_{U_k'}(X') \\
= \int_{U_k} \int_{U_k'} d(\hat{\theta}, D_{\tau,k}) d\hat{\mu}_r(x) d\hat{\nu}_r(x') + \int_{U_k} \int_{U_k'} d(\hat{\theta}, D_{\tau,k}) d\hat{\mu}_r(x) d\hat{\nu}_r(x') \\
\geq \int_{U_k} \int_{U_k'} \left( d(\hat{\theta}, D_{\tau,k}) + d(\hat{\theta}, D_{\tau,k}) \right) d\hat{\mu}_r(x) d\hat{\nu}_r(x') \\
\geq \Delta \left( \int_{U_k} d\hat{\mu}_r(x) \right) \left( \int_{U_k'} d\hat{\nu}_r(x') \right) \\
\geq \Delta (1 - \alpha),
\]
where we used that \(d(\hat{\theta}, D_{\tau,k}) + d(\hat{\theta}, D_{\tau,k}) \geq \Delta\). The result follows by summing the bound above \(|\{1, \ldots, m\} \times \{0, 1\}^m| = m2^m\) times.

### C.2 Construction of Generic Hypotheses

In this section we prove Lemma 18 and Lemma 19.

**Proof of Lemma 18.** It is clear from the definition that \(Q_r^{(i)}(x, Z) \in \overline{\text{Conv}}(\eta_{\tau})^{\otimes n}\). By construction of the \(\Phi_r^{A,A}^{(i)}\)'s, these maps leave the sets

\[ B_{\mathbb{R}^d \times \{0\}^{D-d}}(x_k, \delta) + B_{\text{span}(x)}(0, \tau_{\min}/2) \]

unchanged for all \(A, L\). 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 \(X_1, A_k, A_k,\) and \(Z_{2:n}\) as a function of the rest of the \(X_{j}'s, A_k's\) 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 \(z = \Phi_r^{A,A}^{(i)}(x_0) \in U_k \cap M_r^{A,A}^{(i)}\). By construction, one necessarily has \(x_0 \in M_0 \cap B(x_k, \delta/2)\).

- If \(\tau_k = 0\), locally around \(x_0\), \(\Phi_r^{A,A}^{(i)}\) is the translation of vector \(A_ke\).

Therefore, since \(M_0\) satisfies \(T_{x_0} M_0 = \mathbb{R}^d \times \{0\}^{D-d}\) and \(I_{x_0} M_0 = 0\), we have

\[ T_{x} M_r^{A,A}^{(i)} = \mathbb{R}^d \times \{0\}^{D-d} \quad \text{and} \quad \left\|I_{x} M_r^{A,A}^{(i)} \circ T_{x} M_r^{A,A}^{(i)}\right\|_{\text{op}} = 0.\]

Furthermore, by construction, \(z_k = x_k + A_ke\) belongs to \(M_r^{A,A}^{(i)}\). Since \(e\) is orthogonal to \(M_0\), \(d(z_0, M_0) \geq |A_k|\). Thus

\[ d_H(M_0, M_r^{A,A}^{(i)}) \geq |A_k|.\]
• if \( \tau_k = 1 \),

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

\[
\angle \left( T_x M_{\tau,A,k}^{1}, \mathbb{R}^d \times \{ 0 \}^{D-d} \right) \geq (1 + 1/A_k^2)^{-1/2} \geq A_k/2 \geq A_-/2.
\]

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

\[
\left\| T_x M_{\tau,A,k}^{2} \circ \pi_x M_{\tau,A,k}^{2} \right\|_{\text{op}} \geq A_k/2 \geq A_-/2.
\]

\( \Box \)

Proof of Lemma 19. First note that all the distributions involved have support in \( \mathbb{R}^d \times \text{span}(e) \times \{ 0 \}^{D-(d+1)} \). Therefore, we use the canonical coordinate system of \( \mathbb{R}^d \times \text{span}(e) \), 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 \( \tau \) and \( \tau^k \)). Recall that \( \phi \) has been chosen to be constant and equal to 1 on the ball \( B(0,1/2) \).

By definition \( \{10\} \), on the event \( \{ Z \in U_k \} \), a random variable \( Z \) having distribution \( \bar{Q}_{r,1}^{(i)} \) can be represented as \( Z = X + \phi \left( \frac{Z - x_k}{\delta} \right) \Lambda_k e = X + \Lambda_k e \) where \( X \) and \( \Lambda_k \) are independent and have respective distributions \( P_0 \) (the uniform distribution on \( M_0 \)) and the uniform distribution on \( [-\Lambda_+, \Lambda_+] \). Therefore, on \( U_k \), \( \bar{Q}_{r,1}^{(i)} \) has a density with respect to the Lebesgue measure \( \lambda_{d+1} \) on \( \mathbb{R}^d \times \text{span}(e) \) that can be written as

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

Analogously, nearby \( x_k \) a random variable \( Z \) having distribution \( \bar{Q}_{r+1}^{(i)} \) can be represented as \( Z = X + A_k(X - x_k)^2e \) where \( A_k \) has uniform distribution on \( [A_-, A_+] \). Therefore, a straightforward change of variable yields the density

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

We recall that \( \text{Vol}(M_0) = (2\tau_{\text{min}})^d \text{Vol}(M_0^{(0)}) = c_d' \tau_{\text{min}}^d \). Let us now tackle the right-hand side inequality, writing
\[
\int_{U_k} dQ^{(i)}_{\tau,1} \land dQ^{(i)}_{\tau^+1} \\
= \int_{B(x_k,\delta/2)} \left( \frac{1}{2Vol(M_0)\Lambda_+} \right) \land \left( \frac{1}{2^d} \cdot \frac{1}{\Lambda_+} \right) dydx_1dx_2; d
\]

\[
\geq \int_{B_{\delta/4}(0,1)} \int_{-\delta/4}^{\delta/4} \int_{B_{\delta/4}(0,1)} \left( \frac{1}{2^d} \cdot \frac{1}{\Lambda_+} \right) \land \left( \frac{1}{2^d} \cdot \frac{1}{\Lambda_+} \right) dydx_1dx_2; d
\]

It follows that

\[
\int_{U_k} dQ^{(i)}_{\tau,1} \land dQ^{(i)}_{\tau^+1} \\
\geq \frac{c_d}{\tau_{min}^d} \int_{0}^{\delta/4} \int_{0}^{\delta/4} \frac{1}{\Lambda_+} \frac{2}{\Lambda_+} dydx_1 \\
\geq \frac{c_d}{\tau_{min}^d} \int_{0}^{\delta/4} \int_{0}^{\delta/4} \frac{1}{\Lambda_+} \frac{2}{\Lambda_+} dydx_1 \\
= \frac{c_d}{\tau_{min}^d} \int_{0}^{\delta/4} \int_{0}^{\delta/4} \frac{1}{\Lambda_+} \frac{2}{\Lambda_+} dydx_1 \\
\geq \frac{c_d}{\tau_{min}^d} \frac{\delta}{\tau_{min}} \frac{d}{\tau_{min}} 
\]

For the integral on \( U'_k \), notice that by definition, \( Q^{(i)}_{\tau,n-1} \) and \( Q^{(i)}_{\tau^+n-1} \) 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 \( \mathbb{R}^D \setminus \{ B_{\delta^2} \times \{0\} \} \) unchanged. Therefore,

\[
\int_{U'_k} dQ^{(i)}_{\tau,n-1} \land dQ^{(i)}_{\tau^+n-1} \\
= P_0 \otimes \frac{1}{\tau_{min}^d} (U'_k) \\
= \left( 1 - P_0 \left( B_{\delta^2} \times \{0\} \right) \right) \left( 1 - \frac{\omega d^d}{Vol(M)} \right)^{n-1}
\]

\[\text{hence the result.} \]

**C.3 Minimax Inconsistency Results**

This section is devoted to the proof of lower bound for tangent space estimation (Theorem 5): we build hypotheses \( P, P' \) and apply Theorem 15. For \( \delta \geq \Lambda > 0 \), let \( C', C' \subset \mathbb{R}^3 \) be closed curves of the Euclidean space as in Figure 4 and such that outside the figure, \( C' \) and \( C' \) coincide and are \( C' \). The bumped parts are
obtained with a smooth diffeomorphism similar to [9], centered at \( x \). Here, \( \delta \) and \( \Lambda \) can be chosen arbitrarily small.

Let \( S^{d-1} \subset \mathbb{R}^d \) be a \( d - 1 \)-sphere of radius \( 1/L_\perp \). Consider the Cartesian products \( M = \mathcal{C} \times S^{d-1} \) and \( M' = \mathcal{C}' \times S^{d-1} \). \( M \) and \( M' \) 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 5.2.3, yielding, from Proposition 4, that \( P_1, P'_1 \) belong to \( P_k^{(x)} \) provided that \( L_3/L_\perp, \ldots, L_k/L_\perp, L_1^d/f_{\min} \) and \( f_{\max}/L_\perp^d \) are large enough (depending only on \( d \) and \( k \)), and that \( \Lambda, \delta \) and \( \Lambda^k/\delta \) are small enough. From Le Cam’s Lemma 15, we have for all \( n \geq 1, \)

\[
\inf \sup_{\overline{T}, \delta} \mathbb{E}_{P^{\perp n}} \angle(T_x M, \overline{T}) \geq \frac{1}{2} \angle(T_x M_1, T_x M'_1) (1 - TV(P_1, P'_1))^n.
\]

By construction, \( \angle(T_x M_1, T_x M'_1) = 1 \), and since \( \mathcal{C} \) and \( \mathcal{C}' \) coincide outside \( \mathcal{B}_{\mathbb{R}^3}(0, \delta) \),

\[
TV(P_1, P'_1) = Vol((\mathcal{B}_{\mathbb{R}^3}(0, \delta) \cap \mathcal{C}) \times S^{d-1}) / Vol(\mathcal{C} \times S^{d-1})
\]

\[
= \frac{\text{Length}(\mathcal{B}_{\mathbb{R}^3}(0, \delta) \cap \mathcal{C})}{\text{Length}(\mathcal{C})} \leq c_{L_\perp, \delta}.
\]

Hence, letting \( \Lambda, \delta \) go to 0 with \( \Lambda^k/\delta \) small enough, we get the announced bound.

We now tackle the lower bound on second fundamental form estimation with the same strategy. Let \( M_2, M'_2 \subset \mathbb{R}^D \) be \( d \)-dimensional submanifolds as in Figure 2; 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 [9], centered at \( x \). Outside \( \mathcal{B}(x, \delta) \), \( M_2, M'_2 \) 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 \( P_k^{(x)} \) according to Proposition 4. From Le Cam’s Lemma 15, we deduce

\[
\inf \sup_{\overline{T}, \delta} \mathbb{E}_{P^{\perp n}} \| II_x^M \circ \pi_{T_x M} - \overline{T} \|_{op} \geq \frac{1}{2} \| II_x^{M_2} \circ \pi_{T_x M_2} - II_x^{M'_2} \circ \pi_{T_x M'_2} \|_{op} (1 - TV(P_2, P'_2))^n.
\]

By construction, \( \| II_x^{M_2} \circ \pi_{T_x M_2} \|_{op} = 0 \), and since \( M'_2 \) is a part of a sphere of radius \( 2/L_\perp \) nearby \( x \), \( \| II_x^{M'_2} \circ \pi_{T_x M'_2} \|_{op} = L_\perp/2 \). Hence,

\[
\| II_x^{M_2} \circ \pi_{T_x M_2} - II_x^{M'_2} \circ \pi_{T_x M'_2} \|_{op} \geq L_\perp/2.
\]

Moreover, since \( P_2 \) and \( P'_2 \) coincide on \( \mathbb{R}^D \setminus \mathcal{B}(x, \delta) \),

\[
TV(P_2, P'_2) = P_{\mathbb{R}^2}(\mathcal{B}(x, \delta)) \leq c_{d, L_\perp, \delta} d^d.
\]

Letting \( \Lambda, \delta \) go to 0 with \( \Lambda^k/\delta \) small enough, we have the desired result.
References


